Chapter 4

The Fundamental Group, Orientability

4.1 The Fundamental Group

If we want to somehow classify surfaces, we have to deal with the issue of deciding when we consider two surfaces to be equivalent. It seems reasonable to treat homeomorphic surfaces as equivalent, but this leads to the problem of deciding when two surfaces are not homeomorphic, which is a very difficult problem. One way to approach this problem is to forget some of the topological structure of a surface and look for more algebraic objects that can be associated with a surface. For example, we can consider closed curves on a surface, and see how they can be deformed. It is also fruitful to give an algebraic structure to appropriate sets of closed curves on a surface, for example, a group structure. Two important tools for studying surfaces were invented by Poincaré, the fundamental group, and the homology groups. In this section, we take a look at the fundamental group.

Roughly speaking, given a topological space E and some chosen point $a \in E$, a group $\pi(E, a)$ called the fundamental group of E based at a is associated with (E, a), and to every continuous map $f\colon (X, x) \to (Y, y)$ such that $f(x) = y$, is associated a group homomorphism $f_*\colon \pi(X, x) \to \pi(Y, y)$. Thus, certain topological questions about the space E can translated

Figure 4.1: Henri Poincaré, 1854-1912

into algebraic questions about the group $\pi(E, a)$. This is the paradigm of algebraic topology. In this section, we will focus on the concepts rather than dwelve into technical details. For a thorough presentation of the fundamental group and related concepts, the reader is referred to Massey [32, 33], Munkres [37], Bredon [7], Hatcher [20], Dold [13], Fulton [18] and Rotman [41]. We also recommend Sato [42] for an informal and yet very clear presentation.

The intuitive idea behind the fundamental group is that closed paths on a surface reflect some of the main topological properties of the surface. Actually, the idea applies to any topological space E. Let us choose some point a in E (a *base point*), and consider all closed curves $\gamma\colon [0,1] \to E$ based at a, that is, such that $\gamma(0) = \gamma(1) = a$. We can compose closed curves γ_1, γ_2 based at a, and consider the inverse γ^{-1} of a closed curve, but unfortunately, the operation of composition of closed curves is not associative, and $\gamma\gamma^{-1}$ is not the identity in general. In order to obtain a group structure, we define a notion of equivalence of closed curves under continuous deformations. Actually, such a notion can be defined for any two paths with the same origin and extremity, and even for continuous maps.

Definition 4.1 Given any two paths $\gamma_1\colon [0,1] \to E$ and $\gamma_2\colon [0,1] \to E$ with the same intial point a and the same terminal point b, i.e., such that $\gamma_1(0) = \gamma_2(0) = a$, and $\gamma_1(1) = \gamma_2(1) = b$, a map $F\colon [0,1] \times [0,1] \to E$ is a *(path) homotopy* between γ_1 and γ_2 if F is continuous, and if

$$F(t,0) = \gamma_1(t),$$
$$F(t,1) = \gamma_2(t),$$

for all $t \in [0,1]$, and

$$F(0,u) = a,$$
$$F(1,u) = b,$$

for all $u \in [0,1]$. In this case, we say that γ_1 and γ_2 are *homotopic*, and this is denoted as $\gamma_1 \approx \gamma_2$.

Given any two continuous maps $f_1\colon X \to Y$ and $f_2\colon X \to Y$ between two topological spaces X and Y, a map $F\colon X \times [0,1] \to Y$ is a *homotopy* between f_1 and f_2 iff F is continuous and

$$F(t,0) = f_1(t),$$
$$F(t,1) = f_2(t),$$

for all $t \in X$. We say that f_1 and f_2 are homotopic, and this is denoted as $f_1 \approx f_2$.

Intuitively, a (path) homotopy F between two paths γ_1 and γ_2 from a to b is a continuous family of paths $F(t,u)$ from a to b, giving a deformation of the path γ_1 into the path γ_2. It is easily shown that homotopy is an equivalence relation on the set of paths from a to b. A simple example of homotopy is given by reparameterizations. A continuous nondecreasing function $\tau\colon [0,1] \to [0,1]$ such that $\tau(0) = 0$ and $\tau(1) = 1$ is called a *reparameterization*. Then, given a path $\gamma\colon [0,1] \to E$, the path $\gamma \circ \tau\colon [0,1] \to E$ is homotopic to $\gamma\colon [0,1] \to E$, under the homotopy

$$(t, u) \mapsto \gamma((1-u)t + u\tau(t)).$$

As another example, any two continuous maps $f_1\colon X \to \mathbb{A}^2$ and $f_2\colon X \to \mathbb{A}^2$ with range the affine plane $\mathbb{A}^2$ are homotopic under the homotopy defined such that

$$F(t, u) = (1-u)f_1(t) + uf_2(t).$$

However, if we remove the origin from the plane $\mathbb{A}^2$, we can find two paths γ_1 and γ_2 in $\mathbb{A}^2 - \{(0,0)\}$, from $(-1, 0)$ to $(1, 0)$ that are not homotopic. For example, we can consider the upper half unit circle, and the lower half unit circle. The problem is that the "hole" created by the missing origin prevents continuous deformation of one path into the other. Thus, we should expect that homotopy classes of closed curves on a surface contain information about the presence or absence of "holes" in a surface.

It is easily verified that if $\gamma_1 \approx \gamma_1'$ and $\gamma_2 \approx \gamma_2'$, then $\gamma_1\gamma_2 \approx \gamma_1'\gamma_2'$, and that $\gamma_1^{-1} \approx \gamma_1'^{-1}$. Thus, it makes sense to define the composition and the inverse of homotopy classes.

Definition 4.2 Given any topological space, E, for any choice of a point $a \in E$ (a *base point*), the *fundamental group (or Poincaré group), $\pi(E, a)$, at the base point a* is the set of homotopy classes of closed curves, $\gamma\colon [0,1] \to E$, such that $\gamma(0) = \gamma(1) = a$, under the multiplication operation, $[\gamma_1][\gamma_2] = [\gamma_1\gamma_2]$, induced by the composition of closed paths based at a.

One actually needs to prove that the above multiplication operation is associative, has the homotopy class of the constant path equal to a as an identity, and that the inverse of the homotopy class $[\gamma]$ is the class $[\gamma^{-1}]$. The first two properties are left as an exercise, and the third property uses the homotopy

$$F(t, u) = \begin{cases} \gamma(2t) & \text{if } 0 \leq t \leq (1-u)/2; \\ \gamma(1-u) & \text{if } (1-u)/2 \leq t \leq (1+u)/2; \\ \gamma(2-2t) & \text{if } (1+u)/2 \leq t \leq 1. \end{cases}$$

As defined, the fundamental group depends on the choice of a base point. Let us now assume that E is arcwise connected (which is the case for surfaces). Let a and b be any two distinct base points. Since E is arcwise connected, there is some path α from a to b. Then, to every closed curve γ based at a corresponds a close curve $\gamma' = \alpha^{-1}\gamma\alpha$ based at b. It is easily verified that this map induces a homomorphism $\varphi\colon \pi(E, a) \to \pi(E, b)$ between the groups $\pi(E, a)$ and $\pi(E, b)$. The path α^{-1} from b to a induces a homomorphism $\psi\colon \pi(E, b) \to \pi(E, a)$ between the groups $\pi(E, b)$ and $\pi(E, a)$. Now, it is immediately verified that $\varphi \circ \psi$ and $\psi \circ \varphi$ are both the identity, which shows that the groups $\pi(E, a)$ and $\pi(E, b)$ are isomorphic.

Thus, when the space E is arcwise connected, the fundamental groups $\pi(E,a)$ and $\pi(E,b)$ are isomorphic for any two points $a,b\in E$.

Remarks:

(1) The isomorphism $\varphi\colon \pi(E,a)\to\pi(E,b)$ is not canonical, that is, it depends on the chosen path α from a to b.

(2) In general, the fundamental group $\pi(E,a)$ is not commutative.

When E is arcwise connected, we allow ourselves to refer to any of the isomorphic groups $\pi(E,a)$ as *the* fundamental group of E, and we denote any of these groups by $\pi(E)$.

The fundamental group, $\pi(E,a)$, is in fact one of several homotopy groups, $\pi_n(E,a)$, associated with a space, E, and $\pi(E,a)$ is often denoted by $\pi_1(E,a)$. However, we won't have any use for the more general homotopy groups.

If E is an arcwise connected topological space, it may happen that some fundamental group, $\pi(E,a)$, is reduced to the trivial group, $\{1\}$, consisting of the identity element. It is easy to see that this is equivalent to the fact that for any two points $a,b\in E$, any two paths from a to b are homotopic, and thus, the fundamental groups, $\pi(E,a)$, are trivial for all $a\in E$. This is an important case, which motivates the following definition.

Definition 4.3 A topological space E is *simply-connected* if it is arcwise connected and for every $a\in E$, the fundamental group $\pi(E,a)$ is the trivial one-element group.

For example, the plane and the sphere are simply connected, but the torus is not simply connected (due to its hole).

We now show that a continuous map between topological spaces (with base points) induces a homomorphism of fundamental groups. Given two topological spaces X and Y, given a base point x in X and a base point y in Y, for any continuous map $f\colon (X,x)\to(Y,y)$ such that $f(x)=y$, we can define a map $f_*\colon \pi(X,x)\to\pi(Y,y)$ as follows:

$$f_*([\gamma]) = [f\circ\gamma],$$

for every homotopy class $[\gamma]\in\pi(X,x)$, where $\gamma\colon [0,1]\to X$ is a closed path based at x.

It is easily verified that f_* is well defined, that is, does not depend on the choice of the closed curve γ in the homotopy class $[\gamma]$. It is also easily verified that $f_*\colon \pi(X,x)\to\pi(Y,y)$ is a homomorphism of groups. The map $f\mapsto f_*$ also has the following important two properties. For any two continuous maps $f\colon (X,x)\to(Y,y)$ and $g\colon (Y,y)\to(Z,z)$, such that $f(x)=y$ and $g(y)=z$, we have

$$(g\circ f)_* = g_*\circ f_*,$$

and if $Id\colon (X,x)\to(X,x)$ is the identity map, then $Id_*\colon \pi(X,x)\to\pi(X,x)$ is the identity homomorphism.

As a consequence, if $f\colon (X, x) \to (Y, y)$ is a homeomorphism such that $f(x) = y$, then $f_*\colon \pi(X, x) \to \pi(Y, y)$ is a group isomorphism. This gives us a way of proving that two spaces are not homeomorphic: show that for some appropriate base points $x \in X$ and $y \in Y$, the fundamental groups $\pi(X, x)$ and $\pi(Y, y)$ are not isomorphic.

