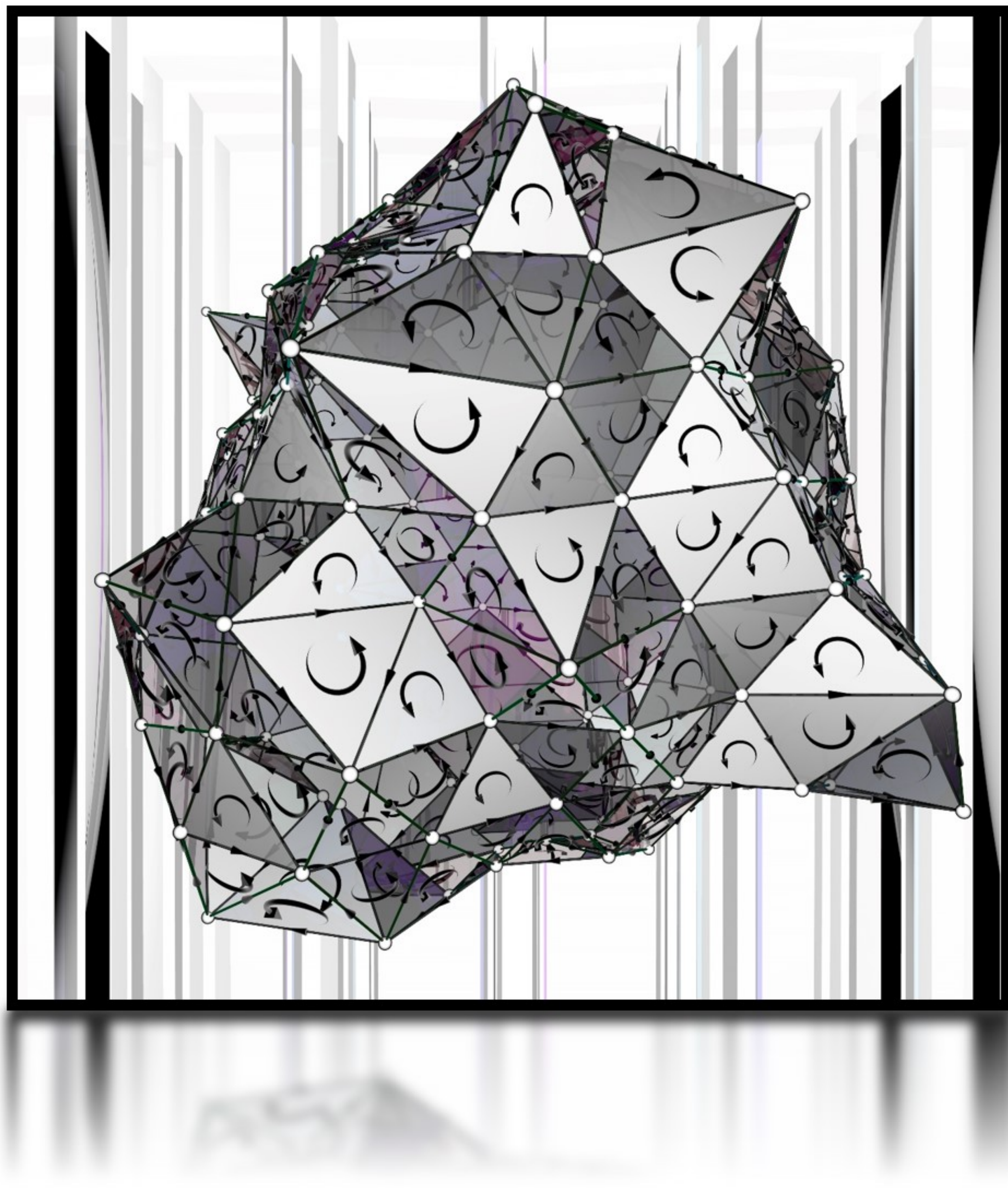
3. HOMOLOGY



3.1 EULER CHARACTERISTIC

Despite all the lemmas and theorems described in the previous Chapter, it remains difficult to explicitly compute – by hand or machine – whether two simplicial complexes are homotopy equivalent or not. As a consequence, we are forced to seek computable **homotopy invariants**; these are assignments $K \mapsto I(K)$ sending each simplicial complex K to some algebraic object $I(K)$ so that the following crucial property is satisfied. If K is homotopy equivalent to some other simplicial complex L, then $I(K)$ is equal (or at least isomorphic in a suitable sense) to $I(L)$. This invariance is necessarily a one-way street: we can not require $I(K) = I(L)$ to imply that K and L are homotopy equivalent, otherwise I-equivalence would be just as difficult to compute as homotopy equivalence. The oldest and simplest homotopy invariant is the **Euler characteristic**, defined as follows.

DEFINITION 3.1. The **Euler characteristic** of a simplicial complex K is the integer $\chi(K) \in \mathbb{Z}$ given by the alternating sum of cardinalities

$$\chi(K) = \sum_{i=0}^{\dim K} (-1)^i \cdot \#K_i.$$

(Here $\#K_i$ indicates the number of all i-dimensional simplices in K).

Reducing an entire simplicial complex to a single integer might seem absurd at first glance, but this simple definition conceals several interesting mysteries. For one thing, it will turn out that χ is a homotopy invariant: if K and L are homotopy equivalent simplicial complexes, then $\chi(K) = \chi(L)$. This fact is by no means obvious, and even verifying it in the case where $L = \mathbf{Sd}\, K$ from the above definition appears painful. Algebraic topology was created to explain why so crude a summary of K should remain entirely unaffected by (combinatorially) enormous perturbations of K which lie within the same homotopy class. For instance, a simple computation reveals that the solid 0-simplex satisfies $\chi(\Delta(0)) = 1$, so by homotopy invariance we can immediately conclude that all contractible simplicial complexes (including $\Delta(k)$ for other choices of $k > 0$) also have $\chi = 1$.

