Homology, Homotopy and Applications, vol.5(2), 2003, pp.233–256

COMPUTING HOMOLOGY

TOMASZ KACZYNSKI, KONSTANTIN MISCHAIKOW AND MARIAN MROZEK

(*communicated by Gunnar Carlsson*)

Abstract

The aim of this paper is to provide a short introduction to computational homology based on cubical complexes. The discussed topics include cubical complexes, a reduction algorithm for computing homology of finitely generated chain complexes, and an algorithmic construction of homology of continuous maps via multivalued acyclic representations.

1. Introduction

The aim of this paper is to provide a synthesis of research by the authors and their collaborators on the subject of computing homology of spaces and maps. Our initial motivation came from dynamical systems, a subject with a history rich in entanglements with algebraic topology. It is not a coincidence that Poincar'e was a founding figure in both disciplines. Furthermore, much of modern applied dynamics is closely tied to numerical simulations. Thus, it should not come as a surprise that attempts to understand the structure of specific dynamical systems naturally leads to problems which require homological computations of large high dimensional numerically or experimentally generated complexes and maps thereof (see [**21, 20, 18, 19**] and references therein). However as one can see e.g. from the expository paper by Dey, Edelsbrunner and Guha [**7**], we can expect considerably more applications in the future.

From our perspective, there are three fundamental issues that need to be resolved in order to have an effective tool for computing homology. The first involves the underlying combinatorial structure. Obviously, one would like to encode this information in an efficient data structure. However, especially in the setting of dynamics, it is important that this combinatorial structure have nice properties with respect to the dimension of both the objects of interest and the ambient space. Furthermore, it should be noted that in many applications the input takes the form of discrete data points. As we will describe below it is with these constraints in mind that we have

---

Research by Tomasz Kaczynski is supported by grants from NSERC and FCAR.
Research by Konstantin Mischaikow is supported in part by NSF grant DMS-0107396.
Research by Marian Mrozek is supported by KBN, grant 2 P03A 011 18.
Received October 15, 2001, revised September 9, 2002; published on April 22, 2003.
2000 Mathematics Subject Classification: Primary 55-04; Secondary 52B05, 54C60, 65D17.
Key words and phrases: homology computation, algorithm, cubical complex, representable set, mulitivalued map.

chosen to develop a computational homology theory based on cubical complexes, rather than the classical approach involving simplicies [**24**].

At first glance the simplicial theory appears more elegant than the cubical theory described in Section 2. Furthermore, triangulations are at the heart of much of computational geometry. Thus, our insistence on using the cubical structure may appear strange. Consider, however, that it is hard to think of a scientific or engineering discipline which does not generate computational simulations or make use of recording devices or sensors to produce or collect image data. For example, it is possible, using X-ray computed tomography, to visualize cardiovascular tissue with a resolution on the order of 10 μm. [**4**]. Because this can be done at a high speed, timed sequences of 3-dimensional images can be constructed, essentially producing a 4-dimensional object. Observe that such images involve large amounts of data since the data sets contain information on the level of tetrals (the four dimensional equivalent of pixels). Obviously, the size and complexity of this data will grow as the sophistication of the sensors or simulations increase. Of course, one can try to render these objects in terms of triangulations thereby producing images that require smaller data sets and even perhaps are more intelligible to the human eye. However, given our current capabilities such rendering is typically quite time consuming and by its very nature results in modifications to the original data. For those applications which only require topological data (for example finding defects in hearts, which unfortunately is an all too common birth defect) the cubical approach allows one in principle to work directly with the raw data. [1]

Another typical example involves geometric structures arising from numerical simulations of partial differential equations. If one employs finite element methods based on a regular cubical lattice, then the output once again takes the form of a cubical complex. Furthermore, such computations can very rapidly produce complexes where memory constraints become a serious issue. If as above we think in terms of structures in space and time, then we have four dimensional complexes. Simply subdividing each cube into a simplex results in a simplicial complex with 24 times as many highest dimensional cells, which we claim is an unnecessary cost. Of course, one can imaging more intelligent methods of triangulating the cubical complex, but at a higher computational cost.

Returning to dynamical systems we have two additional problems to contend with. The first is that for many applications one needs to compute not onlythe homology of a space, but also the homology of a continuous function f. The standard approach of the simplicial theory is to approximate f by a simplicial map $f_{\#}$ which then induces f_* the map on homology. Unfortunately, the cost of computing $f_{\#}$ for a typical nonlinear map f appears to be excessive. To circumvent this problem we essentially compute f_* by making direct use of the geometric acyclic carrier theorem. We will describe this procedure in greater detail shortly, for the moment the important point is that in our applications we need to construct geometric acyclic carriers. Of course this can be done using either simplicies or cubes. However,

[1] For the purpose of this discussion we are ignoring the issue of noise in the data. In applications this is a serious problem which must be dealt with regardless of whether one uses simplicial or cubical homology.

these carriers are obtained using numerical approximations of the original map f. In higher dimensional problems cubes provide a very simple method data structure for encoding the images of points along with their errors, especially if one uses the l^∞ norm.

The second problem that is encountered in dynamical systems has to do with the difference between the dimensions of the objects of interest, e.g. a periodic orbit, and the dimension of the phase space which, in the case of a partial differential equation, can be infinite dimensional. A standard method of numerically solving such problems involves the Galerkin method, whereby one projects the infinite dimensional problem onto a finite dimensional subspace and then computes. Topological methods associated with nonlinear functional analysis provide an ideal means of using the information from the finite dimensional computations to deduce information about the infinite dimensional problem. However, to do so efficiently it is useful to have complexes which are compatible with orthogonal projections. As will become clear in Sections 2 and 3, this is an extremely nice feature of the cubical homology.

With this brief justification for using a cubical combinatorial structure we now turn to the problem of extracting algebraic information, i.e. the computation of homology of spaces. In Section 3, we present a procedure based on *elementary reductions*. The idea is to remove one pair of coinciding generators per time, without changing the homology of the complex, until one gets a complex which has a trivial boundary map, so it is isomorphic to its own homology. Although examples are given among cubical chain complexes, the procedure applies to any finitely generated free chain complex. An algorithm for computing homology of a chain map is a simple by-product. Results presented in this section first appeared in [**14**] and in somewhat refreshed form in [**12**]. We present the algorithm for computing homology with coefficients in a field, while we refer to [**25**] for coefficients in any Euclidean domain. A reduction procedure in the context of cubical complexes was presented in [**15**]. Homology programs based on our algorithms have been designed in C^{++} by Pilarczyk [**25**], Mazur and Szybowski [**17**], and they are currently available at the Web site *http://www.math.gatech.edu/ chom/*. The same idea of reducing a given complex to a minimal one with the help of the Smith normal form algorithm is explored by Gonzalez-Diez and Real [**9**] in the context of cohomology computation. Among algorithms of computing homology in low dimensions (simplicial complexes in $\mathbf{R}^3$ and S^3), probably the most efficient one is given by Delfinado and Edelsbrunner [**6**].

Finally, we deal in Section 4 with the most complex project of computing homology of continuous maps. This project is especially motivated by the previously mentioned applications to dynamical systems but there are also many problems in nonlinear analysis where homology of a map provides a useful information. In this context it is important to have an algorithm which will work for huge complexes coming from fine grids of the space and in arbitrarily high dimensions. The classical approach via barycentric subdivisions and simplicial approximations leads to a high complexity computation. In [**2**] a new approach is launched. It is proposed to inscribe the graph of a given continuous map f defined on a compact cubical domain into a graph of a finitely representable multivalued map F with acyclic cubical values. The construction of F is based on numerical approximation of f:

we only need to know a Lipschitz constant of f (more generally, the modulus of uniform continuity) and to be able to compute its approximative values on fine grid points. The proof of the central Theorem 4.4 provides the construction of a *chain selector* φ of F whose quotient map φ_* is the homology map $\varphi_* = F_* =$ of F. In a sense, the multivalued map F is a geometric realization of so called acyclic carrier presented e.g. in [**24**] thus we bypass the simplicial approximation. In [**13**], our construction is presented as the definition of the homology map f_* of f and the cubical homology theory is developed in a systematic self-contained way. In this paper, we only present the most important steps needed for understanding the explicit formula for f_*. In the end of Section 4.3 we present a new approach to grid refinements: rather than discussing grid subdivisions we introduce a concept of grid rescaling which allows to avoid generalizing previously introduced algebra to many grids. The final formula for computing f_* is given by Theorem 4.11. In Section 4.4 we present an algorithm for constructing so-called *coboundary* of a cycle z in an acyclic set, i.e. a chain c (usually not unique) which solves the equation $\partial(c) = z$. This was the major obstacle from making the construction of the homology map algorithmic. The problem was previously treated in [**1, 3, 12, 17**]. Although this was not our initial motivation, it is an interesting observation that our cobounday problem is an algebraic analogy of a geometric problem of minimal surfaces [**5**]. Another approach to computing homology of maps presented in [**23**] better utilizing the outcome of numerical algorithms for Poincaré maps in differential equations, is under development.

Many results discussed here were previously presented in several research papers [**2, 3, 12, 14, 15**] and in student dissertations [**1, 17, 25**]. A systematic presentation of the material was a subject of courses in computational algebraic topology simultaneously taught by the authors at their departments, and it constitutes a part of the book [**13**] which is in progress now.

2. Cubical Complexes

In numerical and graphical analysis one needs to consider very fine cubical grids. With appropriate units we may assume here that each cube is unitary i.e. it has sides of length 1 and vertices with integer coordinates. Later on in Section 4.3 we will discuss the effects of rescaling on the algebra extracted from a cubical grid.

