

Partial Report for Deliverable Nr. 3.2

Certified Computation of Homology Groups

Responsible: Julio Rubio, julio.rubio@unirioja.es

Written by: Members of La Rioja Node

Site: Universidad de La Rioja, Spain

Deliverable Date: February 2011

Forecast Date: February 2012

In order to get a certified computation of homology groups two separated subtasks are needed. First to construct matrices associated with geometrical (combinatorial) objects. Then, to diagonalise those matrices (computing, for instance, the Smith Normal Form of each matrix, see [20]) and extract a presentation of the homology groups. The first part is dealt with by the La Rioja node, and the second one by the Gothenburg node. The relationship between both nodes has been methodological (the methodological approach was proposed by Thierry Coquand, and then adopted by La Rioja team) and material (a collection of examples was tested simultaneously in La Rioja and Gothenburg, using respectively Kenzo [2] and a new Haskell program [11] to check the coherence of the results).

1 First Task: Matrix construction

For the first subtask, two alternatives are possible:

- to construct from a simplicial complex: (1) a corresponding simplicial set, (2) a chain complex associated with it and (3) the matrices representing the chain complex over given basis.
- to construct from a simplicial complex directly its *incidence* matrices.

Let us present the two different approaches.

1.1 Mathematical preliminaries

In this subsection, we briefly provide the minimal mathematical background needed in the rest of the report. We mainly focus on definitions. Many good textbooks are available for these definitions and results about them, the main one being maybe [17].

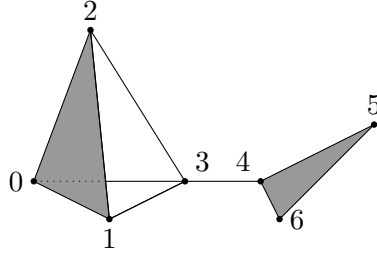


Figure 1: Butterfly Simplicial Complex

Simplicial Complexes: Let us start with the basic terminology. Let V be an ordered set, called the *vertex set*. An (*ordered abstract*) *simplex* over V is any ordered finite subset of V . An (*ordered abstract*) *n-simplex* over V is a simplex over V whose cardinality is equal to $n + 1$. Given a simplex α over V , we call *faces* of α to all the subsets of α .

Definition 1.1 An (*ordered abstract*) *simplicial complex* over V is a set of simplexes \mathcal{K} over V such that it is closed by taking faces (subsets); that is to say:

$$\forall \alpha \in \mathcal{K}, \text{ if } \beta \subseteq \alpha \Rightarrow \beta \in \mathcal{K}$$

Let \mathcal{K} be a simplicial complex. Then the set $S_n(\mathcal{K})$ of n -simplexes of \mathcal{K} is the set made of the simplexes of cardinality $n + 1$ of \mathcal{K} .

Example 1.1 Let us consider $V = (0, 1, 2, 3, 4, 5, 6)$.

The small simplicial complex drawn in Figure 1 is mathematically defined as the object:

$$\mathcal{K} = \left\{ \begin{array}{l} \emptyset, (0), (1), (2), (3), (4), (5), (6), \\ (0, 1), (0, 2), (0, 3), (1, 2), (1, 3), (2, 3), (3, 4), (4, 5), (4, 6), (5, 6), \\ (0, 1, 2), (4, 5, 6) \end{array} \right\}.$$

Note that, because the vertex set is ordered the list of vertices of a simplex is also ordered, which allows us to use a sequence notation (\dots) and not a subset notation $\{\dots\}$ for a simplex and also for the vertex set V . It is also worth noting that simplicial complexes can be infinite. For instance if $V = \mathbb{N}$ and the simplicial complex \mathcal{K} is $\{(n)\}_{n \in \mathbb{N}} \cup \{(n-1, n)\}_{n \geq 1}$, the simplicial complex obtained can be seen as an infinite bunch of segments.

Simplicial Sets: In spite of being a powerful tool, many common constructions in topology are difficult to make explicit in the framework of simplicial complexes. It soon became clear in the forties that the notion of simplicial set is much better.

Simplicial sets were first introduced by Eilenberg and Zilber [3], who called them *semi-simplicial complexes*. They can be used to express some topological properties of spaces by means of combinatorial notions.