Setting aside this mystery of homotopy invariance for the moment, one may wish to take a moment to marvel at how easily we can compute $\chi(K)$ for a single K. But even here, there are good reasons to tread with caution as described below.

EXAMPLE 3.2. Let $S \subset \mathbb{N}$ denote the set of all **square-free** natural numbers — this consists of all those $n > 1$ which can be expressed as a product of *distinct* prime numbers. Thus, the first few numbers in S are $(2, 3, 5, 6 = 2 \times 3, \ldots)$; note that 4 is excluded because it equals 2^2.

For each $n \geq 1$, let $K(n)$ be the simplicial complex defined on the vertex set $V(n) = \{s \in S \mid s \leq n\}$ by the following rule: the k-simplices for $k > 0$ are all subsets $\{s_0, \ldots, s_k\} \subset V(n)$ of vertices so that each s_i divides the subsequent s_{i+1}. This is clearly a simplicial complex for each integer n, since the divisibility property is preserved when passing to subsets. Moreover, we have an inclusion $K(n) \subset K(n+1)$ for all n, so these simplicial complexes $K(n)$ constitute a filtration of unbounded length. The statement

$$\lim_{n \to \infty} \frac{|\chi(K(n))|}{n^{1/2+\epsilon}} = 0 \text{ for all } \epsilon > 0$$

is equivalent to the Riemann hypothesis. For more information on these simplicial complexes $K(n)$, see A. Bjorner's 2011 paper *A Cell Complex in Number Theory.*

The best way to see that χ is homotopy invariant is to recast it as a numerical reduction of a richer homotopy invariant; this richer invariant is called **homology**, and it forms the main theme of this Chapter.

3.2 Orientations and Boundaries

An **orientation** of a simplicial complex K is an injective function $o : K_0 \to \mathbb{N}$ which assigns unique natural numbers to vertices. The number assigned to each vertex will not be as important as the relative ordering of vertices induced by o, so we may as well require o to take values in the first $\#K_0$ natural numbers. Given an orientation of K, we will always write simplices as *ordered* subsets of K_0 — rather than writing each k-simplex as an unstructured set of vertices, we can uniquely write it as a tuple $(v_0, v_1, \ldots, v_k)$ inside $K_0 \times \cdots \times K_0$ satisfying

$$o(v_0) < o(v_1) < \cdots < o(v_k).$$

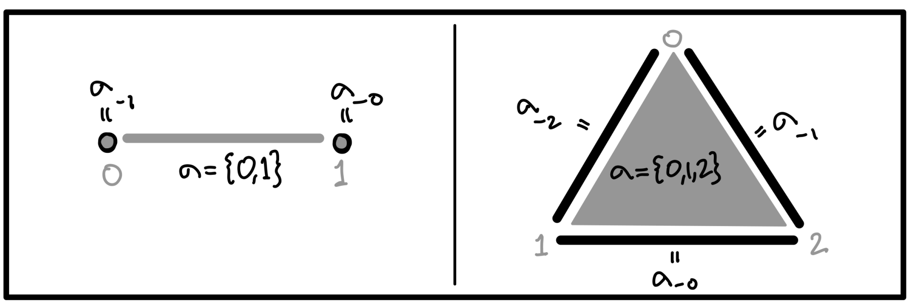
We call this ordered-tuple of vertices an *oriented simplex*.

Definition 3.3. Let K be an oriented simplicial complex and let $\sigma = (v_0, \ldots, v_k)$ be an oriented k-simplex in K. For each i in $\{0, 1, \ldots, k\}$, the ***i*-th face** of σ is the $(k-1)$-dimensional simplex

$$\sigma_{-i} = (v_0, \ldots, v_{i-1}, \cancel{v_i}, v_{i+1}, \ldots, v_d)$$

obtained by removing the i-th vertex.

In the absence of an orientation, there is no coherent way to identify the i-the vertex of σ, so σ_{-i} is not a well-defined simplex — when being explicit about the choice of orientation, one may wish to write the i-th face of σ as σ^o_{-i}. Here are the ordered faces of the top-dimensional simplices of $\Delta(1)$ and $\Delta(2)$ if we assume the standard orientation on the vertices, i.e., $0 < 1 < 2$:

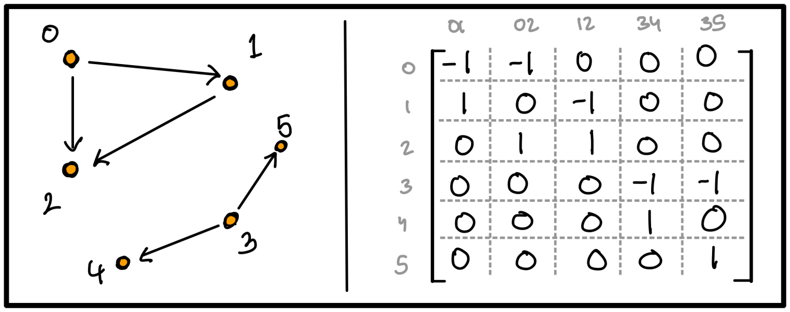


Oriented simplicial complexes form higher-dimensional generalizations of certain types of *directed graphs* — recall that a graph $G = (V, E)$ is directed if each edge comes with a preferred direction, usually indicated as an arrow from one vertex (the source) to the other (the target). The **incidence matrix** $I = I(G)$ of such a graph has the vertices V indexing its rows and edges E indexing its columns; the entry in row v and column e of I is given by the pleasant rule

$$I_{v,e} = \begin{cases} -1 & \text{if } u \text{ is the source of } e, \\ +1 & \text{if } u \text{ is the target of } e, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, each column of I contains exactly one $+1$ and one -1, with all other entries necessarily being zero. In contrast, the rows of I are far less structured; the number of ± 1 entries in each rows depends on the number of edges for which the associated vertex forms a source or target. If we label the vertices of an undirected graph by distinct natural numbers, we automatically induce a directed structure on the edges by forcing sources to have smaller values than targets.

