Chapter 4

COVERING PROJECTIONS AND FUNDAMENTAL GROUP

4.1 Basic Theory and Examples

4.1.1 The Fundamental Group

By a *path* in a topological space X we mean a continuous map $\gamma : I = [0,1] \to X$. To avoid pathologies, all topological spaces are assumed to be locally path connected and Hausdorf, and all maps between topological spaces are continuous. Given paths γ and δ with terminal point $\gamma(1)$ of γ equal to the initial point $\delta(0)$ of δ, it makes sense to consider the product $\delta\gamma$ as the path γ followed by the path δ. We scale the parameter so that we can still regard $\delta\gamma$ as a mapping of I into X. A path whose initial point and terminal points are identical is called a *loop*. We fix a point $x \in X$ and call it the *base point*. The set of loops with initial point x will be denoted by $\Omega(X, x)$. For a loop $\gamma \in \Omega(X, x)$ we define γ^{-1} by $\gamma^{-1}(t) = \gamma(1-t)$. On $\Omega(X, x)$ we define the equivalence relation $\sim$ by the requirement $\gamma \sim \delta$ if γ and δ are homotopic relative to ∂I, i.e., there is a continuous map $G : I \times I \to X$ with $G(t, 0) = \gamma(t),\ G(t, 1) = \delta(t), G(0, s) = \gamma(0) = \delta(0) = \gamma(1) = \delta(1) = G(1, s)$. The quotient $\Omega(X, x)/\sim$ has a group structure with multiplication defined by the product of paths as defined above. The inverse of the equivalence class $[\gamma]$, represented by the loop γ, is $[\gamma^{-1}]$. It is easy to verify that the loop $\gamma^{-1}\gamma$ is homotopic, relative to ∂I , to the constant loop $\underline{x} : t \to x$ which is the identity of the group. The homotopy is given by

$$G(t, s) = \begin{cases} \gamma(st) & \text{if } 0 \leq t \leq 1/2; \\ \gamma(s(1-t)) & \text{if } 1/2 \leq t \leq 1. \end{cases} \tag{4.1.1}$$

We often simply write γ for the equivalence class $[\gamma]$. We omit the simple verification of the fact that the product $[\gamma][\delta] = [\gamma\delta]$ is well-defined, i.e., is independent of the choice of representatives for homotopy classes of loops. This group is denoted by $\pi_1(X, x)$ and is called the *fundamental group*, *first homotopy group*, or the *Poincaré group* of X (with base point x). A space whose first homotopy group is trivial is called *simply-connected*.

The groups $\pi_1(X, x)$ and $\pi_1(X, y)$ with $x \neq y$ are isomorphic, but not canonically. In fact, since X is path-connected, there is $\lambda : I \to X$ with $\lambda(0) = x$ and $\lambda(1) = y$. Then the mapping $\gamma \to \lambda^{-1}\gamma\lambda$ induces an isomorphism $\pi_1(X, y) \simeq \pi_1(X, x)$ which depends on λ and is therefore not canonical. At any rate, it makes sense to talk about the fundamental group of a space. Also note that a mapping $f : (X, x) \to (Y, y)$ induces a homomorphism $f_\sharp : \pi_1(X, x) \to \pi_1(Y, y)$ by composition.

It is trivial that if a space is contractible, then its fundamental group is trivial. Furthermore, if $f : (X, x) \to (Y, y)$ is a homotopy equivalence, then $f_\sharp$ is an isomorphism. The spheres S^n are simply connected for $n > 1$ since every loop is homotopic relative to the ∂I to the constant loop through the obvious deformation (see Chapter 3, Example 1.7). In the next section we show that $\pi_1(S^1, x) \simeq \mathbf{Z}$. Intuitively, we assign to each loop in S^1 the number of times it winds around the circle, with the sign being positive or negative according as it is counterclockwise or clockwise.

The algebraic notion of free product of groups plays an important role in the computation of fundamental groups. We denote the elements of (abstract) groups G and H by g_j and h_k respectively. The *free product* $G \star H$ is the set of expressions of the forms (l any integer)

$$g_1h_1g_2h_2 \cdots g_lh_l, \quad h_1g_1h_2 \cdots h_lg_l, \quad g_1h_1g_2h_2 \cdots h_{l-1}g_l, \quad h_1g_1h_2 \cdots g_{l-1}h_l$$

with multiplication defined in the obvious manner, for example,

$$(g_{i_1} \cdots h_{i_l})(h_{j_1}g_{j_1} \cdots) = g_{i_1} \cdots (h_{i_l}h_{j_1})g_{j_1} \cdots, \quad \text{etc.}$$

It is straightforward to verify that $G \star H$ is a group. $G \star H$ is an infinite group unless both G and H are finite groups and one of them is the trivial group of one element. Given a group A and homomorphisms $\rho_1 : A \to G$ and $\rho_2 : A \to H$ we define $G \star_A H$ to be the quotient of $G \star H$ by its normal subgroup generated by elements of the form $\rho_1(a)\rho_2(a^{-1})$ (and their conjugates). Clearly $G \star_A H$ depends on the homomorphisms ρ_i.

Exercise 4.1.1 *Show that* $\underbrace{\mathbf{Z} \star \cdots \star \mathbf{Z}}_{n \text{ copies}}$ *is the free group on* n *generators.*

Exercise 4.1.2 *Let* $\rho_2 : A \to H$ *be an isomorphism. Show that* $G \star_A H \simeq G$.

Exercise 4.1.3 *Let $x \in \mathbf{Z}/4$ and $y \in \mathbf{Z}/3$ be generators. Show that*

$$x \longrightarrow \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad y \longrightarrow \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$$

gives a surjective homomorphism of $\mathbf{Z}/4 \star \mathbf{Z}/3$ *onto* $SL(2, \mathbf{Z})$. (It is in fact an isomorphism.)

The following simple proposition is useful in the computation of fundamental groups:

Proposition 4.1.1 *Let X and Y be simplicial or cell complexes. Then*

1. $\pi_1(X \times Y, (x, y)) \simeq \pi_1(X, x) \times \pi_1(Y, y)$, *(also true for infinite products);*

2. *Let $X \vee Y$ be the space obtained by joining X and Y at the points $x \in X$ and $y \in Y$. Then $\pi_1(X \vee Y, x = y)$ is isomorphic to the free product of $\pi_1(X, x)$ and $\pi_1(Y, y)$.* (Some mild assumptions on X and Y are necessary for the validity of this result. See [Gri] and remark 4.2.1 below. This result is a special case of van Kampen's theorem given below.)

The proof of (1) is a straightforward application of the definition to the projections $X \times Y \to X$ and $X \times Y \to Y$. (2) is intuitively reasonable since there is no relation between loops in X and Y. It is also a special case of theorem 4.2.1 below.

