Church-Rosser property and homology of monoids *
(revised version)

Yves LAFONT †
Laboratoire d'Informatique
Ecole Normale Supérieure
(URA 1327 du CNRS)

Alain PROUTÉ ‡
UFR de Mathématiques
Université de Paris 7
(URA 212 du CNRS)

May 13, 1991

Abstract

We present a result of C. C. Squier relating two topics:

- the canonical rewriting systems, which have been widely studied by computer scientists as a tool for solving word problems,
- the homology of groups (more generally of monoids), which belongs to the core of pure mathematics, on the borders of algebra and topology.

It is natural to ask whether finite canonical systems are the only way to solve word problems in the case of finitely presented monoids [Jantzen 1985]. Squier proved that, if a monoid M is presented by a finite canonical rewriting system, then the homology group $H_3(M)$ is finitely generated [Squier 1987]. Therefore, he could exhibit a finitely presented monoid with a decidable word problem, but which cannot be presented by a finite canonical system, because its H_3 is too big. So he answered negatively to Jantzen's question.

Our aim is to make this result accessible to computer scientists who may be unfamiliar with homological algebra. In particular we manage to be self-contained, assuming only some elementary algebra, we simplify Squier's original proof and we stress the underlying geometrical intuition.

We thank F. Otto who pointed out to us the recent article of Y. Kobayashi, and the anonymous referee for his helpful suggestions.

*reprint of *Mathematical Structures in Computer Science*, vol 1-3, 1991.

†New address: CNRS - Laboratoire de mathématiques discrètes, 163 avenue de Luminy - Case 930, 13288 Marseille Cedex 9, France. Email: lafont@lmd.univ-mrs.fr

‡Address: 2 place Jussieu, 75251 Paris Cedex 05, France. Email: AP@FRMAP711.BITNET

1 Rewriting

Here we introduce some terminology about rewriting systems.

1.1 Presentations of monoids

A *presentation* of a monoid M consists of

- a set Σ of *generators*,
- a set $\mathcal{R}$ of *relations* $r \approx s$ where r and s are words in the free monoid Σ^*

such that M is (isomorphic to) the quotient of Σ^* by the congruence associated with $\mathcal{R}$. If Σ is finite, we say that M is *finitely generated*, and if $\mathcal{R}$ is finite, M is *finitely related.* Finally, if both Σ and $\mathcal{R}$ are finite, we say that M is *finitely presented.*

Multiplication in Σ^* is concatenation of words, with unit 1 (the empty word). If w is a word, we write $\overline{w}$ for its *congruence class* in M.

A monoid M can always be presented by means of its *standard presentation*:

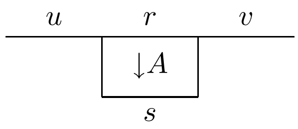
- Σ contains one generator α_u for each $u \in M - \{1\}$,
- $\mathcal{R}$ contains all relations $\alpha_u\alpha_v \approx \alpha_w$ (if $uv = w \neq 1$) and $\alpha_u\alpha_v \approx 1$ (if $uv = 1$).

But of course, only *finite* monoids have a finite standard presentation.

1.2 Rewriting systems

A *rewriting system* is a presentation $(\Sigma, \mathcal{R})$ where each relation is explicitly oriented and considered as a *rewrite rule.* We shall give names $A, B, C, \ldots$ to those rules and write $A : r \to s$. We assume that the left member r is never empty.

A word w is called *reducible* if it is of the form urv where u, v are (possibly empty) words and r is the left member of some rule $A : r \to s$. Then, we say that urv *reduces in one step* to usv and we write $uAv : urv \to usv$ or $urv \overset{uAv}{\to} usv$. The expression uAv is called an *elementary reduction*, and can be pictured as follows:



If no rule applies, w is called *irreducible.* For example, the empty word is irreducible, since left members of rules are nonempty.

A *reduction path* is a finite sequence W of elementary reductions:

$$u_0 \overset{F_1}{\to} u_1 \overset{F_2}{\to} \cdots \overset{F_n}{\to} u_n.$$

In this case, we say that u_0 *reduces to* u_n and we write $W : u_0 \to^* u_n$. The *concatenation* of two reduction paths U and V is written $U * V$. In particular, the above path is $F_1 * \cdots * F_n$.

If W is a reduction path and u, v are words, uWv is defined in the obvious way:

$$u(F_1 * \cdots * F_n)v = uF_1v * \cdots * uF_nv.$$

1.3 Noetherianity

$(\Sigma, \mathcal{R})$ is a *noetherian system* if there is no infinite sequence

$$u_0 \to u_1 \to \cdots \to u_n \to \cdots$$

In that case, it is possible to reason by *noetherian induction*: a property Φ on words is proved by checking that for any u, if Φ holds for any v such that u reduces in at least one step to v, then Φ holds on u.

A *strategy* consists in choosing an elementary reduction $\Gamma(w) : w \to w'$ for each reducible word w. If the system is noetherian, then by iterating a strategy we get a reduction path

$$\Lambda(w) : w \stackrel{\Gamma(w)}{\to} w' \stackrel{\Gamma(w')}{\to} \cdots \to \widehat{w}$$

where the word $\widehat{w}$ is irreducible. In general $\widehat{w}$ depends on the strategy, but it is in the same congruence class as w.

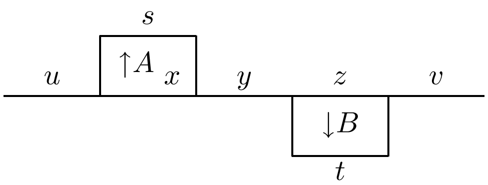
1.4 Confluence

Let $F : t \to u$ and $G : t \to v$ be elementary reductions of a same word. We say that the pair $\{F, G\}$ is *confluent* if it is possible to pursue F and G with *residual paths*[1] $G\backslash F$ and $F\backslash G$ ending up in a same word:

$$\begin{tikzcd}
 & t \arrow[dl, "F"'] \arrow[dr, "G"] & \\
u \arrow[dr, "G\backslash F"', "*"] & & v \arrow[dl, "F\backslash G", "*"'] \\
 & w &
\end{tikzcd}$$

In particular, $\{F, F\}$ is always confluent: take the empty path for $F\backslash F$.

We say that two elementary reductions F and G of a same word are *orthogonal* and we write $F \perp G$ if they reduce disjoint subwords, namely F is of the form $uAyzv$ with $A : x \to s$ and G is of the form $uxyBv$ with $B : z \to t$ (or vice versa):



In particular F does not modify z, which can still be reduced, and similarly for G and x. So $\{F, G\}$ is confluent:

[1]This notion comes from [Huet-Levy 1979].

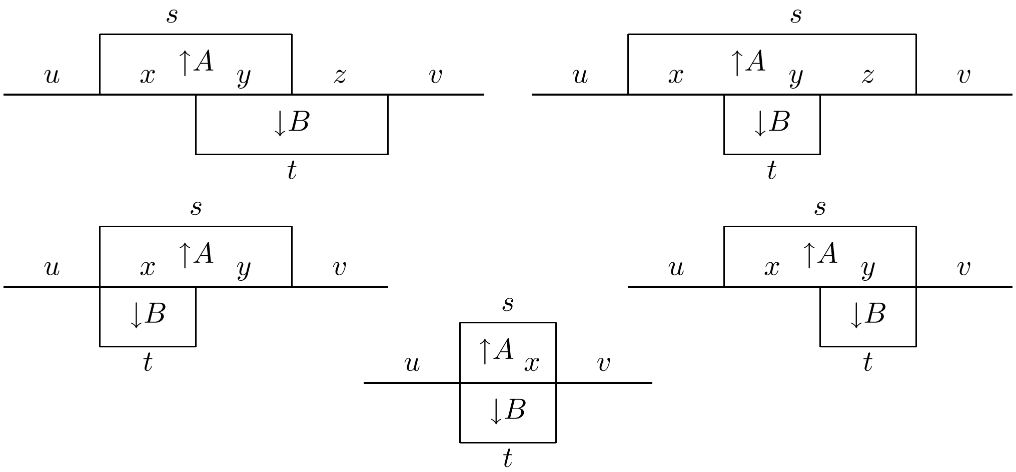
$$\begin{tikzcd}[column sep=small]
 & uxyzv \arrow[dl, "uAyzv"'] \arrow[dr, "uxyBv"] & \\
usyzv \arrow[dr, "usyBv"'] & & uxytv \arrow[dl, "uAytv"] \\
 & usytv &
\end{tikzcd}$$

Here the residual paths happen to be elementary:

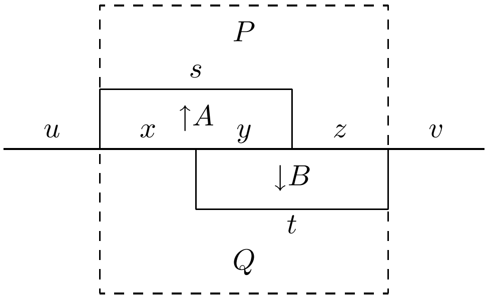
$$G\backslash F = usyBv, \qquad F\backslash G = uAytv.$$

1.5 Critical pairs

Otherwise, when F and G are distinct and reduce *nondisjoint* subwords, the following configurations are possible (up to exchange of A and B):



In all those cases, we write $F = uPv$ and $G = uQv$ where u and v are maximal. For example, in the first case $P = Az$ and $Q = xB$:



The pair $\{P, Q\}$ consists of elementary reductions of the same word, and is called *critical pair*. Note that if $\mathcal{R}$ is finite, there is only a finite number of critical pairs, which are obtained by a simple pattern matching algorithm. Furthermore, if $\{P, Q\}$ is confluent, so is $\{F, G\}$ with

$$G\backslash F = u(Q\backslash P)v, \qquad F\backslash G = u(P\backslash Q)v.$$

The residual paths $Q\backslash P$ and $P\backslash Q$ need not be unique, but we assume that some choice is made once for all. Therefore, in all cases where $F\backslash G$ is defined, we have the following identity:

$$uFv\backslash uGv = u(F\backslash G)v.$$

1.6 Canonical systems

Theorem 0 *(Knuth-Bendix)*

Let $(\Sigma, \mathcal{R})$ be a noetherian system and assume that some strategy has been fixed. The following statements are equivalent:

i) all critical pairs are confluent,

ii) all pairs of elementary reductions of the same word are confluent (local confluence),

iii) if $u \to v$, then $\widehat{u} = \widehat{v}$,

iv) if $\overline{u} = \overline{v}$ then $\widehat{u} = \widehat{v}$,

v) if $\overline{u} = \overline{v}$ then there is w such that $u \overset{*}{\to} w$ and $v \overset{*}{\to} w$ (Church-Rosser property).