A small example of a directed graph constructed via this vertex-labelling method is depicted below along with its adjacency matrix.



At first glance, the matrix I appears to simply be an alternate way to encode the structure of G in a manner that is specifically tailored to the needs of computers (or equivalently, algebraists). The advantages of this encoding become clearer when I is treated as a linear map — if we let $\mathbb{R}[V]$ and $\mathbb{R}[E]$ be the real vector spaces obtained by treating the vertices and edges of G respectively as orthonormal bases, then I prescribes a linear map $\mathbb{R}[E] \to \mathbb{R}[V]$ defined by the following action on every basis edge e. If e has source vertex u and target vertex v, then

$$I(e) = v - u.$$

Algebraic properties of $I : \mathbb{R}[E] \to \mathbb{R}[V]$ reflect geometric properties of the graph G — for instance, the number of connected components of G is $\#V - \text{rank}(I)$, and the number of undirected cycles in G is $\#E - \text{rank}(I)$.

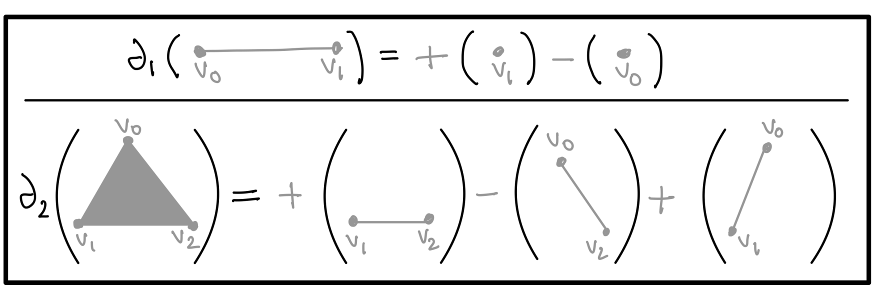
With a simplicial complex, it becomes necessary to define an incidence matrix not only from 1-dimensional simplices to vertices as described above, but from k-dimensional simplices to $(k-1)$-dimensional simplices for *every* dimension k in $\{1, 2, \ldots, \dim K\}$. We therefore require a formula sending each k-simplex to a ± 1 linear combination of its codimension one faces.

DEFINITION 3.4. Let σ be a k-dimensional oriented simplex. The **algebraic boundary** of σ is the linear combination

$$\partial_k \sigma = \sum_{i=0}^{k} (-1)^i \sigma_{-i},$$

where σ_{-i} denotes the i-th face of σ as in Definition 3.3.

Since zero-dimensional simplices have no lower-dimensional faces, we require by convention that ∂_0 of every vertex is 0. The following figure illustrates algebraic boundaries for oriented simplices of dimension one and two:



It will be convenient for the moment to not dwell too much on *where* these ± 1 coefficients are supposed live — we do not demand to know whether they are integers, rational numbers, real numbers, etc. For now we seek solace in the fact that ± 1 are defined in every unital ring (i.e., ring with a multiplicative identity). We will simply treat the algebraic boundary of each simplex

σ as a formal sum of its faces. The first purely algebraic miracle of this subject occurs when we try to compute boundaries of boundaries under the assumption that each ∂_k is a linear map.

PROPOSITION 3.5. *For any oriented simplex σ of dimension $k \geq 0$, we have*
$$\partial_{k-1} \circ \partial_k \sigma = 0.$$

PROOF. The proof of the full statement has been assigned as an exercise, but let's at least compute everything in the case $k = 2$. Consider an oriented 2-simplex $\sigma = (v_0, v_1, v_2)$ and note by Definition 3.4 that
$$\begin{aligned}\partial_2 \sigma &= (\sigma_{-0}) - (\sigma_{-1}) + (\sigma_{-2}) \\ &= (v_1, v_2) - (v_0, v_2) + (v_0, v_1).\end{aligned}$$
Assuming linearity of ∂_1, we have
$$\begin{aligned}\partial_1 \circ \partial_2 \sigma &= \partial_1(v_1, v_2) - \partial_1(v_0, v_2) + \partial_1(v_0, v_1) \\ &= [(v_2) - (v_1)] - [(v_2) - (v_0)] + [(v_1) - (v_0)].\end{aligned}$$
Now the desired conclusion follows by noticing that every vertex has appeared twice, but with opposite signs. □

Since ∂_0 is identically zero, the proposition above has non-trivial content only when $k \geq 2$; thus, this miraculous cancellation remains entirely un-witnessed in the realm of graphs and their incidence matrices.

3.3 CHAIN COMPLEXES

In order to take full advantage of Proposition 3.5, we must fix a *coefficient ring* to give precise meaning to the formal sums obtained when we take algebraic boundaries of simplices. The simplest choice, in terms of computation, is to work with a *field* $\mathbb{F}$ — typical choices include

- $\mathbb{F} = \mathbb{Q}$, the rational numbers;
- $\mathbb{F} = \mathbb{Z}/p$, integers modulo a prime number p, (often $p = 2$), and
- $\mathbb{F} = \mathbb{R}$, the real numbers.