In general, it is difficult to determine the fundamental group of a space. We will determine the fundamental group of $\mathbb{A}^n$ and of the punctured plane. For this, we need the concept of the winding number of a closed curve in the plane.

4.2 The Winding Number of a Closed Plane Curve

Consider a closed curve, $\gamma\colon [0, 1] \to \mathbb{A}^2$, in the plane, and let z_0 be a point not on γ. In what follows, it is convenient to identify the plane $\mathbb{A}^2$ with the set $\mathbb{C}$ of complex numbers. We wish to define a number, $n(\gamma, z_0)$, which counts how many times the closed curve γ winds around z_0.

We claim that there is some real number $\rho > 0$ such that $|\gamma(t) - z_0| > \rho$ for all $t \in [0, 1]$. If not, then for every integer $n \geq 0$, there is some $t_n \in [0, 1]$ such that $|\gamma(t_n) - z_0| \leq 1/n$. Since $[0, 1]$ is compact, the sequence (t_n) has some convergent subsequence (t_{n_p}) having some limit $l \in [0, 1]$. But then, by continuity of γ, we have $\gamma(l) = z_0$, contradicting the fact that z_0 is not on γ. Now, again since $[0, 1]$ is compact and γ is continuous, γ is actually uniformly continuous. Thus, there is some $\epsilon > 0$ such that $|\gamma(t) - \gamma(u)| \leq \rho$ for all $t, u \in [0, 1]$, with $|u - t| \leq \epsilon$. Letting n be the smallest integer such that $n\epsilon > 1$, and letting $t_i = i/n$, for $0 \leq i \leq n$, we get a subdivision of $[0, 1]$ into subintervals, $[t_i, t_{i+1}]$, such that $|\gamma(t) - \gamma(t_i)| \leq \rho$ for all $t \in [t_i, t_{i+1}]$, with $0 \leq i \leq n - 1$.

For every $i, 0 \leq i \leq n - 1$, if we let

$$w_i = \frac{\gamma(t_{i+1}) - z_0}{\gamma(t_i) - z_0},$$

it is immediately verified that $|w_i - 1| < 1$, and thus, w_i has a positive real part. Thus, there is a unique angle, θ_i, with $-\frac{\pi}{2} < \theta_i < \frac{\pi}{2}$, such that $w_i = \lambda_i(\cos\theta_i + i\sin\theta_i)$, where $\lambda_i > 0$. Furthermore, because γ is a closed curve,

$$\prod_{i=0}^{n-1} w_i = \prod_{i=0}^{n-1} \frac{\gamma(t_{i+1}) - z_0}{\gamma(t_i) - z_0} = \frac{\gamma(t_n) - z_0}{\gamma(t_0) - z_0} = \frac{\gamma(1) - z_0}{\gamma(0) - z_0} = 1,$$

and the angle $\sum \theta_i$ is an integral multiple of 2π. Thus, for every subdivision of $[0, 1]$ into intervals $[t_i, t_{i+1}]$ such that $|w_i - 1| < 1$, with $0 \leq i \leq n - 1$, we define the *winding number, $n(\gamma, z_0)$, or index, of γ with respect to z_0*, as

$$n(\gamma, z_0) = \frac{1}{2\pi} \sum_{i=0}^{i=n-1} \theta_i.$$

Actually, in order for $n(\gamma, z_0)$ to be well defined, we need to show that it does not depend on the subdivision of $[0,1]$ into intervals $[t_i, t_{i+1}]$ (such that $|w_i - 1| < 1$). Since any two subdivisions of $[0,1]$ into intervals $[t_i, t_{i+1}]$ can be refined into a common subdivision, it is enough to show that nothing is changed is we replace any interval $[t_i, t_{i+1}]$ by the two intervals $[t_i, \tau]$ and $[\tau, t_{i+1}]$. Now, if θ_i' and θ_i'' are the angles associated with

$$\frac{\gamma(t_{i+1}) - z_0}{\gamma(\tau) - z_0},$$

and

$$\frac{\gamma(\tau) - z_0}{\gamma(t_i) - z_0},$$

we have

$$\theta_i = \theta_i' + \theta_i'' + k2\pi,$$

where k is some integer. However, since $-\frac{\pi}{2} < \theta_i < \frac{\pi}{2}$, $-\frac{\pi}{2} < \theta_i' < \frac{\pi}{2}$, and $-\frac{\pi}{2} < \theta_i'' < \frac{\pi}{2}$, we must have $|k| < \frac{3}{4}$, which implies that $k = 0$, since k is an integer. This shows that $n(\gamma, z_0)$ is well defined.

The next two propositions are easily shown using the above technique. Proofs can be found in Ahlfors and Sario [1].

Proposition 4.1 *For every plane closed curve, $\gamma\colon [0,1] \to \mathbb{A}^2$, for every z_0 not on γ, the index $n(\gamma, z_0)$ is continuous on the complement of γ in $\mathbb{A}^2$ and, in fact, constant in each connected component of the complement of γ. We have $n(\gamma, z_0) = 0$ in the unbounded component of the complement of γ.*

Proposition 4.2 *For any two plane closed curve, $\gamma_1\colon [0,1] \to \mathbb{A}^2$ and $\gamma_2\colon [0,1] \to \mathbb{A}^2$, for every homotopy, $F\colon [0,1] \times [0,1] \to \mathbb{A}^2$, between γ_1 and γ_2, for every z_0 not on any $F(t,u)$, for all $t, u \in [0,1]$, we have $n(\gamma_1, z_0) = n(\gamma_2, z_0)$.*

Proposition 4.2 shows that the index of a closed plane curve is not changed under homotopy (provided that none the curves involved go through z_0). We can now compute the fundamental group of the punctured plane, i.e., the plane from which a point is deleted.

4.3 The Fundamental Group of the Punctured Plane

First, we note that the fundamental group of $\mathbb{A}^n$ is the trivial group. Indeed, consider any closed curve $\gamma\colon [0,1] \to \mathbb{A}^n$ through $a = \gamma(0) = \gamma(1)$, take a as base point, and let a be the constant closed curve reduced to a. Note that the map

$$(t, u) \mapsto (1 - u)\gamma(t)$$

is a homotopy between γ and a. Thus, there is a single homotopy class $[a]$, and $\pi(\mathbb{A}^n, a) = \{1\}$.

The above reasoning also shows that the fundamental group of an open ball, or a closed ball, is trivial. However, the next proposition shows that the fundamental group of the punctured plane is the infinite cyclic group $\mathbb{Z}$.

Proposition 4.3 *The fundamental group of the punctured plane is the infinite cyclic group $\mathbb{Z}$.*

Proof. Assume that the origin $z = 0$ is deleted from $\mathbb{A}^2 = \mathbb{C}$, and take $z = 1$ as base point. The unit circle can be parameterized as $t \mapsto \cos t + i \sin t$, and let α be the corresponding closed curve. First of all, note that for every closed curve $\gamma\colon [0,1] \to \mathbb{A}^2$ based at 1, there is a homotopy (central projection) $F\colon [0,1] \times [0,1] \to \mathbb{A}^2$ deforming γ into a curve β lying on the unit circle. By uniform continuity, any such curve β can be decomposed as $\beta = \beta_1\beta_2 \cdots \beta_n$, where each β_k either does not pass through $z = 1$, or does not pass through $z = -1$. It is also easy to see that β_k can deformed into one of the circular arcs δ_k between its endpoints. For all k, $2 \leq k \leq n$, let σ_k be one of the circular arcs from $z = 1$ to the initial point of δ_k, and let $\sigma_1 = \sigma_{n+1} = 1$. We have

$$\gamma \approx (\sigma_1\delta_1\sigma_2^{-1}) \cdots (\sigma_n\delta_n\sigma_{n+1}^{-1}),$$

and it is easily seen that each arc $\sigma_k\delta_k\sigma_{k+1}^{-1}$ is homotopic either to α, or α^{-1}, or 1. Thus, $\gamma \approx \alpha^m$, for some integer $m \in \mathbb{Z}$.

It remains to prove that α^m is not homotopic to 1 for $m \neq 0$. This is where we use Proposition 4.2. Indeed, it is immediate that $n(\alpha^m, 0) = m$, and $n(1, 0) = 0$, and thus α^m and 1 are not homotopic when $m \neq 0$. But then, we have shown that the homotopy classes are in bijection with the set of integers. □

The above proof also applies to a cicular annulus, closed or open, and to a circle. In particular, the circle is not simply connected.

We will need to define what it means for a surface to be orientable. Perhaps surprisingly, a rigorous definition is not so easy to obtain but can be given using the notion of degree of a homeomorphism from a plane region. First, we need to define the degree of a map in the plane.

4.4 The Degree of a Map in the Plane

Let $\varphi\colon D \to \mathbb{C}$ be a continuous function to the plane, where the plane is viewed as the set $\mathbb{C}$ of complex numbers and with domain some open set D in $\mathbb{C}$. We say that φ is *regular at* $z_0 \in D$ if there is some open set $V \subseteq D$ containing z_0 such that $\varphi(z) \neq \varphi(z_0)$, for all $z \in V$. Assuming that φ is regular at z_0, we will define the *degree of* φ *at* z_0.

Let Ω be a punctured open disk $\{z \in V \mid |z - z_0| < r\}$ contained in V. Since φ is regular at z_0, it maps Ω into the punctured plane Ω' obtained by deleting $w_0 = \varphi(z_0)$. Now, φ induces a homomorphism $\varphi_*\colon \pi(\Omega) \to \pi(\Omega')$. From Proposition 4.3, both groups $\pi(\Omega)$ and

Figure 4.2: L E B Brouwer, 1881-1966

$\pi(\Omega')$ are isomorphic to $\mathbb{Z}$. Thus, it is easy to determine exactly what the homorphism φ_* is. We know that $\pi(\Omega)$ is generated by the homotopy class of some circle α in Ω with center a, and that $\pi(\Omega')$ is generated by the homotopy class of some circle β in Ω' with center $\varphi(a)$. If $\varphi_*([\alpha]) = [\beta^d]$, then the homomorphism φ_* is completely determined. If $d = 0$, then $\pi(\Omega') = 1$, and if $d \neq 0$, then $\pi(\Omega')$ is the infinite cyclic subgroup generated by the class of β^d. We let d be the *degree of φ at z_0*, and we denote it as $d(\varphi)_{z_0}$. It is easy to see that this definition does not depend on the choice of a (the center of the circle α) in Ω, and thus, does not depend on Ω.