Proof: We just showed that $(i) \Rightarrow (ii)$ and it is easy to check that $(iii) \Rightarrow (iv) \Rightarrow (v) \Rightarrow (i)$. Finally, we assume (ii) and we prove (iii) by noetherian induction on u.

Let $F : u \to v$ be an elementary reduction. Since u is a reducible word, the strategy gives $\Gamma(u) : u \to u'$, and by definition $\widehat{u} = \widehat{u'}$. By (ii) we get a confluence diagram

$$\begin{tikzcd} & u \arrow[dl] \arrow[dr] & \\ v \arrow[dr, "*"'] & & u' \arrow[dl, "*"] \\ & v' & \end{tikzcd}$$

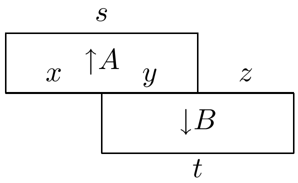
Therefore, we have a reduction sequence $v = v_0 \to v_1 \to \cdots \to v_n = v'$, and the induction hypothesis on $v_0, v_1, \ldots, v_{n-1}$ gives $\widehat{v} = \widehat{v_0} = \widehat{v_1} = \cdots = \widehat{v_n} = \widehat{v'}$. Similarly, we see that $\widehat{u'} = \widehat{v'}$, so that $\widehat{u} = \widehat{u'} = \widehat{v'} = \widehat{v}$. $Q.e.d.Q.e.d.$

A noetherian system which satisfies the Church-Rosser property is called a *canonical system.* By (iv), the *canonical form* $\widehat{u}$ does not depend on the strategy: it is the only irreducible word in the class of u. In case of a *finite* canonical system, it is always possible to compute canonical forms. So, given u and v, we can decide whether $\overline{u} = \overline{v}$ (the *word problem* for the finitely presented monoid M) by comparing the canonical forms $\widehat{u}$ and $\widehat{v}$. By the way, we can also decide if a finite noetherian system is canonical by testing the confluence of all critical pairs.

1.7 Minimal systems

A canonical system is *minimal* if for each rule $A : r \to s$, the left member r is only reducible by A. Otherwise we can always remove A from $\mathcal{R}$: we get exactly the same irreducible words, the same congruence classes, and the system is still canonical. So, from a finite canonical system, we extract a minimal system presenting the same monoid M by successively eliminating superfluous rules.

In a minimal system, critical pairs are necessarily of the following form:



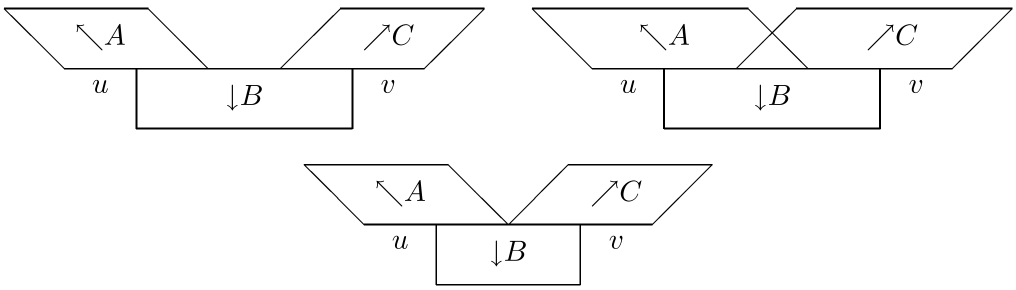
So we have a clear notion of the *leftmost reduction* of a reducible word, and a *leftmost strategy* which is characterised by the following property:

$$\Lambda(uv) = \Lambda(u)v * \Lambda(\widehat{u}v).$$

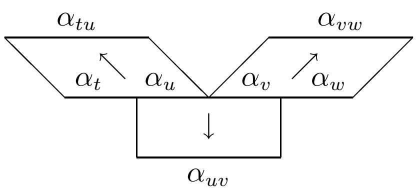
In other words, reducing uv to its canonical form by the leftmost strategy amounts to first reducing the prefix u without changing v and then reducing $\widehat{u}v$.

1.8 Critical triples

We shall need the ternary analogue of critical pairs. In the case of a minimal system, a *critical triple* consists of three elementary reductions Pv, uBv and uQ such that $\{P, uB\}$ and $\{Bv, Q\}$ are critical pairs. The following configurations are possible:



For example the *standard presentation* of a monoid defines a minimal canonical system, with a lot of generators, rules, critical pairs and critical triples. Those critical triples are of the third kind:



2 Discussion

Knuth-Bendix theorem suggests the following method for solving the word problem of a monoid given by a finite presentation $(\Sigma, \mathcal{R})$:

1. Try to orient the relations so that $(\Sigma, \mathcal{R})$ becomes a noetherian system.

2. Let $\{P : t \to u,\ Q : t \to v\}$ be a critical pair: if $\widehat{u} = \widehat{v}$ then it is confluent, and if all critical pairs are confluent the system is canonical.

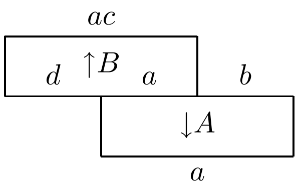
3. If some critical pair is not confluent, it means that $\widehat{u} \neq \widehat{v}$. But since $\overline{u} = \overline{t} = \overline{v}$, the relation $\widehat{u} \approx \widehat{v}$ can be added without changing the monoid. Try to orient this relation so that $(\Sigma, \mathcal{R})$ remains noetherian and keep on with the others nonconfluent critical pairs, if any.

Two difficulties arise. Firstly, it is not possible to decide in general if a system is noetherian, although there are good heuristics. Secondly, this *completion algorithm* may loop.

Consider the following example:

$$\Sigma = \{a, b, c, d\}, \qquad \mathcal{R} = \{A : ab \to a,\ B : da \to ac\}.$$

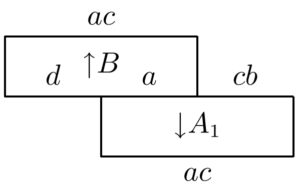
This system is noetherian because the total number of occurrences of b and d decreases at each step, but there is one critical pair



which is not confluent. Indeed, acb is irreducible, whereas da reduces only to ac which is irreducible. Therefore we add the rule

$$A_1 : acb \to ac,$$

so that the critical pair becomes confluent. The system is still noetherian but we get a new critical pair



which is not confluent. Again we add a rule

$$A_2 : accb \to acc,$$

which creates a new critical pair, and so on. By iterating this process we get in fact an *infinite* canonical system:

$$\Sigma = \{a, b, c, d\}, \qquad \mathcal{R} = \{A_n : ac^n b \to ac^n;\ n \in N\} \cup \{B : da \to ac\},$$

Here, the set $\mathcal{R}$ is given in such a way that canonical forms are obviously computable, and the word problem is decidable. But maybe it is possible to find a *finite* canonical system for the same monoid.

Kapur and Narendran considered the presentation consisting of two generators a, b and one relation $aba \approx bab$. Since aba and bab have the same length, the congruence class of a given word is computed in finite time, so that the word problem is decidable. By a simple combinatorial argument [Kapur-Narendran 1985], there is no finite equivalent canonical system with the *same* generators a, b. Therefore the process of completion will inevitably loop, whatever choice you make when you orient the relations. But introducing an extra generator c as an abbreviation for ab, you get an equivalent system

$$\Sigma = \{a, b, c\}, \qquad \mathcal{R} = \{ab \to c,\ ca \to bc\}$$

for which the completion succeeds (if you make the good choices) producing a finite canonical system with only two extra rules:

$$bcb \to cc, \qquad ccb \to acc.$$

So it may be useful to enrich a given presentation with new (definable) generators to get a terminating completion. But is it enough?

From a universal Turing machine, it is easy to build a (rather complicated) finite presentation $(\Sigma, \mathcal{R})$ for which the word problem is undecidable. Of course, it is useless to look for a finite canonical system in that case! But in fact, there are also finitely presented monoids which cannot be presented by a finite canonical system although they have a decidable word problem: we shall exhibit a simple one in the next section.[2] The hard point consists in proving a negative result, namely that there is no finite canonical system for a certain monoid. This is the place where homological algebra is needed.

3 Homology of rewriting

Here we define the homological invariants of a minimal canonical system and we use them to build counterexamples.

3.1 Z-modules

If I is a set, we write $Z[I]$ for the *free Z-module* (= abelian group) generated by I. It can be described as the set of formal sum $\sum_{i \in I} n_i[i]$ where the coefficients n_i are in Z and only a finite number of them are $\neq 0$. The $[i]$ form a *basis* of $Z[I]$. Addition and opposite are defined in the obvious way:

$$\sum_{i \in I} p_i[i] + \sum_{i \in I} q_i[i] = \sum_{i \in I} (p_i + q_i)[i], \qquad -\sum_{i \in I} n_i[i] = \sum_{i \in I} (-n_i)[i].$$

In particular, if I is finite with cardinality n, then $Z[I]$ is (isomorphic to) Z^n.

A Z-module is *finitely generated* if it is of the form Z^n/R for some submodule $R \subset Z^n$. It is clear that $Z[I]$ is finitely generated if and only if I is finite. We shall also use the following result (see for example [Lang 1984]):

Theorem 1

Every sub-Z-module of Z^n is isomorphic to Z^p for some p such that $p \leq n$.

3.2 Basic construction

Now, given a minimal canonical system $(\Sigma, \mathcal{R})$, one constructs a sequence

$$C_4 \xrightarrow{\partial_4} C_3 \xrightarrow{\partial_3} C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0$$

where the C_i are free Z-modules and the ∂_i are Z-linear maps satisfying $\partial_i \partial_{i+1} = 0$. More precisely,

$$C_4 = Z[\mathcal{T}], \qquad C_3 = Z[\mathcal{P}], \qquad C_2 = Z[\mathcal{R}], \qquad C_1 = Z[\Sigma], \qquad C_0 = Z,$$

[2] Our counterexample is a simplification of the one proposed in [Squier 1987].

where $\mathcal{P}$ (resp. $\mathcal{T}$) is the set of critical pairs (resp. critical triples). Note that if the system is finite, those Z-modules are finitely generated.

We extend the bracket notation to words and paths:

- if $w = a_1 \dots a_n$ is a word, let $[w] = [a_1] + \cdots + [a_n] \in Z[\Sigma]$, in particular $[1] = 0$,
- if $F = uAv$ is an elementary reduction, let $[F] = [A] \in Z[\mathcal{R}]$,
- if $W = F_1 * \cdots * F_n$ is a reduction path, let $[W] = [F_1] + \cdots + [F_n] \in Z[\mathcal{R}]$.