The main advantage when using field coefficients is that we get to work with vector spaces and matrices, so all the standard machinery of linear algebra is at our disposal. With non-field coefficients (even the ring $\mathbb{Z}$ of integers), the algebraic objects at hand become considerably more intricate.

Let K be an oriented simplicial complex and $\mathbb{F}$ a field; both will remain fixed throughout this section.

DEFINITION 3.6. For each dimension $k \geq 0$, the k-th **chain group** of K is the vector space $\mathbf{C}_k(K)$ over $\mathbb{F}$ generated by treating the k-simplices of K as a basis.

Thus, every element γ in $\mathbf{C}_k(K)$ – which is called a k**-chain** of K – can be uniquely expressed as a linear combination of the form
$$\gamma = \sum_\sigma \gamma_\sigma \cdot \sigma,$$
where σ ranges over the k-simplices of K while the coefficients γ_σ are chosen from $\mathbb{F}$. Each k-simplex σ in K constitutes a basis vector in $\mathbf{C}_k(K)$, namely the chain γ whose coefficients are all zero except γ_σ, which equals the multiplicative identity $1 \in \mathbb{F}$. When k exceeds $\dim X$, there are no simplices to serve as basis elements, so $\mathbf{C}_k(X)$ is the trivial (i.e., zero-dimensional) vector space for all large k.

Having described chains as linear combinations of simplices, we are able to reinterpret the algebraic boundaries of Definition 3.4 as linear maps between consecutive chain groups.

DEFINITION 3.7. For each dimension $k \geq 0$, the k-th **boundary operator** of K is the $\mathbb{F}$-linear map $\partial_k^K : \mathbf{C}_k(K) \to \mathbf{C}_{k-1}(K)$ which sends each basis k-chain σ to the $(k-1)$-chain

$$\partial_k^K(\sigma) = \sum_{i=0}^{k} (-1)^i \cdot \sigma_{-i}.$$

In contrast to the formal sums of Definition 3.4, the $(-1)^i$ coefficients appearing in the boundary operator formula above do have a precise meaning — they are simply coefficients chosen from the field $\mathbb{F}$. Note that each σ_{-i} is a $(k-1)$-simplex of K, so the boundary $\partial_k^K(\sigma)$ of each k-simplex σ is automatically a $(k-1)$-chain as expected. To evaluate ∂_k^K on an arbitrary k-chain γ rather than a basis simplex, one simply exploits linearity:

$$\partial_k^K(\gamma) = \sum_{\sigma} \gamma_\sigma \cdot \partial_k^K(\sigma).$$

Here is an immediate consequence of Proposition 3.5

COROLLARY 3.8. *For every dimension $k \geq 0$, the composite*

$$\partial_k^K \circ \partial_{k+1}^K : \mathbf{C}_{k+1}(K) \to \mathbf{C}_{k-1}(K)$$

is the zero map. In other words, for each dimension k the image of ∂_{k+1}^K lies in the kernel of ∂_k^K.

Thus, we now have the ability to build (starting from an oriented simplicial complex K and a coefficient field $\mathbb{F}$) a sequence of finite-dimensional vector spaces connected by linear maps:

$$\cdots \xrightarrow{\partial_{k+1}^K} \mathbf{C}_k(K) \xrightarrow{\partial_k^K} \mathbf{C}_{k-1}(K) \xrightarrow{\partial_{k-1}^K} \cdots \xrightarrow{\partial_2^K} \mathbf{C}_1(K) \xrightarrow{\partial_1^K} \mathbf{C}_0(K) \longrightarrow 0$$

And moreover, this sequence has the magic property that whenever we compose two adjacent maps, the result is always zero. Such sequences play an enormous role in homology theory and beyond, so they have a special name.

DEFINITION 3.9. A **chain complex** $(C_\bullet, d_\bullet)$ over the field $\mathbb{F}$ is a collection of $\mathbb{F}$-vector spaces C_k (indexed by integers $k \geq 0$) and $\mathbb{F}$-linear maps $d_k : C_k \to C_{k-1}$ which satisfy the condition $d_k \circ d_{k+1} = 0$ for all k.

Chain complexes of the form $(\mathbf{C}_\bullet(K), \partial_\bullet^K)$ which arise from a simplicial complex K will be called *simplicial chain complexes* in order to distinguish them from the arbitrary chain complexes $(C_\bullet, d_\bullet)$ of Definition 3.9.

REMARK 3.10. Definition 3.9 works verbatim when we replace the field $\mathbb{F}$ with a commutative, unital ring. For many reasons, the most common choice of non-field coefficients is the ring of integers $\mathbb{Z}$. When working with $\mathbb{Z}$ coefficients, the chain groups C_k are abelian groups rather than vector spaces, and the boundary operators d_k are abelian group homomorphisms; the chain complex condition $d_k \circ d_{k+1} = 0$ makes sense in this context. But now one finds a stark difference between simplicial chain complexes and arbitrary ones: in a simplicial chain complex, each chain group C_k is always *free*, i.e., it has the form $\mathbb{Z}^n$ for some $n \geq 0$. On the other hand, arbitrary chain complexes over $\mathbb{Z}$ can have *torsion* in their chain groups, e.g., $C_k = \mathbb{Z} \oplus \mathbb{Z}/2$ is allowed. Torsion plays no role whatsoever in chain complexes (simplicial or otherwise) when we work with field coefficients.