Assuming knowledge of $\pi_1(S^1, x) \simeq \mathbf{Z}$, the proposition implies that the fundamental group of the n-torus is isomorphic to $\mathbf{Z}^n$. Similarly,

Exercise 4.1.4 *Show that fundamental group of n circles joined at a point (bouquet of n circles), is the free group $\mathbf{F}_n$ on n generators.*

Exercise 4.1.5 *Let $L_1, \cdots, L_m$ be m lines passing through the origin in $\mathbf{R}^3$, and $\Delta = \cup L_j$. Show that the sphere S^2 with $2m$ points removed is a deformation retract of $\mathbf{R}^3 \setminus \Delta$. The former space has the homotopy type of the bouquet of $2m - 1$ circles, and consequently $\pi_1(\mathbf{R}^3 \setminus \Delta) \simeq \mathbf{F}_{2m-1}$.*

More will be said about the computation of π_1 in the following sections.

Example 4.1.1 Consider the half-space $\bar{\mathcal{H}}_3 = \{(x, y, z) |\ z \geq 0\}$. Let $C_1, ..., C_n$ be half-circles or $\sqcap$-shaped curves with end points on the plane $z = 0$ and lying in planes orthogonal to this plane. It will be clear that the analysis in this example is applicable to much more general curves than C_i's, but this simple case demonstrates the idea clearly. Assume C_i's are disjoint, and let $C = \cup C_i$. Let $v = (v_1, v_2, v_3) \in \bar{\mathcal{H}}_3$ with v_3 large. It is a simple matter to see that $\pi_1(\bar{\mathcal{H}}_3 \setminus C; v)$ is isomorphic to the free group $\mathbf{F}_n$ on n generators. A set of generators $\chi_1, ..., \chi_n$ for the fundamental group are the loops given in Figure 1.1 with given orientations. We also assign orientations to the curves C_i as shown in Figure 1.1. Let P be a plane perpendicular to $z = 0$ such that the orthogonal projection of each curve C_i into P is one-to-one. It is no loss of generality to assign orientations to C_i's and χ_i's in such a way whenever the curve χ_i goes behind C_i the the direction of χ_i followed by the direction of C_i is a positively oriented basis for the plane P. Note that since we have fixed the plane P it makes sense to say whether χ_i goes behind or in front C_i at a point where their orthogonal projections on P intersect. Now consider an oriented loop γ with base-point v. For example, let γ be the dotted curve in Figure 1.1. We want to express the homotopy class of $[\gamma]$ in terms of the generators $\chi_1, ..., \chi_n$. We look at the projection of γ on the plane P, and determine whether at each intersection of the projection of γ with those of the χ_i's whether γ is behind χ_i or not. We only consider the points where γ goes behind one of the C_i's. Following along γ at the first such point γ goes behind C_2. The ordered pair consisting of the positively oriented tangents to γ and C_2 form a positively oriented basis for $\mathbf{R}^2$. We assign to this point χ_2. At the next relevant point γ goes behind C_1 but the ordered pair of tangents will be negatively oriented. We multiply χ_2 by χ_1^{-1} to obtain $\chi_1^{-1}\chi_2$. Following along the curve γ we finally obtain $\chi_1\chi_2\chi_3\chi_1^{-1}\chi_2$. This is the expression for $[\gamma]$ in terms of the generators $\{\chi_i\}$ of $\pi_1(\bar{\mathcal{H}}_3 \setminus C; v)$. ♠

4.1.2 Covering Spaces

A triple (E, p, B) is called a *covering projection* (or a *covering space*) if $p : E \rightarrow B$, and for every $x \in B$ there is a neighborhood U of x such that $p^{-1}(U)$ is a disjoint union $\cup V_i$ with restriction of p to each V_i a homeomorphism onto U. The *fibre* over b is the set $p^{-1}(b)$ which is a discrete set. It is a simple matter to show that if $p^{-1}(b)$ is finite for some b, then it is finite for all b (B is path-wise connected), and all the fibres have the same cardinality. If the cardinality of $p^{-1}(b)$ is n, then we say that E is an *n-sheeted covering* of B. The simplest example of a covering space is

$$E = \mathbf{R}, \quad B = S^1 \subset \mathbf{C}, \text{ and } p(t) = e^{it}.$$

By a *lift(ing)* of a mapping $f : X \rightarrow B$ we mean a mapping $f' : X \rightarrow E$ such that $pf' = f$. Similarly, if $f : (X, x) \rightarrow (B, b)$ and $e \in p^{-1}(b)$, then by a *lift* of f to $f' : (X, x) \rightarrow (E, e)$ we mean a lift of $f : X \rightarrow B$ to E with the additional requirement $f'(x) = e$. The theory of covering spaces depends on two important properties which are described by the following definitions:

1. (*Homotopy Lifting Property*) - A triple (E, p, B) with $p : E \rightarrow B$ (not necessarily a covering projection) has the homotopy lifting property with respect to a space X, if given a homotopy $F : X \times I \rightarrow B$, $(I = [0, 1])$, and $f(x) = F(x, 0)$ admitting of a lifting $f' : X \rightarrow E$, then the homotopy F admits of a lifting F'. If the homotopy lifting property holds for all X, we say that (E, p, B) has the homotopy lifting property. Triples (E, p, B) having the homotopy lifting property are called *fibrations*.

2. (*Unique Path Lifting Property*) - A triple (E, p, B) with $p : E \rightarrow B$ has the unique path lifting property if given $\gamma : I \rightarrow B$, and $e \in p^{-1}(\gamma(0))$, then there is a unique lift $\gamma' : I \rightarrow E$ of γ with $\gamma'(0) = e$.

The theorems below, whose proofs are postponed to the end of this section, describe the basic properties of covering spaces.

Theorem 4.1.1 *Covering projections are fibrations with unique path lifting. Furthermore, if $\gamma, \gamma' : I \times I \rightarrow B$, relative to ∂I, then for any lifting $G' : I \times I \rightarrow E$, $G'(0, s)$ and $G'(1, s)$ are independent of s (i.e., homotopy relative to ∂I lifts to homotopy relative ∂I), and we can choose $G'(0, s)$ to be any point in the fibre over $\gamma(0)$.*

Theorem 4.1.2 *Every locally simply connected topological space B admits of a unique (up to equivalence) covering space (E, p, B) with E simply connected.*

A covering projection (E, p, B) with E simply connected is called the *universal covering space* of B.

While the definition of a fibration $E \rightarrow B$ requires the homotopy lifting property for every space X, in practice it is adequate to establish this property for only a class of simple spaces. This point will be clear from the applications of the homotopy lifting property and there is no need for further elaboration at this point.