If you prefer, $[w] = \sum_{a \in \Sigma} w_a[a]$ where w_a is the number of occurrences of the letter a in w, and $[W] = \sum_{A \in \mathcal{R}} W_A[A]$ where W_A is the number of occurrences of the rule A in W.

The ∂_i are given by their values on the basis:

- $\partial_1[a] = 0$ (and so $\partial_1 = 0$),
- $\partial_2[A : r \to s] = [r] - [s]$,
- $\partial_3[P, Q] = [P] + [Q \backslash P] - [Q] - [P \backslash Q]$.

The latter expression is obtained by orienting the confluence diagram

$$\begin{tikzcd}
& t \arrow[dl, "P"'] \arrow[dr, "Q"] \arrow[dd, phantom, "\mathrm{x}"] & \\
u \arrow[dr, "*", "Q\backslash P"'] & & v \arrow[dl, "*"', "P\backslash Q"] \\
& w &
\end{tikzcd}$$

and counting each path positively or negatively, according to this orientation. As for ∂_4, it is not so simple to describe: anyways, we shall not use its explicit definition, but only the fact that $\partial_3 \partial_4 = 0$.

Note that if $W : u \xrightarrow{*} v$ is a reduction path, then $\partial_2[W] = [u] - [v]$ (by induction on W). Applying this identity to our confluence diagram, we get

$$\partial_2 \partial_3[P, Q] = [t] - [u] + [u] - [w] - [t] + [v] - [v] + [w] = 0,$$

and so $\partial_2 \partial_3 = 0$ as promised. We have also $\partial_1 \partial_2 = 0$ because $\partial_1 = 0$.

3.3 Homological invariants

Since $\partial_i \partial_{i+1} = 0$, *i.e.* $\operatorname{im} \partial_{i+1} \subset \ker \partial_i$, we can introduce the Z-modules

$$H_i(\Sigma, \mathcal{R}) = \ker \partial_i / \operatorname{im} \partial_{i+1},$$

for $i = 1, 2, 3$. Their interest is all in the following fact:

$H_i(\Sigma, \mathcal{R})$ does not depend on the presentation, but only on the monoid M.

$H_i(\Sigma,\mathcal{R})$ is in fact the *homology* $H_i(M)$ of the monoid M, which is defined in section 4 for all $i \in N$, without using any canonical system for M.

If $(\Sigma,\mathcal{R})$ is a finite system, the $H_i(\Sigma,\mathcal{R})$ are finitely generated by theorem 1. For $i=1$ and 2, this is not astounding, since it follows from a classical theorem of homological algebra (which does not use canonical systems):

Theorem 2

i) If M is finitely generated (as a monoid) then $H_1(M)$ is finitely generated.

ii) If M is finitely related then $H_2(M)$ is finitely generated.

But the interesting case is $i=3$:

Theorem 3 *(Squier)*

If M is presented by a finite canonical system, then $H_3(M)$ is finitely generated.

All those results are proved in section 5.

3.4 Preliminary counterexample

Here we construct *a finitely generated monoid which cannot be finitely presented.*

For that purpose it is enough to find a canonical system with a finite Σ, an infinite $\mathcal{R}$ and an empty $\mathcal{P}$. Indeed, the sequence associated with such a canonical system is of the form

$$0 \xrightarrow{0} 0 \xrightarrow{0} Z[\mathcal{R}] \xrightarrow{\partial_2} Z[\Sigma] \xrightarrow{0} Z.$$

But the dimensions of $Z[\mathcal{R}]$ and $Z[\Sigma]$ are respectively infinite and finite. So the dimension of $H_2(M) = \ker \partial_2$ is infinite, and by theorem 2 the monoid M cannot be finitely presented.

Take for example

$$\Sigma = \{a,b,c\}, \qquad \mathcal{R} = \{A_n : ac^n b \to 1;\ n \in N\}.$$

The role of a and b on each side of c^n is to prevent the creation of critical pairs. This system is obviously noetherian, and furthermore $\mathcal{R}$ is given in such a way that the word problem is clearly decidable. By definition

$$\partial_2[A_n] = [ac^n b] - [1] = [a] + n[c] + [b].$$

So we have $H_1(M) = Z[\Sigma]/\operatorname{im}\partial_2 \simeq Z$ and $H_2(M) = \ker\partial_2$ is the free Z-module generated by the infinite basis

$$\xi_n = [A_n] - n[A_1] + (n-1)[A_0], \qquad n \geq 2.$$

We invite the reader to build a similar example with two generators instead of three, and to prove that, on the other hand, any monoid with a single generator has a finite presentation.

3.5 Main counterexample

Here we construct *a finitely presented monoid which cannot be presented by a finite canonical system*. By theorem 3, it is enough to find a finitely presented M such that $H_3(M)$ has infinite dimension.

Consider the example of section 2 with the following two relations:

$$ab \approx a, \qquad da \approx ac.$$

We have seen that M can be presented by an infinite canonical system, namely

$$\Sigma = \{a, b, c, d\}, \qquad \mathcal{R} = \{A_n : ac^n b \to ac^n;\ n \in N\} \cup \{B : da \to ac\},$$

with an infinity of confluent critical pairs:

$$\begin{tikzcd}
& dac^n b \arrow[dl, "Bc^n b"'] \arrow[dr, "dA_n"] & \\
ac^{n+1}b \arrow[dr, "A_{n+1}"'] & & dac^n \arrow[dl, "Bc^n"] \\
& ac^{n+1} &
\end{tikzcd}$$

This system has no critical triple, so $C_4 = 0$ and the sequence is of the form

$$0 \xrightarrow{0} Z[\mathcal{P}] \xrightarrow{\partial_3} Z[\mathcal{R}] \xrightarrow{\partial_2} Z[\Sigma] \xrightarrow{0} Z.$$

where

$$\partial_2[A_n] = [ac^n b] - [ac^n] = [b], \qquad \partial_2[B] = [da] - [ac] = [d] - [c],$$
$$\partial_3[Bc^n b, dA_n] = [Bc^n b] + [A_{n+1}] - [dA_n] - [Bc^n] = [A_{n+1}] - [A_n].$$

So we get $H_1(M) = Z[\Sigma]/\operatorname{im}\partial_2 \simeq Z^2$, $H_2(M) = \ker\partial_2/\operatorname{im}\partial_3 \simeq 0$, $H_3(M) = \ker\partial_3 \simeq 0$ … *and we have lost*!

In fact, our infinity of critical pairs is "compensated" by an infinity of rules. Moreover, looking back to our initial presentation, we discover that the second equation could have been oriented in the opposite way, yielding a finite canonical system for M, namely

$$\Sigma = \{a, b, c, d\}, \qquad \mathcal{R} = \{A : ab \to a,\ B : ac \to da\},$$

with no critical pair!

3.6 Dénouement

But wait a moment: if we add a generator d' and a relation $d'a \approx ac$, we obtain a new monoid M' which can be presented by an infinite canonical system

$$\Sigma' = \{a, b, c, d, d'\}, \qquad \mathcal{R}' = \{A_n : ac^n b \to ac^n;\ n \in N\} \cup \{B : da \to ac,\ B' : d'a \to ac\},$$

with "two infinities" of confluent critical pairs:

$$\begin{tikzcd}
& dac^n b \arrow[dl, "Bc^n b"'] \arrow[dr, "dA_n"] & \\
ac^{n+1}b \arrow[dr, "A_{n+1}"'] & & dac^n \arrow[dl, "Bc^n"] \\
& ac^{n+1} &
\end{tikzcd}
\qquad
\begin{tikzcd}
& d'ac^n b \arrow[dl, "B'c^n b"'] \arrow[dr, "d'A_n"] & \\
ac^{n+1}b \arrow[dr, "A_{n+1}"'] & & d'ac^n \arrow[dl, "B'c^n"] \\
& ac^{n+1} &
\end{tikzcd}$$

The sequence is similar to the previous case, but now $H_3(M') = \ker \partial_3$ is the free Z-module generated by the infinite basis

$$\xi_n = [Bc^n b, dA_n] - [B'c^n b, d'A_n], \qquad n \in N.$$

By Squier's theorem, M' cannot be presented by a finite canonical system. Yet it has a decidable word problem (canonical forms are computable) and a finite presentation consisting of the five generators a, b, c, d, d' and the following three relations:

$$ab \approx a, \qquad da \approx ac, \qquad d'a \approx ac.$$

Of course, the trick which consisted in reversing B does not work here: it would create nasty critical pairs.

4 Homology of monoids

Here we introduce the notions of homological algebra needed in this article. For a larger panorama we refer to [MacLane 1963, Spanier 1966, Brown 1982].

4.1 ZM-modules

Let M be an arbitrary monoid. We write ZM for the *ring*[3] of M, that is the free Z-module $Z[M]$ with a bilinear multiplication extending the multiplication of M:

$$(\sum_{u\in M} p_u u)(\sum_{v\in M} q_v v) = \sum_{w\in M} n_w w \qquad \text{where } n_w = \sum_{uv=w} p_u q_v$$

(we omit square brackets here) and with the same unit as M. For example, if M is N with its additive structure, ZM is the ring of polynomials with coefficients in Z.

A *left linear action* of the monoid M on a Z-module C is a map $u, x \mapsto ux$ from $M \times C$ to C satisfying:

$$u0 = 0 \qquad u(x+y) = ux + uy \qquad 1x = x \qquad (uv)x = u(vx)$$

This action extends uniquely to a bilinear multiplication $ZM \times C \to C$ so that C is a *left ZM-module* and conversely, any structure of left ZM-module restricts to a left linear action of M on C. For example, the *trivial left action* $ux = x$ defines a structure of left ZM-module on any Z-module.

If I is a set, the *free left ZM-module* $ZM[I]$ is defined in the same way as $Z[I]$, but with coefficients in ZM instead of Z. Clearly, it can be seen as the set of formal sums

$$\sum_{(u,i)\in M\times I} n_{u,i} u[i]$$

where the coefficients $n_{u,i}$ are in Z and only a finite number of them are $\neq 0$. Note that:

- the $[i]$ form a basis of $ZM[I]$ considered as ZM-module,
- the $u[i]$ form a basis of $ZM[I]$ considered as Z-module.

[3] In [Lang 1984], ZM is called the "monoid algebra" of M over Z.