An *elementary cube* Q is a finite product

$$Q = I_1 \times I_2 \times \cdots \times I_d \subset \mathbf{R}^d$$

where I_i is an interval of the form $I = [l, l+1]$ or $I = [l, l]$ for some $l \in \mathbf{Z}$. To simplify the notation we will write $[l] = [l, l]$ for an interval that contains only one point, such an interval is called *degenerate*. When $Q \subset \mathbf{R}^d$, the dimension d is called the *embedding number* of Q and denoted by $\operatorname{emb} Q$. The interval I_i is referred to as the i-th component of Q and written as $I_i(Q)$. The number of non degenerate components in Q is the *dimension* of Q and denoted by $\dim Q$. The set of all elementary cubes in $\mathbf{R}^d$ is denoted by $\mathcal{K}^d$. The set of all elementary cubes is denoted by $\mathcal{K}$, i.e.

$$\mathcal{K} := \bigcup_{d=1}^{\infty} \mathcal{K}^d.$$

It is easy to see that if $Q \in \mathcal{K}_k^d$ and $P \in \mathcal{K}_{k'}^{d'}$, then

$$Q \times P \in \mathcal{K}_{k+k'}^{d+d'}.$$

A set $X \subset \mathbf{R}^d$ is *cubical* if X can be written as a finite union of elementary cubes. We shall adopt the following notation.

$$\mathcal{K}(X) := \{Q \in \mathcal{K} \mid Q \subset X\}$$

and

$$\mathcal{K}_k(X) := \{Q \in \mathcal{K}(X) \mid \dim Q = k\}.$$

The elements of $\mathcal{K}_0(X)$ are the *vertices* of X and the elements of $\mathcal{K}_1(X)$ are the *edges* of X. More generally, the elements of $\mathcal{K}_k(X)$ are the *k-cubes* of X.

The group C_k of k-chains is the free abelian group generated by elements of $\mathcal{K}_k$, i.e. $C_k := \mathbf{Z}(\mathcal{K}_k)$. More explicitly, C_k is formally defined as the group of all those integer-valued functions on $\mathcal{K}_k$ which are non-zero on finitely many elements of $\mathcal{K}_k$. This is a free abelian group generated by functions $\widehat{Q} : \mathcal{K}_k \to \mathbf{Z}$, called *duals* of Q, given by the formula

$$\widehat{Q}(P) = \begin{cases} 1 & \text{if } P = Q \\ 0 & \text{otherwise} \end{cases}$$

for any Q and P in $\mathcal{K}_k$. The set of all duals of elementary cubes denoted by $\widehat{\mathcal{K}}$ is a basis of C_k. In the literature, it is customary to identify the dual of Q with Q but, for a better understanding of data structures, it is important to distinguish the algebraic object $\widehat{Q}$ from the geometric object Q. For that reason, we shall continue using the "hat" notation for duals of elementary cubes. The elements of C_k called *k-dimensional chains* (k-chains for short) are finite sums of the form

$$c = \alpha_1 \widehat{Q_1} + \alpha_2 \widehat{Q_2} \dots \alpha_m \widehat{Q_m},$$

where $\{Q_1, Q_2, \dots, Q_m\} \subset \mathcal{K}_k$. If $c \in C_k$ then $\dim c := k$. Obviously, for $k < 0$ and $k > d$ the set $\mathcal{K}_k = \emptyset$ and $C_k = 0$.

Let $c_1, c_2 \in C_k$ and let $a = \sum_{i=1}^m \alpha_i \widehat{Q}_i$, $b = \sum_{i=1}^m \beta_i \widehat{Q}_i$ Given chains c and d, we use the scalar product notation

$$\langle c_1, c_2 \rangle := \sum_{i=1}^{m} \alpha_i \beta_i.$$

The product of two elementary cubes is again an elementary cube. This motivates the following definition. Given two elementary cubes $P \in \mathcal{K}_k$ and $Q \in \mathcal{K}_{k'}$ put

$$\widehat{P} \diamond \widehat{Q} := \widehat{P \times Q}.$$

This formula extends to arbitrary chains $c_1 \in \widehat{\mathcal{K}}_k$ and $c_2 \in \widehat{\mathcal{K}}_{k'}$ by linearity:

$$c_1 \diamond c_2 := \sum_{P \in \mathcal{K}_k, Q \in \mathcal{K}_{k'}} \langle c_1, \widehat{P} \rangle \langle c_2, \widehat{Q} \rangle \widehat{P_i \times Q_j}.$$

The element $c \diamond c_2 \in C_{k+k'}$ is called the *cubical product* of c_1 and c_2. It it easy to show that

Proposition 2.1. *Let Q be an elementary cube such that* $\operatorname{emb} Q = d > 1$. *Then, there exist unique elementary cubes I and P with* $\operatorname{emb} I = 1$ *and* $\operatorname{emb} P = d-1$ *such that*

$$\widehat{Q} = \widehat{I} \diamond \widehat{P}.$$

Let $X \subset \mathbf{R}^d$ be a cubical set. Let $\widehat{\mathcal{K}}_k(X) := \{\widehat{Q} \mid Q \in \mathcal{K}_k(X)\}$. $C_k(X)$ is the finitely dimensional subgroup of C_k generated by the elements of $\widehat{\mathcal{K}}_k(X)$ and is referred to as the set of *k-chains of X*.

Let $c \in C_k$ and let $c = \sum_{i=1}^m a_i \widehat{Q}_i$ where $a_i \neq 0$ for $i = 1, \dots, m$. The *support* of the chain c is the cubical set

$$|c| := \bigcup_{i=1}^m Q_i \subset \mathbf{R}^d.$$

It is clear that $C_k(X) = \{c \in C_k \mid |c| \subset X\}$.

Given $k \in \mathbf{Z}$, the *cubical boundary operator*

$$\partial_k : C_k \to C_{k-1}$$

is defined on generators by induction on the embedding number. Let $\widehat{Q} \in \widehat{\mathcal{K}}^1_k$, then Q is an elementary interval and hence $Q = [l] \in \mathcal{K}^1_0$ or $Q = [l, l+1] \in \mathcal{K}^1_1$ for some $l \in \mathbf{Z}$. Define

$$\partial \widehat{Q} := \begin{cases} 0 & \text{if } Q = [l], \\ \widehat{[l+1]} - \widehat{[l]} & \text{if } Q = [l, l+1]. \end{cases}$$

Now assume that $\widehat{Q} \in \widehat{\mathcal{K}}^d_k$ where $d > 1$. By Proposition 2.1 there exist unique elementary cubical chains $\widehat{I}, \widehat{P}$ with $\operatorname{emb} I = 1$ and $\operatorname{emb} P = d-1$ such that

$$\widehat{Q} = \widehat{I} \diamond \widehat{P}.$$

Define

$$\partial \widehat{Q} := \partial \widehat{I} \diamond \widehat{P} + (-1)^{\dim I} \widehat{I} \diamond \partial \widehat{P}. \tag{1}$$

Finally, we extend the definition to all chains by linearity.

Proposition 2.2. *Let c and c' be cubical chains, then*

$$\partial (c \diamond c') = \partial c \diamond c' + (-1)^{\dim c} c \diamond \partial c'. \tag{2}$$

Proof. By linearity, it is sufficient to prove the proposition for elementary cubical chains, i.e. to show that

$$\partial (\widehat{Q} \diamond \widehat{Q'}) = \partial \widehat{Q} \diamond \widehat{Q'} + (-1)^{\dim Q} \widehat{Q} \diamond \partial \widehat{Q'}.$$

The proof goes by induction on $d := \operatorname{emb} Q$ and the decomposition of Q given by Proposition 2.1. □

Here is the explicit formula for the boundary map

Corollary 2.3. *If $\widehat{Q}_1, \widehat{Q}_2, \ldots, \widehat{Q}_m$ are elementary cubical chains, then*

$$\partial(\widehat{Q}_1 \diamond \widehat{Q}_2 \diamond \cdots \diamond \widehat{Q}_m) = \sum_{j=1}^{m} (-1)^{\sum_{i=1}^{j-1} \dim Q_i} \widehat{Q}_1 \diamond \cdots \diamond \widehat{Q}_{j-1} \diamond \partial \widehat{Q}_j \diamond \widehat{Q}_{j+1} \diamond \cdots \diamond \widehat{Q}_m.$$

Again by induction on the embedding number and an application of Propositions 2.1 and 2.2 one proves the fundamental property $\partial \circ \partial = 0$ for the cubical boundary operator. It is also useful to observe that $|\partial c| \subset |c|$, so $\partial_k(C_k(X)) \subset C_{k-1}(X)$. This implies that, given a cubical set $X \subset \mathbf{R}^d$, the *cubical chain complex*

$$\mathcal{C}(X) := \{C_k(X), \partial_k^X\},$$

is well defined. In the sequel we will omit the superscript X in ∂_k^X.

The subgroups of *cycles* and *boundaries* in $C_k(X)$ are respectively denoted by

$$Z_k(X) := \ker \partial_k = C_k(X) \cap \ker \partial_k \subset C_k(X). \tag{3}$$

and

$$B_k(X) := \operatorname{im} \partial_{k+1} = \partial_k(C_{k+1}(X)) \subset C_k(X). \tag{4}$$

The k-th *cubical homology group* of X is $H_k(X) := Z_k(X)/B_k(X)$. The homology of X is the graded group $H_*(X) := \{H_k(X)\}_{k \in \mathbf{Z}}$.

3. Computing Homology of Chain Complexes

3.1. Examples of Elementary Reductions

The aim of this section is to develop geometric intuitions behind the general reduction procedure presented next. We emphasize the importance of seeing the geometry behind the algebra since we believe that this will help to render our general reduction algorithm more efficient in particular cases of chain complexes coming from applications, as it was the case with cubical complexes in [**3, 15**].