Next, if we have a second mapping ψ regular at $w_0 = \varphi(z_0)$, then $\psi \circ \varphi$ is regular at z_0, and it is immediately verified that

$$d(\psi \circ \varphi)_{z_0} = d(\psi)_{w_0} d(\varphi)_{z_0}.$$

Let us now assume that D is a region (a connected open set) and that φ is a homeomorphism between D and $\varphi(D)$. By a theorem of Brouwer (the invariance of domain), it turns out that $\varphi(D)$ is also open and, thus, we can define the degree of the inverse mapping φ^{-1}, and since the identity clearly has degree 1, we get that $d(\varphi)d(\varphi^{-1}) = 1$, which shows that $d(\varphi)_{z_0} = \pm 1$.

In fact, Ahlfors and Sario [1] prove that if $\varphi(D)$ has a nonempty interior, then the degree of φ is constant on D. The proof is not difficult, but not very instructive.

Proposition 4.4 *Given a region, D, in the plane, for every homeomorphism φ between D and $\varphi(D)$, if $\varphi(D)$ has a nonempty interior, then the degree $d(\varphi)_z$ is constant for all $z \in D$, and in fact, $d(\varphi) = \pm 1$.*

When $d(\varphi) = 1$ in Proposition 4.4, we say that φ is *sense-preserving*, and when $d(\varphi) = -1$, we say that φ is *sense-reversing*. We can now define the notion of orientability.

4.5 Orientability of a Surface

Given a surface, F, we will call a region V on F a *planar region* if there is a homeomorphism, $h \colon V \to U$, from V onto an open set in the plane. From Proposition 4.4, the homeomorphisms

$h\colon V \to U$ can be divided into two classes, by defining two such homeomorphisms h_1, h_2 as equivalent iff $h_1 \circ h_2^{-1}$ has degree 1, i.e., is sense-preserving. Observe that for any h as above, if $\overline{h}$ is obtained from h by conjugation (i.e., for every $z \in V$, $\overline{h}(z) = \overline{h(z)}$, the complex conjugate of $h(z)$), then $d(h \circ \overline{h^{-1}}) = -1$, and thus h and $\overline{h}$ are in different classes. For any other such map g, either $h \circ g^{-1}$ or $\overline{h} \circ g^{-1}$ is sense-preserving, and thus, there are exactly two equivalence classes.

The choice of one of the two classes of homeomorphims h as above constitutes an *orientation* of V. An orientation of V induces an orientation on any subregion W of V, by restriction. If V_1 and V_2 are two planar regions and these regions have received an orientation, we say that these orientations are *compatible* if they induce the same orientation on all common subregions of V_1 and V_2.

Definition 4.4 A surface, F, is *orientable* if it is possible to assign an orientation to all planar regions in such a way that the orientations of any two overlapping planar regions are compatible.

Clearly, orientability is preserved by homeomorphisms. Thus, there are two classes of surfaces, the orientable surfaces, and the nonorientable surfaces. An example of a nonorientable surface is the Klein bottle. Because we defined a surface as being connected, note that an orientable surface has exactly two orientations. It is also easy to see that to orient a surface it is enough to orient all planar regions in some open covering of the surface by planar regions.

We will also need to consider bordered surfaces.

4.6 Bordered Surfaces

Consider a torus, and cut out a finite number of small disks from its surface. The resulting space is no longer a surface but is certainly of geometric interest. It is a surface with boundary, or bordered surface. In this section, we extend our concept of surface to handle this more general class of bordered surfaces. In order to do so, we need to allow coverings of surfaces using a richer class of open sets. This is achieved by considering the open subsets of the half-space, in the subset topology.

Definition 4.5 The *half-space* $\mathbb{H}^m$ is the subset of $\mathbb{R}^m$ defined as the set

$$\{(x_1, \ldots, x_m) \mid x_i \in \mathbb{R},\, x_m \geq 0\}.$$

For any $m \geq 1$, a *(topological) m-manifold with boundary* is a second-countable, topological Hausdorff space M, together with an open cover $(U_i)_{i\in I}$ of open sets and a family $(\varphi_i)_{i\in I}$ of homeomorphisms $\varphi_i\colon U_i \to \Omega_i$, where each Ω_i is some open subset of $\mathbb{H}^m$ in the subset topology. Each pair (U, φ) is called a *coordinate system*, or *chart*, of M, each homeomorphism $\varphi_i\colon U_i \to \Omega_i$ is called a *coordinate map*, and its inverse $\varphi_i^{-1}\colon \Omega_i \to U_i$ is called a *parameterization* of U_i. The family $(U_i, \varphi_i)_{i\in I}$ is often called an *atlas* for M. A *(topological) bordered surface* is a connected 2-manifold with boundary.

Note that an m-manifold is also an m-manifold with boundary.

If $\varphi_i \colon U_i \to \Omega_i$ is some homeomorphism onto some open set Ω_i of $\mathbb{H}^m$ in the subset topology, some $p \in U_i$ may be mapped into $\mathbb{R}^{m-1} \times \mathbb{R}_+$, or into the "boundary" $\mathbb{R}^{m-1} \times \{0\}$ of $\mathbb{H}^m$. Letting $\partial\mathbb{H}^m = \mathbb{R}^{m-1} \times \{0\}$, it can be shown using homology that if some coordinate map φ defined on p maps p into $\partial\mathbb{H}^m$, then every coordinate map ψ defined on p maps p into $\partial\mathbb{H}^m$. For $m = 2$, Ahlfors and Sario prove it using Proposition 4.4.

Thus, M is the disjoint union of two sets ∂M and $\text{Int}M$, where ∂M is the subset consisting of all points $p \in M$ that are mapped by some (in fact, all) coordinate map φ defined on p into $\partial\mathbb{H}^m$, and where $\text{Int}\, M = M - \partial M$. The set ∂M is called the *boundary* of M, and the set $\text{Int}\, M$ is called the *interior* of M, even though this terminology clashes with some prior topological definitions. A good example of a bordered surface is the Möbius strip. The boundary of the Möbius strip is a circle.

The boundary ∂M of M may be empty but $\text{Int}M$ is nonempty. Also, it can be shown using homology that the integer m is unique. It is clear that $\text{Int}\, M$ is open and an m-manifold and that ∂M is closed. If $p \in \partial M$, and φ is some coordinate map defined on p, since $\Omega = \varphi(U)$ is an open subset of $\partial\mathbb{H}^m$, there is some open half ball B^m_{o+} centered at $\varphi(p)$ and contained in Ω which intersects $\partial\mathbb{H}^m$ along an open ball B^{m-1}_o, and if we consider $W = \varphi^{-1}(B^m_{o+})$, we have an open subset of M containing p which is mapped homeomorphically onto B^m_{o+} in such that way that every point in $W \cap \partial M$ is mapped onto the open ball B^{m-1}_o. Thus, it is easy to see that ∂M is an $(m-1)$-manifold.

In particular, in the case $m = 2$, the boundary ∂M is a union of curves homeomorphic either to circles of to open line segments. In this case, if M is connected but not a surface, it is easy to see that M is the topological closure of $\text{Int}\, M$. We also claim that $\text{Int}\, M$ is connected, i.e. a surface. Indeed, if this was not so, we could write $\text{Int}\, M = M_1 \cup M_2$, for two nonempty disjoint sets M_1 and M_2. But then, we have $M = \overline{M_1} \cup \overline{M_2}$, and since M is connected, there is some $a \in \partial M$ also in $\overline{M_1} \cap \overline{M_2} \neq \emptyset$. However, there is some open set V containing a whose intersection with M is homeomorphic with an open half-disk, and thus connected. Then, we have

$$V \cap M = (V \cap M_1) \cup (V \cap M_2),$$

with $V \cap M_1$ and $V \cap M_2$ open in V, contradicting the fact that $M \cap V$ is connected. Thus, $\text{Int}\, M$ is a surface.

When the boundary ∂M of a bordered surface M is empty, M is just a surface. Typically, when we refer to a bordered surface, we mean a bordered surface with a nonempty border, and otherwise, we just say surface.

A bordered surface M is orientable iff its interior Int M is orientable. It is not difficult to show that an orientation of Int M induces an orientation of the boundary ∂M. The components of the boundary ∂M are called *contours*.

The concept of triangulation of a bordered surface is identical to the concept defined for a surface in Definition 3.4, and Proposition 3.5 also holds. However, a small change needs to made to Proposition 3.6, see Ahlfors and Sario [1].

Proposition 4.5 *A 2-complex $K = (V, \mathcal{S})$ is a triangulation $\sigma\colon \mathcal{S} \to 2^M$ of a bordered surface M such that $\sigma(s) = s_g$ for all $s \in \mathcal{S}$ iff the following properties hold:*

(D1) Every edge a such that a_g contains some point in the interior Int M *of M is contained in exactly two triangles A. Every edge a such that a_g is inside the border ∂M of M is contained in exactly one triangle A. The border ∂M of M consists of those a_g which belong to only one A_g. A border vertex or border edge is a simplex σ such that $\sigma_g \subseteq \partial M$.*

(D2) For every non-border vertex α, the edges a and triangles A containing α can be arranged as a cyclic sequence $a_1, A_1, a_2, A_2, \ldots, A_{m-1}, a_m, A_m$, in the sense that $a_i = A_{i-1} \cap A_i$ for all i, with $2 \leq i \leq m$, and $a_1 = A_m \cap A_1$, with $m \geq 3$.

(D3) For every border vertex α, the edges a and triangles A containing α can be arranged in a sequence $a_1, A_1, a_2, A_2, \ldots, A_{m-1}, a_m, A_m, a_{m+1}$, with $a_i = A_i \cap A_{i-1}$ for of all i, with $2 \leq i \leq m$, where a_1 and a_{m+1} are border vertices only contained in A_1 and A_m respectively.

(D4) K is connected, in the sense that it cannot be written as the union of two disjoint nonempty complexes.

A 2-complex K which satisfies the conditions of Proposition 4.5 will also be called a *bordered triangulated 2-complex* and its geometric realization a *bordered polyhedron*. Thus, bordered triangulated 2-complexes are the complexes that correspond to triangulated bordered surfaces. Actually, it can be shown that every bordered surface admits some triangulation and thus the class of geometric realizations of the bordered triangulated 2-complexes is the class of all bordered surfaces.

We will now give a brief presentation of simplicial and singular homology, but first, we need to review some facts about finitely generated abelian groups.

Chapter 5

Homology Groups

5.1 Finitely Generated Abelian Groups

Given a topological space, X, besides its fundamental group (a topological invariant), there is another useful kind of topological invariant, namely, its family of *homology groups*. One of the advantages of the homology groups is that they are abelian groups and, in the case of finite simplicial complexes, finitely generated abelian groups. Fortunately, the structure of finitely abelian groups is very well understood and this knowledge can be used to better understand the structure of polyhedras in terms of their homology. We begin by reviewing the structure theorem for finitely generated abelian groups.