Definition 1.2 A *simplicial set* K , is a union $K = \bigcup_{q \geq 0} K^q$, where the K^q are disjoint sets, together with functions:

$$\begin{aligned} \partial_i^q : K^q &\rightarrow K^{q-1}, & q > 0, & \quad i = 0, \dots, q, \\ \eta_i^q : K^q &\rightarrow K^{q+1}, & q \geq 0, & \quad i = 0, \dots, q, \end{aligned}$$

subject to the relations:

$$\begin{aligned} (1) \quad \partial_i^{q-1} \partial_j^q &= \partial_{j-1}^{q-1} \partial_i^q & \text{if} & \quad i < j, \\ (2) \quad \eta_i^{q+1} \eta_j^q &= \eta_j^{q+1} \eta_{i-1}^q & \text{if} & \quad i > j, \\ (3) \quad \partial_i^{q+1} \eta_j^q &= \eta_{j-1}^{q-1} \partial_i^q & \text{if} & \quad i < j, \\ (4) \quad \partial_i^{q+1} \eta_i^q &= \text{identity} & = & \quad \partial_{i+1}^{q+1} \eta_i^q, \\ (5) \quad \partial_i^{q+1} \eta_j^q &= \eta_j^{q-1} \partial_{i-1}^q & \text{if} & \quad i > j + 1. \end{aligned}$$

The ∂_i^q and η_i^q are called *face* and *degeneracy* operators respectively.

The elements of K^q are called *q-simplexes*. A simplex x is called *degenerate* if $x = \eta_i y$ for some simplex y and some degeneracy operator η_i ; otherwise x is called *non degenerate*.

Chain Complexes: Now, we are going to introduce a central notion in Algebraic Topology. We assume as known the notions of ring, module over a ring and module morphism (see [10] for details).

Definition 1.3 Given a ring R , a *graded module* M is a family of left R -modules $(M_n)_{n \in \mathbb{Z}}$.

Definition 1.4 Given a pair of graded modules M and M' , a *graded module morphism* f of degree k between them is a family of module morphisms $(f_n)_{n \in \mathbb{Z}}$ such that $f_n : M_n \rightarrow M'_{n+k}$ for all $n \in \mathbb{Z}$.

Definition 1.5 Given a graded module M , a *differential* $(d_n)_{n \in \mathbb{Z}}$ is a family of module endomorphisms of M of degree -1 such that $d_{n-1} \circ d_n = 0$ for all $n \in \mathbb{Z}$.

The previous definitions define a graded structure and a way of going from a level of the structure to the inferior one. From the previous definitions, the notion of chain complex is defined as follows.

Definition 1.6 A *chain complex* C_* is a family of pairs $(C_n, d_n)_{n \in \mathbb{Z}}$ where $(C_n)_{n \in \mathbb{Z}}$ is a graded module and $(d_n)_{n \in \mathbb{Z}}$ is a differential on $(C_n)_{n \in \mathbb{Z}}$.

The module C_n is called the module of *n-chains*. The image $B_n = \text{im } d_{n+1} \subseteq C_n$ is the (sub)module of *n-boundaries*. The kernel $Z_n = \text{ker } d_n \subseteq C_n$ is the (sub)module of *n-cycles*.

Definition 1.7 The *n-homology group* of C_* , denoted by $H_n(C_*)$, is defined as the quotient $\text{Ker } d_n / \text{Im } d_{n+1}$.

In an intuitive sense, homology groups measure n -dimensional holes in topological spaces. For instance, H_0 measures the number of connected components of a space. The homology groups H_n measure higher dimensional connectedness. For instance, the n -sphere, S^n , has exactly one connected component, one n -dimensional hole and no m -dimensional holes if $0 \neq m \neq n$.

There are two different ways of computing homology groups depending on the type of the object. On the one hand, the task of calculating homology groups of a finite object is translated to a problem of diagonalizing certain matrices, see [20]. On the other hand, in the case of non-finite type objects, Sergeraert's effective homology theory [19], implemented in Kenzo, provides a framework where this question can be handled. Roughly speaking, the effective homology method links a non-finite type object, X , with a finite type object, Y , with the same homology groups; then the problem of computing the homology groups of X is reduced to the task of diagonalizing some matrices associated with Y .

Let us present now the two different alternatives to construct matrices associated with finite geometrical (combinatorial) objects encoded as simplicial complexes.