We discuss two types of reductions for the cubical chain complex $\mathcal{C} := \mathcal{C}(X)$ where $X = [0,1] \times [0,2]$. Since elementary reductions may give complexes which are not necessarily cubical and duals of elementary cubes may be replaced by generators of a different geometric meaning, the "hat" notation of the previous section will be cumbersome here. We denote elementary cubes by upper case letters and dual chains by lower case, e.g. $v_i := \widehat{V_i}$ are duals of vertices, $a_i := \widehat{A_i}$ are duals of edges and $b_i := \widehat{B_i}$ are duals of squares. We indicate on Figure 1 dual chains rather geometric cubes so to point out their orientation.

With this in mind, C_0 is generated by $E_0 := \widehat{\mathcal{K}}_0(X) = \{v_1, v_2, \ldots v_6\}$, C_1 is generated by $E_1 := \widehat{\mathcal{K}}_1(X) = \{a_1, a_2, \ldots a_7\}$, and C_2 is generated by $E_2 := \widehat{\mathcal{K}}_2(X) = \{b_1, b_2\}$. The only nontrivial boundary maps ∂_1 and ∂_2. Their definitions can be deduced from fig:collapses (i). e.g. $\partial_1(a_1) = v_2 - v_1$, $\partial_2(b_1) = a_1 + a_2 - a_4 - a_3$. Note that the edge a_4 is an *interior* edge in the sense that it is a common edge of more then one elementary square of X. The remaining edges are *exterior* or *free* edges, in the sense any one of them is an edge of an exactly one square.

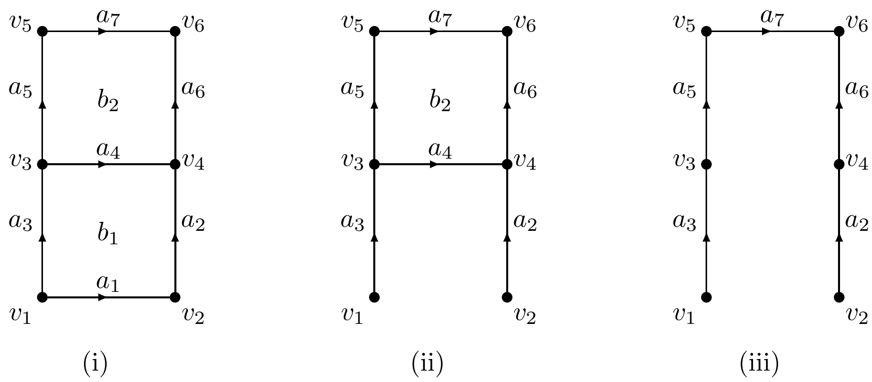

Figure 1: Elementary collapses by exterior edges.

Example 3.1. *(Exterior face collapses.)* Imagine that we push A_1 from outside so that B_1 is projected onto the remaining part of its boundary formed of A_2 A_3 and A_4. By taking account of orientations in Figure 1 (i), a_1 is projected onto the chain $\bar{a}_1 := a_3+a_4-a_2$ and b_1 disappears. No new homologically nontrivial cycle is created by the fact that b_1 disappears because the image of the cycle $\partial b_1 = a_1+a_2-a_4-a_3$ is $(a_3+a_4-a_2)+a_2-a_4-a_3 = 0$. This indicates that the homology of the complex $\bar{\mathcal{C}}$ obtained by projecting the generators

$$a_1 \mapsto \bar{a}_1 = a_3 + a_4 - a_2 \ , \ b_1 \mapsto 0,$$

and keeping the remaining generators fixed should be the same as that of $\mathcal{C}$. The new complex $\bar{\mathcal{C}}$ illustrated in Figure 1 (ii) is a subcomplex of $\mathcal{C}$ with bases $\bar{E}_0 := E_0$, $\bar{E}_1 := \{a_2, a_3, \dots a_6\}$ and $\bar{E}_2 := \{b_2\}$.

By repeating the same procedure with the edge a_4 and square b_2, we get a one-dimensional complex $\bar{\bar{\mathcal{C}}}$ illustrated in Figure 1 (iii). The same procedure can be repeated for the free vertex v_2 and its unique edge a_2 and so on. Thus $\bar{\bar{\mathcal{C}}}$ can be reduced in the next five elementary collapses to the single vertex v_1.

Example 3.2. *(Interior vertex reduction.)* Face reductions may also be performed starting from the lowest dimension. Imagine that we push the vertex V_4 along the edge A_4 so it projects to the vertex V_3. The edge A_4 disappears. If we want to do that without moving interiors of B_1 and B_2, we must drag the edges A_2 and A_6 with the vertex V_4 along A_4. With some care about orientations of edges we get the projections

$$v_4 \mapsto v_3, a_4 \mapsto 0, a_2 \mapsto \bar{a}_2 := a_2 - a_4, a_6 \mapsto \bar{a}_6 := a_4 + a_6 \ .$$

Figure 2 (i) shows a geometric interpretation of the new complex $\bar{\mathcal{C}}$. The topological structure realizing these reductions is the CW-complex (see [**24**]). Note that the new oriented edges $\bar{a}_2$ and $\bar{a}_6$ follow the same path in reverse directions for a

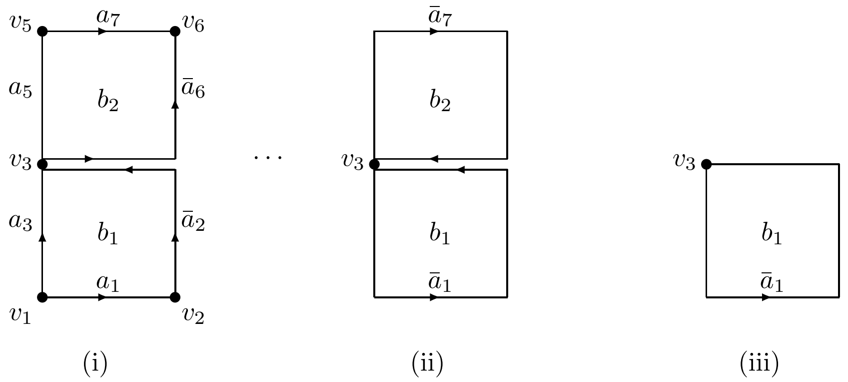

Figure 2: Starting from interior vertex reduction.

while. The whole complex can be visualized as two leafs B_1 and B_2 having v_3 as the only common vertex. By performing vertex reductions in each leaf as in the previous example we arrive at the complex $\bar{\bar{\mathcal{C}}}$ shown in Figure 2 (ii). Next, $\bar{a}_7$ is a free face of b_2 so an elementary collapse brings the complex to one leaf pictured in Figure 2 (iii) and another projection of $\bar{a}_1$ through b_1 brings the complex to $\mathcal{C}^f := \mathcal{C}(\{V_3\})$.

3.2. Reduction Procedure

Let now $(\mathcal{C}, \partial) = (\{C_q\}_{q\in\mathbf{Z}}, \{\partial_q\}_{q\in\mathbf{Z}})$ be a finitely generated free chain complex. Let n_0 and $d(\mathcal{C})$ be, respectively, the least and the greatest value of q such that $C_q \neq 0$. Most commonly, $n_0 = 0$ but, for example, if we want to compute the reduced homology (see [**24**]), we study an augmented chain complex and then $n_0 = -1$. The number $d(\mathcal{C})$ is called the dimension of $\mathcal{C}$. Let $d_q := \dim C_q := \mathrm{card}\,(E_q)$ (the number of elements in E_q) for each q.

Assume that a fixed base E_q of C_q is given for each q such that $C_q \neq 0$. We shall call it the *canonical basis* of C_q. Let $\langle\cdot,\cdot\rangle$ denote the scalar product associated to that canonical basis. Given $c \in E_q$, the generators $e \in E_{q-1}$ such that $\langle\partial c, e\rangle \neq 0$ are *faces* of c. The number $\langle\partial c, e\rangle$ is the *incidence number* of e in ∂c. It is simply the coefficient of e in the expansion of ∂c as a linear combination of elements of E_q.

Let $m \in \mathbf{Z}$ be a fixed number. Assume that $a \in E_{m-1}$ and $b \in E_m$ are two fixed elements such that

$$\lambda := \langle\partial b, a\rangle \text{ is invertible.} \tag{5}$$

So far we have been using coefficients in $\mathbf{Z}$ so the condition λ invertible means $\lambda = 1$ or $\lambda = -1$. Later on, we will use coefficients in a field where λ is invertible if and only if $\lambda \neq 0$ so this condition will be less restrictive.

Consequently, ∂b can be written as $\partial b = \lambda a + r$, where $\langle a, r\rangle = 0$. Let

$$E_{m-1} = \{a_1, a_2, \ldots a_{d_{m-1}-1}, a\},$$

$$E_m = \{b_1, b_2, \dots b_{d_m - 1}, b\}\ .$$

We shall write r as

$$r = \sum_{i=1}^{d_{m-1}-1} \alpha_i a_i\ .$$

For any $q \in \mathbf{Z}$, define a map $\pi_q : C_q \to C_q$ by the formula

$$\pi_q v := \begin{cases} v - \lambda^{-1}\langle v, a\rangle \partial b & \text{if } q = m-1\ , \\ v - \lambda^{-1}\langle \partial v, a\rangle b & \text{if } q = m\ , \\ v & \text{otherwise,} \end{cases} \tag{6}$$

where $v \in C_q$. We let $\bar{C}_q := \operatorname{im} \pi_q$ be the image of C_q in π_q.