An abelian group is a commutative group. We will denote the identity element of an abelian group by 0, and the inverse of an element, a, by $-a$. Given any natural number $n \in \mathbb{N}$, we denote

$$\underbrace{a + \cdots + a}_{n}$$

by na, and let $(-n)a$ be defined as $n(-a)$ (with $0a = 0$). Thus, we can make sense of finite sums of the form, $\sum n_i a_i$, where $n_i \in \mathbb{Z}$. Given an abelian group, G, and a family, $A = (a_j)_{j \in J}$, of elements, $a_j \in G$, we say that G is *generated by* A if every $a \in G$ can be written (in possibly more than one way) as

$$a = \sum_{i \in I} n_i a_i,$$

for some finite subset, I, of J, and some $n_i \in \mathbb{Z}$. If J is finite, we say that G is *finitely generated by* A. If every $a \in G$ can be written in a *unique manner* as

$$a = \sum_{i \in I} n_i a_i,$$

as above, we say that G is *freely generated by* A, and we call A a *basis of* G. In this case, it is clear that the a_j are all distinct. We also have the following familiar property:

If G is a free abelian group generated by $A = (a_j)_{j\in J}$, for every abelian group, H, for every function, $f\colon A \to H$, there is a unique homomorphism, $\widehat{f}\colon G \to H$, such that $\widehat{f}(a_j) = f(a_j)$, for all $j \in J$.

Remark: If G is a free abelian group, one can show that the cardinality of all bases is the same. When G is free and finitely generated by $(a_1, \ldots, a_n)$, this can be proved as follows: Consider the quotient of the group G modulo the subgroup $2G$ consisting of all elements of the form $g + g$, where $g \in G$. It is immediately verified that each coset of $G/2G$ is of the form

$$\epsilon_1 a_1 + \cdots + \epsilon_n a_n + 2G,$$

where $\epsilon_i = 0$ or $\epsilon_i = 1$, and thus, $G/2G$ has 2^n elements. Thus, n only depends on G. The number n is called the *dimension* of G.

Given a family, $A = (a_j)_{j\in J}$, we will need to construct a free abelian group generated by A. This can be done easily as follows: Consider the set, $F(A)$, of all functions $\varphi\colon A \to \mathbb{Z}$, such that $\varphi(a) \neq 0$ for only finitely many $a \in A$. We define addition on $F(A)$ pointwise, that is, $\varphi + \psi$ is the function such that $(\varphi + \psi)(a) = \varphi(a) + \psi(a)$, for all $a \in A$.

It is immediately verified that $F(A)$ is an abelian group and if we identify each a_j with the function, $\varphi_j\colon A \to \mathbb{Z}$, such that $\varphi_j(a_j) = 1$ and $\varphi_j(a_i) = 0$ for all $i \neq j$, it is clear that $F(A)$ is freely generated by A. It is also clear that every $\varphi \in F(A)$ can be uniquely written as

$$\varphi = \sum_{i\in I} n_i \varphi_i,$$

for some finite subset I of J such that $n_i = \varphi(a_i) \neq 0$. For notational simplicity, we write φ as

$$\varphi = \sum_{i\in I} n_i a_i.$$

Given an abelian group, G, for any $a \in G$, we say that a has *finite order* if there is some $n \neq 0$ in $\mathbb{N}$ such that $na = 0$. If $a \in G$ has finite order, there is a least $n \neq 0$ in $\mathbb{N}$ such that $na = 0$, called the *order of* a. It is immediately verified that the subset T of G consisting of all elements of finite order is a subroup of G, called the *torsion subgroup of* G. When $T = \{0\}$, we say that G is *torsion-free*. One should be careful that a torsion-free abelian group is not necessarily free. For example, the field $\mathbb{Q}$ of rationals is torsion-free, but not a free abelian group.

Clearly, the map $(n, a) \mapsto na$ from $\mathbb{Z} \times G$ to G satisfies the properties

$$\begin{aligned}
(m + n)a &= ma + na, \\
m(a + b) &= ma + nb, \\
(mn)a &= m(na), \\
1a &= a,
\end{aligned}$$

which hold in vector spaces. However, $\mathbb{Z}$ is not a field. The abelian group G is just what is called a *$\mathbb{Z}$-module*. Nevertheless, many concepts defined for vector spaces transfer to $\mathbb{Z}$-modules. For example, given an abelian group G and some subgroups $H_1, \ldots, H_n$, we can define the *(internal) sum*

$$H_1 + \cdots + H_n$$

of the H_i as the abelian group consisting of all sums of the form $a_1 + \cdots + a_n$, where $a_i \in H_i$. If in addition, $G = H_1 + \cdots + H_n$ and $H_i \cap H_j = \{0\}$ for all i, j, with $i \neq j$, we say that G is the *direct sum of the* H_i, and this is denoted as

$$G = H_1 \oplus \cdots \oplus H_n.$$

When $H_1 = \ldots = H_n = H$, we abbreviate $H \oplus \cdots \oplus H$ as H^n. Homomorphims between abelian groups are $\mathbb{Z}$-linear maps. We can also talk about linearly independent families in G, except that the scalars are in $\mathbb{Z}$. The *rank* of an abelian group is the maximum of the sizes of linearly independent families in G. We can also define (external) direct sums.

Given a family, $(G_i)_{i\in I}$, of abelian groups, the *(external) direct sum* $\bigoplus_{i\in I} G_i$ is the set of all functions, $f: I \to \bigcup_{i\in I} G_i$, such that $f(i) \in G_i$, for all $\in I$ and $f(i) = 0$ for all but finitely many $i \in I$. An element, $f \in \bigoplus_{i\in I} G_i$, is usually denoted by $(f_i)_{i\in I}$. Addition is defined component-wise, that is, given two functions $f = (f_i)_{i\in I}$ and $g = (g_i)_{i\in I}$ in $\bigoplus_{i\in I} G_i$, we define $(f + g)$ such that

$$(f + g)_i = f_i + g_i,$$

for all $i \in I$. It is immediately verified that $\bigoplus_{i\in I} G_i$ is an abelian group. For every $i \in I$, there is an injective homomorphism, $in_i : G_i \to \bigoplus_{i\in I} G_i$, defined such that, for every $x \in G_i$, $in_i(x)(i) = x$ and $in_i(x)(j) = 0$ iff $j \neq i$. If $G = \bigoplus_{i\in I} G_i$ is an external direct sum, it is immediately verified that $G = \bigoplus_{i\in I} in_i(G_i)$, as an internal direct sum. The difference is that G must have been already defined for an internal direct sum to make sense. For notational simplicity, we will usually identify $in_i(G_i)$ with G_i.

The structure of finitely generated abelian groups can be completely described. Actually, the following result is a special case of the structure theorem for finitely generated modules over a principal ring. Recall that $\mathbb{Z}$ is a principal ring, which means that every ideal $\mathcal{I}$ in $\mathbb{Z}$ is of the form $d\mathbb{Z}$, for some $d \in \mathbb{N}$. For the sake of completeness, we present the following result, whose neat proof is due to Pierre Samuel.

Proposition 5.1 *Let G be a free abelian group finitely generated by $(a_1, \ldots, a_n)$ and let H be any subroup of G. Then, H is a free abelian group and there is a basis, $(e_1, ..., e_n)$, of G, some $q \leq n$, and some positive natural numbers, $n_1, \ldots, n_q$, such that $(n_1e_1, \ldots, n_qe_q)$ is a basis of H and n_i divides n_{i+1} for all i, with $1 \leq i \leq q - 1$.*

Proof. The proposition is trivial when $H = \{0\}$ and thus, we assume that H is nontrivial. Let $L(G, \mathbb{Z})$ we the set of homomorphisms from G to $\mathbb{Z}$. For any $f \in L(G, \mathbb{Z})$, it is immediately verified that $f(H)$ is an ideal in $\mathbb{Z}$. Thus, $f(H) = n_h\mathbb{Z}$, for some $n_h \in \mathbb{N}$, since every ideal in $\mathbb{Z}$ is a principal ideal. Since $\mathbb{Z}$ is finitely generated, any nonempty family of ideals has a maximal element so let f be a homomorphism such that $n_h\mathbb{Z}$ is a maximal ideal in $\mathbb{Z}$. Let $\pi\colon G \to \mathbb{Z}$ be the i-th projection, i.e., π_i is defined such that $\pi_i(m_1a_1 + \cdots + m_na_n) = m_i$. It is clear that π_i is a homomorphism and since H is nontrivial, one of the $\pi_i(H)$ is nontrivial, and $n_h \neq 0$. There is some $b \in H$ such that $f(b) = n_h$.

We claim that, for every $g \in L(G, \mathbb{Z})$, the number n_h divides $g(b)$.

Indeed, if d is the gcd of n_h and $g(b)$, by the Bézout identity, we can write

$$d = rn_h + sg(b),$$

for some $r, s \in \mathbb{Z}$, and thus

$$d = rf(b) + sg(b) = (rf + sg)(b).$$

However, $rf + sg \in L(G, \mathbb{Z})$, and thus,

$$n_h\mathbb{Z} \subseteq d\mathbb{Z} \subseteq (rf + sg)(H),$$

since d divides n_h and, by maximality of $n_h\mathbb{Z}$, we must have $n_h\mathbb{Z} = d\mathbb{Z}$, which implies that $d = n_h$, and thus, n_h divides $g(b)$. In particular, n_h divides each $\pi_i(b)$ and let $\pi_i(b) = n_hp_i$, with $p_i \in \mathbb{Z}$.

Let $a = p_1a_1 + \cdots + p_na_n$. Note that

$$b = \pi_1(b)a_1 + \cdots + \pi_n(b)a_n = n_hp_1a_1 + \cdots + n_hp_na_n,$$

and thus, $b = n_ha$. Since $n_h = f(b) = f(n_ha) = n_hf(a)$, and since $n_h \neq 0$, we must have $f(a) = 1$.

Next, we claim that

$$G = a\mathbb{Z} \oplus f^{-1}(0)$$

and

$$H = b\mathbb{Z} \oplus (H \cap f^{-1}(0)),$$

with $b = n_ha$.