1.2 Differential Map Matrices

As we have said previously in this approach we proceed as follows. Given a *finite* simplicial complex, the first step consists in constructing a corresponding simplicial set.

Definition 1.8 Let \mathcal{SC} be an (ordered abstract) simplicial complex over V . Then the *simplicial set* $K(\mathcal{SC})$ *canonically associated* with \mathcal{SC} is defined as follows. The set $K^n(\mathcal{SC})$ is $S_n(\mathcal{SC})$, that is, the set made of the simplexes of cardinality $n + 1$ of \mathcal{SC} . In addition, let (v_0, \dots, v_q) be a q -simplex, then the *face* and *degeneracy* operators of the simplicial set $K(\mathcal{SC})$ are defined as follows:

$$\begin{aligned} \partial_i^q((v_0, \dots, v_i, \dots, v_q)) &= (v_0, \dots, v_{i-1}, v_{i+1}, \dots, v_q), \\ \eta_i^q((v_0, \dots, v_i, \dots, v_q)) &= (v_0, \dots, v_i, v_i, \dots, v_q). \end{aligned}$$

That is, the face operator ∂_i^q removes the vertex in the position i of a q -simplex, and the degeneracy operator η_i^q duplicates the vertex in the position i of a q -simplex.

Subsequently, in the second step, the chain complex associated with the simplicial set is built.

Definition 1.9 Let K be a simplicial set, we define the *chain complex associated with* K , $C_*(K) = (C_n(K), d_n)_{n \in \mathbb{N}}$, in the following way:

- $C_n(K) = \mathbb{Z}[K^n]$ is the free \mathbb{Z} -module generated by K^n . Therefore an n -chain $c \in C_n(K)$ is a combination $c = \sum_{i=1}^m \lambda_i x_i$ with $\lambda_i \in \mathbb{Z}$ and $x_i \in K^n$ for $1 \leq i \leq m$;
- the differential map $d_n : C_n(K) \rightarrow C_{n-1}(K)$ is given by

$$d_n(x) = \sum_{i=0}^n (-1)^i \partial_i(x) \text{ for } x \in K^n$$

and it is extended by linearity to the combinations $c = \sum_{i=1}^m \lambda_i x_i \in C_n(K)$.

It is worth noting that the homology groups of a simplicial set K are the ones of the chain complex $C_*(K)$; and the homology groups of a simplicial complex \mathcal{SC} are the ones of the simplicial set $K(\mathcal{SC})$.

Finally, as we are working with chain complexes coming from finite simplicial complexes, the differential maps can be represented as matrices over the basis of the chain complex.

This first line has been formalised in the ACL2 Theorem Prover [15]. Step (1) was presented in the LOPSTR 2010 conference¹, and produced the publication [8]. Step (2), together with a reduction between the two chain complexes which can be associated to a simplicial set, produced [16] which will be presented in the ITP 2011 conference².

1.3 Incidence Simplicial Matrices

In the second line, we directly construct *incidence matrices* from finite simplicial complexes. There are two different definitions depending on the ground ring.

Definition 1.10 Let \mathcal{K} be a simplicial complex over V and let n be an integer such that $n \geq 1$ and (v_0, \dots, v_n) be an n -simplex of \mathcal{K} , the differential of this simplex is defined as:

$$d_n := \sum_{i=0}^n (-1)^i \partial_i^n.$$

Let \mathcal{K} be a simplicial complex over V and let n be an integer such that $n \geq 1$. The n -th *incidence matrix* of \mathcal{K} over the ring \mathbb{Z} , denoted by $M_n(\mathcal{K}, \mathbb{Z})$, represents the $(n-1)$ -simplices of \mathcal{K} as rows and the n -simplices of \mathcal{K} as columns. Assuming an ordering on the simplices of the same dimension (in the rest of the paper we assume that the simplices of the same dimension will be ordered), $M_n(\mathcal{K}, \mathbb{Z})$ is $[a_i^j]$ where i ranges from 1 to the cardinality of $S_{n-1}(\mathcal{K})$, j ranges from 1 to the cardinality of $S_n(\mathcal{K})$ and the value of a_i^j is the coefficient of the i -th $(n-1)$ -simplex in the differential of the j -th n -simplex; then a_i^j is a value in $\{0, \pm 1\}$.