It is easy to show that the map π_q is a linear projection of C_q onto its image $\bar{C}_q$. More precisely, it is a linear map with the property $\pi_q w = w$ for all $w \in \bar{C}_q$. We should explicitly identify the images of the basic elements under π_{m-1} and π_m. For simplicity of notation we put $\bar{v} := \pi_q v$ and $\gamma_i := \lambda^{-1}\langle \partial b_i, a\rangle$. It is easily seen that

$$\bar{a_i} = a_i\ ,\ \bar{a} = -\lambda^{-1} r \tag{7}$$

$$\bar{b_i} = b_i - \gamma_i b\ ,\ \bar{b} = 0\ . \tag{8}$$

We put

$$\bar{E}_q = \begin{cases} \{\bar{b_1}, \bar{b_2}, \dots \bar{b}_{d_m - 1}\} & \text{if } q = m\ , \\ \{a_1, a_2, \dots a_{d_{m-1}-1}\} & \text{if } q = m-1\ , \\ E_q & \text{otherwise.} \end{cases} \tag{9}$$

For the proof of the following proposition and theorem, we refer the reader to [**12**].

Proposition 3.3. *$\bar{E}_q$ is a basis for $\bar{C}_q$ for all $q \in \mathbf{Z}$ and $\bar{E}_m \cup \{b\}$ is a basis for C_m.*

Theorem 3.4. *The function $\pi : \mathcal{C} \to \mathcal{C}$ consisting of the sequence of projections $\{\pi_q\}_{q \in \mathbf{Z}}$ is a chain map.*

Corollary 3.5. *The sequence of groups $(\bar{\mathcal{C}}, \partial) := \{(\bar{C}_q, \partial_q)_{q \in \mathbf{Z}}\}$ is a chain subcomplex of $(\mathcal{C}, \partial)$ and the restriction $\bar{\pi} : \mathcal{C} \to \bar{\mathcal{C}}$ of the codomain of π to its image is a chain map.*

The most important feature of the new chain complex is expressed in the following theorem. We present a part of its proof which contains the construction of maps useful in the sequel.

Theorem 3.6. *$H_*(\bar{\mathcal{C}}) \cong H_*(\mathcal{C})$.*

Proof. We will show that $\bar{\pi} : \mathcal{C} \to \bar{\mathcal{C}}$ is a chain equivalence with the inclusion $i : \bar{\mathcal{C}} \hookrightarrow \mathcal{C}$ as a homotopical inverse. Indeed, as mentioned before, $\bar{\pi}$ is a projection so $\bar{\pi} i = \mathrm{id}_{\bar{\mathcal{C}}}$. Hence it is sufficient to find a chain homotopy between $i\bar{\pi} = \pi$ and $\mathrm{id}_{\mathcal{C}}$. Let $D_q : C_q \to C_{q+1}$ be given by

$$D_q v = \begin{cases} \lambda^{-1}\langle v, a\rangle b & \text{if } q = m-1\ , \\ 0 & \text{otherwise} \end{cases}$$

for any $v \in C_q$. It remains to show the identity

$$\mathrm{id}_{C_q} - i_q \bar{\pi}_q = \partial_{q+1} D_q + D_{q-1} \partial_q \ . \tag{10}$$

We refer the reader to [**12, 14**] for details. □

Let now

$$\mathcal{C}^0 \xrightarrow{\bar{\pi}^1} \mathcal{C}^1 \xrightarrow{\bar{\pi}^2} \mathcal{C}^2 \cdots \tag{11}$$

be a sequence of chain subcomplexes and projections obtained from $(\mathcal{C}, \partial)$ by iterating the above construction as long as it is possible to choose $m \in \{n_0, n_0 + 1, \ldots, d(\mathcal{C})\}, a \in C_{m-1}$ and $b \in C_m$ such that $\lambda := \langle \partial b, a \rangle$ is invertible. Thus $\mathcal{C}^0 := \mathcal{C}$, $\mathcal{C}^{k+1} := \bar{\mathcal{C}}^k$ and $E_q^{k+1} := \bar{E}_q^k$ for all $q \in \mathbf{Z}$. We denote by ∂^k the restriction of ∂ to $\mathcal{C}^k$.

Denote by $M(k) = \sum_q \mathrm{card}\,(E_q^k)$, for $k = 0, 1, 2, \ldots$ Since $\mathcal{C}$ is finitely generated, $M(k) < \infty$ and $M(k+1) = M(k) - 2$, therefore there exists a final element of that sequence denoted by $(\mathcal{C}^f, \partial^f)$, beyond which the construction cannot be extended. The following corollary is deduced from Theorem 3.6 by induction.

Corollary 3.7.

$$H_*(\mathcal{C}) \cong H_*(\mathcal{C}^f) \ .$$

It seems that $H_*(\mathcal{C}^f)$ should now be easier to compute than $H_*(\mathcal{C})$, because $\mathcal{C}^f$ has less generators left. In general, one should incorporate the Smith normal form reduction arguments for judicious choice of pairs (b, a) of generators to eliminate. This has been done in [**25**]. In many applications however we are only interested in detecting nontrivial homologies or estimating Betti numbers. In that case one can reduce a task by computing homology with coefficients in a field $\mathbf{F}$, e.g. $\mathbf{F} = \mathbf{Z}_p$ where p is prime, or $F = Q$, the field of rational numbers. This means in practice that C_q is defined as a vector space over $\mathbf{F}$ with the same E_q as the canonical basis and the chain complex $\mathcal{C}$ is a graded vector space with ∂ the graded linear map defined as previously on generators. The relationship between homologies for field coefficients and those for integer coefficients are described by the Universal Coefficient Theorem (see [**24**]).

Theorem 3.8. *Suppose now that the computation is performed for coefficients in a field* $\mathbf{F}$*. Then* $\partial^f = 0$ *and*

$$H(\mathcal{C}, \mathbf{F}) \cong H(\mathcal{C}^f, \mathbf{F}) = \mathcal{C}^f \ .$$

Proof. The first identity is proved in Corollary 3.7. Since $\mathbf{F}$ is a field then λ is invertible if and only if and only if it is non-zero. Hence the construction can be iterated as long as there exist m and two elements $a \in E_{m-1}^k$ and $b \in E_m$ satisfying $\langle \partial b, a \rangle \neq 0$ i.e. as long as $\partial \neq 0$. Therefore $\partial^f = 0$. But this means that $H(\mathcal{C}^f, \mathbf{Z}_p) = \mathcal{C}^f$. □

We outline below an algorithm due to Kaczynski-Mrozek-Ślusarek [**14**], which we call here KMS algorithm, based on the above reduction procedure. This algorithm computes homology of chain complexes with coefficients in a field but we shall restrict the analysis to mod p computation, i.e. for coefficients in the field $\mathbf{Z}_p$, because this choice is the most efficient due to a low cost od division operation.

For simplicity we assume that $n_0 = 0$, this is the case when $\mathcal{C}$ is a cubical complex. We put $N = \max\{d_q \mid q = 0, 1, \ldots, d(\mathcal{C})\}$. We assume that there is a fixed upper bound for $d(\mathcal{C})$ and we shall estimate the *time* or *complexity* of calculations as a function of N.

A direct implementation of a classical homology computation by linear algebra methods ([**24**]) when applied to computations over the field $\mathbf{Z}_p$ yields an $O(N^3)$ algorithm. This means that if the number of arithmetic operations is comparable with N^3 for large values of N. For some special cases KMS algorithm achieves significant improvement of that bound.

For a fixed dimension the process is controlled by procedure COLLAPSE and a single reduction step is executed by procedure REDUCE.

Algorithm 3.9. *(Algorithm HOMOLOGY)*
input: finite chain complex $\mathcal{C}$ *with coefficients in* Z_p.
output: nonnegative integers $\beta(0), \beta(1), \ldots \beta(d(C))$
(these are dimensions of $H_q(\mathcal{C}, \mathbf{F})$*).*
for $q = d(\mathcal{C})$ *downto* 1 *do*
COLLAPSE(q);
$\beta(q) = \operatorname{card} E(q)$;
end.
procedure COLLAPSE(q);
while $E(q) \neq \emptyset$ *and there exist* $b \in E(q)$ *and* $a \in E(q-1)$ *such that* $\langle \partial b, a \rangle \neq 0$ *do*
REDUCE(b, a);
end.

Procedure REDUCE operates on the following data structures:
1. An array T_m of all m-generators;
2. An array T_{m-1} containing all faces of m-generators;
3. For each m-generator b a list $F(b)$ of its faces sorted by faces with their incidence numbers in ∂b.

Algorithm 3.10. *procedure REDUCE*(b, a);
for each b' *in* T_m *do*
if a *appears in* $F(b')$ *then*
compute new list $F(b')$ *by means of synchronous scan (i.e. merge) of* $F(b)$ *and* $F(b')$
end REDUCE.

A single execution of REDUCE costs $O(N^2)$ and assuming that $d(C)$ is fixed one shows that homology groups with coefficients in $\mathbf{Z}_p$ of a finitely generated chain complex can be computed in $O(N^3)$ arithmetical operations.

As already mentioned there are special cases when the algorithm runs much faster. Observe that 0-dimensional homologies can be computed by finding the number of connected components of the graph which can be performed in linear time by classical graph scan methods like depth-first search. Next, 1-dimensional homologies can be specified by computing the number of fundamental cycles in the graph. For a connected graph this needs no computation at all because this number is $d_1 - d_0 + 1$ by the Euler characteristic. When X is a cubical set in $\mathbf{R}^2$ it can be proved that there exists an edge A which is a face of exactly one elementary square Q. The procedure REDUCE applied to $(\widehat{Q}, \widehat{A})$ is in fact an elementary collapse, so the resulting complex is again a cubical complex and incidence numbers are 0, 1 or -1. We may iterate this procedure until we get a 1-dimensional complex. The following result is derived.