Example 4.1.2 In this example we show that fibre bundles are fibrations, i.e., satisfy the homotopy lifting property. Let $\pi : E \rightarrow M$ be a fibre bundle, with typical fibre F and base a manifold M. Let X be a finite simplicial complex and $F : X \times I \rightarrow M$ a continuous mapping such that $F_\circ = F(., 0) = f : X \rightarrow M$ admits of a lifting to $F'_\circ : X \rightarrow E$. We want to establish the existence of $F' : X \times I \rightarrow E$ such that

$$\pi \cdot F'(x, t) = F(x, t), \quad \text{and} \quad F'(x, 0) = F'_\circ(x).$$

After a sufficiently fine subdivision of X and a partition $0 = t_\circ < t_1 < \cdots < t_r < t_{r+1} = 1$ we may assume for every simplex c of X and every j, $F(c \times [t_j, t_{j+1}])$ lies in a neighborhood $U_{(c,j)} \subset M$ which trivializes the bundle $\pi : E \to M$. To fix notation we let the maps

$$\psi_U : U \times F \xrightarrow{\simeq} \pi^{-1}(U)$$

give the desired local trivializations and $\pi_2 : U \times F \to F$ denote the projection on the second coordinate. Let X^s denote the s-skeleton of X. We inductively construct the desired extension of F. The induction is by assuming that the required extension has been constructed for for $(X^{n-1} \times [0,1]) \cup (X^n \times [0, t_j])$ and extending it to $(X^{n-1} \times [0,1]) \cup (X^n \times [0, t_{j+1}])$. The initial step of the induction is clearly valid. Let c be an n-simplex of X so that an extension of F, by induction hypotheses, has been constructed for $(c \times [0, t_j]) \cup (\partial c \times [0, t_{j+1}])$. Since the inclusion

$$\jmath : (c \times [t_j]) \cup (\partial c \times [t_j, t_{j+1}]) \longrightarrow c \times [t_j, t_{j+1}]$$

is a retract, there is a map $\rho : c \times [t_j, t_{j+1}] \to (c \times [t_j]) \cup (\partial c \times [t_j, t_{j+1}])$ such that $\rho \cdot \jmath = \mathrm{id}$. Now, for $x \in c$ and $t \in [t_j, t_{j+1}]$ define the extension

$$F'(x,t) = \psi_{U_{(c,j)}}(F(x,t), \pi_2 \cdot \psi^{-1}_{U_{(c,j)}} \cdot F'(\rho(x,t))).$$

Note that $\rho(x,t) \in (c \times [t_j]) \cup (\partial c \times [t_j, t_{j+1}])$ so that $F'(\rho(x,t))$ is defined. This extension fulfills the requirements. ♠

4.1.3 Structure of Covering Spaces

In this subsection we relate the fundamental group to covering spaces and in essence develop a methodology for understanding the structure of covering spaces. This will be accomplished in a series of corollaries to theorem 4.1.1 and the uniqueness part of theorem 4.1.2 will also be proven. For a covering projection (E, p, B) there is the induced map of fundamental groups $p_\sharp : \pi_1(E, e) \to \pi_1(B, b)$ where $p(e) = b$, defined by $p_\sharp([\gamma]) = [p\gamma]$. We have

Corollary 4.1.1 *$p_\sharp$ is one-to-one.*

Proof - Let $[\gamma] \in \pi_1(E, e)$ and assume that $p_\sharp([\gamma]) = e \in \pi_1(B, b)$, then $p_\sharp([\gamma])$ is homotopic to the constant map $\underline{b} : I \to b$. Lifting the homotopy to E (which is possible since γ is a lift of $p\gamma$ to E), and noting that the constant map necessarily lifts to a constant map in view of discreteness of the fibre, we the desired homotopy between γ and a constant map. ♣

The unique path lifting property ensures that any loop $\gamma : I \to B$, with $\gamma(0) = b$, has a unique lift $\gamma' : I \to E$ with $\gamma'(0) = e$, however, the lift γ' may not be a loop. Even if a lift γ' is a loop, another lift $\gamma'' : I \to E$ with $\gamma''(0) = e' \in p^{-1}(b)$ may not be a loop (see example 2.4 below). A covering projection with the property that for any given loop, either all its lifts are loops or none is a loop is called *regular*.

Corollary 4.1.2 *$p_\sharp(\pi_1(E,e))$ consists precisely of homotopy classes of loops $\gamma : (I, \partial I) \to (B, b)$ whose lifts to (E, e) are loops. $p_\sharp(\pi_1(E,e))$ is a normal subgroup of $\pi_1(B,b)$ if and only if (E,p,B) is a regular covering space. Hence for regular coverings $\pi_1(B,b)/p_\sharp(\pi_1(E,e))$ is a group.*

Proof - The first assertion is trivial. Let $\delta : I \to B$ with $[\delta] \in \pi_1(B,b)$, and $\gamma : I \to E$ with $[\gamma] \in \pi_1(E,e)$. We may lift δ to a map $\delta' : I \to E$ with $\delta'(0) = e$. Now lift $p\gamma$ to $(p\gamma)' : I \to E$ with $(p\gamma)'(0) = \delta'(1)$. By the regularity assumption, $(p\gamma)'$ is a loop. Then the composition $\delta'^{-1}(p\gamma)'\delta'$ defines an element of $\pi_1(E,e)$, and $p_\sharp([\delta'^{-1}(p\gamma)'\delta']) = [\delta]^{-1}[\gamma][\delta]$ proving the normality of $p_\sharp(\pi_1(E,e))$ for regular covering spaces. Conversely, note that if $\delta : I \to E$ with $\delta(0) = e$ and $\delta(1) = e'$, $(e' \in p^{-1}(b))$, then $p_\sharp(\pi_1(E,e')) = [p\delta]p_\sharp(\pi_1(E,e))[p\delta]^{-1}$. Therefore if $p_\sharp(\pi_1(E,e))$ is normal, then the set of homotopy classes of loops whose lifts to (E,e') are loops coincides with $p_\sharp(\pi_1(E,e))$, and the regularity of (E,p,B) follows. ♣

A group Γ acts *properly discontinuously*[1] (on left) on a space X if for every $x \in X$ there is a neighborhood U of x such that for all $e \neq \gamma \in \Gamma$, $U \cap \gamma(U) = \emptyset$. Notice that if Γ acts properly discontinuously, then $X \to \Gamma \setminus X$ is a covering projection. We now show that all regular covering projections are of this form. For the regular covering projection (E,p,B), let $\Gamma = \pi_1(B,b)/p_\sharp(\pi_1(E,e))$. Let $\gamma \in \Gamma$ and $\gamma' : I \to B$ be a loop with $\gamma'(0) = \gamma'(1) = b$ representing γ. For $x \in E$, let $\lambda : I \to E$ be any path with $\lambda(0) = e$ and $\lambda(1) = x$. Let γ'' be a lift of γ' with $\gamma''(0) = e$, and $L(x,\gamma') : I \to E$ be the lift of $p\lambda$ with $L(x,\gamma')(0) = \gamma''(1)$. Define $\gamma(x) = L(x,\gamma')(1)$. We have to check that this definition is meaningful, i.e., the choice of the representative for γ in $\pi_1(B,b)$ and the subsequent choice of the loop γ' representing the element of the fundamental group, and the choice of the path $\lambda : I \to E$ will not affect the value of $\gamma(x)$. These assertions all follow easily from theorem 4.1.1. For example, if we replace λ by another path λ', then $\lambda'^{-1}\lambda$ defines an element of $\pi_1(E,e)$ and so the lift of $p\lambda'^{-1}\lambda$ with initial point $\gamma''(1)$, is a loop by the regularity assumption on (E,p,B). In particular, the lifts $L(x,\gamma')$ and δ of $p\lambda$ and $p\lambda'$ with $L(x,\gamma')(0) = \delta(0) = \gamma''(1)$ have the same end-points $\delta(1) = \delta'(1)$. Similarly, one shows independence from the choice of loop

[1] Our definition of properly discontinuous is more restrictive than the conventional one where one allows $U \cap \gamma(U) \neq \emptyset$ for finitely many γ's. We refer to the action as *discontinuous* when one allows finitely many γ's for which $U \cap \gamma(U) \neq \emptyset$.

representing $\gamma \in \Gamma$. The action of Γ is properly discontinuous since if $e \neq \gamma \in \Gamma$ then $\gamma(x) \in p^{-1}(p(x))$ and is distinct from x.