Example 1.2 If we impose a lexicographical order on the simplices of the same dimension of the simplicial complex depicted in Figure 1 (if $v = (a_0, \dots, a_n)$ and $w = (b_0, \dots, b_n)$ are n -simplices of the simplicial complex, then $v < w$ if $a_0 < b_0$, or $a_0 = b_0$ and $a_1 < b_1$, or $a_0 = b_0$ and $a_1 = b_1$ and $a_2 < b_2, \dots$, or $a_0 = b_0, \dots, a_{n-1} = b_{n-1}$ and $a_n < b_n$), then its first incidence matrix is:

$$\begin{array}{c} \begin{matrix} (0,1) & (0,2) & (0,3) & (1,2) & (1,3) & (2,3) & (3,4) & (4,5) & (4,6) & (5,6) \end{matrix} \\ \begin{matrix} (0) \\ (1) \\ (2) \\ (3) \\ (4) \\ (5) \\ (6) \end{matrix} \begin{pmatrix} -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{matrix}$$

¹<http://www.risc.jku.at/conferences/lopstr2010/index.html>

²<http://itp2011.cs.ru.nl/ITP2011/Home.html>

Definition 1.11 The n -th incidence matrix of \mathcal{K} over the ring $\mathbb{Z}/2\mathbb{Z}$, denoted by $M_n(\mathcal{K})$, is a matrix of size $m \times p$, where m is the cardinality of $S_{n-1}(\mathcal{K})$ and p is cardinality of $S_n(\mathcal{K})$. Its coefficients $[a_i^j]$ are 1 if the i -th $(n - 1)$ -simplex is a face of the j -th n -simplex and 0 otherwise.

Note that the n -th incidence matrix of \mathcal{K} over the ring $\mathbb{Z}/2\mathbb{Z}$ is the absolute value of the n -th incidence matrix of \mathcal{K} over the ring \mathbb{Z} .

Example 1.3 If we impose a lexicographical order on the simplices of the same dimension of the simplicial complex depicted in Figure 1, then its first incidence matrix over the ring $\mathbb{Z}/2\mathbb{Z}$ is:

$$\begin{array}{c} (0) \\ (1) \\ (2) \\ (3) \\ (4) \\ (5) \\ (6) \end{array} \begin{pmatrix} (0,1) & (0,2) & (0,3) & (1,2) & (1,3) & (2,3) & (3,4) & (4,5) & (4,6) & (5,6) \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Using these definitions of incidence matrices, it is not necessary to use chain complexes to compute homology groups of simplicial complexes.

This second approach was formalized in Coq/SSReflect [1, 5] for the case $\mathbb{Z}/2\mathbb{Z}$ and was a joint work of the La Rioja and INRIA-Sophia Antipolis sites. The results have been presented in [9].

2 Second Task: Homology Computation

With respect to the second subtask (obtaining the Smith Normal Form of a matrix), the computational part is finished: a Haskell program has been written doing that computation using *coherent rings*, and specially designed to be easily verified using Coq (this is the method that was presented by Anders Mörtberg at Map 2010³, see [18]). The formalisation of this approach using Coq is still on-going work. This is the reason why the deliverable associated to this task 3.2, and due in the first 12 months, is not finished yet. It will be presented later, in the next year.

We present here some remarks about the formalisation of the computation of homology groups.

2.1 Specification of the Smith Normal Form

Here we show some notes about how to specify the correctness of an algorithm computing the Smith Normal Form of a matrix. We assume that all our matrices are over a fixed principal ideal domain.

Let us suppose that we have an algorithm called `SmithNormalForm` which takes as input a matrix M and returns as output a new matrix S (in the literature we can find several versions

³<http://www.unirioja.es/dptos/dmc/MAP2010/index.shtml>

$H_1(X, \mathbb{Z}/2\mathbb{Z})$: To compute $H_1(X, \mathbb{Z}/2\mathbb{Z})$ we need d_1 and d_2 :

- we have computed previously the Smith Normal Form of d_1 : $k = 6$;
- we compute the Smith Normal Form of d_2 :

$$SNF(d_2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix};$$

so, $m = 2$;

- in addition, $f = 10$.

Therefore, $H_1(X, \mathbb{Z}/2\mathbb{Z}) = (\mathbb{Z}/2\mathbb{Z})^{10-6-2} = \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$.

This result must be interpreted as stating that the butterfly simplicial complex has two “holes” in the topological sense.