Theorem 3.11. *Let X be a cubical set in the plane $\mathbf{R}^2$. Then the homology of $\mathcal{C}(X)$ is be computed in linear time. Moreover this also holds true for computation in integer coefficients.*

The fact that Betti numbers of polyhedra in $\mathbf{R}^2$ can be computed in linear time has already been observed by Delfinado and Edelsbrunner [**6**]. We refer the reader to [**12, 14, 15**] for further discussion of cases when the complexity of the reduction algorithm can be improved.

3.3. Computing Homology of a Chain Map

Given a triple $(\mathcal{C}, \mathcal{D}, \varphi)$ where $\mathcal{C}$, $\mathcal{D}$ are finite chain complexes with coefficients in a field $\mathbf{F}$ and $\varphi : \mathcal{C} \to \mathcal{D}$ is a chain map, a reduction performed either on C or D induces a reduction on the whole triple.

Indeed, if $\bar{\mathcal{C}}$ is obtained from $\mathcal{C}$ by an elementary reduction, we define $\bar{\varphi} : \bar{\mathcal{C}} \to \bar{\mathcal{D}}$ by $\bar{\varphi} := \varphi i$, where $i : \bar{\mathcal{C}} \hookrightarrow \mathcal{C}$ is the inclusion chain map given in the proof of Theorem 3.6. If $\bar{\mathcal{D}}$ is obtained from $\mathcal{D}$, we define $\bar{\varphi} : \mathcal{C} \to \bar{\mathcal{D}}$ by $\bar{\varphi} := \bar{\pi}\varphi$ where $\bar{\pi} : \mathcal{D} \to \bar{\mathcal{D}}$ is the projection chain map given for $\mathcal{D}$ as in Corollary 3.5. Since we showed in the proof of Theorem 3.6 that i and $\bar{\pi}$ induce isomorphisms in homology, this procedure can be iterated so that in the final stage we get a chain map $\varphi^f : C^f \to D^f$ such that the diagram

$$\begin{tikzcd} \mathcal{C} \arrow[r, "\varphi"] & \mathcal{D} \arrow[d, "p^f"] \\ \mathcal{C}^f \arrow[u, "j^f"] \arrow[r, "\varphi^f"] & \mathcal{D}^f \end{tikzcd}$$

commutes and the vertical arrows induce isomorphisms in homology. Thus Corollary 3.7 and the above discussion imply the following

Corollary 3.12. *If field coefficients are assumed then the triple $(\mathcal{C}^f, \mathcal{D}^f, \varphi^f)$ is isomorphic to the triple $(H(\mathcal{C}), H(\mathcal{D}), \varphi_*)$ in the sense that the vertical arrows in the following commuting diagram are isomorphisms:*

$$\begin{tikzcd} H(\mathcal{C}) \arrow[r, "\varphi_*"] & H(\mathcal{D}) \arrow[d, "p^f_*"] \\ \mathcal{C}^f \arrow[u, "i^f_*"] \arrow[r, "\varphi^f"] & \mathcal{D}^f \end{tikzcd}$$

The above discussion permits one to obtain an algorithm of computing homology of a chain map. Assume that $\mathcal{C}$, $\mathcal{D}$ are finite chain complexes with coefficients in a field $\mathbf{F}$ and $\varphi : \mathcal{C} \to \mathcal{D}$ is a chain map. For each $q \in \mathbf{Z}$ let A_q be the matrix of $\varphi_q : C_q \to D_q$ with respect to fixed bases E_q and F_q in C_q and D_q, respectively.

Given the reduction of elements a and b in C_{m-1} and C_m we define the matrix $\bar{A}_q$ of $\bar{\varphi}_q : \bar{C}_q \to D_q$, for each q, by specifying its columns, using the map i discussed in Section 3.3. If $q \neq m$ then $\bar{A}_q = A_q$. Assume then $q = m$. Let $\mathrm{Col}(e)$ be the column of A_q which corresponds to $e \in E_m$, let $\lambda = \langle \partial b, a \rangle \neq 0$, and let $c_e = \langle \partial e, a \rangle$, $e \in \bar{E}_m = E_m \setminus \{b\}$.

If $c_e \neq 0$, put $\overline{\mathrm{Col}}_e = \mathrm{Col}_e$, otherwise $\overline{\mathrm{Col}}_e = \mathrm{Col}_e - c_e \lambda^{-1} \mathrm{Col}_b$.

Analogically, given the reduction of elements a and b in D_{m-1}, D_m we define the matrix $\bar{A}_q$ of $\bar{\varphi}_q : C_q \to \bar{D}_q$, for each q, by specifying its rows, using the definition of the map $\bar{p}$ from discussed in Section 3.3.

If $q \neq m-1, m$ then $\bar{A}_q = A_q$. If $q = m$ we obtain $\bar{A}_q$ by deleting the row Row_b which corresponds to b in A_q. Assume then $q = m-1$ and let $f \in \bar{F}_{m-1} = F_{m-1} \setminus \{a\}$. Let Row_f be the row of A_{m-1} which corresponds to f. Let Row_a be the row which corresponds to a and let $d_f = \langle \partial b, f \rangle$. It is easy to check that the row in the matrix $\bar{A}_{m-1}$ which corresponds to f is given by

$$\overline{\mathrm{Row}}_f = \mathrm{Row}_f - \lambda^{-1} d_f \mathrm{Row}_a$$

For a fixed q the computation of $\bar{\varphi}_q : \bar{\mathcal{C}} \to \bar{\mathcal{D}}$ consists of two iterated transformations of the matrix A_q, each one taking $O(n^2)$ time. We can iterate the above matrix transformations until there is no possibility for a one-step reduction, finally obtaining the homomorphism $\varphi^* : H(\mathcal{C}) \to H(\mathcal{D})$ corresponding to the chain map $\varphi : \mathcal{C} \to \mathcal{D}$. Since the dimension of the complexes is fixed we get an algorithm for computing the homology of a chain map of complexity $O(n^3)$.

4. Towards Computing Homology of Maps

As it was indicated in the introduction, we shall present here the approach to computing homology of a continuous map between cubical sets via its multivalued representation. We shall present only those proofs which contain constructions important for the algorithm design. The detailed proofs of all statements will be presented in [**13**].

4.1. Topology of Representable Sets

Cubical sets are compact. It is often useful to have a locally finite class of sets which contains cubical sets but is closed under set-theoretical and topological operations. For this reason we present the notion of elementary cells and representable sets introduced in [**22**].

With any elementary cube $Q = I_1 \times I_2 \times \ldots \times I_d \subset \mathbf{R}^d$ we associate the associated *elementary cell*

$$\overset{\circ}{Q} := \overset{\circ}{I}_1 \times \overset{\circ}{I}_2 \times \ldots \times \overset{\circ}{I}_d.$$

where

$$\overset{\circ}{I} := \begin{cases} (l, l+1) & \text{if } I = [l, l+1], \\ [l] & \text{if } I = [l, l]. \end{cases}$$

We purposely avoided the temptation of calling $\overset{\circ}{Q}$ an open cube since it is not open in $\mathbf{R}^d$, unless its dimension is equal to the embedding number.

A set $Y \subset \mathbf{R}^d$ is *representable* if it is a finite union of elementary cells. The family of representable sets in $\mathbf{R}^d$ is denoted by $\mathcal{R}^d$.

The *closed hull* of a set $A \subset \mathbf{R}^d$ is

$$\operatorname{ch}(A) := \bigcup\{Q \mid Q \in \mathcal{K}, \overset{\circ}{Q} \cap A \neq \emptyset\}. \tag{12}$$

Example 4.1. If $x, y \in \mathbf{R}$, $x \leqslant y$ are two arbitrary real numbers then

$$\operatorname{ch}([x, y]) = [p, q],$$

where $p := \max\{m \in \mathbf{Z} \mid m \leqslant x\}$ and $q := \min\{m \in \mathbf{Z} \mid m \geqslant y\}$ denote the floor of x and ceiling of y respectively. This in turn implies that given bounded intervals $A_1, A_2, \ldots, A_d$ we have

$$\operatorname{ch}(A_1 \times A_2 \times \ldots \times A_d) = \operatorname{ch}(A_1) \times \operatorname{ch}(A_2) \times \ldots \times \operatorname{ch}(A_d).$$

In particular, $\operatorname{ch}(A_1 \times A_2 \times \ldots \times A_d)$ is a rectangle in $\mathbf{R}^d$.

For a more systematic presentation of properties of representable sets we refer the reader to [**22**].

4.2. Cubical Multivalued Maps

Let X and Y be cubical sets. A *cubical multivalued map* $F : X \overrightarrow{\to} Y$ from X to Y is a function from X to subsets of Y (i.e. for every $x \in X$, $F(x) \subset Y$) which has the following properties.

1. For every $x \in X$, $F(x)$ is a cubical set.
2. For every $Q \in \mathcal{K}(X)$, $F|_{\overset{\circ}{Q}}$ is constant, i.e. if $x, x' \in \overset{\circ}{Q}$, then $F(x) = F(x')$.

Given $A \subset X$ and $B \subset Y$ we let $F(A) := \bigcup_{x \in A} F(x)$. A multivalued map F is *upper semicontinuous* if $\{x \in X \mid F(x) \subset U\}$ is open for any open $U \subset Y$ and it is *lower semicontinuous* if the set $\{x \in X \mid F(x) \cap U \neq \emptyset\}$ is open for any open $U \subset Y$.