3.4 HOMOLOGY

Fix a chain complex $(C_\bullet, d_\bullet)$ of vector spaces and linear maps over some field $\mathbb{F}$:

$$\cdots \xrightarrow{d_{k+1}} C_k \xrightarrow{d_k} C_{k-1} \xrightarrow{d_{k-1}} \cdots \xrightarrow{d_2} C_1 \xrightarrow{d_1} C_0 \xrightarrow{d_0} 0$$

We recall from Definition 3.9 that $d_k \circ d_{k+1}$ is always the zero map from C_{k+1} to C_{k-1}, which means that the kernel $\ker d_k$ admits the image $\operatorname{img} d_{k+1}$ as a subspace for each $k \geq 0$.

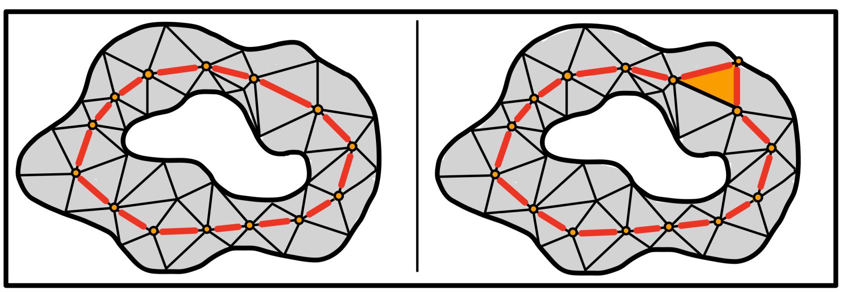
DEFINITION 3.11. For each dimension $k \geq 0$, the k-th **homology group** of $(C_\bullet, d_\bullet)$ is defined to be the quotient vector space

$$\mathbf{H}_k(C_\bullet, d_\bullet) = {\ker d_k}/{\operatorname{img} d_{k+1}}$$

It is customary to refer to $\ker d_k \subset C_k$ as the subspace of k-**cycles** and to $\operatorname{img} d_{k+1}$ as the subspace of k-**boundaries**, so the mantra to chant is

Homology is cycles modulo boundaries.

The best way to become familiar with cycles and boundaries (and hence, with homology groups) is to try drawing them as subsets of simplicial complexes. One can avoid algebraic impediments when building geometric intuition by using $\mathbb{F} = \mathbb{Z}/2$ coefficients, so that it suffices to highlight which simplices have coefficient 1 in a given chain. Illustrated below are two 1-cycles γ and γ' in a triangulated annulus — the key point is that each vertex in sight must be a face of an even number of edges lying in γ (otherwise it will appear with a nonzero coefficient when we take the boundary $\partial_1\gamma$). The cycles γ and γ' represent the same element in the first homology group since they differ only by the boundary of the 2-simplex which has been shaded in the right panel.



When $(C_\bullet, d_\bullet) = (\mathbf{C}_\bullet(K), \partial^K_\bullet)$ is the chain complex associated to a simplicial complex K, the associated homology groups are called the **simplicial homology groups** of K and denoted either $\mathbf{H}_k(K)$ or $\mathbf{H}_k(K;\mathbb{F})$ depending on how emphatically one is trying to showcase the choice of coefficient field. Simplicial homology groups are always finite-dimensional (since we require K to be finite), and for each $k \geq 0$ the dimension

$$\beta_k(K;\mathbb{F}) = \dim \mathbf{H}_k(K;\mathbb{F})$$

is called the k-th **Betti number** of K. As we have chosen to work with field coefficients, this single number completely determines $\mathbf{H}_k(K;\mathbb{F})$ up to isomorphism as a vector space, but it doesn't actually give us a basis of k-chains which generate $\mathbf{H}_k(K;\mathbb{F})$ as a vector space.

EXAMPLE 3.12. The Betti numbers of the solid 0-simplex over any field $\mathbb{F}$ are

$$\beta_k(\Delta_0) = \begin{cases} 1 & k = 0 \\ 0 & k > 0. \end{cases}$$

This can be seen by directly building the simplicial chain complex, which only admits a non-trivial chain group C_k for $k = 0$:

$$\cdots \to 0 \to 0 \to \mathbb{F} \to 0;$$

so all the homology groups are trivial except $\mathbf{H}_0$ which equals $\mathbb{F}$. On the other hand, the hollow 2-simplex has Betti numbers

$$\beta_k(\partial\Delta(2)) = \begin{cases} 1 & k \in \{0,1\} \\ 0 & k > 1. \end{cases}$$

This time the chain complex is non-trivial in dimensions 0 and 1:

$$\cdots \to 0 \to \mathbb{F}^3 \to \mathbb{F}^3 \to 0;$$

and the only non-trivial boundary map $d_1 : \mathbb{F}^3 \to \mathbb{F}^3$ has rank 2. Thus, its kernel has dimension 1 while its image has dimension 2, which means $\beta_0 = 3 - 2 = 1$ and $\beta_1 = 1 - 0 = 1$.

As suggested by these computations, the Betti numbers of K can be determined entirely by the ranks of boundary operators ∂_k^K which appear in the simplicial chain complex.