If C is a ZM-module, we consider the Z-module $\widetilde{C}$ obtained by trivialising the action of M, *i.e.* C quotiented by all relations $ux \approx x$ for $u \in M$ and $x \in C$.[4] This *trivialisation* defines a functor: if $\varphi : C \to C'$ is a ZM-linear map, then $\widetilde{\varphi} : \widetilde{C} \to \widetilde{C'}$ is the unique Z-linear map sending the class of x to the class of $\varphi(x)$.

In fact we shall only use this construction in the case of free ZM-modules: if $C = ZM[I]$ then $\widetilde{C}$ is (isomorphic to) $Z[I]$, and the class of $u[i] \in ZM[I]$ is simply $[i] \in Z[I]$.

4.2 Free resolutions

We call *free resolution* of Z by left ZM-modules an infinite sequence

$$\cdots \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_n} \cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\varepsilon} Z \longrightarrow 0,$$

where

- the C_n $(n \in N)$ are free left ZM-modules,
- Z is seen as a trivial left ZM-module, *i.e.* $un = n$ for all $u \in M$ and $n \in Z$,
- the ∂_n and ε are ZM-linear maps,
- they form an *exact sequence*, *i.e.* the image of each arrow is the kernel of the following one:

$$\operatorname{im}\partial_{n+1} = \ker\partial_n, \qquad \operatorname{im}\partial_1 = \ker\varepsilon, \qquad \operatorname{im}\varepsilon = 0 \ (\textit{i.e. } \varepsilon \text{ is surjective}),$$

  in particular $\partial_n\partial_{n+1} = 0$ and $\varepsilon\partial_1 = 0$.

Such a free resolution exists anyway. Take indeed $C_0 = ZM$ and let $\varepsilon : ZM \to Z$ be the unique ZM-linear map sending the unit of ZM to the unit of Z. This map is clearly surjective and is given by the following formula:

$$\varepsilon(\sum_{u\in M} n_u u) = \sum_{u\in M} n_u.$$

Take then for C_1 the free left ZM-module generated by the elements of $\ker\varepsilon$ and for $\partial_1 : C_1 \to C_0$ the associated canonical map. You can now take for C_2 the free left ZM-module generated by the elements of $\ker\partial_1$, and so on. Of course this resolution is monstrous, and in particular, useless for calculation.[5]

By the same argument, a *partial free resolution*

$$C_k \xrightarrow{\partial_k} \cdots \xrightarrow{\partial_1} C_0 \xrightarrow{\varepsilon} Z \longrightarrow 0$$

can always be extended to a full one. Furthermore, in the case where ∂_k is injective, it can be extended with zeros:

$$\cdots \longrightarrow 0 \longrightarrow \cdots \longrightarrow 0 \longrightarrow C_k \xrightarrow{\partial_k} \cdots \xrightarrow{\partial_1} C_0 \xrightarrow{\varepsilon} Z \longrightarrow 0.$$

[4] $\widetilde{C}$ is usually defined as the tensor product $Z \otimes_{ZM} C$ where Z is considered as a trivial *right* ZM-module. But there is no need for a tensor product here.

[5] A better solution is the *bar resolution* (appendix B) which in many cases is still too big for calculation, but at least is functorial.

4.3 A fundamental lemma

Consider two free resolutions of Z by left ZM-modules:

$$\cdots \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_n} \cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\varepsilon} Z \longrightarrow 0,$$

$$\cdots \xrightarrow{\partial'_{n+1}} C'_n \xrightarrow{\partial'_n} \cdots \xrightarrow{\partial'_2} C'_1 \xrightarrow{\partial'_1} C'_0 \xrightarrow{\varepsilon'} Z \longrightarrow 0.$$

A *morphism* of resolutions (from the first one to the second one) is a family of ZM-linear maps f_n making the following diagram commutative:

$$\begin{tikzcd}
\cdots \arrow[r, "\partial_{n+1}"] & C_n \arrow[r, "\partial_n"] \arrow[d, "f_n"] & \cdots \arrow[r, "\partial_2"] & C_1 \arrow[r, "\partial_1"] \arrow[d, "f_1"] & C_0 \arrow[r, "\varepsilon"] \arrow[d, "f_0"] & Z \arrow[r] \arrow[d, "\mathrm{id}"] & 0 \\
\cdots \arrow[r, "\partial'_{n+1}"] & C'_n \arrow[r, "\partial'_n"] & \cdots \arrow[r, "\partial'_2"] & C'_1 \arrow[r, "\partial'_1"] & C'_0 \arrow[r, "\varepsilon'"] & Z \arrow[r] & 0
\end{tikzcd}$$

Now, if f and g are two such morphisms, we say that f is *homotopic* to g if there is a family of ZM-linear maps h_n pictured in the following diagram

$$\begin{tikzcd}
\cdots \arrow[r, "\partial_{n+1}"] & C_n \arrow[r, "\partial_n"] \arrow[dl, "h_n"'] \arrow[d, shift right, "g_n"'] \arrow[d, shift left, "f_n"] & \cdots \arrow[r, "\partial_2"] \arrow[dl, "h_{n-1}"] & C_1 \arrow[r, "\partial_1"] \arrow[dl, "h_1"'] \arrow[d, shift right, "g_1"'] \arrow[d, shift left, "f_1"] & C_0 \arrow[r, "\varepsilon"] \arrow[dl, "h_0"'] \arrow[d, shift right, "g_0"'] \arrow[d, shift left, "f_0"] & Z \arrow[r] \arrow[d, "\mathrm{id}"] & 0 \\
\cdots \arrow[r, "\partial'_{n+1}"] & C'_n \arrow[r, "\partial'_n"] & \cdots \arrow[r, "\partial'_2"] & C'_1 \arrow[r, "\partial'_1"] & C'_0 \arrow[r, "\varepsilon'"] & Z \arrow[r] & 0
\end{tikzcd}$$

such that $f_0 - g_0 = \partial'_1 h_0$ and $f_n - g_n = \partial'_{n+1} h_n + h_{n-1}\partial_n$ for $n \geq 1$.

Lemma 0 *Between two free resolutions, there is always a morphism, and two such morphisms are homotopic.*

Proof: We construct f_0 using the fact that ε' is surjective and C_0 is a free ZM-module. Then, assuming that f_n is constructed, we must check that for every x in the basis of C_{n+1}, the element $f_n(\partial_{n+1}(x))$ is in $\operatorname{im}\partial'_{n+1} = \ker\partial'_n$. But this follows immediately from the fact that $\partial_n\partial_{n+1} = 0$ and $\partial'_n f_n = f_{n-1}\partial_n$. The construction of a homotopy between two given morphisms is similar. *Q.e.d.*

4.4 Homology

Let us now truncate a resolution by removing what follows C_0 and trivialise it:

$$\cdots \xrightarrow{\widetilde{\partial_{n+1}}} \widetilde{C_n} \xrightarrow{\widetilde{\partial_n}} \cdots \xrightarrow{\widetilde{\partial_2}} \widetilde{C_1} \xrightarrow{\widetilde{\partial_1}} \widetilde{C_0}.$$

This new sequence consists of free Z-modules (the structure of ZM-module has been killed by trivialisation) and is no more an exact sequence, but just a *chain complex*, which means that the image of each arrow is only *included* in the kernel of the following one. Indeed, from $\partial_n\partial_{n+1} = 0$ we deduce $\widetilde{\partial_n}\widetilde{\partial_{n+1}} = 0$. So it makes sense to consider the sequence of Z-modules

$$H_n(M) = \ker\widetilde{\partial_n}/\operatorname{im}\widetilde{\partial_{n+1}},$$

which is called the *homology* of M.[6] This appellation and this notation is justified by the following fact:

Theorem 0

$H_n(M)$ does not depend on a particular choice of a free resolution, but only on M.

Proof: First notice that a morphism of resolutions induces a diagram

$$\begin{tikzcd} \cdots \arrow[r, "\widetilde{\partial_{n+1}}"] & \widetilde{C_n} \arrow[r, "\widetilde{\partial_n}"] \arrow[d, "\widetilde{f_n}"] & \cdots \arrow[r, "\widetilde{\partial_2}"] & \widetilde{C_1} \arrow[r, "\widetilde{\partial_1}"] \arrow[d, "\widetilde{f_1}"] & \widetilde{C_0} \arrow[d, "\widetilde{f_0}"] \\ \cdots \arrow[r, "\widetilde{\partial'_{n+1}}"] & \widetilde{C'_n} \arrow[r, "\widetilde{\partial'_n}"] & \cdots \arrow[r, "\widetilde{\partial'_2}"] & \widetilde{C'_1} \arrow[r, "\widetilde{\partial'_1}"] & \widetilde{C'_0} \end{tikzcd}$$

which itself defines a family of Z-linear maps from $\ker\widetilde{\partial_n}/\mathrm{im}\,\widetilde{\partial_{n+1}}$ to $\ker\widetilde{\partial'_n}/\mathrm{im}\,\widetilde{\partial'_{n+1}}$. Furthermore, if f and g are homotopic, this property of purely equational nature is preserved by trivialisation, so that f and g induce the *same* morphism on the homology. This comes from the fact that if $x \in \ker\widetilde{\partial_n}$ then $\widetilde{f_n}(x) - \widetilde{g_n}(x) = \widetilde{\partial'_{n+1}}(\widetilde{h_n}(x)) \in \mathrm{im}\,\widetilde{\partial'_{n+1}}$.

Now, consider two free resolutions of Z by ZM-modules. By the previous lemma, there is a morphism f from the first resolution to the second one, and g in the opposite direction. By the lemma again, $g \circ f$ and $f \circ g$ are homotopic to the respective identities of the resolutions, and so we get an isomorphism on the homology. Since any two morphisms are homotopic, this isomorphism does not depend on a particular choice of f, so that the homology is well defined, not only its class of isomorphism. *Q.e.d.Q.e.d.*

An element of $\ker\widetilde{\partial_n}$ is called an *n-cycle* and an element of $\mathrm{im}\,\widetilde{\partial_{n+1}}$ is called an *n-boundary*, so that homology is "cycles modulo boundaries".

4.5 Contracting homotopies

In practice, in order to calculate the homology of M, one looks for a free resolution with reasonably small C_n. In most cases, it is easy to establish that such a sequence is a chain complex (*i.e.* $\partial_n\partial_{n+1} = 0$ and $\varepsilon\partial_1 = 0$) but exactness may be more problematic. For that purpose, one exhibits a *contracting homotopy* which consists of a sequence of Z-linear maps

$$\cdots \xleftarrow{s_n} C_n \xleftarrow{s_{n-1}} \cdots \xleftarrow{s_1} C_1 \xleftarrow{s_0} C_0 \xleftarrow{\eta} Z,$$

satisfying the following equations:

$$\partial_{n+1}s_n + s_{n-1}\partial_n = \mathrm{id}_{C_n}, \qquad \partial_1 s_0 + \eta\varepsilon = \mathrm{id}_{C_0}, \qquad \varepsilon\eta = \mathrm{id}_Z.$$

It is clear that a chain complex with such a homotopy is an exact sequence, because each n-cycle is then an n-boundary. Conversely, in the case where the C_n are free as Z-modules, exactness implies the existence of a contracting homotopy, which can be constructed by induction (as in the proof of the lemma).