Indeed, every $x \in G$ can be written as

$$x = f(x)a + (x - f(x)a),$$

and since $f(a) = 1$, we have $f(x - f(x)a) = f(x) - f(x)f(a) = f(x) - f(x) = 0$. Thus, $G = a\mathbb{Z} + f^{-1}(0)$. Similarly, for any $x \in H$, we have $f(x) = rn_h$, for some $r \in \mathbb{Z}$, and thus,

$$x = f(x)a + (x - f(x)a) = rn_ha + (x - f(x)a) = rb + (x - f(x)a),$$

we still have $x - f(x)a \in f^{-1}(0)$, and clearly, $x - f(x)a = x - rn_ha = x - rb \in H$, since $b \in H$. Thus, $H = b\mathbb{Z} + (H \cap f^{-1}(0))$.

To prove that we have a direct sum, it is enough to prove that $a\mathbb{Z} \cap f^{-1}(0) = \{0\}$. For any $x = ra \in a\mathbb{Z}$, if $f(x) = 0$, then $f(ra) = rf(a) = r = 0$, since $f(a) = 1$ and, thus, $x = 0$. Therefore, the sums are direct sums.

We can now prove that H is a free abelian group by induction on the size, q, of a maximal linearly independent family for H.

If $q = 0$, the result is trivial. Otherwise, since

$$H = b\mathbb{Z} \oplus (H \cap f^{-1}(0)),$$

it is clear that $H \cap f^{-1}(0)$ is a subgroup of G and that every maximal linearly independent family in $H \cap f^{-1}(0)$ has at most $q - 1$ elements. By the induction hypothesis, $H \cap f^{-1}(0)$ is a free abelian group and, by adding b to a basis of $H \cap f^{-1}(0)$, we obtain a basis for H, since the sum is direct.

The second part is shown by induction on the dimension n of G.

The case $n = 0$ is trivial. Otherwise, since

$$G = a\mathbb{Z} \oplus f^{-1}(0),$$

and since, by the previous argument, $f^{-1}(0)$ is also free, it is easy to see that $f^{-1}(0)$ has dimension $n-1$. By the induction hypothesis applied to its subgroup, $H \cap f^{-1}(0)$, there is a basis $(e_2, \ldots, e_n)$ of $f^{-1}(0)$, some $q \leq n$, and some positive natural numbers $n_2, \ldots, n_q$, such that, $(n_2e_2, \ldots, n_qe_q)$ is a basis of $H \cap f^{-1}(0)$, and n_i divides n_{i+1} for all i, with $2 \leq i \leq q-1$. Let $e_1 = a$, and $n_1 = n_h$, as above. It is clear that $(e_1, \ldots, e_n)$ is a basis of G, and that that $(n_1e_1, \ldots, n_qe_q)$ is a basis of H, since the sums are direct, and $b = n_1e_1 = n_ha$. It remains to show that n_1 divides n_2. Consider the homomorphism $g \colon G \to \mathbb{Z}$ such that $g(e_1) = g(e_2) = 1$, and $g(e_i) = 0$, for all i, with $3 \leq i \leq n$. We have $n_h = n_1 = g(n_1e_1) = g(b) \in g(H)$, and thus, $n_h\mathbb{Z} \subseteq g(H)$. Since $n_h\mathbb{Z}$ is maximal, we must have $g(H) = n_h\mathbb{Z} = n_1\mathbb{Z}$. Since $n_2 = g(n_2e_2) \in g(H)$, we have $n_2 \in n_1\mathbb{Z}$, which shows that n_1 divides n_2. $\square$

Using Proposition 5.1, we can also show the following useful result:

Proposition 5.2 *Let G be a finitely generated abelian group. There is some natural number, $m \geq 0$, and some positive natural numbers, $n_1, \ldots, n_q$, such that H is isomorphic to the direct sum*

$$\mathbb{Z}^m \oplus \mathbb{Z}/n_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/n_q\mathbb{Z},$$

and where n_i divides n_{i+1} for all i, with $1 \leq i \leq q - 1$.

Proof. Assume that G is generated by $A = (a_1, \ldots, a_n)$ and let $F(A)$ be the free abelian group generated by A. The inclusion map $i \colon A \to G$ can be extended to a unique homomorphism $f \colon F(A) \to G$ which is surjective since A generates G and thus, G is isomorphic to $F(A)/f^{-1}(0)$. By Proposition 5.1, $H = f^{-1}(0)$ is a free abelian group and there is a basis $(e_1, ..., e_n)$ of G, some $p \leq n$, and some positive natural numbers $k_1, \ldots, k_p$, such that

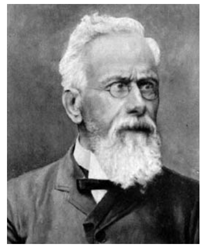

Figure 5.1: Enrico Betti, 1823-1892

$(k_1e_1, \ldots, k_pe_p)$ is a basis of H, and k_i divides k_{i+1} for all i, with $1 \leq i \leq p-1$. Let r, $0 \leq r \leq p$, be the largest natural number such that $k_1 = \ldots = k_r = 1$, rename k_{r+i} as n_i, where $1 \leq i \leq p-r$, and let $q = p - r$. Then, we can write

$$H = \mathbb{Z}^{p-q} \oplus n_1\mathbb{Z} \oplus \cdots \oplus n_q\mathbb{Z},$$

and since $F(A)$ is isomorphic to $\mathbb{Z}^n$, it is easy to verify that $F(A)/H$ is isomorphic to

$$Z^{n-p} \oplus \mathbb{Z}/n_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/n_q\mathbb{Z},$$

which proves the proposition. $\square$

Observe that $\mathbb{Z}/n_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/n_q\mathbb{Z}$ is the torsion subgroup of G. Thus, as a corollary of Proposition 5.2, we obtain the fact that every finitely generated abelian group G is a direct sum, $G = Z^m \oplus T$, where T is the torsion subroup of G and Z^m is the free abelian group of dimension m. It is easy to verify that m is the rank (the maximal dimension of linearly independent sets in G) of G and it is called the *Betti number* of G. It can also be shown that q and the n_i only depend on G.

One more result will be needed to compute the homology groups of (two-dimensional) polyhedras. The proof is not difficult and can be found in most books (a version is given in Ahlfors and Sario [1]). Let us denote the rank of an abelian group G as $r(G)$.

Proposition 5.3 *If*

$$0 \longrightarrow E \xrightarrow{f} F \xrightarrow{g} G \longrightarrow 0$$

is a short exact sequence of homomorphisms of abelian groups and F has finite rank, then $r(F) = r(E) + r(G)$. In particular, if G is an abelian group of finite rank and H is a subroup of G, then $r(G) = r(H) + r(G/H)$.

We are now ready to define the simplicial and the singular homology groups.

5.2 Simplicial and Singular Homology

There are several kinds of homology theories. In this section, we take a quick look at two such theories, *simplicial homology*, one of the most computational theories, and *singular homology theory*, one of the most general and yet fairly intuitive. For a comprehensive treatment of homology and algebraic topology in general, we refer the reader to Massey [33], Munkres [38], Bredon [7], Hatcher [20], Fulton [18], Dold [13], Rotman [41], Amstrong [3] and Kinsey [26]. An excellent overview of algebraic topology, following a more intuitive approach, is presented in Sato [42].

Let $K = (V, \mathcal{S})$ be a complex. The essence of simplicial homology is to associate some abelian groups, $H_p(K)$, with K. This is done by first defining some free abelian groups, $C_p(K)$, made out of oriented p-simplices. One of the main new ingredients is that every oriented p-simplex, σ, is assigned a *boundary*, $\partial_p\sigma$. Technically, this is achieved by defining homomorphisms,

$$\partial_p\colon C_p(K) \to C_{p-1}(K),$$

with the property that $\partial_{p-1} \circ \partial_p = 0$. If we let $Z_p(K)$ be the kernel of ∂_p and

$$B_p(K) = \partial_{p+1}(C_{p+1}(K))$$

be the image of ∂_{p+1} in $C_p(K)$, since $\partial_p \circ \partial_{p+1} = 0$, the group $B_p(K)$ is a subgroup of the group $Z_p(K)$ and we define the homology group, $H_p(K)$, as the quotient group

$$H_p(K) = Z_p(K)/B_p(K).$$

What makes the homology groups of a complex interesting is that they only depend on the geometric realization K_g of the complex K and not on the various complexes representing K_g. Proving this fact requires relatively hard work, and we refer the reader to Munkres [38] or Rotman [41], for a proof.

The first step in defining simplicial homology groups is to define oriented simplices. Given a complex, $K = (V, \mathcal{S})$, recall that an n-simplex is a subset, $\sigma = \{\alpha_0, \ldots, \alpha_n\}$, of V that belongs to the family $\mathcal{S}$. Thus, the set σ corresponds to $(n+1)!$ linearly ordered sequences, $s\colon \{1, 2, \ldots, n+1\} \to \sigma$, where each s is a bijection. We define an equivalence relation on these sequences by saying that two sequences $s_1\colon \{1, 2, \ldots, n+1\} \to \sigma$ and $s_2\colon \{1, 2, \ldots, n+1\} \to \sigma$ are *equivalent* iff $\pi = s_2^{-1} \circ s_1$ is a permutation of even signature (π is the product of an even number of transpositions)

The two equivalence classes associated with σ are called *oriented simplices*, and if $\sigma = \{\alpha_0, \ldots, \alpha_n\}$, we denote the equivalence class of s as $[s(1), \ldots, s(n+1)]$, where s is one of the sequences $s\colon \{1, 2, \ldots, n+1\} \to \sigma$. We also say that the two classes associated with σ are the *orientations of* σ. Two oriented simplices σ_1 and σ_2 are said to have *opposite orientation* if they are the two classes associated with some simplex σ. Given an oriented simplex, σ, we denote the oriented simplex having the opposite orientation by $-\sigma$, with the convention that $-(-\sigma) = \sigma$.

For example, if $\sigma = \{a_1, a_2, a_3\}$ is a 3-simplex (a triangle), there are six ordered sequences, the sequences $\langle a_3, a_2, a_1\rangle$, $\langle a_2, a_1, a_3\rangle$, and $\langle a_1, a_3, a_2\rangle$, are equivalent, and the sequences $\langle a_1, a_2, a_3\rangle$, $\langle a_2, a_3, a_1\rangle$, and $\langle a_3, a_1, a_2\rangle$, are also equivalent. Thus, we have the two oriented simplices, $[a_1, a_2, a_3]$ and $[a_3, a_2, a_1]$. We now define p-chains.

Definition 5.1 Given a complex, $K = (V, \mathcal{S})$, a *p-chain* on K is a function c from the set of oriented p-simplices to $\mathbb{Z}$, such that,

(1) $c(-\sigma) = -c(\sigma)$, iff σ and $-\sigma$ have opposite orientation;

(2) $c(\sigma) = 0$, for all but finitely many simplices σ.