PROPOSITION 3.13. *Let K be a simplicial complex with K_k denoting the set of all k-dimensional simplices in K. Writing r_k for the rank of the boundary map $\partial_k^K : \mathbf{C}_k(K) \to \mathbf{C}_{k-1}(K)$ from Definition 3.7, for each dimension $k \geq 0$ we have*

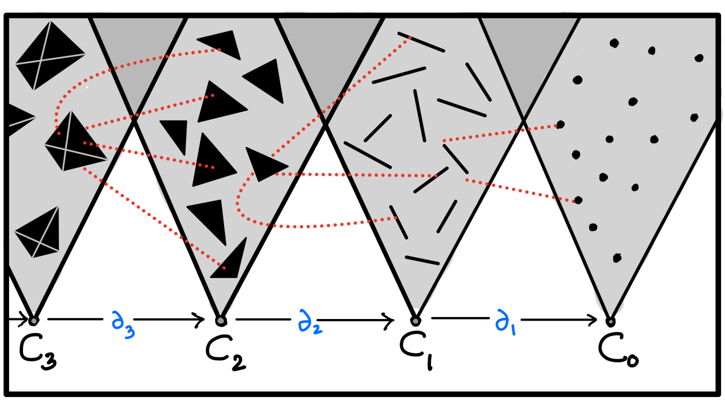
$$\beta_k(K) = \#K_k - (r_k + r_{k+1})$$

PROOF. Since $\beta_k(K)$ is the dimension of the quotient $\ker d_k / \operatorname{img} d_{k+1}$, we have

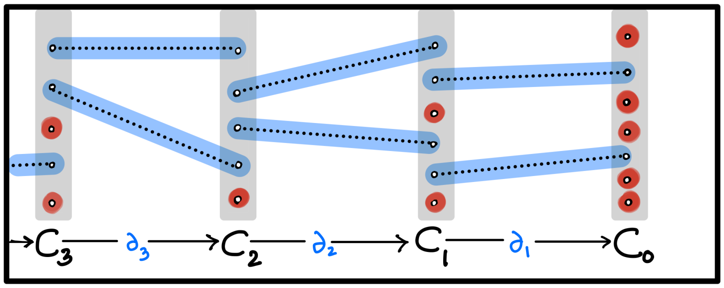
$$\begin{aligned} \beta_k(K) &= \dim(\ker d_k) - \dim(\operatorname{img} d_{k+1}) \\ &= \dim(\ker d_k) - r_{k+1} \end{aligned}$$

By the **rank-nullity theorem** from linear algebra, we have $\dim(\ker d_k) = \dim \mathbf{C}_k(K) - r_k$, and from Definition 3.6 we know $\dim \mathbf{C}_k(K)$ is precisely the number of k-simplices in K. □

In order to go beyond Betti numbers and actually extract *basis elements* for the vector spaces $\mathbf{H}_k(K; \mathbb{F})$, we require more serious linear algebra than the rank-nullity theorem. Before delving into algebraic manipulations, it will be helpful to have the following pictures firmly in mind. The first depicts a simplicial chain complex: for each dimension $k \geq 0$ we have a vector space generated by simplices of dimension k:



The dotted lines in this picture describe nonzero entries in matrix representations of the boundary operators: each basis 1-simplex has exactly two such nonzero entries under its column in ∂_1^K; these lie in the rows corresponding to the two 0-simplices which form its faces. Similarly, each basis 2-simplex has three nonzero entries in its ∂_2^K column, and so on. In this picture, the vector spaces $\mathbf{C}_k(K)$ have a very convenient and geometric description — basis elements are simplices σ, and these clearly form subspaces $|\sigma|$ of the geometric realization $|K|$. But the matrix representations of the boundary operators are a mess — there are lots of dotted lines flying around all over the place, linking each simplex to all of its faces. The key to computing the homology groups of K is to change the bases of all chain groups so that there is at most one incoming or one outgoing dotted line. This produces the following new picture of the same chain complex:



Now the basis elements which generate the chain groups do not have a convenient geometric description — they are weird linear combinations of simplices, and it is not clear how to attach a coherent geometric interpretation within $|K|$ to a linear combination $\gamma = 3\sigma - 5\tau$ of k-dimensional simplices. On the other hand, the boundary matrices have now become gloriously straightforward — there is a trichotomy for each basis element γ of $\mathbf{C}_k(K)$: either

(1) there is a single incoming dotted line to γ from a unique basis element of $\mathbf{C}_{k+1}(K)$, or
(2) there is a single outgoing dotted line from γ to a basis element of $\mathbf{C}_{k-1}(K)$, or
(3) γ remains entirely untouched by dotted lines.

In the first case, γ lies in the image of ∂^K_{k+1} while in the second case γ lies outside the kernel of ∂^K_k. The third case is the most interesting to us, since the basis elements which miss the dotted arrows entirely will simultaneously lie inside the kernel of ∂^K_k *and* outside the image $\operatorname{img} \partial^K_{k+1}$ — these are the desired generators of the homology group $\mathbf{H}_k(K)$.

In the next section, we will describe the algebraic operations which diagonalize the boundary matrices, thus turning the first chain complex picture into the second chain complex picture.

3.5 The Smith Decomposition

Fix a field $\mathbb{F}$ and consider a linear map $A : \mathbb{F}^m \to \mathbb{F}^n$ for some integers $m, n \geq 0$. We fix bases for the domain and codomain so that A has an explicit representation as an $n \times m$ matrix with each entry A_{ij} an element of $\mathbb{F}$.