A contracting homotopy is like a homotopy between id_C and 0, except that it has an extra component η, and the s_n, as well as η, are *not* supposed to be ZM-linear, but only Z-linear: otherwise, they would induce a contracting homotopy on the trivialised complex,

[6] $H_n(M)$ is defined for all $n \in N$, with the convention that $\partial_0 = 0$, but in fact $H_0(M)$ is always Z.

and the homology would be trivial. For example, if $\varepsilon : ZM \to Z$ is the map defined above, one can start by defining

$$\eta : Z \to ZM$$

as the unique Z-linear map sending the unit of Z to the unit of ZM, so that $\varepsilon\eta = \mathrm{id}_Z$. But unless M is the trivial monoid, η is not ZM-linear.

5 A free resolution

Here we investigate the homology of a monoid M presented by a minimal canonical system $(\Sigma, \mathcal{R})$. For that purpose we construct a partial free resolution

$$C_3 \xrightarrow{\partial_3} C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\varepsilon} Z \longrightarrow 0,$$

where $C_3 = ZM[\mathcal{P}]$, $C_2 = ZM[\mathcal{R}]$, $C_1 = ZM[\Sigma]$ and $C_0 = ZM$.

5.1 Filling up 0-cycles

We start with the map ε introduced in section 4.2. By definition, $\varepsilon(\overline{w}) = 1$ for all $\overline{w} \in M$. Since ε is surjective, we have an exact sequence:

$$ZM \xrightarrow{\varepsilon} Z \longrightarrow 0.$$

Now we must find a free ZM-module, as small as possible, which can be mapped *onto* the kernel of ε. Since the latter is clearly the sub-Z-module of ZM generated by the $\overline{w} - \overline{1}$ for $\overline{w} \in M - \{\overline{1}\}$, we could define C_1 as the free ZM-module on $M - \{\overline{1}\}$. The *normalised bar resolution* (appendix B) starts precisely in this way.

However, we know here a system of generators for M, namely the alphabet Σ. As a ZM-module, $\ker\varepsilon$ happens to be generated by the $\overline{a} - \overline{1}$ for $a \in \Sigma$. This is easily proved by induction on the length of words, using the identity

$$\overline{wa} - \overline{1} = \overline{w}(\overline{a} - \overline{1}) + \overline{w} - \overline{1}.$$

So we can take for C_1 the free ZM-module $ZM[\Sigma]$ which in general is much smaller (and may have finite dimension even when the other one has infinite dimension) and define the ZM-linear map

$$\begin{array}{rcl} ZM[\Sigma] & \xrightarrow{\partial_1} & ZM \\ [a] & \mapsto & \overline{a} - \overline{1}, \end{array}$$

so that the following sequence is exact:

$$ZM[\Sigma] \xrightarrow{\partial_1} ZM \xrightarrow{\varepsilon} Z \longrightarrow 0.$$

In particular, if Σ is finite, the dimension of $\widetilde{C_1} = Z[\Sigma]$ is finite, so that $H_1(M)$ is finitely generated: we just proved the first point of theorem 2.

Note also that $\widetilde{\partial_1} = 0$ since $\overline{a}$ and $\overline{1}$ become identical through the process of trivialisation.

5.2 The contracting homotopy

As we have seen, it is quite easy to construct the resolution up to dimension 1. But things become more and more complicated in higher dimensions. So we shall use the technique of the *contracting homotopy* introduced in section 4.5. This means that we shall construct a diagram

$$ZM[\mathcal{P}] \underset{s_2}{\overset{\partial_3}{\rightleftarrows}} ZM[\mathcal{R}] \underset{s_1}{\overset{\partial_2}{\rightleftarrows}} ZM[\Sigma] \underset{s_0}{\overset{\partial_1}{\rightleftarrows}} ZM \underset{\eta}{\overset{\varepsilon}{\rightleftarrows}} Z,$$

such that

- the ∂_n and ε are ZM-linear,
- the s_n and η are only Z-linear,
- the following equations are satisfied:

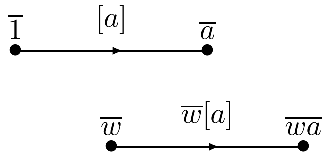
$$\begin{array}{lcll@{\qquad\qquad}lcll}
 & & & & \varepsilon\eta & = & \mathrm{id}_Z & (0) \\
\varepsilon\partial_1 & = & 0 & (1) & \partial_1 s_0 + \eta\varepsilon & = & \mathrm{id}_{ZM} & (2) \\
\partial_1\partial_2 & = & 0 & (3) & \partial_2 s_1 + s_0\partial_1 & = & \mathrm{id}_{ZM[\Sigma]} & (4) \\
\partial_2\partial_3 & = & 0 & (5) & \partial_3 s_2 + s_1\partial_2 & = & \mathrm{id}_{ZM[\mathcal{R}]} & (6)
\end{array}$$

The maps ε and η satisfying (0) have been introduced in sections 4.2 and 4.5. Now we have also defined ∂_1 satisfying (1).

5.3 Geometrical interpretation

Before going on, we shall explain the underlying geometry, and for that purpose, assume that M is a group.[7] From a geometric viewpoint, we are constructing a contractible space X on which M acts freely on the left (see appendix A). It is reasonable to start with one point, since a contractible space is nonempty. Of course, unless M is trivial, it cannot act freely on a single point: we must add at least one point for each element of M distinct from $\overline{1}$, so that our space is in one-to-one correspondence with M. Unfortunately, we have lost contractibility, because the resulting space is no more connected.

To make it connected, we must add edges between points, and those edges must be added is such a way that M acts freely on them. It is enough to connect each generator to $\overline{1}$ by an edge. By making M act on those edges, we get the other ones. For example, if a generator $\overline{a}$ is connected to $\overline{1}$ by the edge $[a]$, the element $\overline{w}$ will be connected to $\overline{wa}$ by the edge $\overline{w}[a]$:



The boundary of an edge is the formal difference of its vertices:

$$\partial_1[a] = \overline{a} - \overline{1}, \qquad \partial_1(\overline{w}[a]) = \overline{wa} - \overline{w} = \overline{w}(\partial_1[a]).$$

[7]This extra assumption is only required for a perfect adequacy between topology and algebra, but we shall not use it in the proof, which is of purely algebraic nature.

This gives a geometric interpretation of the ZM-module $ZM[\Sigma]$ together with the ZM-linear map ∂_1. Furthermore, if for example a, b, c, d are generators, $\overline{abcd}$ is connected to $\overline{1}$ by a *chain* of edges

$$\begin{tikzcd} \overline{1} \arrow[r, "{[a]}"] & \overline{a} \arrow[r, "{\overline{a}[b]}"] & \overline{ab} \arrow[r, "{\overline{ab}[c]}"] & \overline{abc} \arrow[r, "{\overline{abc}[d]}"] & \overline{abcd} \end{tikzcd}$$

and X is now connected.

5.4 Contracting 0-cells

It is therefore natural to extend the bracket notation to words by

$$[1] = 0, \qquad [wa] = [w] + \overline{w}[a],$$

so that for example, $[abcd] = [a] + \overline{a}[b] + \overline{ab}[c] + \overline{abc}[d]$. This extended bracket must not be confused with the one of section 3.2, which is obtained by trivialising this one: here we have an extra coefficient $\overline{w}$. Note also that for any words u and v

$$[uv] = [u] + \overline{u}[v].$$

Geometrically, $[w]$ is a path connecting $\overline{w}$ to $\overline{1}$. This corresponds to the following identity

$$\partial_1[w] = \overline{w} - \overline{1},$$

which is easily proved by induction on the length of w.

Now, assuming that $(\Sigma, \mathcal{R})$ is a minimal canonical system, we define:

$$\begin{array}{rcl} ZM & \stackrel{s_0}{\to} & ZM[\Sigma] \\ \overline{w} & \mapsto & [\widehat{w}], \end{array}$$

so that $\partial_1 s_0(\overline{w}) = \partial_1[\widehat{w}] = \overline{w} - \overline{1} = \overline{w} - \eta\varepsilon(\overline{w})$ and (2) holds. This means that for each point $\overline{w} \in X$, the operator s_0 *chooses* a path connecting $\overline{w}$ to $\overline{1}$. Remember indeed that the family s_n must be a homotopy between the identity of X and a constant map. At the level of 0-*cells* (points of X), it must give a path connecting each point to a fixed one (here $\overline{1}$).

We have

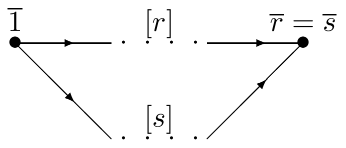
$$s_0\partial_1(\overline{w}[a]) = s_0(\overline{wa} - \overline{w}) = [\widehat{wa}] - [\widehat{w}].$$

In particular, if there is no rule, $\widehat{w} = w$ and $\widehat{wa} = wa$ so that $s_0\partial_1(\overline{w}[a]) = [wa] - [w] = \overline{w}[a]$. This means that if M is a free monoid, ∂_1 is injective and $H_1(M) = \ker \partial_1 = Z[\Sigma]$ is a free Z-module.[8]

5.5 Filling up 1-cycles

Now, assume there is some rule $A : r \to s$ in $\mathcal{R}$. Then $\overline{r} = \overline{s}$, but since r and s are distinct words, there are two distinct paths connecting $\overline{1}$ to $\overline{r}$ in X:

[8] Both facts hold when M is a free group too. In fact, there is a (highly nontrivial) reciprocal for this result [Swan 1969].



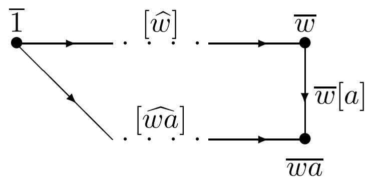
We have added edges to X in order to make it connected. The picture above shows that, doing this, we have lost *simple connectedness.* So we must now add 2-cells to X in order to take a step further towards contractibility. In our case, the 2-cell must have as boundary the cycle consisting of the two paths, namely $[r] - [s]$ (the sign takes the orientation into account).