We define addition of p-chains pointwise, i.e., $c_1 + c_2$ is the p-chain such that $(c_1 + c_2)(\sigma) = c_1(\sigma) + c_2(\sigma)$, for every oriented p-simplex σ. The group of p-chains is denoted by $C_p(K)$. If $p < 0$ or $p > \dim(K)$, we set $C_p(K) = \{0\}$.

To every oriented p-simplex σ is associated an *elementary p-chain* c, defined such that,

$c(\sigma) = 1$,

$c(-\sigma) = -1$, where $-\sigma$ is the opposite orientation of σ, and

$c(\sigma') = 0$, for all other oriented simplices σ'.

We will often denote the elementary p-chain associated with the oriented p-simplex σ also by σ.

The following proposition is obvious, and simply confirms the fact that $C_p(K)$ is indeed a free abelian group.

Proposition 5.4 *For every complex, $K = (V, \mathcal{S})$, for every p, the group $C_p(K)$ is a free abelian group. For every choice of an orientation for every p-simplex, the corresponding elementary chains form a basis for $C_p(K)$.*

The only point worth elaborating is that except for $C_0(K)$, where no choice is involved, there is no canonical basis for $C_p(K)$ for $p \geq 1$, since different choices for the orientations of the simplices yield different bases.

If there are m_p p-simplices in K, the above proposition shows that $C_p(K) = \mathbb{Z}^{m_p}$.

As an immediate consequence of Proposition 5.4, for any abelian group G and any function f mapping the oriented p-simplices of a complex K to G and such that $f(-\sigma) = -f(\sigma)$ for every oriented p-simplex σ, there is a unique homomorphism, $\widehat{f}\colon C_p(K) \to G$, extending f.

We now define the boundary maps $\partial_p\colon C_p(K) \to C_{p-1}(K)$.

Definition 5.2 Given a complex, $K = (V, \mathcal{S})$, for every oriented p-simplex,

$$\sigma = [\alpha_0, \ldots, \alpha_p],$$

we define the *boundary, $\partial_p \sigma$, of σ* by

$$\partial_p \sigma = \sum_{i=0}^{p} (-1)^i [\alpha_0, \ldots, \widehat{\alpha_i}, \ldots, \alpha_p],$$

where $[\alpha_0, \ldots, \widehat{\alpha_i}, \ldots, \alpha_p]$ denotes the oriented $p-1$-simplex obtained by deleting vertex α_i. The *boundary map, $\partial_p \colon C_p(K) \to C_{p-1}(K)$,* is the unique homomorphism extending ∂_p on oriented p-simplices. For $p \leq 0$, ∂_p is the null homomorphism.

One must verify that $\partial_p(-\sigma) = -\partial_p \sigma$, but this is immediate. If $\sigma = [\alpha_0, \alpha_1]$, then

$$\partial_1 \sigma = \alpha_1 - \alpha_0.$$

If $\sigma = [\alpha_0, \alpha_1, \alpha_2]$, then

$$\partial_2 \sigma = [\alpha_1, \alpha_2] - [\alpha_0, \alpha_2] + [\alpha_0, \alpha_1] = [\alpha_1, \alpha_2] + [\alpha_2, \alpha_0] + [\alpha_0, \alpha_1].$$

If $\sigma = [\alpha_0, \alpha_1, \alpha_2, \alpha_3]$, then

$$\partial_3 \sigma = [\alpha_1, \alpha_2, \alpha_3] - [\alpha_0, \alpha_2, \alpha_3] + [\alpha_0, \alpha_1, \alpha_3] - [\alpha_0, \alpha_1, \alpha_2].$$

If σ is the chain

$$\sigma = [\alpha_0, \alpha_1] + [\alpha_1, \alpha_2] + [\alpha_2, \alpha_3],$$

shown in Figure 5.2 (a), then

$$\begin{aligned} \partial_1 \sigma &= \partial_1[\alpha_0, \alpha_1] + \partial_1[\alpha_1, \alpha_2] + \partial_1[\alpha_2, \alpha_3] \\ &= \alpha_1 - \alpha_0 + \alpha_2 - \alpha_1 + \alpha_3 - \alpha_2 \\ &= \alpha_3 - \alpha_0. \end{aligned}$$

On the other hand, if σ is the closed cycle,

$$\sigma = [\alpha_0, \alpha_1] + [\alpha_1, \alpha_2] + [\alpha_2, \alpha_0],$$

shown in Figure 5.2 (b), then

$$\begin{aligned} \partial_1 \sigma &= \partial_1[\alpha_0, \alpha_1] + \partial_1[\alpha_1, \alpha_2] + \partial_1[\alpha_2, \alpha_0] \\ &= \alpha_1 - \alpha_0 + \alpha_2 - \alpha_1 + \alpha_0 - \alpha_2 \\ &= 0. \end{aligned}$$

We have the following fundamental property:

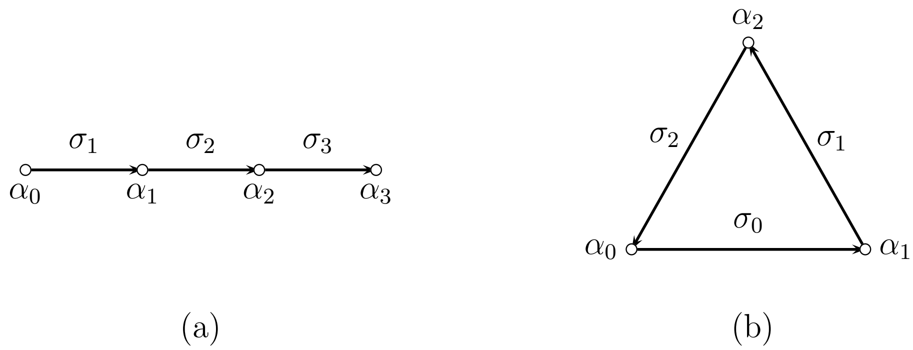

Figure 5.2: (a) A chain with boundary $\alpha_3 - \alpha_0$. (b) A chain with 0 boundary.

Proposition 5.5 *For every complex, $K = (V, \mathcal{S})$, for every p, we have $\partial_{p-1} \circ \partial_p = 0$.*

Proof. For any oriented p-simplex, $\sigma = [\alpha_0, \ldots, \alpha_p]$, we have

$$
\begin{aligned}
\partial_{p-1} \circ \partial_p \sigma &= \sum_{i=0}^{p} (-1)^i \partial_{p-1}[\alpha_0, \ldots, \widehat{\alpha_i}, \ldots, \alpha_p], \\
&= \quad \sum_{i=0}^{p} \sum_{j=0}^{i-1} (-1)^i (-1)^j [\alpha_0, \ldots, \widehat{\alpha_j}, \ldots, \widehat{\alpha_i}, \ldots, \alpha_p] \\
&\quad + \sum_{i=0}^{p} \sum_{j=i+1}^{p} (-1)^i (-1)^{j-1} [\alpha_0, \ldots, \widehat{\alpha_i}, \ldots, \widehat{\alpha_j}, \ldots, \alpha_p] \\
&= 0.
\end{aligned}
$$

The rest of the proof follows from the fact that $\partial_p \colon C_p(K) \to C_{p-1}(K)$ is the unique homomorphism extending ∂_p on oriented p-simplices. □

In view of Proposition 5.5, the image $\partial_{p+1}(C_{p+1}(K))$ of $\partial_{p+1} \colon C_{p+1}(K) \to C_p(K)$ is a subgroup of the kernel $\partial_p^{-1}(0)$ of $\partial_p \colon C_p(K) \to C_{p-1}(K)$. This motivates the following definition:

Definition 5.3 Given a complex, $K = (V, \mathcal{S})$, the kernel, $\partial_p^{-1}(0)$, of the homomorphism, $\partial_p \colon C_p(K) \to C_{p-1}(K)$, is denoted by $Z_p(K)$ and the elements of $Z_p(K)$ are called *p-cycles*. The image, $\partial_{p+1}(C_{p+1})$, of the homomorphism, $\partial_{p+1} \colon C_{p+1}(K) \to C_p(K)$, is denoted by $B_p(K)$, and the elements of $B_p(K)$ are called *p-boundaries*. The *p-th homology group, $H_p(K)$*, is the quotient group

$$H_p(K) = Z_p(K)/B_p(K).$$

Two p-chains c, c' are said to be *homologous* if there is some $(p+1)$-chain, d, such that $c = c' + \partial_{p+1} d$.

We will often omit the subscript p in ∂_p.

If $K = (V, \mathcal{S})$ is a finite dimensional complex, as each group, $C_p(K)$, is free and finitely generated, the homology groups, $H_p(K)$, are all finitely generated. At this stage, we could determine the homology groups of the finite (two-dimensional) polyhedras. However, we are really interested in the homology groups of geometric realizations of complexes, in particular, compact surfaces and, so far, we have not defined homology groups for topological spaces.

It is possible to define homology groups for arbitrary topological spaces, using what is called *singular homology*. Then, it can be shown, although this requires some hard work, that the homology groups of a space, X, which is the geometric realization of some complex, K, are independent of the complex, K, such that $X = K_g$ and equal to the homology groups of any such complex.

The idea behind singular homology is to define a more general notion of an n-simplex associated with a topological space, X, and it is natural to consider continuous maps from some standard simplices to X. Recall that given any set, I, we defined the real vector space, $\mathbb{R}^{(I)}$, freely generated by I (just before Definition 3.3). In particular, for $I = \mathbb{N}$ (the natural numbers), we obtain an infinite dimensional vector space, $\mathbb{R}^{(\mathbb{N})}$, whose elements are the countably infinite sequences, $(\lambda_i)_{i\in\mathbb{N}}$, of reals, with $\lambda_i = 0$ for all but finitely many $i \in \mathbb{N}$. For any $p \in \mathbb{N}$, we let $e_i \in \mathbb{R}^{(\mathbb{N})}$ be the sequence such that $e_i(i) = 1$ and $e_i(j) = 0$ for all $j \neq i$ and we let Δ_p be the p-simplex spanned by $(e_0, \ldots, e_p)$, that is, the subset of $\mathbb{R}^{(\mathbb{N})}$ consisting of all points of the form

$$\sum_{i=0}^{p} \lambda_i e_i, \quad \text{with} \quad \sum_{i=0}^{p} \lambda_i = 1, \text{ and } \lambda_i \geq 0.$$

We call Δ_p the *standard p-simplex*. Note that Δ_{p-1} is a face of Δ_p.

Definition 5.4 Given a topological space, X, a *singular p-simplex* is any continuous map, $T\colon \Delta_p \to X$. The free abelian group generated by the singular p-simplices is called the *p-th singular chain group* and is denoted by $S_p(X)$.

