19 Homology

In topology, the main focus is not on geometric size but rather on how a space is connected. The most elementary notion distinguishes whether we can go from one place to another. If not then there is a gap we cannot bridge. Next we would ask whether there is a loop going around an obstacle, or whether there is a void missing in the space. Homology is a formalization of these ideas. It gives a way to define and count holes using algebra.

The cyclomatic number of a graph. To motivate the more general concepts, consider a connected graph, G, with n vertices and m edges. A spanning tree has $n-1$ edges and every additional edge forms a unique cycle together with edges in this tree; see Figure 86. Every other

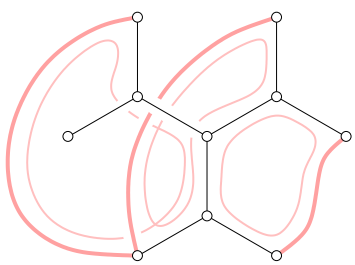

Figure 86: A tree with three additional edges defining the same number of cycles.

cycle in G can be written as a sum of these $m-(n-1)$ cycles. To make this concrete, we define a *cycle* as a subset of the edges such that every vertex belongs to an even number of these edges. A cycle does not need to be connected. The *sum* of two cycles is the symmetric difference of the two sets such that multiple edges erase each other in pairs. Clearly, the sum of two cycles is again a cycle. Every cycle, γ, in G contains some positive number of edges that do not belong to the spanning tree. Calling these edges $e_1, e_2, \ldots, e_k$ and the cycles they define $\gamma_1, \gamma_2, \ldots, \gamma_k$, we claim that

$$\gamma \quad = \quad \gamma_1 + \gamma_2 + \ldots + \gamma_k.$$

To see this assume that $\delta = \gamma_1 + \gamma_2 + \ldots + \gamma_k$ is different from γ. Then $\gamma + \delta$ is again a cycle but it contains no edges that do not belong to the spanning tree. Hence $\gamma + \delta = \emptyset$ and therefore $\gamma = \delta$, as claimed. This implies that the $m-n+1$ cycles form a basis of the group of cycles which motivates us to call $m-n+1$ the *cyclomatic number* of the graph. Note that the basis depends on the choice of spanning tree while the cyclomatic number is independent of that choice.

Simplicial complexes. We begin with a combinatorial representation of a topological space. Using a finite ground set of vertices, V, we call a subset $\sigma \subseteq V$ an *abstract simplex*. Its *dimension* is one less than the cardinality, $\dim \sigma = |\sigma| - 1$. A *face* is a subset $\tau \subseteq \sigma$.

DEFINITION. An *abstract simplicial complex* over V is a system $K \subseteq 2^V$ such that $\sigma \in K$ and $\tau \subseteq \sigma$ implies $\tau \in K$.

The *dimension* of K is the largest dimension of any simplex in K. A graph is thus a 1-dimensional abstract simplicial complex. Just like for graphs, we sometimes think of K as an abstract structure and at other times as a geometric object consisting of geometric simplices. In the latter interpretation, we glue the simplices along shared faces to form a *geometric realization* of K, denoted as $|K|$. We say K *triangulates* a space $\mathbb{X}$ if there is a homeomorphism $h : \mathbb{X} \to |K|$. We have seen 1- and 2-dimensional examples in the preceding sections. The *boundary* of a simplex σ is the collection of co-dimension one faces,

$$\partial\sigma \quad = \quad \{\tau \subseteq \sigma \mid \dim \tau = \dim \sigma - 1\}.$$

If $\dim \sigma = p$ then the boundary consists of $p+1$ $(p-1)$-simplices. Every $(p-1)$-simplex has p $(p-2)$-simplices in its own boundary. This way we get $(p+1)p$ $(p-2)$-simplices, counting each of the $\binom{p+1}{p-1} = \binom{p+1}{2}$ $(p-2)$-dimensional faces of σ twice.

Chain complexes. We now generalize the cycles in graphs to cycles of different dimensions in simplicial complexes. A *p-chain* is a set of p-simplices in K. The *sum* of two p-chains is their symmetric difference. We usually write the sets as formal sums,

$$\begin{aligned} c &= a_1\sigma_1 + a_2\sigma_2 + \ldots + a_n\sigma_n; \\ d &= b_1\sigma_1 + b_2\sigma_2 + \ldots + b_n\sigma_n, \end{aligned}$$

where the a_i and b_i are either 0 or 1. Addition can then be done using modulo 2 arithmetic,

$$c +_2 d \quad = \quad (a_1 +_2 b_1)\sigma_1 + \ldots + (a_n +_2 b_n)\sigma_n,$$

where $a_i +_2 b_i$ is the exclusive or operation. We simplify notation by dropping the subscript but note that the two plus signs are different, one modulo two and the other a formal notation separating elements in a set. The p-chains