It is therefore natural to define

$$\begin{array}{rcl} ZM[\mathcal{R}] & \stackrel{\partial_2}{\to} & ZM[\Sigma] \\ [A : r \to s] & \mapsto & [r] - [s], \end{array}$$

so that $\partial_1\partial_2[A] = \partial_1([r] - [s]) = \overline{r} - \overline{1} - \overline{s} + \overline{1} = 0$ and (3) holds, which was already clear from the picture.

We must now construct a map s_1 satisfying (4), namely $\partial_2 s_1(\xi) = \xi - s_0\partial_1(\xi)$ for all $\xi \in ZM[\Sigma]$. Let $\xi = \overline{w}[a]$, *i.e.* any element of the canonical Z-base of $ZM[\Sigma]$. Geometrically, ξ is an edge connecting $\overline{w}$ to $\overline{wa}$. We have already seen that $s_0\partial_1(\overline{w}[a]) = [\widehat{wa}] - [\widehat{w}]$. So we have the following picture:



In other words, $\xi - s_0\partial_1(\xi)$ is precisely the *cycle* above (note that backward arrows occur negatively) and $s_1(\xi)$ has to be a 2-chain whose boundary is this 1-cycle. When $\widehat{w}a$ is irreducible, *i.e.* $\widehat{wa} = \widehat{w}a$, the two paths above are identical and there is no hole to fill up. But in general there is a nontrivial cycle.

5.6 An example

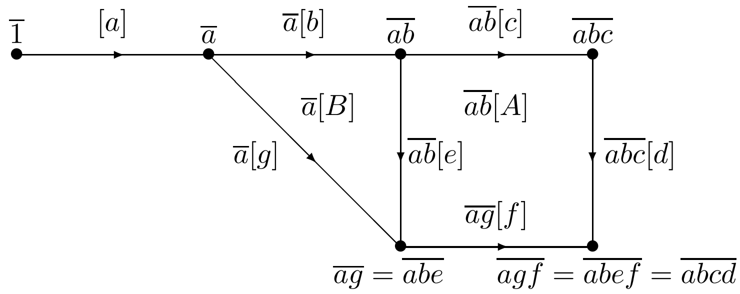
In order to make the next definitions understandable, we shall first consider a concrete example. Assume we have two rules

$$A : cd \to ef, \qquad B : be \to g,$$

and consider $\xi = \overline{abc}[d]$ where abc is irreducible. Starting with the word $abcd$, two successive reductions give:

$$abcd \stackrel{abA}{\to} abef \stackrel{aBf}{\to} agf,$$

and we get the following picture



where the square and the triangle can be filled up with the 2-cells $\overline{ab}[A]$ and $\overline{a}[B]$. If agf is irreducible, the bottom path is $s_0(\overline{abcd})$ and we can define:

$$s_1(\overline{abc}[d]) = \overline{ab}[A] + \overline{a}[B].$$

5.7 Contracting 1-cells

Our example shows that the cycle $\overline{w}[a] - s_0\partial_1(\overline{w}[a])$ can be filled up by using a reduction path from the word $\widehat{w}a$ to its canonical form. So we shall first extend the bracket notation to reduction paths, and s_1 will be definable in terms of a particular path, namely the leftmost reduction of $\widehat{w}a$.

Bracket is extended to elementary reductions, and then to reduction paths by

$$[uAv] = \overline{u}[A], \qquad [F_1 * \cdots * F_n] = [F_1] + \cdots + [F_n].$$

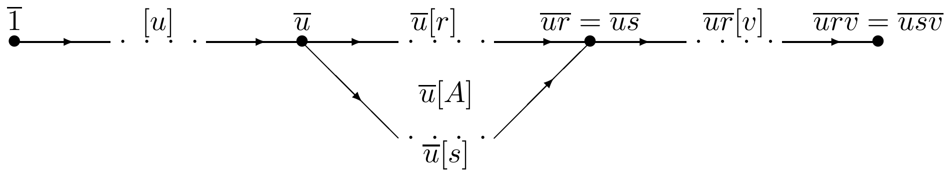
Once again, there is an extra coefficient $\overline{u}$ which is killed by trivialisation in the extended bracket of section 3.2. Note also the following obvious identities:

$$[uWv] = \overline{u}[W], \qquad [U * V] = [U] + [V].$$

Our definitions have been devised in order to get the identity:

$$\partial_2[W : u \to^* v] = [u] - [v].$$

Indeed, for an elementary reduction $uAv : urv \to usv$, we have the following picture



and $\partial_2[uAv] = \partial_2(\overline{u}[A]) = \overline{u}([r] - [s]) = [urv] - [usv]$, because $\overline{ur} = \overline{us}$. The case of a composed reduction is straightforward.

Now, using the leftmost strategy, we can introduce

$$\begin{array}{rcl} ZM[\Sigma] & \xrightarrow{s_1} & ZM[\mathcal{R}] \\ \overline{w}[a] & \mapsto & [\Lambda(\widehat{w}a) : \widehat{w}a \to^* \widehat{wa}], \end{array}$$

so that $\partial_2 s_1(\overline{w}[a]) = \partial_2[\Lambda(\widehat{w}a)] = [\widehat{w}a] - [\widehat{wa}] = [\widehat{w}] + \overline{w}[a] - [\widehat{wa}] = \overline{w}[a] - s_0\partial_1(\overline{w}[a])$ and (4) holds. This means that

$$ZM[\mathcal{R}] \xrightarrow{\partial_2} ZM[\Sigma] \xrightarrow{\partial_1} ZM \xrightarrow{\varepsilon} Z$$

is a partial resolution. In particular, we just proved the second point of theorem 2 in the case where the relations are the rules of a canonical system.

5.8 Finitely related monoids

Up to now, the condition of $(\Sigma, \mathcal{R})$ being canonical was not essential. What is indeed required to construct s_0 and s_1?

- For any congruence class $\overline{w}$ in M, we must be able to choose a representative (we took the canonical form $\widehat{w}$).
- For any word w, we must be able to choose a reduction path from w to the representative of $\overline{w}$ (we took the leftmost reduction path $\Lambda(w)$).

In general, it is always possible to choose a representative, but the fact that $\overline{u} = \overline{v}$ does not mean that there is a reduction path from u to v. Even so, s_1 can be defined if we allow *backward reductions* by adding an inverse $A^{-1} : s \to r$ for each rule $A : r \to s$, with the convention that $[A^{-1}] = -[A]$. So theorem 2 can be proved without extra assumption.

5.9 Contracting words

Let us now return to the case of a canonical system. First notice that for any word w

$$s_1[w] = [\Lambda(w)].$$

This is proved by induction on the length of w:

- $s_1[1] = 0 = [\Lambda(1)]$,
- $s_1[wa] = s_1([w] + \overline{w}[a]) = [\Lambda(w)] + [\Lambda(\widehat{w}a)] = [\Lambda(w)a * \Lambda(\widehat{w}a)] = [\Lambda(wa)]$, assuming that $s_1[w] = [\Lambda(w)]$ and using the fact that $\Lambda(wa) = \Lambda(w)a * \Lambda(\widehat{w}a)$ (section 1.7).

Therefore we get

$$s_1\partial_2(\overline{w}[A : r \to s]) = s_1(\overline{w}([r] - [s])) = s_1([\widehat{w}r] - [\widehat{w}s]) = [\Lambda(\widehat{w}r)] - [\Lambda(\widehat{w}s)].$$

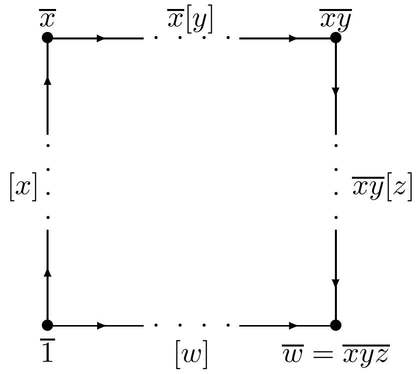
In particular, if there is no *critical pair*, $\widehat{w}A$ is obviously the leftmost elementary reduction of $\widehat{w}r$ so that $\Lambda(\widehat{w}r) = \widehat{w}A * \Lambda(\widehat{w}s)$ and $s_1\partial_2(\overline{w}[A]) = [\widehat{w}A] = \overline{w}[A]$, which means that ∂_2 is injective.

5.10 Filling up 2-cycles

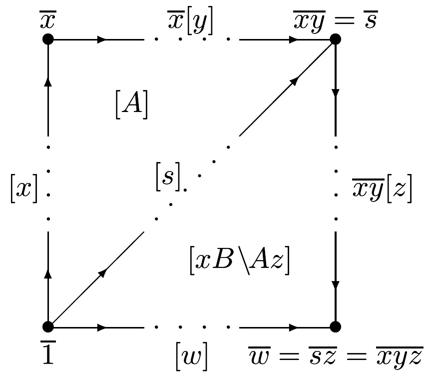
The space X is now connected and simply connected, but in general, critical pairs may introduce 2-cycles, so that ∂_2 is not injective. Indeed, if there are rules $A : xy \to s$ and $B : yz \to t$ with $y \neq 1$, we get a confluence diagram:

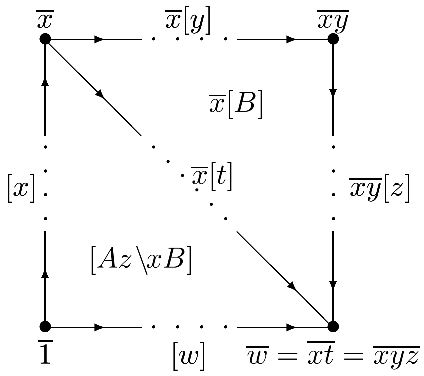
$$\begin{tikzcd}
& xyz \arrow[dl, "Az"'] \arrow[dr, "xB"] & \\
sz \arrow[dr, "xB\backslash Az"', "*"] & & xt \arrow[dl, "Az\backslash xB", "*"'] \\
& w &
\end{tikzcd}$$

We have in X the following picture of the 1-cycle $[xyz] - [w]$



which can be filled up in two different ways:





The formal difference of those 2-chains is a 2-cycle (topologically, a sphere) which has to be filled up with a 3-cell.

Therefore it is natural to introduce

$$\begin{array}{rcl} ZM[\mathcal{P}] & \stackrel{\partial_3}{\to} & ZM[\mathcal{R}] \\ [P,Q] & \mapsto & [P] + [Q\backslash P] - [Q] - [P\backslash Q]. \end{array}$$

Note that $\partial_3[P,Q]$ is precisely the 2-cycle described above, so that (5) is clearly satisfied.