Given any $p+1$ points, $a_0, \ldots, a_p$, in $\mathbb{R}^{(\mathbb{N})}$, there is a unique affine map, $f\colon \Delta_p \to \mathbb{R}^{(\mathbb{N})}$, such that $f(e_i) = a_i$, for all i, $0 \leq i \leq p$, namely, the map such that

$$f(\sum_{i=0}^{p} \lambda_i e_i) = \sum_{i=0}^{p} \lambda_i a_i,$$

for all λ_i such that $\sum_{i=0}^{p} \lambda_i = 1$, and $\lambda_i \geq 0$. This map is called the *affine singular simplex* determined by $a_0, \ldots, a_p$ and it is denoted by $l(a_0, \ldots, a_p)$. In particular, the map

$$l(e_0, \ldots, \widehat{e_i}, \ldots, e_p),$$

where the hat over e_i means that e_i is omited, is a map from Δ_{p-1} onto a face of Δ_p. We can consider it as a map from Δ_{p-1} to Δ_p (although it is defined as a map from Δ_{p-1} to $\mathbb{R}^{(\mathbb{N})}$) and call it the *i-th face of* Δ_p.

Then, if $T\colon \Delta_p \to X$ is a singular p-simplex, we can form the map

$$T \circ l(e_0, \ldots, \widehat{e_i}, \ldots, e_p)\colon \Delta_{p-1} \to X,$$

which is a singular $p-1$-simplex, which we think of as the i-th face of T. Actually, for $p = 1$, a singular p-simplex, $T\colon \Delta_p \to X$, can be viewed as curve on X and its faces are its two endpoints. For $p = 2$, a singular p-simplex, $T\colon \Delta_p \to X$, can be viewed as triangular surface patch on X and its faces are its three boundary curves. For $p = 3$, a singular p-simplex, $T\colon \Delta_p \to X$, can be viewed as tetrahedral "volume patch" on X and its faces are its four boundary surface patches. We can give similar higher-order descriptions when $p > 3$.

We can now define the boundary maps, $\partial_p\colon S_p(X) \to S_{p-1}(X)$.

Definition 5.5 Given a topological space, X, for every, singular p-simplex, $T\colon \Delta_p \to X$, we define the *boundary,* $\partial_p T$*, of* T by

$$\partial_p T = \sum_{i=0}^{p} (-1)^i\, T \circ l(e_0, \ldots, \widehat{e_i}, \ldots, e_p).$$

The *boundary map,* $\partial_p\colon S_p(X) \to S_{p-1}(X)$, is the unique homomorphism extending ∂_p on singular p-simplices. For $p \leq 0$, ∂_p is the null homomorphism. Given a continuous map, $f\colon X \to Y$, between two topological spaces X and Y, the homomorphism, $f_{\sharp,p}\colon S_p(X) \to S_p(Y)$, is defined such that

$$f_{\sharp,p}(T) = f \circ T,$$

for every singular p-simplex, $T\colon \Delta_p \to X$.

The next easy proposition gives the main properties of ∂.

Proposition 5.6 *For every continuous map,* $f\colon X \to Y$*, between two topological spaces,* X *and* Y*, the maps* $f_{\sharp,p}$ *and* ∂_p *commute for every* p*, i.e.,*

$$\partial_p \circ f_{\sharp,p} = f_{\sharp,p-1} \circ \partial_p.$$

We also have $\partial_{p-1} \circ \partial_p = 0$.

Proof. For any singular p-simplex, $T\colon \Delta_p \to X$, we have

$$\partial_p f_{\sharp,p}(T) = \sum_{i=0}^{p} (-1)^i\, (f \circ T) \circ l(e_0, \ldots, \widehat{e_i}, \ldots, e_p),$$

and

$$f_{\sharp,p-1}(\partial_p T) = \sum_{i=0}^{p} (-1)^i\, f \circ (T \circ l(e_0, \ldots, \widehat{e_i}, \ldots, e_p)),$$

and the equality follows by associativity of composition. We also have

$$\begin{aligned} \partial_p l(a_0, \ldots, a_p) &= \sum_{i=0}^{p} (-1)^i\, l(a_0, \ldots, a_p) \circ l(e_0, \ldots, \widehat{e_i}, \ldots, e_p) \\ &= \sum_{i=0}^{p} (-1)^i\, l(a_0, \ldots, \widehat{a_i}, \ldots, a_p), \end{aligned}$$

since the composition of affine maps is affine. Then, we can compute $\partial_{p-1}\partial_p l(a_0, \ldots, a_p)$ as we did in Proposition 5.5 and the proof is similar, except that we have to insert an l at appropriate places. The rest of the proof follows from the fact that

$$\partial_{p-1}\partial_p T = \partial_{p-1}\partial_p(T_\sharp(l(e_0, \ldots, e_p))),$$

since $l(e_0, \ldots, e_p)$ is simply the inclusion of Δ_p in $\mathbb{R}^{(\mathbb{N})}$, and that ∂ commutes with $T_\sharp$. $\square$

In view of Proposition 5.6, the image $\partial_{p+1}(S_{p+1}(X))$ of $\partial_{p+1} \colon S_{p+1}(X) \to S_p(X)$ is a subgroup of the kernel $\partial_p^{-1}(0)$ of $\partial_p \colon S_p(X) \to S_{p-1}(X)$. This motivates the following definition:

Definition 5.6 Given a topological space, X, the kernel, $\partial_p^{-1}(0)$, of the homomorphism, $\partial_p \colon S_p(X) \to S_{p-1}(X)$, is denoted by $Z_p(X)$ and the elements of $Z_p(X)$ are called *singular p-cycles*. The image, $\partial_{p+1}(S_{p+1})$, of the homomorphism, $\partial_{p+1} \colon S_{p+1}(X) \to S_p(X)$, is denoted by $B_p(X)$ and the elements of $B_p(X)$ are called *singular p-boundaries*. The *p-th singular homology group, $H_p(X)$,* is the quotient group

$$H_p(X) = Z_p(X)/B_p(X).$$

If $f \colon X \to Y$ is a continuous map, the fact that

$$\partial_p \circ f_{\sharp,p} = f_{\sharp,p-1} \circ \partial_p$$

allows us to define homomorphisms, $f_{*,p} \colon H_p(X) \to H_p(Y)$, and it it easily verified that

$$(g \circ f)_{*,p} = g_{*,p} \circ f_{*,p}$$

and that $Id_{*,p} \colon H_p(X) \to H_p(Y)$ is the identity homomorphism when $Id \colon X \to Y$ is the identity. As a corollary, if $f \colon X \to Y$ is a homeomorphism, then each $f_{*,p} \colon H_p(X) \to H_p(Y)$ is a group isomorphism. This gives us a way of showing that two spaces are not homeomorphic, by showing that some homology groups $H_p(X)$ and $H_p(Y)$ are not isomorphic.

It is fairly easy to show that $H_0(X)$ is a free abelian group and that if the path components of X are the family $(X_i)_{i\in I}$, then $H_0(X)$ is isomorphic to the direct sum $\bigoplus_{i\in I} \mathbb{Z}$. In particular, if X is arcwise connected, then $H_0(X) = \mathbb{Z}$.

The following important theorem shows the relationship between simplicial homology and singular homology. The proof is fairly involved, and can be found in Munkres [38], or Rotman [41].


Figure 5.3: Leonhard Euler, 1707-1783 (left) and Henri Poincaré, 1854-1912 (right)

Theorem 5.7 *Given any polytope, X, if $X = K_g = K'_g$ is the geometric realization of any two complexes, K and K', then*

$$H_p(X) = H_p(K) = H_p(K'),$$

for all $p \geq 0$.

Theorem 5.7 implies that $H_p(X)$ is finitely generated for all $p \geq 0$. It is immediate that if K has dimension m, then $H_p(X) = 0$ for $p > m$, and it can be shown that $H_m(X)$ is a free abelian group.

A fundamental invariant of finite complexes is the Euler-Poincaré characteristic.

Definition 5.7 Given a finite complex, $K = (V, \mathcal{S})$, of dimension m, letting m_p be the number of p-simplices in K, we define the *Euler-Poincaré characteristic, $\chi(K)$, of K* by

$$\chi(K) = \sum_{p=0}^{m} (-1)^p \, m_p.$$

The following remarkable theorem holds:

Theorem 5.8 *Given a finite complex, $K = (V, \mathcal{S})$, of dimension m, we have*

$$\chi(K) = \sum_{p=0}^{m} (-1)^p \, r(H_p(K)),$$

the alternating sum of the Betti numbers (the ranks) of the homology groups of K.

Proof. We know that $C_p(K)$ is a free group of rank m_p. Since $H_p(K) = Z_p(K)/B_p(K)$, by Proposition 5.3, we have

$$r(H_p(K)) = r(Z_p(K)) - r(B_p(K)).$$

Since we have a short exact sequence

$$0 \longrightarrow Z_p(K) \longrightarrow C_p(K) \xrightarrow{\partial_p} B_{p-1}(K) \longrightarrow 0,$$

again, by Proposition 5.3, we have

$$r(C_p(K)) = m_p = r(Z_p(K)) + r(B_{p-1}(K)).$$

Also, note that $B_m(K) = 0$, and $B_{-1}(K) = 0$. Then, we have

$$\begin{aligned}\chi(K) &= \sum_{p=0}^{m} (-1)^p\, m_p \\ &= \sum_{p=0}^{m} (-1)^p\, (r(Z_p(K)) + r(B_{p-1}(K))) \\ &= \sum_{p=0}^{m} (-1)^p\, r(Z_p(K)) + \sum_{p=0}^{m} (-1)^p\, r(B_{p-1}(K)).\end{aligned}$$

Using the fact that $B_m(K) = 0$, and $B_{-1}(K) = 0$, we get

$$\begin{aligned}\chi(K) &= \sum_{p=0}^{m} (-1)^p\, r(Z_p(K)) + \sum_{p=0}^{m} (-1)^{p+1}\, r(B_p(K)) \\ &= \sum_{p=0}^{m} (-1)^p\, (r(Z_p(K)) - r(B_p(K))) \\ &= \sum_{p=0}^{m} (-1)^p\, r(H_p(K)).\end{aligned}$$

□

A striking corollary of Theorem 5.8 (together with Theorem 5.7) is that the Euler-Poincaré characteristic, $\chi(K)$, of a complex of finite dimension m only depends on the geometric realization, K_g, of K, since it only depends on the homology groups, $H_p(K) = H_p(K_g)$, of the polytope K_g. Thus, the Euler-Poincaré characteristic is an invariant of all the finite complexes corresponding to the same polytope, $X = K_g$, and we can say that it is *the* Euler-Poincaré characteristic of the polytope, $X = K_g$, and denote it by $\chi(X)$. In particular, this is true of surfaces that admit a triangulation and, as we shall see shortly, the Euler-Poincaré characteristic in one of the major ingredients in the classification of the compact surfaces. In this case, $\chi(K) = m_0 - m_1 + m_2$, where m_0 is the number of vertices, m_1 the number of edges, and m_2 the number of triangles, in K. We warn the reader that Ahlfors and Sario have flipped the signs and define the Euler-Poincaré characteristic as $-m_0 + m_1 - m_2$.