Proposition 4.2. *Assume $F : X \overrightarrow{\to} Y$ is a cubical map. Then F is lower semicontinuous if and only if it has the following property:*

$$\textit{For any } P, Q \in \mathcal{K}(X) \textit{ such that } P \textit{ is a face of } Q,\ F(\overset{\circ}{P}) \subset F(\overset{\circ}{Q}). \tag{13}$$

Similarly, upper semicontinuous cubical maps are characterized by the property:

Proposition 4.3. *Assume $F : X \overrightarrow{\to} Y$ is a cubical map. Then F is upper semicontinuous if and only if it has the following property:*

$$\textit{For any } P, Q \in \mathcal{K}(X) \textit{ such that } P \textit{ is a face of } Q,\ F(\overset{\circ}{Q}) \subset F(\overset{\circ}{P}). \tag{14}$$

The above propositions show how to construct lower semicontinuous and upper semicontinuous maps if their values on elementary cells are given. Indeed, if $F(\overset{\circ}{Q})$

is defined for all maximal Q and we put

$$F(\overset{\circ}{P}) = \bigcap\{F(\overset{\circ}{Q}) \mid P \prec Q \text{ and } Q \text{ is maximal}\}, \tag{15}$$

we get a lower semicontinuous map. Similarly, if we put

$$F(\overset{\circ}{P}) = \bigcup\{F(\overset{\circ}{Q}) \mid P \prec Q \text{ and } Q \text{ is maximal}\},$$

we get an upper semicontinuous map. Lower semicontiuous maps are those which we need here for computing the homology. Upper semicontinous maps, however, are those whose graphs are easier to draw since they are closed.

A cubical multivalued map $F : X \overrightarrow{\to} Y$ is called *acyclic* if for every $x \in X$ the set $F(x)$ is acyclic. The next theorem is the central result of this section. Although it can be found in various formulations e.g. in [**1, 2**] and in [**24**] under the name Acyclic Career Theorem, we present it here with a detailed proof since the proof directly serves for the algorithm design.

Theorem 4.4. *Assume $F : X \overrightarrow{\to} Y$ is a lower semicontinuous, acyclic, cubical map. Then, there exists a chain map $\varphi : C(X) \to C(Y)$ with the two properties*

$$|\varphi(\widehat{Q})| \subset F(\overset{\circ}{Q}) \textit{ for all } Q \in \mathcal{K}(X), \tag{16}$$

$$\varphi(\widehat{Q}) \in \widehat{\mathcal{K}}_0(F(Q)) \textit{ for all } Q \in \mathcal{K}_0(X), \tag{17}$$

Proof. We will construct the homomorphisms $\varphi_k : C_k(X) \to C_k(Y)$ by induction in k. For $k < 0$, $C_k(X) = 0$, therefore there is no choice but to define $\varphi_k := 0$.

Consider $k = 0$. For each $Q \in \mathcal{K}_0$, choose $P \in \mathcal{K}_0(F(Q))$ and set

$$\varphi_0(\widehat{Q}) := \widehat{P}. \tag{18}$$

Clearly, $|\varphi_0\widehat{Q}| = P \in F(Q)$. Since, $Q \in \mathcal{K}_0$, $\overset{\circ}{Q} = Q$ and hence $F(Q) = F(\overset{\circ}{Q})$. Therefore $|\varphi_0\widehat{Q}| \subset F(\overset{\circ}{Q})$. Furthermore, $\varphi_{-1}\partial_0 = 0 = \partial_0\varphi_0$.

Suppose now that $\varphi_i : C_i(X) \to C_i(Y)$, $i = 0, 1, 2, \ldots, k-1$, are constructed in such a way that

$$|\varphi_i\widehat{Q}| \subset F(\overset{\circ}{Q}) \text{ for all } Q \in \mathcal{K}_i(K),$$

and

$$\varphi_{i-1}\partial_i = \partial_i\varphi_i. \tag{19}$$

Let $Q \in \mathcal{K}_k(X)$. Then $\partial\widehat{Q} = \sum_{j=1}^{m} \alpha_j\widehat{Q}_j$ for some $\alpha_j \in \mathbf{Z}$ and $Q_j \in \mathcal{K}_{k-1}(X)$. Since F is lsc, we have by Proposition 4.2

$$|\varphi_{k-1}\widehat{Q}_j| \subset F(\overset{\circ}{Q}_j) \subset F(\overset{\circ}{Q})$$

for all $j = 1, \ldots, m$. Thus $|\varphi_{k-1}\partial\widehat{Q}| \subset F(\overset{\circ}{Q})$. Since $F(\overset{\circ}{Q}) = F(x)$ for any $x \in \overset{\circ}{Q}$, the set $F(\overset{\circ}{Q})$ is acyclic. By the induction assumption (19)

$$\partial_{k-1}\varphi_{k-1}\partial_k\widehat{Q} = \varphi_{k-2}\partial_{k-1}\partial_k\widehat{Q} = 0,$$

i.e. $\varphi_{k-1}\partial\widehat{Q}$ is a cycle.

The case $k = 1$ should be distinguished because all 0-chains are 0-cycles and this information is not useful. In that case, Q is an interval and $\partial\widehat{Q} = \widehat{B} - \widehat{A}$ where A and B are vertices of Q. We showed above that the vertices $\varphi_0(A)$ and $\varphi_0(B)$ are supported in $F(\overset{\circ}{Q})$. Since $F(\overset{\circ}{Q})$ is acyclic, it is connected and that this implies the existence of a chain of edges $c \in C_1(F(\overset{\circ}{Q}))$ such that $\partial c = \varphi_0(B) - \varphi_0(A)$. We put

$$\varphi_1\widehat{Q} := c.$$

When $k > 1$, the acyclicity of $F(\overset{\circ}{Q})$ implies that there exists a chain $c \in C_k(F(\overset{\circ}{Q}))$ such that $\partial c = \varphi_{k-1}\partial\widehat{Q}$. Define

$$\varphi_k\widehat{Q} := c.$$

By definition, the homomorphism φ_k satisfies the property

$$\partial_k\varphi_k = \varphi_{k-1}\partial_k.$$

Also, if $Q \in \mathcal{K}_k(X)$, then $\varphi_k\widehat{Q} \in C_k(F(\overset{\circ}{Q}))$, hence $|\varphi_k\widehat{Q}| \subset F(\overset{\circ}{Q})$. This completes the induction step. □

Observe that in the first nontrivial step (18) of the inductive construction of φ we were allowed to choose any $P \in \mathcal{K}_0(F(Q))$. Thus, this procedure allows us to produce many chain maps of the type described in Theorem 4.4. Any such chain map is called a *chain selector* of F.

By carrying over the inclusion (16) to linear combinations of generators of a chain, one proves that, for any $c \in C(X)$ w have $|\varphi(c)| \subset F(|c|)$. The following theorem is proved by standard chain homotopy arguments and induction.

Theorem 4.5. *Let $\varphi, \psi : C(X) \to C(Y)$ be chain selectors for the lower semicontinuous, acyclic, cubical map $F : X \overrightarrow{\to} Y$. Then, φ is chain homotopic to ψ, and hence, they induce the same homomorphism in homology.*

The above theorem permits now to give the following definition. Let $F : X \overrightarrow{\to} Y$ be a lower semicontinuous, acyclic, cubical map. Let $\varphi : C(X) \to C(Y)$ be a chain selector of F. The *homology map* $F_* : H_*(X) \to H_*(Y)$ of F is defined by

$$F_* := \varphi_*.$$

This definition coincides with a classical definition of homology of a multivalued map given via the Vietoris-Begle Theorem e.g. in [**10**]. Here is a conditional functoriality property of homology in the class of acyclic multivalued maps:

Theorem 4.6. *Let $F : X \overrightarrow{\to} Y$ and $G : Y \overrightarrow{\to} Z$ be lower semicontinuous, acyclic, cubical maps. If $G \circ F$ also is acyclic then*

$$(G \circ F)_* = G_* \circ F_* \,,$$

where $G \circ F(x) := G(F(x))$.

4.3. Homology of Continuous Maps Via Cubical Representations

Let X and Y be cubical sets and let $f : X \to Y$ be a continuous function. A *cubical representation* to f or simply a *representation* to f is a lower semicontinuous multivalued cubical map $F : X \overrightarrow{\to} Y$ such that

$$f(x) \in F(x) \tag{20}$$

for every $x \in X$. The *minimal cubical representation* M_f of f is given by

$$M_f(x) := \mathrm{ch}\,(f(\mathrm{ch}\,(x))) \,. \tag{21}$$

It is easily proved that M_f is in fact a representation of f and that it is minimal in the sense that given any other representation F of f, M_f is a submap of F. If F admits an acyclic representation, the Vietoris-Begle Theorem (c.f. [**10**]) implies that the homomorphism $f_* : H_*(X) \to H_*(Y)$ induced by f in homology is $f_* = F_*$. Of course, this is not the case in general as the following example shows.

Example 4.7. Let X be the boundary of the unit square $Q = [0,1]^2$. Consider its sides $K_1 := [0] \times [0,1]$, $K_2 := [0,1] \times [1]$ $K_3 := [1] \times [0,1]$, and $K_4 := [0,1] \times [0]$. Consider the map $\lambda : [0,1] \to X$ given for any $t \in [0,1]$ by

$$\lambda(t) := \begin{cases} (0, 4t) & \text{if } t \in [0, 1/4] \\ (4t-1, 1) & \text{if } t \in [1/4, 1/2] \\ (1, 3-4t) & \text{if } t \in [1/2, 3/4] \\ (4-4t, 0) & \text{if } t \in [3/4, 1] \end{cases}$$

Let $f : X \to X$ be given at $(x_1, x_2) \in X$ by

$$f(x_1, x_2) := \begin{cases} \lambda(x_2) & \text{if } (x_1, x_2) \in K_1 \cup K_3 \\ \lambda(x_1) & \text{if } (x_1, x_2) \in K_2 \cup K_4. \end{cases}$$

Then f is continuous and for $(x_1, x_2) \in \overset{\circ}{K}_i$

$$M_f(x_1, x_2) = \mathrm{ch}\,(f(\mathrm{ch}\,(x_1, x_2))) = \mathrm{ch}\,(f(K_i)) = \mathrm{ch}\,(X) = X.$$

Obviously M_f is not acyclic.