Corollary 4.1.3 *All regular covering spaces are of the form $E \rightarrow \Gamma \setminus E$ where the group Γ acts properly discontinuously on E. Conversely, if the group Γ acts properly discontinuously on E, then $p : E \rightarrow \Gamma \setminus E = B$ is a regular covering space, and $\pi_1(B, b)/p_\sharp(\pi_1(E, e))$ is isomorphic to Γ. Here $e \in p^{-1}(b)$ is any point.*

Proof - It only remains to prove the isomorphism $\pi_1(B, b)/p_\sharp(\pi_1(E, e)) \simeq \Gamma$ in the converse statement. Consider the mapping $Q : \pi_1(B, b) \rightarrow \Gamma$ defined by $Q([\delta]) = \gamma^{-1}$ where $\delta'(1) = \gamma(e)$, and δ' is the lift of δ with $\delta'(0) = e$. Clearly Q is well-defined and surjective. We show that Q is a homomorphism. Assume $Q([\lambda]) = \gamma'^{-1}$. The lift of the loop $\lambda\delta$ with $(\lambda\delta)'(0) = e$, is the path $\gamma(\lambda')\delta'$, where λ' is the lift of λ with $\lambda'(0) = e$, and so we have $Q([\lambda\delta]) = (\gamma\gamma')^{-1} = \gamma'^{-1}\gamma^{-1} = Q([\lambda])Q([\delta])$. Since the kernel of Q is $p_\sharp(\pi_1(E, e))$ (corollary 4.1.2), the proof is complete. ♣

Example 4.1.3 Since $\mathbf{Z}^n$ acts properly discontinuously on $\mathbf{R}^n$, the fundamental group of the n-torus $T^n = \mathbf{R}^n/\mathbf{Z}^n$ is $\mathbf{Z}^n$. Furthermore, the images in T^n of the straight line segments joining the origin to the points $(1, 0, \cdots, 0), \cdots, (0, \cdots, 0, 1)$, form a basis for its fundamental group. ♠

Example 4.1.4 As another application of group actions we construct, for every finitely generated abelian group G, a compact manifold M with $\pi_1(M) \simeq G$. For every integer q the cyclic group of order q can be written as $\mathbf{Z}/q \simeq \{e^{\frac{2\pi ij}{q}} | j = 0, 1, \cdots, q-1\}$. Now $S^3 = \{(z_1, z_2) | |z_1|^2 + |z_2|^2 = 1\}$ is simply connected. Therefore to construct a compact manifold whose fundamental group is isomorphic to $\mathbf{Z}/q$, it suffices to define a properly discontinuous action of $\mathbf{Z}/q$ on S^3. Such an action is given by

$$e^{\frac{2\pi ij}{q}} : (z_1, z_2) \longrightarrow (e^{\frac{2\pi ij}{q}} z_1, e^{\frac{2\pi ijk}{q}} z_2),$$

where $k \not\equiv 0 \bmod q$ is any integer. This action has no fixed point from which it easily follows that it is properly discontinuous. We denote the quotient space of S^3 under this action of $\mathbf{Z}/q$ by $L(q; k)$ and we have shown $\pi_1(L(q; k)) \simeq \mathbf{Z}/q$. For a general finitely generated abelian group G we have the decomposition

$$G \simeq \mathbf{Z}^k \times \mathbf{Z}/q_1 \times \mathbf{Z}/q_2 \times \cdots \times \mathbf{Z}/q_n$$

into a product of cyclic groups. It follows that

$$\pi_1(\underbrace{S^1 \times \cdots \times S^1}_{k \text{ copies}} \times L(q_1) \times \cdots \times L(q_n)) \simeq G.$$

$L(q;k)$ is called a *Lens* space. ♠

Corollary 4.1.4 *Let (E,p,B) be a covering projection and $f:(X,x) \to (B,b)$. Then f can be lifted to a map $\varphi:(X,x) \to (E,e)$, $(p(e)=b)$ such that $p\varphi = f$ if and only if $f_\sharp(\pi_1(X,x)) \subseteq p_\sharp(\pi_1(E,e))$. The lift is unique once the base point $e \in E$ is specified.*

Proof - The necessity follows from $p_\sharp\varphi_\sharp = f_\sharp$. Let $y \in X$ and $\gamma : I \to X$ be any path with $\gamma(0) = x$ and $\gamma(1) = y$. Define $\varphi(y) = (f\gamma)'(1)$ where $(f\gamma)'$ is the unique lift of $f\gamma$ to a path in E with $(f\gamma)'(0) = e$. We have to show that φ is well-defined. If δ is any other path joining x to y, then $\delta^{-1}\gamma = \tau$ defines an element of $\pi_1(X,x)$ and since $f_\sharp(\pi_1(X,x)) \subseteq p_\sharp(\pi_1(E,e))$, the lift of $f\tau$ with initial point e is a loop. Hence $(f\gamma)'(1) = (f\delta)'(1)$ as desired. Uniqueness follows uniqueness of path lifting. ♣

Example 4.1.5 To appreciate the significance of corollary 4.1.4 we prove the fact stated in chapter 1, section (XXX) that the winding number of a simple closed curve in the plane is ± 1. Suppose we have a mapping $\psi : T^2 \to S^1$, where T^2 is the two dimensional torus. Corollary 4.1.4 implies that ψ lifts to a mapping $\tilde{\psi} : \mathbf{R}^2 \to \mathbf{R}$ of their universal covering spaces. Now consider a C^1 closed curve $\gamma : [0,L] \to \mathbf{R}^2$ which we assume is parametrized by arc length and L is the length of the curve. Regarding γ as a mapping of a circle of radius $\frac{L}{2\pi}$ into $\mathbf{C} = \mathbf{R}^2$ we define $\psi : T^2 \to S^1$, for $s_1 \neq s_2$, by

$$\psi(s_1,s_2) = e^{i\mathrm{Arg}(\gamma(s_2)-\gamma(s_1))} = \frac{\gamma(s_2)-\gamma(s_1)}{|\gamma(s_2)-\gamma(s_1)|}.$$

For the limiting value $s_1 \to s_2$ we obtain the unit tangent vector field to γ:

$$\psi(s,s) = \gamma'(s).$$

Notice that ψ is defined except at the points of self intersection of the curve γ and it takes values in S^1. Thus for γ a simple closed curve, ψ is defined everywhere on T^2. The lift $\tilde{\psi}$ of ψ is a real valued function on $\mathbf{R}^2$. Let us try to understand the behavior of the function $\tilde{\psi}$. We assume, with no loss of generality, that $\gamma(0)$ is the origin in $\mathbf{R}^2$, $\gamma'(0) = (1,0)$ and the entire curve lies in the upper half plane. It is clear from the definition of winding number that

$$\frac{1}{2\pi}(\tilde{\psi}(L,L)-\tilde{\psi}(0,0)) = \pm W(\gamma)$$

is the winding number of γ. To compute this number we move along the s_2-axis from $(0,0)$ to $(0,L)$, then we move from $(0,L)$ to (L,L) by moving parallel to the s_1-axis. It follows easily that (draw a picture and use the hypotheses on how the curve is situated in $\mathbf{R}^2$)

$$\tilde{\psi}(0,0)=0, \quad \tilde{\psi}(0,L)=-\pi, \quad \tilde{\psi}(L,L)-\tilde{\psi}(0,L)=-\pi.$$