THEOREM 3.14. **[Smith Normal Form]** *If $A : \mathbb{F}^m \to \mathbb{F}^n$ has rank r, then there exist invertible matrices $P : \mathbb{F}^n \to \mathbb{F}^n$ and $Q : \mathbb{F}^m \to \mathbb{F}^m$ satisfying the matrix equation*

$$D = PAQ,$$

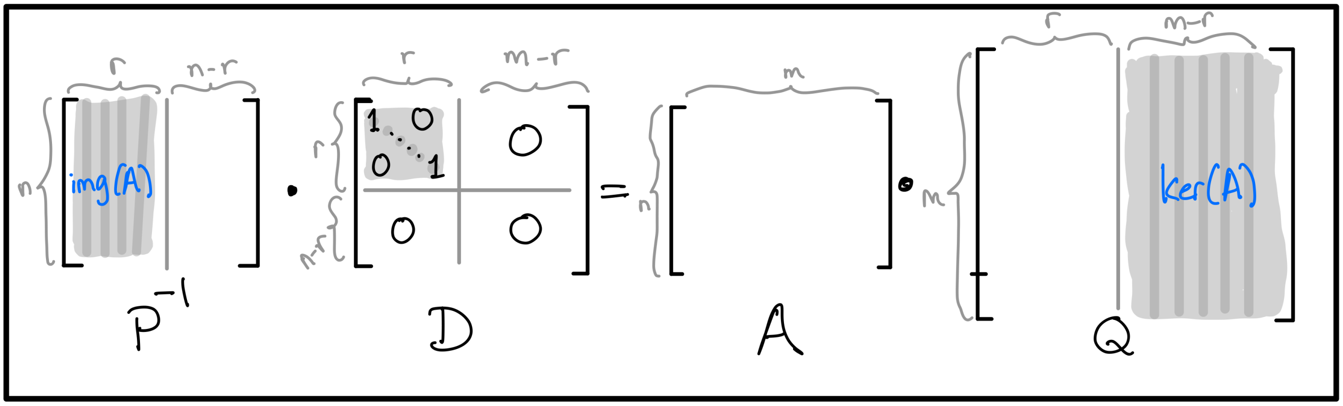
where D is an $n \times m$ diagonal matrix whose entries are given by

$$D_{ij} = \begin{cases} 1 & \text{if } i = j \leq r, \\ 0 & \text{otherwise.} \end{cases}$$

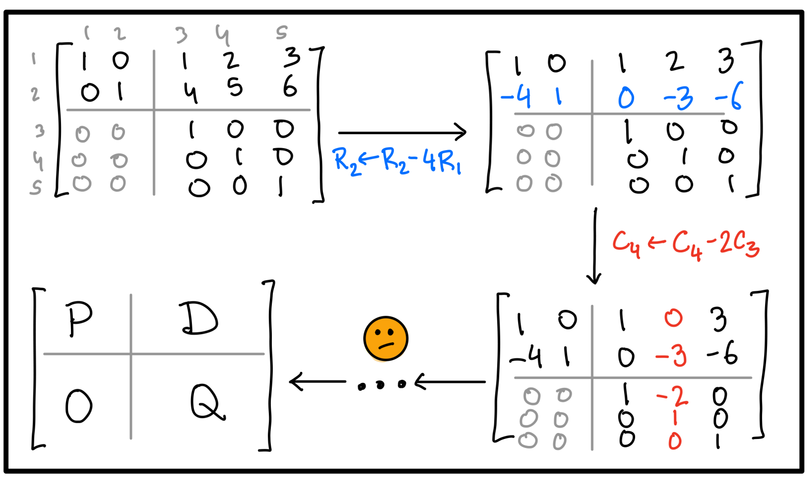
(This matrix D is called the **Smith normal form** *of A.)*

It takes very little imagination (or computation) to produce the matrix D from A, since it only makes use of the numbers m, n and r — one simply creates an $n \times m$ matrix whose leading $r \times r$

minor is the identity matrix $\mathrm{Id}_{r\times r}$ and every other entry is zero. The information that we seek to extract from the *Smith decomposition* $D = PAQ$ is hidden in the invertible matrices P and Q: the first r columns of the inverse P^{-1} form a basis for $\operatorname{img} A \subset \mathbb{F}^n$ while the last $(m-r)$ columns of Q form a basis of $\ker A \subset \mathbb{F}^m$:



The good news is that computing these matrices P and Q from A requires nothing more sophisticated than the sorts of (hopefully familiar) row and column operations which one might use to put matrices in echelon form, namely: (1) add an $\mathbb{F}$-multiple of one row to another row, (2) scale a row by some nonzero element of $\mathbb{F}$, and (3) interchange two rows, plus the three corresponding operations for columns. You can compute everything at once by starting with the block matrix $\begin{bmatrix} \mathrm{Id}_{n\times n} & A \\ 0_{m\times n} & \mathrm{Id}_{m\times m}\end{bmatrix}$. As we perform row and column operations to diagonalize A, the identity matrices below and to the left will evolve accordingly, but the zero block remains unmolested. When A is fully diagonalized, we are conveniently left with $\begin{bmatrix} P & D \\ 0_{m\times n} & Q\end{bmatrix}$. Here are the first two moves of this computation for the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6\end{bmatrix}$, assuming that $\mathbb{F}$ is the field of rational numbers:



Let $(C_\bullet, d_\bullet)$ be a chain complex over $\mathbb{F}$ so that each C_k is finite-dimensional. Knowledge of the Smith decomposition of all the boundary operators d_k allows us to find basis vectors for the homology groups $\mathbf{H}_k(C_\bullet, d_\bullet)$ as follows.

PROPOSITION 3.15. *For each dimension $k \geq 0$, let $D_k = P_k d_k Q_k$ be the Smith decomposition of the boundary operator $d_k : C_k \to C_{k-1}$. For $n_k = \dim C_k$ and $r_k = \operatorname{rank} d_k$, let G_k be the matrix of size $(r_{k+1} + n_k - r_k) \times n_k$ given by the block decomposition*

$$G_k = [B_k \mid Z_k],$$

where B_k equals the first r_{k+1} columns of P_{k+1}^{-1} while Z_k equals the last $(n_k - r_k)$ columns of Q_k. If $E_k = [B_k' \mid Z_k']$ is the reduced row echelon form of G_k obtained by performing row (but not column) operations over $\mathbb{F}$, then the columns of Z_k corresponding to the pivot columns of Z_k' form a basis for $\mathbf{H}_k(C_\bullet, d_\bullet)$.