In order to overcome the lack of acyclic representation we could introduce subdivisions of the unitary cubical grid we considered. However, that would require a generalization of all what we have done until now to fractional grids. That is not necessary. Instead we take a natural approach based on rescaling the domain of the function to a larger size by changing units.

A *scaling vector* is a vector of positive integers

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \in \mathbf{Z}^d$$

The corresponding *scaling* is the linear coordinate preserving isomorphism $\Lambda^\alpha : \mathbf{R}^d \to \mathbf{R}^d$ defined by

$$\Lambda^\alpha(x) := (\alpha_1 x_1, \alpha_2 x_2, \ldots \alpha_d x_d).$$

It is easy to see that Λ^α maps cubical sets onto cubical sets. Let $X \subset \mathbf{R}^d$ be a cubical set and let $\alpha \in \mathbf{Z}^d$ be a scaling vector. Define $\Lambda^\alpha_X := \Lambda^\alpha|_X$. The *scaling of*

X by α is

$$X^\alpha := \Lambda^\alpha_X(X) = \Lambda^\alpha(X).$$

The inverse of the map $\Lambda^\alpha_X : X \to X^\alpha$ is $\Omega^\alpha_X : X^\alpha \to X := (\Lambda^\alpha_X)^{-1}$ given by

$$\Omega^\alpha_X(x) := (\alpha_1^{-1}x_1, \alpha_2^{-1}x_2, \ldots, \alpha_d^{-1}x_d).$$

Scalings, their compositions, and inverses are convenient maps in the sense that their minimal cubical representations are acyclic:

Proposition 4.8. *Let X be a cubical set and let α be a scaling vector. The maps $M_{\Lambda^\alpha_X}$ and $M_{\Omega^\alpha_X}$ are acyclic.*
Moreover let X, Y, and Z be cubical sets and let α and β be scaling vectors. If $\Lambda^\alpha(X) \subset Y$ and $\Lambda^\beta(Y) \subset Z$, then $M_{\Lambda^\beta_Y \circ \Lambda^\alpha_X}$ and $M_{\Lambda^\beta_Y} \circ M_{\Lambda^\alpha_X}$ are acyclic.

Since the map $\Lambda^\alpha_X : X \to X^\alpha$ is isomorphism and Ω^α_X its inverse, the maps induced in homology

$$(\Lambda^\alpha_X)_* : H_*(X) \to H_*(X^\alpha) \quad \text{and} \quad (\Omega^\alpha_X)_* : H_*(X^\alpha) \to H_*(X)$$

are also isomorphisms. What is important that, in the view of Proposition 4.8, we can compute the inverse $((\Lambda^\alpha_X)_*)^{-1}$ directly from $M_{\Omega^\alpha_X}$ thus avoiding a complex algorithm of finding an inverse of a homomorphism.

Given a continuous map $f : X \to Y$ and a scaling vector α put

$$f^\alpha := f \circ \Omega^\alpha_X$$

Observe that $f^\alpha : X^\alpha \to Y$. We shall head towards showing that M_{f^α} is acyclic for a sufficiently large α.

Example 4.9. Return to Example 4.7. Consider $\alpha = (2, 2)$, and let $Q = [0, 1] \times [2]$. Let $(x_1, x_2) \in \overset{\circ}{Q}$. Then

$$M_{f^\alpha}(x_1, x_2) = \mathrm{ch}\,(f^\alpha(Q)) = \mathrm{ch}\,(\lambda([0, 1/2])) = [0] \times [0, 1] \cup [0, 1] \times [1]$$

which is acyclic. Similar checks at all segments on X^α shows that M_{f^α} is acyclic.

Proposition 4.10. *Let X and Y be cubical sets and $f : X \to Y$ be continuous. Then there exists a scaling vector α such that M_{f^α} is acyclic.*

Proof. The continuity of f lets us choose $\delta > 0$ such that for $x, y \in X$

$$\mathrm{dist}\,(x, y) \leqslant \delta \quad \Rightarrow \quad \mathrm{dist}\,(f(x), f(y)) \leqslant \frac{1}{2} \tag{22}$$

Let α be a scaling vector such that $\min\{\alpha_i \mid i = 1, \ldots, n\} \geqslant 1/\delta$. Since $\mathrm{diam}\,\mathrm{ch}\,(x) \leqslant 1$, we get from (22) that

$$\mathrm{diam}\, f^\alpha(\mathrm{ch}\,(x)) \leqslant \frac{1}{2}.$$

Let $A = f^\alpha(\mathrm{ch}\,(x))$. We need to show that $\mathrm{ch}\,(A)$ is acyclic. Let

$$\mathcal{C} := \{Q \in \mathcal{K}(Y) \mid \overset{\circ}{Q} \cap A \neq \emptyset\}.$$

Since Y is cubical $\mathrm{ch}\,(A) = \bigcup_{Q\in\mathcal{C}} Q$. Observe that for any two elementary cubes $P, Q \in \mathcal{C}$ the intersection $P \cap Q$ is non-empty, because $\mathrm{diam}\, A < 1$. From this it can be deduced that $\bigcap \mathcal{C}$ is non-empty. It follows that $\mathrm{ch}\,(A)$ is star-shaped and consequently acyclic. □

We can now give the formula for the homology map of a continuous function on cubical set.

Theorem 4.11. *Let $f : X \to Y$ be a continuous function where X and Y are cubical sets. Let α be a scaling vector such that M_{f^α} is acyclic. Then, the* homology map *of $f_* : H_*(X) \to H_*(Y)$ is given by*

$$f_* = M_{(f^\alpha)_*} \circ M_{(\Lambda_X^\alpha)_*}. \tag{23}$$

Proof. The map $M_{(f^\alpha)_*}$ is well defined by the acyclicity assumption and $M_{(\Lambda_X^\alpha)_*}$ is well defined by Proposition 4.8. Since Ω_X^α is the inverse of Λ_X^α, we have

$$f = f \circ \Omega_X^\alpha \circ \Lambda_X^\alpha = f^\alpha \circ \Lambda_X^\alpha$$

and formula (23) follows. □

We conclude this section with the following remark. In [**13**] the homology theory is developed for the category Cub of cubical sets with continuous maps as morphism. The construction of homology map f_* by means of the formula (23) and the chain selector provided by Theorem 4.4 is presented there as the definition of f_*. It is proved there that so obtained homology is a functor from Cub to Ab, the category of abelian groups. Direct proofs of homology axioms, in particular, the homotopy property, are provided. This functor is next extended to the category Pol of topological polyhedra which is defined in [**13**] as the class of topological spaces which are homeomorphic to cubical sets. Since any simplex is homeomorphic to an elementary cube, this class contains, in particular all polyhedra.

4.4. Coboundary Formula

Computation of homology of a continuous map f relies on three algorithms. One computes a cubical representation F of f. The second one does rescaling. The third one computes a chain selector of a cubical acyclic map obtained by rescaling. The proof of Theorem 4.4 which provides the chain selector goes by induction which is very easy to program, it does however contain one existential quantifier:

"The acyclicity of $F(\overset{\circ}{Q})$ implies that there exists a chain $c \in C_k(F(\overset{\circ}{Q}))$ such that $\partial c = \varphi_{k-1}\partial\widehat{Q}$."

One should therefore provide an explicit algorithm for constructing that c. The problem may be formulated as follows. Let $\mathcal{C}$ will be an acyclic finitely generated chain complex. Let $z \in Z_{m-1}$ be a $(m-1)$-cycle. By the acyclicity assumption, there exists a chain $c \in C_m$ such that

$$\partial c = z\ . \tag{24}$$

In general, c is not unique. We want to provide an explicit algorithm for finding at least one such a chain which will be denoted by $\mathrm{Cob}\,(z)$. We shall provide here a summary of the algorithm presented in [**12**]. We refer to [**12**] for proofs and for a presentation of the formal algorithm. The construction of $\mathrm{Cob}\,(z)$ is derived as a by-product of the reduction algorithm presented in Section 3.2.

Lemma 4.12. *Let $a \in C_{m-1}$, $b \in C_m$ and $p : C \to C$ be as in the elementary reduction step in Section 3.2. Let $z \in Z_{m-1}$ be a given cycle. Put $\alpha := \langle z, a\rangle$. Then*
(a) $\bar{z} := p_{m-1}z \in \bar{Z}_{m-1}$.
(b) $\partial(\alpha\lambda^{-1}b) = z - \bar{z}$.

Note that the conclusion of Lemma 4.12 holds true even if $\alpha = 0$ but in that case $\bar{z} = z$. We may want to avoid repetitions of (a, b)-reductions which are not necessary for computing $\mathrm{Cob}\,(z)$. We introduce the following notation. Given any $c \in C_{m-1}$ let

$$E_{m-1}(c) := \{a \in E_{m-1} : \langle c, a\rangle \neq 0\}\ ,\ E_m(c) := \{b \in E_m : \langle \partial b, c\rangle \neq 0\}\ .$$

Lemma 4.13. *Let $z \in Z_{m-1}$, $z \neq 0$. Then $E_{m-1}(z) \neq \emptyset$ and, for any $a \in E_{m-1}(z)$, $E_m(a) \neq \emptyset$.*

We shall now head towards the construction of $\mathrm{Cob}\,(z)$. We start from $c = 0$. Then use recursively (b, a)-reductions simultaneously adding the term $\alpha\lambda^{-1}b$ which appears in Lemma 4.12 to the previous value of c, and replacing the previous value of z by $\bar{z}$. This procedure must end since, each time we repeat it, the cardinality of $\bar{E}_{m-1}$ and $\bar{E}_m$ decreases by one and those sets are finite. When it ends, the final value of $\bar{z}$ is 0, and the final value of c satisfies $\partial c = z - 0 = z$.