The required result that $W(\gamma)=\pm 1$ follows immediately. Note that the essential point in the argument was that we can lift the circle valued function ψ to a (single-valued) function on $\mathbf{R}^2$ with values in $\mathbf{R}$, and we were able to compute $\tilde{\psi}(L,L)$ by moving along the s_2-axis and then parallel to the s_1 axis, rather than along the diagonal. ♠

Notice that in corollary 4.1.4 we required $\varphi(x)=e$. Let us instead assume that (X,f,B) is a covering projection and ask whether there is a lift $\varphi : X \to E$ such that $p\varphi = f$, but do not require $\varphi(x)=e$. It is clear that if φ exists then $\varphi(x)=p^{-1}(b)$. The subgroups $p_\sharp(\pi_1(E,e'))$, $e'\in p^{-1}(b)$, are all conjugates of $p_\sharp(\pi_1(E,e))$ in $\pi_1(B,b)$. In fact, if λ is a path joining e to e', then $p\lambda$ is a loop, and

$$p_\sharp(\pi_1(E,e)) = [p\lambda]^{-1}p_\sharp(\pi_1(E,e'))[p\lambda].$$

Furthermore, by lifting a loop representing an element of $\pi_1(B,b)$ to E with $\gamma(0)=e$ we see that all conjugates are of this form. Therefore we have shown

Corollary 4.1.5 *Let (E,p,B) and (X,f,B) be covering projections. Then there is a map $\varphi : X\to E$ such that $p\varphi=f$ if and only if $f_\sharp(\pi_1(X,x))$ is contained in a conjugate of $p_\sharp(\pi_1(E,e))$ in $\pi_1(B,b)$. Every conjugate of $p_\sharp(\pi_1(E,e))$ is of the form $p_\sharp(\pi_1(E,e'))$ for some $e'\in p^{-1}(b)$.*

We say that two covering projections (E,p,B) and (E',p',B') are *equivalent* if there is a homeomorphism $f : E\to E'$ such that $pf=p'$. With the aid of corollary 4.1.5 we easily obtain the following:

Corollary 4.1.6 *Two covering spaces (E,p,B) and (E',p',B') are equivalent if and only if $p_\sharp(\pi_1(E,e))$ and $p'_\sharp(\pi_1(E',e'))$ are conjugate in $\pi_1(B,b)$, where $e\in p^{-1}(b)$ and $e'\in p'^{-1}(b)$.*

Proof - The necessity follows easily from corollary 2.5. We prove the sufficiency. By corollary 2.5 we have covering projections $\varphi : E \to E'$ and $\varphi' : E' \to E$ such that $p'\varphi = \varphi'$ and $p\varphi' = \varphi$. Since the map φ is a local homeomorphism, it suffices to show that it is injective. In view of corollary 2.4 and the hypothesis, $p_\sharp(\pi_1(E, e)) = p'_\sharp(\pi_1(E', e'))$ for $e' = \varphi(e)$. Then by another application of corollary 2.4, we have $\varphi'\varphi(e) = e$. If φ were not injective, then there would exist $x \in \varphi^{-1}(e')$ distinct from e. Let λ be a path joining e to x, then λ and $\varphi'\varphi\lambda$ are lifts of $p\lambda$ with initial point e, which contradicts the uniqueness of path lifting unless $x = e$. ♣

Corollary 4.1.6 may be regarded as a uniqueness theorem for covering spaces and proves uniqueness part of theorem 4.1.2, however, it says nothing about the existence of such coverings.

Since $\pi_1(B, b)$ acts properly discontinuously on the universal cover E (corollary 4.1.3) the existence of the universal cover (theorem 4.1.2) implies the existence and uniqueness of a covering space for every conjugacy class of subgroups of $\pi_1(B, b)$. In fact, for every subgroup Γ of $\pi_1(B, b)$ the natural projection $\Gamma \setminus E \to B$ is the desired covering space and the uniqueness follows from corollary 4.1.6. Summarizing:

Corollary 4.1.7 *Covering spaces of B are in one to one correspondence with conjugacy classes of subgroups of $\pi_1(B, b)$.*

Example 4.1.6 The universal cover of figure **8** is given in Figure 2.1. The fundamental group of figure **8** is the free group $\mathbf{F}_2$ on two generators. Its universal cover is the infinite homogeneous tree four edges meeting at each vertex. (By definition a *tree* is a graph X with no loops, or equivalently a contractible graph.) Note the images of the point $\circ$ under the action of the elements x, y, xy, and yx of $\mathbf{F}_2$ (x and y are generators). ♠

Exercise 4.1.6 *Assuming $x \cdot \circ$ and $y \cdot \circ$ are as given in Figure 2.1, and using the procedure described above for the action of Γ on E show that $xy \cdot \circ$ and $yx \cdot \circ$ are as given. Similarly, the universal covering space of a bouquet of k circles is the infinite homogeneous tree with $2k$ vertices meeting at each vertex, and its fundamental group is the free group $\mathbf{F}_k$ on k generators.*

Example 4.1.7 We use example 4.1.6 to construct a non-regular covering space. In Figure 2.2 we have a three-sheeted covering of figure **8**. Let γ denote the loop which is the circle on the left in the counterclockwise direction and $\gamma(0) = b$. Then the lift of γ with initial point e or e' is not closed, while its lift with initial point e'' is closed. One can similarly construct n-fold coverings of a surface M_g of genus g. In Figure 2.3 and 2.4 we have exhibited how copies of a surface of genus 2 could be cut and pasted to obtain 2-fold and 3-fold coverings. It is clear that the procedure generalizes. However, M_g admits of many n-fold coverings (see corollary 4.1.8 exercise 4.1.12 below). ♠