5.11 Contracting 2-cells

We extend bracket to pairs of *elementary reductions of the same word*:

$$[F,F] = 0, \qquad [F,G] = 0 \text{ when } F \perp G, \qquad [uPv, uQv] = -[uQv, uPv] = \overline{u}[P,Q].$$

It is clear from this definition that

$$[uFv, uGv] = \overline{u}[F,G].$$

Furthermore, we have the following identity:

$$\partial_3[F,G] = [F] + [G\backslash F] - [G] - [F\backslash G].$$

Indeed, in the case where $F \perp G$, this follows from the fact that $[F\backslash G] = [F]$, which can be seen from the definition of $F\backslash G$, and symmetrically $[G\backslash F] = [G]$. In other words, $ZM[\mathcal{R}]$ “does not see” the relative order of orthogonal reductions. The other cases are obvious.

Since 3-dimensional pictures are quite hard to grasp, we shall rather use diagrams whose vertices are words, corresponding to 1-cells in X, and whose edges are reductions, corresponding to 2-cells in X. Thus we gain one dimension. For example, the latter identity expresses that the *2-cycle*

$$\begin{tikzcd} & t \arrow[dl, "F"'] \arrow[dr, "G"] & \\ u \arrow[dr, "G\backslash F"', "*"] & & v \arrow[dl, "F\backslash G", "*"'] \\ & w & \end{tikzcd}$$

can be filled up with the *3-cell* $[F, G]$.

For any path $W : u \xrightarrow{*} v$ we shall define, by *noetherian induction* on u, a 3-chain $\langle W \rangle$ filling up the following 2-cycle:

$$\begin{tikzcd} v \arrow[dr, "\Lambda(v)"', "*"] & & u \arrow[ll, "W"', "*"] \arrow[dl, "\Lambda(u)", "*"'] \\ & \widehat{v} = \widehat{u} & \end{tikzcd}$$

This $\langle W \rangle$ may be seen as a "path between reduction paths", or "second order reduction", since the target path $\Lambda(u)$ is "canonical" in the sense that it is given by the strategy. The construction is directly inspired by our proof of the Church-Rosser property in section 1.6.

- In the case of an elementary reduction $F : u \to v$, the word u is reducible, so that $\Lambda(u)$ is of the form $G * \Lambda(u')$ where $G = \Gamma(u) : u \to u'$ and we have the following diagram:

$$\begin{tikzcd} & u \arrow[dl, "F"'] \arrow[dr, "G"] & \\ v \arrow[dr, "G\backslash F", "*"'] \arrow[ddr, "\Lambda(v)"', "*"] & & u' \arrow[dl, "F\backslash G"', "*"] \arrow[ddl, "\Lambda(u')", "*"'] \\ & v' \arrow[d, "\Lambda(v')", "*"'] & \\ & \widehat{u} & \end{tikzcd}$$

  Since both v and u' are obtained from u by one step of reduction, it is licit to define:

$$\langle F \rangle = [F, G] - \langle G\backslash F \rangle + \langle F\backslash G \rangle.$$

  As usual, the sign takes the orientation into account. Note already that if F is leftmost, then $G = F$ and $\langle F \rangle = 0$.

- In the case of a composed reduction, we apply the previous definition to each component:

$$\langle F_1 * \cdots * F_n \rangle = \langle F_1 \rangle + \cdots + \langle F_n \rangle.$$

It is easy to check that $\langle W \rangle$ fills up the wanted 2-cycle, namely:

$$\partial_3 \langle W : u \xrightarrow{*} v \rangle = [W] + [\Lambda(v)] - [\Lambda(u)].$$

We can now introduce

$$\begin{array}{rcl} ZM[\mathcal{R}] & \xrightarrow{s_2} & ZM[\mathcal{P}] \\ \overline{w}[A] & \mapsto & \langle \widehat{w} A \rangle, \end{array}$$

so that $\partial_3 s_2(\overline{w}[A : r \to s]) = \partial_3\langle \widehat{w}A\rangle = [\widehat{w}A] + [\Lambda(\widehat{w}s)] - [\Lambda(\widehat{w}r)] = \overline{w}[A] - s_1\partial_2(\overline{w}[A])$ and (6) holds. This means that

$$ZM[\mathcal{P}] \xrightarrow{\partial_3} ZM[\mathcal{R}] \xrightarrow{\partial_2} ZM[\Sigma] \xrightarrow{\partial_1} ZM \xrightarrow{\varepsilon} Z$$

is a partial resolution. Theorem 3 follows easily, but we shall go a step further in order to show that ∂_3 is injective if there is no critical triple.

5.12 Contracting paths

Lemma 0 *$\langle wF\rangle = \langle \widehat{w}F\rangle$ and $\langle Fw\rangle = \langle F\rangle$ for any word w and for any elementary reduction $F : u \to v$.*

Proof: The first identity is proved by noetherian induction on w. It is true when w is irreducible, *i.e.* $w = \widehat{w}$. Otherwise, if $G = \Gamma(w) : w \to w'$, then $\Gamma(wu) = Gu$ and we get the following confluence diagram:

$$\begin{tikzcd} & wu \arrow[dl, "wF"'] \arrow[dr, "Gu"] & \\ wv \arrow[dr, "Gv"'] & & w'u \arrow[dl, "w'F"] \\ & w'v & \end{tikzcd}$$

So $\langle wF\rangle = [wF, Gu] - \langle Gv\rangle + \langle w'F\rangle = \langle \widehat{w}F\rangle$ by induction hypothesis, since $wF \perp Gu$ and Gv is leftmost.

The second identity is proved by noetherian induction on u. The leftmost elementary reduction of uw is Gw, where $G = \Gamma(u)$. We get the following confluence diagram

$$\begin{tikzcd} & uw \arrow[dl, "Fw"'] \arrow[dr, "Gw"] & \\ vw \arrow[dr, "(G\backslash F)w"', "*"] & & u'w \arrow[dl, "(F\backslash G)w", "*"'] \\ & v'w & \end{tikzcd}$$

and $\langle Fw\rangle = [Fw, Gw] - \langle (G\backslash F)w\rangle + \langle (F\backslash G)w\rangle$. But $G\backslash F$ is of the form $F_1 * \cdots * F_n$ where $\langle F_i w\rangle = \langle F_i\rangle$ by induction hypothesis. So $\langle (G\backslash F)w\rangle = \langle G\backslash F\rangle$ and similarly $\langle (F\backslash G)w\rangle = \langle F\backslash G\rangle$. Therefore $\langle Fw\rangle = [F, G] - \langle G\backslash F\rangle + \langle F\backslash G\rangle = \langle F\rangle$. *Q.e.d.Q.e.d.*

As a consequence, we get that for any reduction path W

$$s_2[W] = \langle W\rangle.$$

Indeed, for an elementary reduction, we have $s_2[uAv] = s_2[\widehat{u}A] = \langle \widehat{u}A\rangle = \langle uAv\rangle$ and the general case follows easily.

Therefore we get

$$\begin{aligned} s_2\partial_3(\overline{w}[P, Q]) &= s_2([\widehat{w}P] + [\widehat{w}(Q\backslash P)] - [\widehat{w}Q] - [\widehat{w}(P\backslash Q)]) \\ &= \langle \widehat{w}P\rangle + \langle \widehat{w}(Q\backslash P)\rangle - \langle \widehat{w}Q\rangle - \langle \widehat{w}(P\backslash Q)\rangle. \end{aligned}$$

In particular, if there is no *critical triple*, then $\widehat{w}P$ is leftmost, so that $\langle \widehat{w}P\rangle = 0$ and by definition $\langle \widehat{w}Q\rangle = [\widehat{w}Q, \widehat{w}P] - \langle \widehat{w}P\backslash \widehat{w}Q\rangle + \langle \widehat{w}Q\backslash \widehat{w}P\rangle = \overline{w}[Q, P] - \langle \widehat{w}(P\backslash Q)\rangle + \langle \widehat{w}(Q\backslash P)\rangle$, hence $s_2\partial_3(\overline{w}[P, Q]) = \overline{w}[P, Q]$ and ∂_3 is injective.

5.13 Going further

The reader can check that the sequence of section 3.2 is obtained by trivialising the one given here. We could pursue with $C_4 = ZM[\mathcal{T}]$ where $\mathcal{T}$ is the set of critical triples. Since in section 3, we did not need the fact that C_4 is $Z[\mathcal{T}]$, but only that it is 0 when $\mathcal{T} = \emptyset$, we can stop here. However, we point out the following generalisation of Squier's theorem:

Theorem 0 *(Kobayashi)*

If M is presented by a finite canonical system, then $H_n(M)$ is finitely generated for all $n \in N$.

The proof [Kobayashi 1990] is based on a generalisation of the *bar resolution* (see appendix B).

Conclusion

We hope the reader is now convinced of the relevance of a geometrical viewpoint in this area. It is not straightforward to extend those methods to *term rewriting*, because the geometry is partly hidden in the handling of variables (copying, garbage collection and exchange). The most promising approach seems to be the one of [Burroni 1991] who proposes a unified framework for higher-dimensional rewriting, including term rewriting as a special case.

References

[Brown 1982] : **K. S. Brown**. *Cohomology of groups.* Springer-Verlag, Berlin.

[Burroni 1991] : **A. Burroni**. *Higher-dimensional word problem*, to appear in *Category Theory and Computer Science.* Lecture Notes in Computer Science.

[Greenberg 1967] : **M. Greenberg**. *Lectures on Algebraic Topology.* Benjamin, New York, Amsterdam.

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

[Huet-Levy 1979] : **G. Huet & J. J. Levy**. *Call by need computations in non-ambiguous linear term rewriting systems.* Rapport de Recherche IRIA **359**.

[Jantzen 1985] : **M. Jantzen**. *A note on a special one-rule semi-Thue system.* Inform. Process. Lett. **21**, 135-140.

[Kapur-Narendran 1985] : **D. Kapur & P. Narendran**. *A finite Thue system with decidable word problem and without equivalent finite canonical system.* Theor. Comput. Sci. **35**, 337-344.

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

[Lang 1984] : **S. Lang**. *Algebra.* Addison-Wesley, Reading, Massachussets.

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

[MacLane 1963] : **S. Mac Lane**. *Homology.* Springer-Verlag, Berlin.

[McDuff 1979] : **D. McDuff**. *On the classifying spaces of discrete monoids.* Topology **18**, 313-320.

[Spanier 1966] : **E. H. Spanier**. *Algebraic Topology.* McGraw-Hill, New York.

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

[Squier-Otto 1987] : **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.

[Swan 1969] : **R. G. Swan**. *Groups of cohomological dimension 1.* Journal of Algebra **12**, 585-601.