References

- [1] COQ development team. The COQ Proof Assistant Reference Manual, version 8.3. Technical report, 2010.
- [2] X. Dousson, J. Rubio, F. Sergeraert, and Y. Siret. The Kenzo program. Institut Fourier, Grenoble, 1998. <http://www-fourier.ujf-grenoble.fr/~sergerar/Kenzo/>.
- [3] S. Eilenberg and J. A. Zilber. Semi-simplicial complexes and singular homology. *Annals of Mathematics*, 51(3):499–513, 1950.
- [4] M. Giesbrecht. Fast Computation of the Smith Normal Form of an Integer Matrix. In *Proceedings 20th International Symposium on Symbolic and Algebraic Computation (ISSAC'1995)*, pages 110–118, 1995.
- [5] G. Gonthier and A. Mahboubi. A Small Scale Reflection Extension for the Coq system. Technical report, Microsoft Research INRIA, 2009. <http://hal.inria.fr/inria-00258384>.
- [6] J. Hafner and K. McCurley. Asymptotically Fast Triangularization of Matrices over Rings. *SIAM J. Computing*, 20(6):1068–1083, 1991.

- [7] B. Hartley and T. Hawkes. *Rings, Modules and Linear Algebra*. Chapman and Hall, 1976.
- [8] J. Heras, V. Pascual, and J. Rubio. Proving with ACL2 the correctness of simplicial sets in the Kenzo system. In *Proceedings of the 20th International Symposium on Logic-Based Program Synthesis and Transformation (LOPSTR'2010)*, volume 6564 of *Lectures Notes in Computer Science*, pages 37–51. Springer-Verlag, 2011.
- [9] J. Heras, M. Poza, M. Dénès, and L. Rideau. Incidence simplicial matrices formalized in Coq/SSReflect. In *Proceedings 18th Symposium on the Integration of Symbolic Computation and Mechanised Reasoning (Calculemus'2011)*, 2011. <http://wiki.portal.chalmers.se/cse/pmwiki.php/ForMath/PapersAndSlides>.
- [10] N. Jacobson. *Basic Algebra II*. W. H. Freeman and Company, 2nd edition, 1989.
- [11] S. P. Jones et al. The Haskell 98 language and libraries: The revised report. *Journal of Functional Programming*, 13(1):0–255, 2003. <http://www.haskell.org>.
- [12] T. Kaczynski, K. Mischaikow, and M. Mrozek. *Computational Homology*, volume 157 of *Applied Mathematical Sciences*. Springer, 2004.
- [13] E. Kaltofen, M. Krishnamoorthy, and B. Saunders. Parallel Algorithms for Matrix Normal Forms. *Linear Algebra and its Applications*, 136:189–208, 1990.
- [14] R. Kannan and A. Bachem. Polynomial Algorithms for Computing the Smith and Hermite Normal Forms of an Integer Matrix. *SIAM J. Computing*, 8(4):499–507, 1979.
- [15] M. Kaufmann and J. S. Moore. ACL2. <http://www.cs.utexas.edu/users/moore/acl2/>.
- [16] L. Lambán, F. J. Martín-Mateos, J. Rubio, and J. L. Ruíz-Reina. Applying ACL2 to the Formalization of Algebraic Topology: Simplicial Polynomials. In *Proceedings of Interactive Theorem Proving 2011 (ITP'2011)*, 2011. <http://wiki.portal.chalmers.se/cse/uploads/ForMath/aafatasp>.
- [17] J. P. May. *Simplicial objects in Algebraic Topology*, volume 11 of *Van Nostrand Mathematical Studies*. University of Chicago Press, 1967.
- [18] A. Mörtberg. Constructive Algebra in Functional Programming and Type Theory. In *Mathematics Algorithms and Proofs*, 2010. <http://wiki.portal.chalmers.se/cse/uploads/ForMathInternalPages/Anders-Nov2010.pdf>.
- [19] J. Rubio and F. Sergeraert. Constructive Homological Algebra and Applications, Lecture Notes Summer School on Mathematics, Algorithms, and Proofs. University of Genova, 2006. <http://www-fourier.ujf-grenoble.fr/~sergerar/Papers/Genova-Lecture-Notes.pdf>.
- [20] O. Veblen. *Analysis Situs*. AMS Coll. Publ., 1931.