Example 4.1.8 (Compare with the proof of theorem 4.1.2 below) We show that all simply connected manifolds are orientable. Recall that by an orientation at $x \in M$ we mean the choice of a generator $\mathbf{m}_x \in H_m(M, M \setminus x) \simeq \mathbf{Z}$. We have the *coherence* condition that every $x \in M$ has a (compact) neighborhood K and a generator $\mathbf{m}_K \in H_m(M, M \setminus K) \simeq \mathbf{Z}$ whose image in $H_m(M, M \setminus x)$ is the generator $\mathbf{m}_x \in H_m(M, M \setminus x)$ for all $x \in K$. Let $\mathbf{U} = \{\mathcal{U}_i\}$ be a covering of M by open relatively compact open sets such that $H_m(M, M \setminus \overline{U_i}) \simeq \mathbf{Z}$. For every i let $\mathbf{m}_i \in H_m(M, M \setminus \overline{U_i})$ be a generator. Let M' be the manifold with covering $\{\mathcal{U}'_i\} \cup \{\mathcal{U}''_i\}$ with $\mathcal{U}'_i = \{(x, \mathbf{m}_i) | x \in \mathcal{U}_i\}$ and $\mathcal{U}''_i = \{(x, -\mathbf{m}_i) | x \in \mathcal{U}_i\}$ and the points $(x, \mathbf{m}_i)$ and $(y, \mathbf{m}_j)$ (resp. $(y, -\mathbf{m}_j)$ are identified if $x = y \in \mathcal{U}_i \cap \mathcal{U}_j$ and the images of $\mathbf{m}_i$ and $\mathbf{m}_j$ (resp. $-\mathbf{m}_j$) are identical in $H_m(M, M \setminus x)$. We define $p : M' \to M$ by defining p on $\mathcal{U}_i$ and $\mathcal{U}'_i$ by $p(x, \pm\mathbf{m}_i) = x$ which is clearly meaningful. We show that if M is not orientable then M' is path-connected, and so M admits of a double covering and this will prove the assertion. If M is not orientable then for every $x \in M$ we can choose a sequence $K_1, \cdots, K_k$ of compact sets with $x \in \breve{K}_1 \cap \breve{K}_k$, $\breve{K}_i \cap \breve{K}_{i+1} \neq \emptyset$ and generators $\mathbf{m}_i \in H_m(M, M \setminus K_i)$ such that the images of $\mathbf{m}_i$ and $\mathbf{m}_{i+1}$ are identical in $H_m(M, M \setminus (K_i \cap K_{i+1})$ for $1 \leq i \leq i - 1$ and $\mathbf{m}_k = -\mathbf{m}_1$. This clearly proves path-connectedness of M' and we are done. ♠

Exercise 4.1.7 *Show that the argument of example 2.4 actually proves more, viz., if M is not orientable then $\pi_1(M, x)$ contains a (normal) subgroup of index 2.*

Exercise 4.1.8 *Show that every vector bundle over a simply connected manifold is orientable.* (The argument in example 4.1.8 can be used with the definition of orientation in terms of choice of bases in which case it adopts to vector bundles.)

Exercise 4.1.9 *Let $E \to M$ be a non-orientable vector bundle. Show that there is a double covering (M', p, M) such that $p^\star(E) \to M'$ is orientable.*

Example 4.1.9 We can now prove the statement in chapter 2, example (XXX) that every compact surface $M \subset \mathbf{R}^3$ of genus $\neq 1$ has at least one umbilical point. Assume the surface M has no umbilical points, then there are two smooth functions κ_i, $i = 1, 2$, defined on M which give the principal curvatures of M, and we may assume $\kappa_1(x) > \kappa_2(x)$ for all $x \in M$. Let $\mathcal{L}_i \to M$ be the real line bundle on M whose fibre at $x \in M$ is the eigenspace (in $\mathcal{T}_x M$) for eigenvalue κ_i of the second fundamental form. If M has genus zero, then it is simply connected and $\mathcal{L}_i \to M$ are trivial bundles and therefore we obtain nowhere vanishing vector fields on the sphere which is not possible. If M has genus > 1, then consider the double covering (M', p, M) such that $p^\star(\mathcal{L}_1) \to M'$ is orientable and therefore trivial (exercise 4.1.9). Since $p^\star(\mathcal{T}M)$ is the tangent bundle of M', $p^\star(\mathcal{L}_1) \to M'$ is a trivial sub-bundle of its tangent bundle and consequently M' has a nowhere vanishing vector field which is not possible since genus of M' is $\neq 1$. ♠

While corollary 4.1.7 gives a complete classification of covering spaces of B, it does not immediately give a practical method for the enumeration and the construction of finite sheeted coverings of B. There is a fairly practical method for the enumeration of finite covers of a space which we now describe. Given a covering projection (E, p, B), $b \in B$, we label $p^{-1}(b) = \{e_1, \cdots, e_n\}$. Clearly $\pi_1(B, b)$ acts as a group of permutations of $\{e_1, \cdots, e_n\}$. In fact, for every loop $\gamma : I \to B$ with $\gamma(0) = b$, let γ'_i be the lift of γ with $\gamma'_i(0) = e_i$, and define $\gamma(e_i) = \gamma'(1)$. (Notice that this action is defined only on the fibre $p^{-1}(b)$ and does not extend to an action on E unless (E, p, B) is a regular covering; see the argument preceding corollary 2.3.) The action of $\pi_1(B, b)$ on $p^{-1}(b) = \{e_1, \cdots, e_n\}$ is transitive,- it has only one orbit. Hence we have a homomorphism $\rho : \pi_1(B, b) \to \mathcal{S}_n$ (=the permutation group on n letters) and $\mathrm{Im}(\rho)$ is a transitive group of permutations of n letters. Relabeling $p^{-1}(b)$ by a permutation $\sigma \in \mathcal{S}_n$ replaces ρ by $\sigma^{-1}\rho\sigma$. We say two homomorphisms $\rho,\ \rho' : \pi_1(B, b) \to \mathcal{S}_n$ are equivalent if $\rho' = \sigma^{-1}\rho\sigma$ for some $\sigma \in \mathcal{S}_n$.

Corollary 4.1.8 *n-sheeted coverings of B are in 1-1 correspondence with equivalence classes of homomorphisms $\rho : \pi_1(B, b) \to \mathcal{S}_n$ such that* $\mathrm{Im}(\rho)$ *is a transitive group of permutation of n letters.*

Proof - We have already shown that an n-sheeted cover of B determines an equivalence class of homomorphisms $\rho : \pi_1(B, b) \to \mathcal{S}_n$ satisfying the transitivity condition. Conversely, given a homomorphism $\rho : \pi_1(B, b) \to \mathcal{S}_n$ satisfying the transitivity condition, let $\Gamma = \{\gamma \in \pi_1(B, b) | \rho(\gamma)(1) = 1\}$ and $E_\rho = \Gamma \setminus E$ where E is the universal cover of B. In view of the transitivity condition, Γ has index n in $\pi_1(B, b)$ and $E_\rho \to B$ is an n-sheeted covering space. All coverings of B equivalent to $E_\rho \to B$ are obtained by replacing Γ with a conjugate subgroup. Let $\Gamma' = \tau\Gamma\tau^{-1}$, and $p' : E' = \Gamma' \setminus E \to B$ be the natural projection. The equivalence $E_\rho \to E'$ is given by $x \to \tau(x)$ which is well-defined since $\gamma(x) \to \tau\gamma(x) = (\tau\gamma\tau^{-1})(\tau(x))$. Therefore $p'^{-1}(b) = \{\tau(e_1), \cdots, \tau(e_n)\}$, and

$$\gamma(\tau(e_j)) = e_{\rho(\gamma)\rho(\tau)(j)} = \tau(e_{\rho(\tau')(j)})$$

where $\tau' = \tau^{-1}\gamma\tau$. Hence $\pi_1(B, b)$ acts on $p'^{-1}(b)$ via the homomorphism

$$\rho'(\gamma) = \rho(\tau^{-1})\rho(\gamma)\rho(\tau),$$

which completes the proof of the corollary. ♣

As an application of corollary 4.1.8 we have

Exercise 4.1.10 *(a) Show that there are three double coverings of figure* **8** *and realize them geometrically. (b) Show that there are seven 3-sheeted coverings of figure* **8** *and realize them geometrically.*