Index

Appendices

A Topology and homology

This appendix is not essential to the proof of Squier's theorem. The terminology of algebraic topology can be found in [Greenberg 1967] or [Spanier 1966].

A.1 Homotopy and homology

Let X and Y be topological spaces, and let f, g be continuous maps from X to Y. A *homotopy* between f and g is a continuous transformation from f to g, *i.e.* a continuous map $h : [0,1] \times X \to Y$ such that

$$h(0,x) = f(x), \qquad h(1,x) = g(x).$$

In that case, we say that f and g are *homotopic*. There is a strong connection between the algebraic and the topological notions of homotopy, as we shall see in appendix B.

A *homotopy equivalence* between X and Y is a pair of continuous maps $f : X \to Y$ and $g : Y \to X$ which are inverse *up to homotopy*, *i.e.* $g \circ f$ is homotopic to the identity of X and $f \circ g$ is homotopic to the identity of Y. In that case, we say that X and Y have the same *homotopy type*. In particular X is *contractible* when it has the homotopy type of the (space consisting of a single) point. This means that there is a point $x_0 \in X$ and a continuous map $h : [0,1] \times X \to X$ called *contracting homotopy* such that

$$h(0,x) = x_0, \qquad h(1,x) = x.$$

This explains the presence of ε and η in the algebraic resolutions. Indeed the free resolution

$$\cdots \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_n} \cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\varepsilon} Z \longrightarrow 0$$

may better be seen as a morphism of complexes in the following way:

$$\begin{tikzcd}
\cdots \arrow[r, "\partial_{n+1}"] & C_n \arrow[r, "\partial_n"] \arrow[d] & \cdots \arrow[r, "\partial_2"] & C_1 \arrow[r, "\partial_1"] \arrow[d] & C_0 \arrow[r] \arrow[d, "\varepsilon"] & 0 \arrow[d] \\
\cdots \arrow[r] & 0 \arrow[r] & \cdots \arrow[r] & 0 \arrow[r] & Z \arrow[r] & 0
\end{tikzcd}$$

The top complex represents a space X, and the bottom complex represents the one point space. Now ε (completed by a family of null maps) represents a continuous map from X to the one point space and η (also completed by null maps) is the homotopy inverse of ε. Saying that we have a homotopy equivalence is just saying that the original "augmented" complex has a contracting homotopy.

This also explains the reason why we drop the module Z by truncating the complex before trivialising. This is just because we do not need the one point space to compute the homology of X.

For any topological space X, one defines a complex of (very big) free Z-modules

$$\cdots \xrightarrow{\partial_{n+1}} C_n(X) \xrightarrow{\partial_n} \cdots \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X)$$

where $C_0(X)$ is generated by all the points of X, $C_1(X)$ by all the continuous arcs in X, $C_2(X)$ by all the *singular triangles* in X (*i.e.* the continuous maps from some fixed triangle towards X) and so on. This *complex of singular chains* is used to define the *singular homology* of X which depends only on its homotopy type.

A.2 Homology of groups

The homology of a group G is also the (singular) homology of the quotient of any contractible space by some free (and sufficiently regular) action of G.

Assume that we managed to construct a contractible topological space X on which G acts freely (on the left) and properly (*i.e.* the canonical map $X \to G \setminus X$ is a *covering projection*). Then the *augmented complex of singular chains*

$$\cdots \xrightarrow{\partial_{n+1}} C_n(X) \xrightarrow{\partial_n} \cdots \xrightarrow{\partial_2} C_1(X) \xrightarrow{\partial_1} C_0(X) \xrightarrow{\varepsilon} Z \longrightarrow 0$$

is a resolution of Z by ZG-modules which happen to be free since the action of G on X is free. Furthermore, by elementary properties of covering spaces, trivialisation corresponds to the geometric operation of quotienting X by the action of G. In other words, the complex

$$\cdots \xrightarrow{\widetilde{\partial_{n+1}}} \widetilde{C_n(X)} \xrightarrow{\widetilde{\partial_n}} \cdots \xrightarrow{\widetilde{\partial_2}} \widetilde{C_1(X)} \xrightarrow{\widetilde{\partial_1}} \widetilde{C_0(X)}$$

is isomorphic to the complex of singular chains of the quotient $G \setminus X$. So the homology of the group G is the singular homology of $G \setminus X$.

Finally, the homology of a *monoid* M is also the singular homology of a topological space, which is called the *classifying space* of M. Conversely any reasonable topological space (*i.e.* a *CW-complex*) has the homotopy type of the classifying space of some monoid M [McDuff 1979], and so it has the same homology as M.

B Bar resolution

Let M be a monoid and let C_n be the free ZM-module over the set M^n. An element of the basis of C_n is written $[a_1|\dots|a_n]$, where $a_1, \dots, a_n \in M$. In particular, the basis of $C_0 \simeq ZM$ consists of a single element $[\,]$. As a Z-module, C_n is generated by the $a_0[a_1|\dots|a_n]$ where $(a_0, a_1, \dots, a_n) \in M^{n+1}$.

B.1 Definition

The *bar resolution* is

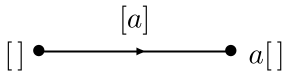
$$\cdots \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_n} \cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\varepsilon} Z \longrightarrow 0,$$

where ε is the map introduced in section 4.2, and ∂_n is the unique ZM-linear map defined by:

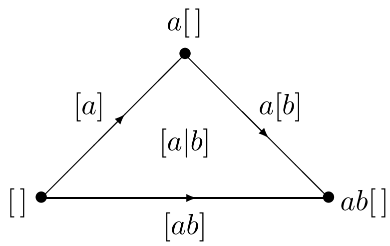
$$\partial_n[a_1|\dots|a_n] = a_1[a_2|\dots|a_n] + \sum_{i=1}^{n-1}(-1)^i[a_1|\dots|a_ia_{i+1}|\dots|a_n] + (-1)^n[a_1|\dots|a_{n-1}].$$

Geometrically, we have a space X with a vertex $a[\,]$ for each $a \in M$, an edge $a[b]$ for each $(a,b) \in M^2$, a triangle $a[b|c]$ for each $(a,b,c) \in M^3$, a tetrahedron $a[b|c|d]$ for each $(a,b,c,d) \in M^4$, an so on. Of course, M acts on X by left multiplication. The formula defining ∂_n becomes clear if we look at the low dimensional cases:

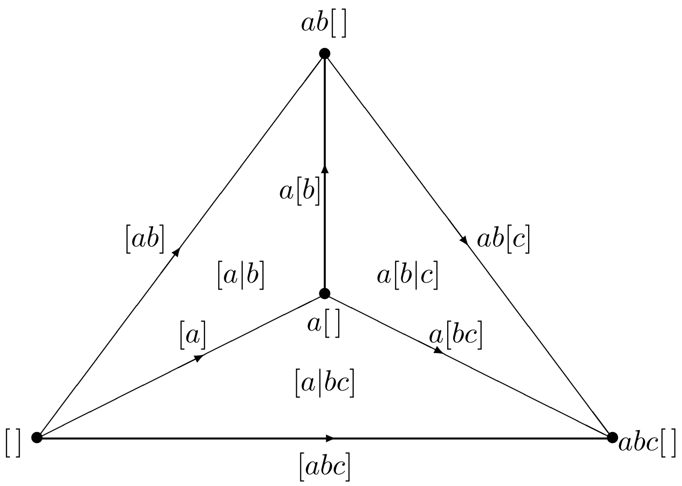
- the boundary of the edge $[a]$ is $a[\,] - [\,]$,



- the boundary of the triangle $[a|b]$ is $a[b] - [ab] + [a]$,



- the boundary of the tetrahedron $[a|b|c]$ is $a[b|c] - [ab|c] + [a|bc] - [a|b]$.



In the latter picture, the back face $[ab|c]$ and the tetrahedron $[a|b|c]$ itself are not indicated. As usual, signs keep track of the orientation.

B.2 The contracting homotopy

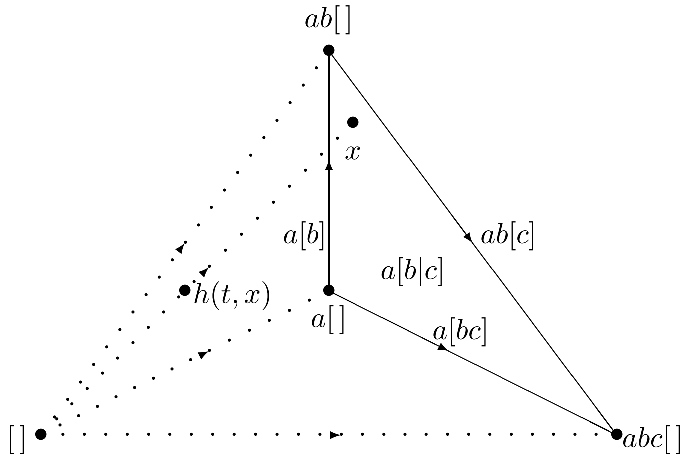
It is easy to check that we have indeed a chain complex. To prove exactness, we consider the *contracting homotopy* consisting of the map η of section 4.5 and the Z-linear maps s_n defined by

$$s_n(a_0[a_1|\dots|a_n]) = [a_0|a_1|\dots|a_n].$$

The reader may check that the equations of a contracting homotopy are indeed satisfied. In fact this corresponds to a contracting homotopy in the topological sense (appendix A.1). Indeed, if Δ is a n-dimensional cell, and $x \in \Delta$, then $s_n(\Delta)$ is a convex $n+1$-dimensional polyhedron, so that it makes sense to define

$$h(t, x) = tx + (1-t)[\,].$$

In particular, $s_n(\Delta)$ is the trajectory $h([0,1] \times \Delta)$. For example, if Δ is the triangle $a[b|c]$, then $s_n(\Delta)$ is the tetrahedron $[a|b|c]$ and you get the following picture:



B.3 Normalised bar resolution

It is possible to restrict the construction to the $[a_1|\dots|a_n]$ such that $a_1, \dots, a_n \neq 1$. The same formulae hold, with the convention that $[a_1|\dots|1|\dots|a_n] = 0$. It is clear that, up to dimension 3, this *normalised bar resolution* coincides with the one of section 5, in the case of the canonical system defined by the *standard presentation*.

The reader can use this resolution to prove that, if G is a group, $H_1(G)$ is (isomorphic to) the *abelianisation* $G/[G, G]$, where $[G, G]$ is the subgroup of G generated by the *commutators* $[x, y] = xyx^{-1}y^{-1}$ for $x, y \in G$.

(Hint: use the miraculous identity $(1-x)(1-y) = (1-x) + (1-y) - (1-xy)$.)