form a group, which we denote as $(\mathsf{C}_p, +)$ or simply C_p. Note that the boundary of a p-simplex is a $(p-1)$-chain, an element of C_{p-1}. Extending this concept linearly, we define the boundary of a p-chain as the sum of boundaries of its simplices, $\partial c = a_1 \partial\sigma_1 + \ldots + a_n \partial\sigma_n$. The boundary is thus a map between chain groups and we sometimes write the dimension as index for clarity,

$$\partial_p : \mathsf{C}_p \to \mathsf{C}_{p-1}.$$

It is a homomorphism since $\partial_p(c+d) = \partial_p c + \partial_p d$. The infinite sequence of chain groups connected by boundary homomorphisms is called the *chain complex* of K. All groups of dimension smaller than 0 and larger than the dimension of K are trivial. It is convenient to keep them around to avoid special cases at the ends. A *p-cycle* is a p-chain whose boundary is zero. The sum of two p-cycles is again a p-cycle so we get a subgroup, $\mathsf{Z}_p \subseteq \mathsf{C}_p$. A *$p$-boundary* is a p-chain that is the boundary of a $(p+1)$-chain. The sum of two p-boundaries is again a p-boundary so we get another subgroup, $\mathsf{B}_p \subseteq \mathsf{C}_p$, Taking the boundary twice in a row gives zero for every simplex and thus for every chain, that is, $(\partial_p(\partial_{p+1} d) = 0$. It follows that B_p is a subgroup of Z_p. We can therefore draw the chain complex as in Figure 87.

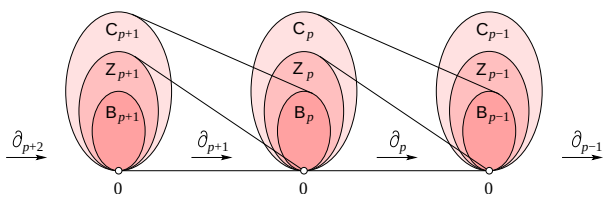

Figure 87: The chain complex consisting of a linear sequence of chain, cycle, and boundary groups connected by homomorphisms.

Homology groups. We would like to talk about cycles but ignore the boundaries since they do not go around a hole. At the same time, we would like to consider two cycles the same if they differ by a boundary. See Figure 88 for a few 1-cycles, some of which are 1-boundaries and some of which are not. This is achieved by taking the quotient of the cycle group and the boundary group. The result is the *p-th homology group*,

$$\mathsf{H}_p = \mathsf{Z}_p / \mathsf{B}_p.$$

Its elements are of the form $[c] = c + \mathsf{B}_p$, where c is a p-cycle. $[c]$ is called a *homology class*, c is a *representative* of $[c]$, and any two cycles in $[c]$ are *homologous* denoted

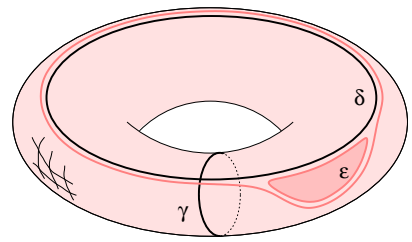

Figure 88: The 1-cycles γ and δ are not 1-boundaries. Adding the 1-boundary ε to δ gives a 1-cycle homologous to δ.

as $c \sim c'$. Note that $[c] = [c']$ whenever $c \sim c'$. Also note that $[c+d] = [c'+d']$ whenever $c \sim c'$ and $d \sim d'$. We use this as a definition of addition for homology classes, so we again have a group. For example, the 1-st homology group of the torus consists of four elements, $[0] = \mathsf{B}_1$, $[\gamma] = \gamma + \mathsf{B}_1$, $[\delta] = \delta + \mathsf{B}_1$, and $[\gamma+\delta] = \gamma + \delta + \mathsf{B}_1$. We often draw the elements as the corners of a cube of some dimension; see Figure 89. If the dimension is β then it has

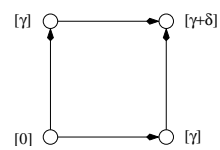

Figure 89: The four homology classes of H_1 are generated by two classes, $[\gamma]$ and $[\delta]$.

2^β corners. The dimension is also the number of classes needed to generate the group, the size of the basis. For the p-th homology group, this number is $\beta_p = \operatorname{rank} \mathsf{H}_p = \log_2 |\mathsf{H}_p|$, the *$p$-th Betti number*. For the torus we have

$$\begin{aligned} \beta_0 &= 1; \\ \beta_1 &= 2; \\ \beta_2 &= 1, \end{aligned}$$

and $\beta_p = 0$ for all $p \neq 0, 1, 2$. Every 0-chain is a 0-cycle. Two 0-cycles are homologous if they are both the sum of an even number or both of an odd number of vertices. Hence $\beta_0 = \log_2 2 = 1$. We have seen the reason for $\beta_1 = 2$ before. Finally, there are only two 2-cycles, namely 0 and the set of all triangles. The latter is not a boundary, hence $\beta_2 = \log_2 2 = 1$.