Exercise 4.1.11 *Which homomorphism $\rho : \mathbf{F}_2 \rightarrow \mathcal{S}_3$ does the 3-sheeted cover of Figure 2.2 correspond to? Describe explicitly the subgroups $p_\sharp(\pi_1(E, e))$ and $p_\sharp(\pi_1(E, e''))$ of $\pi_1(B, b)$ for this covering space.*

Exercise 4.1.12 *Use corollary 2.8 to show that there are $2^{2g} - 1$ double coverings of M_g.*

The homomorphism $\rho : \pi_1(B, b) \rightarrow \mathcal{S}_n$ described above is called the *monodromy representation*. It is to be distinguished from the group of *covering transformations* of (E, p, B) which is the group of those homeomorphismsof E which map every fibre onto itself. The following exercises clarify the relationship between these concepts:

Exercise 4.1.13 *Show that there is at most one covering transformation of (E, p, B) mapping e_i to e_j, where $p^{-1}(b) = \{e_1, \cdots, e_n\}$.* (If the covering transformation τ fixes e_i and maps e_j to e_k, then consider a path γ joining e_i to e_j and contradict the uniqueness of the lift of the path $p\gamma$. It is not necessary to assume n is finite.)

Exercise 4.1.14 *Show that the group of covering transformations acts transitively on a fibre (therefore all fibres) if and only if (E, p, B) is a regular covering space. Therefore the group of covering transformations coincides with $\pi_1(B, b)/p_\sharp(\pi_1(E, e))$ if and only if (E, p, B) is a regular covering.*

Exercise 4.1.15 *Let $\Delta_i = \{\gamma \in \pi_1(B, b) | \rho(\gamma)(e_i) = e_i\}$, and $\Delta = \cap \Delta_i$. Show that Δ is a normal subgroup and $\mathrm{Im}(\rho) \simeq \pi_1(B, b)/\Delta$. Furthermore, $Im(\rho)$ is a group of order n if and only if (E, p, B) is a regular covering (i.e., $\Delta = \Delta_i$ for all i).*

4.1.4 Some Proofs

Now we discuss the proofs of theorems 4.1.1 and 4.1.2.
Proof of theorem 4.1.1 - Let $\mathcal{U} = \{U_i\}$ be a covering of B such that $p^{-1}(U_i)$ is a disjoint union $\cup_k V_{ik}$ with p a homeomorphism of each V_{ik} onto U_i for all i. We fix $e \in p^{-1}(b)$, and may assume that $b \in U_1$ and $e \in V_{11}$. We partition the square into sufficiently small subsquares such that $G(s_n)$ lies entirely within one U_i, $i = i(n)$. Starting from s_1 =(the square with subsquare containing the origin), we may assume $G(s_1)$ is contained in U_1, and since $p : V_{11} \rightarrow U_1$ is a homeomorphism, the restriction of G to s_1 can be uniquely lifted to a mapping into E with $G((0, 0)) = e$. Ordering the subsquares in a way that each is adjacent to the next, we can repeat the process until we have exhausted the square. This argument proves that covering spaces are fibrations with unique path-lifting and notice that $G'(0, s)$ can be chosen to be any point in $p^{-1}(\gamma(0))$. To prove the second assertion assume

$s_n \cap (\partial I \times I) \neq \emptyset$. Since $p^{-1}(U_{i(n)})$ consists of disjoint open sets, G' is constant on $s_n \cap (\partial I \times I)$, and G' is locally constant on $\partial I \times I$ which is precisely the assertion to be proven. ♣

Proof of theorem 4.1.2 - We prove this theorem for the special case where B is a manifold of dimesion b. A technical refinement of the idea of the proof makes it work for general case. Fix a triangulation of B, and assume that the triangulation is sufficiently fine so that for each vertex v, $\mathrm{St}(v) = \{b - \text{simplices containing } v\}$ (called *star of* v has a neighborhood $U(v)$ with $p^{-1}(U(v)) = \cup V_i$ (disjoint union), and the restriction of p to V_i a homeomorphism onto $U(v)$. Let $\{s_i\}$ be an enumeration of the b-simplices of B, and fix a point $x_i \in \mathrm{Int}(s_i)$. Draw a line l_{ji} joining x_i to x_j if s_i and s_j have a face in common, and in this case assume that l_{ji} lies entirely in $\mathrm{Int}(s_i \cup s_j)$. Let $x_\circ$ be the base-point. Obviously the paths l_{ji} and l_{kj} can be composed to obtain a path $l_{kj}l_{ji} = l_{kji}$ joining x_i to x_k. Every path with initial point $x_\circ$ and terminal point x_n is homotopic to one of the form $l_{ni(k)\cdots i(1)\circ}$. Let Y be the space $\{(s_i, l) | I \text{ path of the above form joining } x_\circ \text{ to } x_i\}$. On Y define equivalence relation $\sim$ by the rule $(s_i, l) \sim (s_j, l')$ if $i = j$ and l and l' are homotopic. Let $E = Y/\sim$ be the quotient space. To endow E with the structure of a simplicial complex, we choose representatives for each equivalence class. We say (s_i, l) and (s_j, l') have a face in common if s_i and s_j have a face in common and the loop $l'^{-1}l_{ji}l$ is homotopic to the constant loop. One checks easily that this condition is independent of the choice of the representatives for the equivalence classes in Y. With this provision we can put together the (s_i, l)'s to make a simplicial complex out of E. Notice that if (s_i, l) and (s_j, l') have a face in common then they are joined together in the same manner as s_i and s_j are joined. Points of E have representatives of the form (x, l) where l is a path of the above form joining $x_\circ$ to x_i and $x \in s_i$. Furthermore, the projection $(x, l) \to x$ gives us the covering projection (E, p, B), and clearly E is a manifold since B is one. To see that E is simply connected let γ' be a loop in E with initial point $x'_\circ$ lying above $x_\circ$, and let $\gamma = p\gamma'$. Clearly γ' is the lift of γ with initial point $x'_\circ$. If $\gamma \neq e \in \pi_1(B, x_\circ)$, then from the construction of E, $(s_\circ, e) \cap (s_\circ, \gamma) = \emptyset$. Since $(x_\circ, \gamma)$ is the terminal point of γ', we have contradicted the assumption that γ' is a loop. Hence E is simply connected. As noted earlier, uniqueness follows from corollary 4.1.6. ♣

4.2 Computing Fundamental Groups

4.2.1 Simple Examples

Example 4.2.1 In example 3.5.1 of chapter 1 we used the algebra $\mathbf{H}$ of quaternions to indentify S^3 with the group of unit quaternions. We now use that example to construct the universal cover of $SO(3)$. Let $\mathbf{H}'$ be the subspace of $\mathbf{H}$ spanned by $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ called the subspace of *pure quaternions*. For $v \in \mathbf{H}'$ and $0 \neq q \in \mathbf{H}$, $qvq^{-1} \in \mathbf{H}'$ which yields a representation $\rho : SU(2) \rightarrow GL(3, \mathbf{R})$. (This representation is equivalent to the adjoint representation of $SU(2)$.) Since $||v|| = ||qvq^{-1}||$ the representation ρ takes values in $SO(3)$. Clearly, $\mathrm{Ker}\rho = \pm I$ and for dimension reasons the mapping is onto $SO(3)$. It follows that $\rho : SU(2) \rightarrow SO(3)$ is a covering projection and since $SU(2) \simeq S^3$ is simply connected, $\pi_1(SO(3), e) \simeq \mathbf{Z}/2$. ♠

Exercise 4.2.1 *Show that the fundamental group of the projectivized tangent bundle $P(\mathbf{RP}(2))$ of $\mathbf{RP}(2)$ i.e., the set of tangent lines to $\mathbf{RP}(2)$ is the quaternion group $\mathcal{Q} = \{\pm\mathbf{1}, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}\}$ of order 8.* (Consider the above action of $SU(2)$ on pure quaternions to obtain $\mathbf{RP}(2)$ as a homogeneous space for $SU(2)$.)