To formalize this discussion let us go back to the sequence of projections in Equation 11. Let

$$C^0 \xrightarrow{\bar{p}^1} C^1 \xrightarrow{\bar{p}^2} C^2 \cdots$$

be the sequence of projections and let $(a^k, b^k) \in (E^k_{m-1}, E^k_m)$ be the reduced pairs of generators for a fixed m. Let z^k be defined by

$$z^0 := z\ ,\ z^k := \bar{p}^k z^{k-1}, k = 1, 2, \ldots$$

Since each projection is a chain map, each z^k is a cycle. Thus by Lemma 4.12

$$z^{k-1} - z^k = \partial(\alpha_k\lambda_k^{-1}b^{k-1}) \tag{25}$$

for some $\lambda_k \neq 0$ and some α_k (possibly equal to 0 but then $z^{k-1} = z^k$). We construct a sequence of chains $c^k \in C_m$, $k = 0, 1, 2, \ldots$ as follows.

$$c^0 := 0\ , c^k := c^{k-1} + \alpha_k\lambda_k^{-1}b^k\ , k = 1, 2, \ldots \tag{26}$$

It follows from Equations 25 and 26 by induction that

$$z - z^k = \partial c^k\ , \tag{27}$$

By the same argument as in discussion following Equation 11, there is a finite element C^{f_m} in the sequence beyond which the construction in this dimension cannot be extended and then $\partial^{f_m}_m = 0$ where $\partial^k = 0$ is the restriction of ∂ to C^k. Since C is acyclic, by the proof of Theorem 3.8, C^{f_m} is acyclic too. Since $\partial^{f_m}_m = 0$, we have $\ker \partial^{f_m}_{m-1} = \mathrm{im}\,\partial^{f_m}_m = 0$ hence $z^{f_m} = 0$. From Equation 27 we get the following

Theorem 4.14. $\partial c^{f_m} = z$.

The algorithm based on the above construction is presented in [**12**]. We do not expect that our algorithm, in its most general setting, is more efficient that standard linear algebra programs applied for Equation 24. We expect however that the geometric flavor of our algorithm will help to render it more efficient in particular cases of chain complexes coming from applications. This was the case with the coboundary algorithm for cycles in a rectangular set presented in [**3**]. Let us briefly describe how the improvement is reached there.

Let R be a representable *rectangle* in $\mathbf{R}^d$, i.e. a product of intervals with integer coordinates and $\mathcal{C} := \mathcal{C}(R)$ be the cubical complex of X. The geometric algorithm of finding a coboundary of an $(m-1)$-cycle z given in [**3**] is based on the recurrence with respect to the dimension d of R. Here is the main idea of the recurrence step. Let R' be a $(d-1)$-dimensional face of X. The orthogonal projection of R onto R' induces, in an evident way, a projection $p : \mathcal{C}(R) \to \mathcal{C}(R')$ which is orthogonal with respect to the canonical basis of $\mathcal{C}(R)$ consisting of unitary cubes. We let $z' := pz$ and compute $\operatorname{Cob}(z')$. Then $\operatorname{Cob}(z)$ is obtained by adding to $\operatorname{Cob}(z')$ all m-dimensional unitary cubes through which z is projected (we may visualize them as side-walls of the projection cylinder), with some care about signs of coefficients. It is visible that this construction may be viewed as a major shortcut of what we presented here. Instead of projecting in small steps through one elementary cube per time, we get the whole cylinder enclosed between a face $e \in E_{m-1}(z)$ and a corresponding face of z', in a single operation of replacing an endpoint by an interval in the expression of e as a product of intervals. Due to this simplification, the complexity of the algorithm in [**3**] is linear with respect to the number of elements of $E_{m-1}(z)$ provided an arithmetics on elementary cubes is introduced in a way which permits a convenient presentation of the output. Cubical complexes are very particular but we believe that those type of shortcuts in our universal algorithm will be possible in many situations when a subdivision of a space to cells is chosen in a "custom-fit" way.

We refer the reader to [**12**] for further discussion of cases when the efficiency of the reduction algorithm can be improved.

References

[**1**] M. Allili, Une approche algorithmique pour le calcul de l'homologie de fonctions continues, Ph.D. thesis, Dép. de math. et d'info., Univ. de Sherbrooke, Sherbrooke 1999.

[**2**] M. Allili and T. Kaczynski, An algorithmic approach to the construction of homomorphisms induced by maps in homology, Trans. Amer. Math. Soc., 352 (2000), 2261-2281.

[**3**] M. Allili and T. Kaczynski, Geometric construction of a coboundary of a cycle, Discrete & Comp. Geometry, 25 (2001), 125-140.

[**4**] A. Bovick, ed., *Handbook of Image & Video Processing*, Academic Press, 2000.

[5] Dao Trong Thi and A. T. Fomenko, *Minimal Sufaces, Stratified Multivarifold and the Plateau Problem*, Transl. of Math. Monographs Vol 84, Amer. Math. Soc. 1991.

[6] C. J. A. Delfinado and H. Edelsbrunner, An incremental algorithm for Betti numbers of simplicial complexes on the 3-sphere, Computer Aided Geometric Design 12 (1995), 771-784.

[7] T. K. Dey, H. Edelsbrunner and S. Guha, Computational topology, in *Advances in Computational and Discrete Geometry* (South Hadley, Ma, 1996) Contemp. Math. 223, AMS Providence 1999, 109-143.

[8] R. Ehrenborg and G. Hetyei, Generalizations of Baxter's Theorem and Cubical Homology, J. Combinatorial Theory, Series A, 69 (1995), 233-287.

[9] R. Gonzalez-Diez and P. Real, Computation of cohomology operations on finite simplicial complexes. preprint.

[10] L. Górniewicz, Homological Methods in Fixed Point Theory of Multi-Valued Maps, Dissertationes Math. (129), PWN Warsaw 1976.

[11] C .S. Illiopoulos, Worst case complexity bounds on algorithms for computing the canonical structure of finite abelian groups and Hermite and Smith normal form of an integer matrix, SIAM J. Computing 18 (1989), 658-669.

[12] T. Kaczynski, Recursive Coboundary Formula for Cycles in Acyclic Chain Complexes, Top. Meth. Nonlin. Anal. Appl., 18 (2)(2002), 351-372.

[13] T. Kaczynski, K. Mischaikow, and M. Mrozek, Computational Homology, book in progress.

[14] T. Kaczynski, M. Mrozek, and M. Ślusarek, Homology computation by reduction of chain complexes, Computers and Math. Appl. 35 (1998), 59-70.

[15] W. Kalies, K. Mischaikow, and G. Watson, Cubical approximation and computation of homology, *Conley Index Theory* (eds. K. Mischaikow, M. Mrozek, and P. Zgliczyński), Banach Center Publications 47 Warsaw (1999), 115-131.

[16] W. S. Massey, *A Basic Course in Algebraic Topology*, Springer-Verlag, New York 1991.

[17] M. Mazur, J. Szybowski, *Algorytm Allili-Kaczynskiego Konstrukcji Homomorfizmu £ańcuchowego Generowanego Przez Odwzorowanie Wielowartościowe Reprezentowalne*, M.Sc. Diss. (in polish), Comp. Sci. Inst., Jagellonian Univ. Krakow 1999.

[18] K. Mischaikow, Topological techniques for efficient rigorous computation in dynamics, *Acta Numerica* **11** Cambridge Univ. Press (2002) 435-478.

[19] K. Mischaikow, M. Mrozek, J. Reiss and A. Szymczak, Construction of Symbolic Dynamics from Experimental Time Series, Physical Review Letters, 82(1999), 1144-1148.

[20] K. Mischaikow and M. Mrozek, Conley Index, in *Handbook of Dynamical Systems*, (B. Fiedler, ed.) North-Holland (2002) 393-460.

[21] M. Mrozek, Topological invariants, multivalued maps and computer assisted proofs, Computers and Math. Appl. 32(1996),83-104.

[22] M. Mrozek, An algorithmic approach to the Conley index theory, J. Dyn. Diff. Equ., 11(1999), 711-734.

[23] M. Mrozek, Čech Type Approach to Computing Homology of Maps, in preparation.

[24] J. R. Munkres, *Elements of Algebraic Topology*, Addison-Wesley 1984.

[25] P. Pilarczyk, Algorithmic homology computation over a Euclidean domain, preprint Comp. Sci. Inst. Jagellonian Univ. Krakow 2000.

This article may be accessed via WWW at http://www.rmi.acnet.ge/hha/ or by anonymous ftp at

ftp://ftp.rmi.acnet.ge/pub/hha/volumes/2003/n2a8/v5n2a8.(dvi,ps,pdf)

Tomasz Kaczynski `kaczyn@dmi.usherb.ca`

Département de mathématiques et d'informatique
Université de Sherbrooke
Sherbrooke, Québec, Canada J1K 2R1

Konstantin Mischaikow `mischaik@math.gatech.edu`

Center for Dynamical Systems and Nonlinear Studies
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332-0190

Marian Mrozek `mrozek@ii.uj.edu.pl`

Institute of Computer Science
Jagiellonian University
ul. Nawojki 11
PL-30-072 Kraków,
Poland