Boundary matrices. To compute homology groups and Betti numbers, we use a matrix representation of the simplicial complex. Specifically, we store the boundary homomorphism for each dimension, setting $\partial_p[i,j] = 1$ if

the i-th $(p-1)$-simplex is in the boundary of the j-th p-simplex, and $\partial_p[i,j] = 0$, otherwise. For example, if the complex consists of all faces of the tetrahedron, then the boundary matrices are

$$\begin{aligned}
\partial_0 &= \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}; \\
\partial_1 &= \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}; \\
\partial_2 &= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}; \\
\partial_3 &= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.
\end{aligned}$$

Given a p-chain as a column vector, $\mathbf{v}$, its boundary is computed by matrix multiplication, $\partial_p \mathbf{v}$. The result is a combination of columns in the p-th boundary matrix, as specified by $\mathbf{v}$. Thus, $\mathbf{v}$ is a p-cycle iff $\partial_p \mathbf{v} = 0$ and $\mathbf{v}$ is a p-boundary iff there is $\mathbf{u}$ such that $\partial_{p+1}\mathbf{u} = \mathbf{v}$.

Matrix reduction. Letting n_p be the number of p-simplices in K, we note that it is also the rank of the p-th chain group, $n_p = \operatorname{rank} \mathsf{C}_p$. The p-th boundary matrix thus has n_{p-1} rows and n_p columns. To figure the sizes of the cycle and boundary groups, and thus of the homology groups, we reduce the matrix to normal form, as shown in Figure 90. The algorithm of choice uses column and

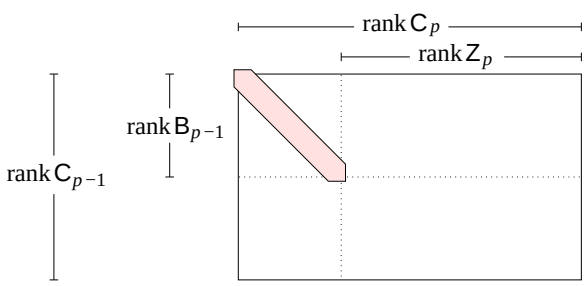

Figure 90: The p-th boundary matrix in normal form. The entries in the shaded portion of the diagonal are 1 and all other entries are 0.

row operations similar to Gaussian elimination for solving a linear system. We write it recursively, calling it with $m = 1$.

```
void REDUCE(m)
  if ∃k, l ≥ m with ∂_p[k, l] = 1 then
    exchange rows m and k and columns m and l;
    for i = m + 1 to n_{p−1} do
      if ∂_p[i, m] = 1 then
        add row m to row i
      endif
    endfor;
    for j = m + 1 to n_p do
      if ∂_p[m, j] = 1 then
        add column m to column j
      endif
    endfor;
    REDUCE(m + 1)
  endif.
```

For each recursive call, we have at most a linear number of row and column operations. The total running time is therefore at most cubic in the number of simplices. Figure 90 shows how we interpret the result. Specifically, the number of zero columns is the rank of the cycle group, Z_p, and the number of 1s in the diagonal is the rank of the boundary group, B_{p-1}. The Betti number is the difference,

$$\beta_p = \operatorname{rank} \mathsf{Z}_p - \operatorname{rank} \mathsf{B}_p,$$

taking the rank of the boundary group from the reduced matrix one dimension up. Working on our example, we get the following reduced matrices.

$$\begin{aligned}
\partial_0 &= \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}; \\
\partial_1 &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}; \\
\partial_2 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \\
\partial_3 &= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.
\end{aligned}$$

Writing $z_p = \operatorname{rank} \mathsf{Z}_p$ and $b_p = \operatorname{rank} \mathsf{B}_p$, we get $z_0 = 4$ from the zeroth and $b_0 = 3$ from the first reduced boundary matrix. Hence $\beta_0 = z_0 = b_0 = 1$. Furthermore,

$z_1 = 3$ and $b_1 = 3$ giving $\beta_1 = 0$, $z_2 = 1$ and $b_2 = 1$ giving $\beta_2 = 0$, and $z_3 = 0$ giving $\beta_3 = 0$. These are the Betti numbers of the closed ball.

Euler-Poincaré Theorem. The *Euler characteristic* of a simplicial complex is the alternating sum of simplex numbers,

$$\chi = \sum_{p \geq 0} (-1)^p n_p.$$

Recalling that n_p is the rank of the p-th chain group and that it equals the rank of the p-th cycle group plus the rank of the $(p-1)$-st boundary group, we get

$$\begin{aligned} \chi &= \sum_{p \geq 0} (-1)^p (z_p + b_{p-1}) \\ &= \sum_{p \geq 0} (-1)^p (z_p - b_p), \end{aligned}$$

which is the same as the alternating sum of Betti numbers. To appreciate the beauty of this result, we need to know that the Betti numbers do not depend on the triangulation chosen for the space. The proof of this property is technical and omitted. This now implies that the Euler characteristic is an invariant of the space, same as the Betti numbers.

EULER-POINCARÉ THEOREM. $\chi = \sum (-1)^p \beta_p$.