Going back to the triangulations of the sphere, the torus, the projective space, and the Klein bottle, it is easy to see that their Euler-Poincaré characteristic is 2 (sphere), 0 (torus), 1 (projective space), and 0 (Klein bottle).

At this point, we are ready to compute the homology groups of finite (two-dimensional) polyhedras.

5.3 Homology Groups of the Finite Polyhedras

Since a polyhedron is the geometric realization of a triangulated 2-complex, it is possible to determine the homology groups of the (finite) polyhedras. We say that a triangulated 2-complex K is orientable if its geometric realization K_g is orientable. We will consider the finite, bordered, orientable, and nonorientable, triangulated 2-complexes. First, note that $C_p(K)$ is the trivial group for $p < 0$ and $p > 2$, and thus, we just have to consider the cases where $p = 0, 1, 2$. We will use the notation $c \sim c'$, to denote that two p-chains are homologous, which means that $c = c' + \partial_{p+1} d$, for some $(p+1)$-chain d.

The first proposition is very easy, and is just a special case of the fact that $H_0(X) = \mathbb{Z}$ for an arcwise connected space X.

Proposition 5.9 *For every triangulated 2-complex (finite or not), K, we have $H_0(K) = \mathbb{Z}$.*

Proof. When $p = 0$, we have $Z_0(K) = C_0(K)$, and thus, $H_0(K) = C_0(K)/B_0(K)$. Thus, we have to figure out what the 0-boundaries are. If $c = \sum x_i \partial a_i$ is a 0-boundary, each a_i is an oriented edge $[\alpha_i, \beta_i]$ and we have

$$c = \sum x_i \partial a_i = \sum x_i \beta_i - \sum x_i \alpha_i,$$

which shows that the sum of all the coefficients of the vertices is 0. Thus, it is impossible for a 0-chain of the form $x\alpha$, where $x \neq 0$, to be homologous to 0. On the other hand, we claim that $\alpha \sim \beta$ for any two vertices α, β. Indeed, since we assumed that K is connected, there is a path from α to β consisting of edges

$$[\alpha, \alpha_1], \ldots, [\alpha_n, \beta],$$

and the 1-chain

$$c = [\alpha, \alpha_1] + \ldots + [\alpha_n, \beta]$$

has boundary

$$\partial c = \beta - \alpha,$$

which shows that $\alpha \sim \beta$. But then, $H_0(K)$ is the infinite cyclic group generated by any vertex. □

Next, we determine the groups $H_2(K)$.

Proposition 5.10 *For every triangulated 2-complex (finite or not), K, either $H_2(K) = \mathbb{Z}$ or $H_2(K) = 0$. Furthermore, $H_2(K) = \mathbb{Z}$ iff K is finite, has no border and is orientable, else $H_2(K) = 0$.*

Proof. When $p = 2$, we have $B_2(K) = 0$ and $H_2(K) = Z_2(K)$. Thus, we have to figure out what the 2-cycles are. Consider a 2-chain, $c = \sum x_i A_i$, where each A_i is an oriented triangle, $[\alpha_0, \alpha_1, \alpha_2]$, and assume that c is a cycle, which means that

$$\partial c = \sum x_i \partial A_i = 0.$$

Whenever A_i and A_j have an edge a in common, the contribution of a to ∂c is either $x_i a + x_j a$, or $x_i a - x_j a$, or $-x_i a + x_j a$, or $-x_i a - x_j a$, which implies that $x_i = \epsilon x_j$, with $\epsilon = \pm 1$. Consequently, if A_i and A_j are joined by a path of pairwise adjacent triangles, A_k, all in c, then $|x_i| = |x_j|$. However, Proposition 3.6 and Proposition 4.5 imply that any two triangles A_i and A_j in K are connected by a sequence of pairwise adjacent triangles. If some triangle in the path does not belong to c, then there are two adjacent triangles in the path, A_h and A_k, with A_h in c and A_k not in c such that all the triangles in the path from A_i to A_h belong to c. But then, A_h has an edge not adjacent to any other triangle in c, so $x_h = 0$ and thus, $x_i = 0$. The same reasoning applied to A_j shows that $x_j = 0$. If all triangles in the path from A_i to A_j belong to c, then we already know that $|x_i| = |x_j|$. Therefore, all x_i's have the same absolute value. If K is infinite, there must be some A_i in the finite sum which is adjacent to some triangle A_j not in the finite sum and the contribution of the edge common to A_i and A_j to ∂c must be zero, which implies that $x_i = 0$ for all i. Similarly, the coefficient of every triangle with an edge in the border must be zero. Thus, in these cases, $c \sim 0$, and $H_2(K) = 0$.

Let us now assume that K is a finite triangulated 2-complex without a border. The above reasoning showed that any nonzero 2-cycle, c, can be written as

$$c = \sum \epsilon_i x A_i,$$

where $x = |x_i| > 0$ for all i, and $\epsilon_i = \pm 1$. Since $\partial c = 0$, $\sum \epsilon_i A_i$ is also a 2-cycle. For any other nonzero 2-cycle, $\sum y_i A_i$, we can subtract $\epsilon_1 y_1 (\sum \epsilon_i A_i)$ from $\sum y_i A_i$, and we get the cycle

$$\sum_{i \neq 1} (y_i - \epsilon_1 \epsilon_i y_1) A_i,$$

in which A_1 has coefficient 0. But then, since all the coefficients have the same absolute value, we must have $y_i = \epsilon_1 \epsilon_i y_1$ for all $i \neq 1$, and thus,

$$\sum y_i A_i = \epsilon_1 y_1 (\sum \epsilon_i A_i).$$

This shows that either $H_2(K) = 0$, or $H_2(K) = \mathbb{Z}$.

It remains to prove that K is orientable iff $H_2(K) = \mathbb{Z}$. The idea is that in this case, we can choose an orientation such that $\sum A_i$ is a 2-cycle. The proof is not really difficult but a little involved and the reader is referred to Ahlfors and Sario [1] for details. □

Finally, we need to determine $H_1(K)$. We will only do so for finite triangulated 2-complexes and refer the reader to Ahlfors and Sario [1] for the infinite case.

Proposition 5.11 *For every finite triangulated 2-complex, K, either $H_1(K) = \mathbb{Z}^{m_1}$ or $H_1(K) = \mathbb{Z}^{m_1} \oplus \mathbb{Z}/2\mathbb{Z}$, the second case occurring iff K has no border and is nonorientable.*

Proof. The first step is to determine the torsion subgroup of $H_1(K)$. Let c be a 1-cycle, and assume that $mc \sim 0$ for some $m > 0$, i.e., there is some 2-chain, $\sum x_i A_i$, such that $mc = \sum x_i \partial A_i$. If A_i and A_j have a common edge, a, the contribution of a to $\sum x_i \partial A_i$ is either $x_i a + x_j a$, or $x_i a - x_j a$, or $-x_i a + x_j a$, or $-x_i a - x_j a$, which implies that either $x_i \equiv x_j (\mathrm{mod}\, m)$, or $x_i \equiv -x_j (\mathrm{mod}\, m)$. Because of the connectedness of K, the above actually holds for all i, j. If K is bordered, there is some A_i which contains a border edge not adjacent to any other triangle and thus, x_i must be divisible by m, which implies that every x_i is divisible by m. Thus, $c \sim 0$. Note that a similar reasoning applies when K is infinite but we are not considering this case. If K has no border and is orientable, by a previous remark, we can assume that $\sum A_i$ is a cycle. Then, $\sum \partial A_i = 0$, and we can write

$$mc = \sum (x_i - x_1)\partial A_i.$$

Due to the connectness of K, the above argument shows that every $x_i - x_1$ is divisible by m, which shows that $c \sim 0$. Thus, the torsion group is 0.

Let us now assume that K has no border and is nonorientable. Then, by a previous remark, there are no 2-cycles except 0. Thus, the coefficients in $\sum \partial A_i$ must be either 0 or ± 2. Let $\sum \partial A_i = 2z$. Then, $2z \sim 0$, but z is not homologous to 0, since from $z = \sum x_i \partial A_i$, we would get $\sum (2x_i - 1)\partial A_i \sim 0$, contrary to the fact that there are no 2-cycles except 0. Thus, z is of order 2.

Consider again $mc = \sum x_i \partial A_i$. Since $x_i \equiv x_j (\mathrm{mod}\, m)$, or $x_i \equiv -x_j (\mathrm{mod}\, m)$, for all i, j, we can write

$$mc = x_1 \sum \epsilon_i \partial A_i + m \sum t_i \partial A_i,$$

with $\epsilon_i = \pm 1$ and at least some coefficient of $\sum \epsilon_i \partial A_i$ is ± 2, since otherwise $\sum \epsilon_i A_i$ would be a nonnull 2-cycle. But then, $2x_1$ is divisible by m, and this implies that $2c \sim 0$. If $2c = \sum u_i \partial A_i$, the u_i are either all odd or all even. If they are all even, we get $c \sim 0$, and if they are all odd, we get $c \sim z$. Hence, z is the only element of finite order, and the torsion group if $\mathbb{Z}/2\mathbb{Z}$.

Finally, having determined the torsion group of $H_1(K)$, by the corollary of Proposition 5.2, we know that $H_1(K) = \mathbb{Z}^{m_1} \oplus T$, where m_1 is the rank of $H_1(K)$, and the proposition follows. □

Recalling Proposition 5.8, the Euler-Poincaré characteristic $\chi(K)$ is given by

$$\chi(K) = r(H_0(K)) - r(H_1(K)) + r(H_2(K)),$$

and we have determined that $r(H_0(K)) = 1$ and either $r(H_2(K)) = 0$ when K has a border or has no border and is nonorientable, or $r(H_2(K)) = 1$ when K has no border and is orientable.

Thus, the rank m_1 of $H_1(K)$ is either

$$m_1 = 2 - \chi(K)$$

if K has no border and is orientable and

$$m_1 = 1 - \chi(K)$$

otherwise. This implies that $\chi(K) \leq 2$.

We will now prove the classification theorem for compact (two-dimensional) polyhedras.