PROOF. As discussed above, for each $k \geq 0$ the matrix B_k contains a basis for $\operatorname{img} d_{k+1}$, and so every column in the left block B_k' of E_k is guaranteed to have a pivot. Recall that $\operatorname{img} d_{k+1}$ is a subspace of $\ker d_k$ by Definition 3.9, and that the right block Z_k contains a basis for $\ker d_k$. Thus, there will be exactly $(n_k - r_k - r_{k+1})$ pivot columns in Z_k' once we are in row echelon form, and the corresponding columns of Z_k provide a basis for the quotient $\ker d_k / \operatorname{img} d_{k+1} = \mathbf{H}_k(C_\bullet, d_\bullet)$. □

There is nothing unique about bases obtained via the procedure above — for one thing, we can always add vectors from $\operatorname{img} d_{k+1}$ to a basis vector to get a new basis vector. More seriously, we could take interesting linear combinations, e.g., replace basis vectors $\{\gamma_1, \gamma_2\}$ by $\{\gamma_1 + \gamma_2, \gamma_1 - \gamma_2\}$.

REMARK 3.16. Theorem 3.14 also holds when A is an integer matrix $\mathbb{Z}^m \to \mathbb{Z}^n$; in this case, both P and Q will be invertible integer matrices. But here the Smith normal form D has a more interesting structure; instead of a leading $r \times r$ identity matrix, we have a diagonal matrix populated by positive integers $a_1 \leq \cdots \leq a_r$, called the **invariant factors** of A. Each a_i divides the subsequent a_{i+1}, and whenever $a_i \neq 1$ we obtain torsion of the form $\mathbb{Z}/a_i$ in the integral homology groups.

We have not yet explained what homology groups have to do with the fact that the Euler characteristic is homotopy invariant. But we are able to take the first steps in the right direction: one of the exercises below asks you to confirm that $\chi(K)$ for a simplicial complex K can be completely recovered from its rational Betti numbers $\{\beta_k(K;\mathbb{Q}) \mid k \geq 0\}$. In the next Chapter, we will develop enough machinery to see that homology groups, and hence Betti numbers, are themseleves homotopy invariant.

EXERCISES

EXERCISE 3.1. What is the Euler characteristic $\chi(\partial\Delta(k))$ of the hollow k-simplex as a function of the dimension $k \geq 1$?

EXERCISE 3.2. For any sequence $a = (a_0, a_1, \ldots)$ of real numbers and simplicial complex K, define the real number $\chi_a(K)$ by

$$\chi_a(K) = \sum_{i=0}^{\dim K} a_i \cdot \#K_i.$$

Show that if $\chi_a(\Delta(k))$ is constant for all $k \geq 0$ then there exists some real number λ satisfying $a_i = \lambda \cdot (-1)^i$, so χ_a must be a scalar multiple of the Euler characteristic.

EXERCISE 3.3. Show that the Euler characteristic of a simplicial complex equals the alternating sum of its Betti numbers over $\mathbb{Q}$, i.e.,

$$\chi(K) = \sum_{i=0}^{\dim K} \beta_i(K;\mathbb{Q}).$$

[Hint: use Proposition 3.13]

EXERCISE 3.4. Prove Proposition 3.5 by extending the argument for $k = 2$ to arbitrary k. [Hint: evaluate $\partial_{k-1} \circ \partial_k \sigma$ as follows:

$$\partial_{k-1}\left(\sum_{i=0}^{k}(-1)^i\sigma_{-i}\right) = \sum_{i=0}^{k}(-1)^i \cdot \partial_{k-1}\sigma_{-i} = \sum_{i=0}^{k}(-1)^i \cdot \sum_{j=0}^{k-1}(-1)^j(\sigma_{-i})_{-j}.$$

Now decompose the double-sum into the parts where $j > i$ and $j < i$.]

EXERCISE 3.5. Let K be a one-dimensional oriented simplicial complex (i.e., a directed graph). Describe the simplicial chain complex of K in terms of its incidence matrix I.

EXERCISE 3.6. For K a one-dimensional oriented simplicial complex and coefficients in $\mathbb{F} = \mathbb{Z}/2$, show that $\ker \partial_1^K$ must consist entirely of cycles, i.e., paths which start and end at the same vertex.

EXERCISE 3.7. Let K be a simplicial complex and L a subcomplex so that $K - L$ only contains simplices of dimension k or above. Prove that $\beta_i(K) = \beta_i(L)$ for all $i < k - 1$.

EXERCISE 3.8. Write down the simplicial chain complex for the solid simplex $\Delta(2)$. Determine the ranks of the boundary operators and hence the Betti numbers of $\Delta(2)$.

EXERCISE 3.9. Compute the Smith decomposition $D = PAQ$ for the matrix $A = \left[\begin{smallmatrix} 1 & 2 & 1 \\ -2 & -4 & -2 \end{smallmatrix}\right]$ over $\mathbb{Q}$; use this to write down basis vectors for its kernel and image.

EXERCISE 3.10. Find bases for $\mathbf{H}_0$ and $\mathbf{H}_1$ of the hollow 2-simplex $\partial\Delta(2)$ with $\mathbb{F} = \mathbb{Z}/2$.

EXERCISE 3.11. A simplicial complex is *connected* if any two vertices u and v can be joined by a path of consecutive edges, i.e.,

$$u \leftrightarrow w_0 \leftrightarrow w_1 \leftrightarrow \cdots \leftrightarrow w_k \leftrightarrow v$$

Show that $\beta_0(K;\mathbb{Q}) = 1$ whenever K is connected.