Example 4.2.2 We show by induction that $\mathbf{CP}(n)$ is simply connected. For $n = 1$ this is just simple connectedness of $\mathbf{CP}(1) = S^2$. Recall that $\mathbf{CP}(n) \supset \mathbf{CP}(n-1)$, and it is obtained from $\mathbf{CP}(n-1)$ by attaching the unit disc D^{2n}, with $\partial D^{2n} = S^{2n-1}$ being mapped to $\mathbf{CP}(n-1)$. Let $x \in \mathbf{CP}(1)$ and $\gamma : I \rightarrow \mathbf{CP}(n)$ with $\gamma(0) = \gamma(1) = x$. By the simplicial approximation theorem, we may assume γ is a simplicial map and therefore misses a point z lying in the cell D^{2n}. By joining z to the points on $\mathrm{Im}(\gamma)$ that lie in D^{2n} and continuing, we can deform γ so that the entire image of γ lies in $\mathbf{CP}(n-1)$ (see Figure 3.1), i.e., γ is homotopic relative ∂I to a loop γ' lying entirely in $\mathbf{CP}(n-1)$. Applying the induction hypothesis to γ' we see that γ is homotopic to the constant loop and so $\mathbf{CP}(n)$ is simply connected. ♠

Example 4.2.3 In this example we show that the fundamental group of an analytic group is abelian. The group operation of an analytic group G allows us to define another operation on the space of paths in G, viz., if $\gamma, \tau : I \rightarrow G$ then we define the path $(\gamma \circ \tau)(t) = \gamma(t)\tau(t)$. (to avoid confusion with the product $\gamma\tau$ of two loops γ and τ in the fundamental group, we use the notation $\gamma \circ \tau$.) Assuming that γ and τ are loops with base point e, then the loops

$\gamma\circ\tau$ and $\gamma\tau$ are homotopic. In fact, the homotopy is explicitly constructed as follows:

$$F(t,s)=\begin{cases}\gamma(\frac{2t}{1+s}) & \text{for } \ 0\leq t\leq\frac{1-s}{2},\\ \gamma(\frac{2t-1+s}{1+s})\tau(\frac{2t}{1+s}) & \text{for } \ \frac{1-s}{2}\leq t\leq\frac{1+s}{2},\\ \tau(\frac{2t-1+s}{1+s}) & \text{for } \ \frac{1+s}{2}\leq t\leq 1.\end{cases}$$

This homotopy can be described as follows: At $s=0$, $F(t,0)=\gamma\tau$. For $0<s<1$ γ and τ are rescaled so that going around the entire loop takes $(1+s)/2$, with γ starting at $t=0$ and τ starting at $t=(1-s)/2$. Furthermore, in the interval $[(1-s)/2,(1+s)/2]$, $F(t,s)$ is the product $\gamma\circ\tau$. At $s=1$, $F(t,1)=\gamma\circ\tau(t)$. Notice that we may as well define $F(t,s)$ to be the product $\tau\circ\gamma$ in the interval $[(1-s)/2,(1+s)/2]$, so that $\gamma\tau$ is also homotopic $\tau\circ\gamma$. In particular, $\gamma\tau$ and $\tau\gamma$ are homotopic and $\pi_1(G,e)$ is abelian. Here we made very little use of the group structure on G. All that we needed was the possibility of multiplying two quantities and the existence of identity. ♠

Example 4.2.4 Let G be an analytic group and $p:\tilde{G}\rightarrow G$ be its universal covering space. In this example we show that $\tilde{G}$ has a natural group stucture and p is a continuous (and therefore analytic) homomorphism of groups. Let $\mu:G\times G\rightarrow G$ denote the map $\mu(g,h)=gh^{-1}$. Consider the diagram

$$\begin{array}{ccc}\tilde{G}\times\tilde{G} & \xrightarrow{\tilde{\mu}} & \tilde{G}\\ \downarrow & & \downarrow\\ G\times G & \xrightarrow{\mu} & G\end{array}$$

where the vertical arrows are $p\times p$ and p. We fix base point $\tilde{e}$ for $\tilde{G}$ such that $p(\tilde{e})=e\in G$ and use $(\tilde{e},\tilde{e})$ as base point for $\tilde{G}\times\tilde{G}$. The map $\tilde{\mu}$ exists and is unique by corollary 4.1.4 and the requirement $\tilde{\mu}(\tilde{e},\tilde{e})=\tilde{e}$. It is straightforward to show that $\tilde{\mu}$ endows $\tilde{G}$ with the structure of an analytic group and p is a continuous homomorphism.

Exercise 4.2.2 *Let G be an analytic group and $p:G'\rightarrow G$ a covering projection. Show that G' has the structure of an analytic group with p a continuous (and therefore analytic) homomorphism.*

Example 4.2.5 Let $n\geq 2$ and $F:I^n\rightarrow S^1$ be a continuous map such that $F(\partial I^n)=x$ where $x\in S^1$ is any fixed point on the circle which we take it to be the point $1\in S^1\subset\mathbf{C}$. Now $F(y)=e^{i\theta(y)}$ for some real valued function θ on I^n. Since ∂I^n is connected for $n\geq 2$, θ is constant on I^n and therefore $\theta(\partial I^n)=0$. Now consider the homotopy

$$G:I\times I^n\longrightarrow S^1,\qquad G(t,y)=e^{i(1-t)\theta(y)}.$$

Therefore for $n \geq 2$, any mapping $F : I^n \to S^1$,constant on the boundary ∂I^n, is homotopic to a constant map relative to the boundary. ♠

Fibrations play a very important role in homotopy theory. In the examples that follow we examine some applications of fibrations to the fundamental group. Their significance will become more evident after the introduction of higher homotopy groups and the long exact sequence for homotopy.

Example 4.2.6 The realization of the sphere S^n as the homogeneous space $S^n = SO(n+1)/SO(n)$ implies that we have a fibration $p : SO(n+1) \to S^n$ with fibre $SO(n)$. Assume $n \geq 3$. We show by induction on n that $\pi_1(SO(n+1), e)$ is either $\mathbf{Z}/2$ or is the trivial group. We have already shown that $\pi_1(SO(3), e) \simeq \mathbf{Z}/2$. Let $\gamma : I \to SO(n+1)$ be a loop with $\gamma(0) = \gamma(1) = e$. Then $p\gamma$ is a loop in S^n which is simply connected. Therefore $p\gamma$ is homotopic to the constant map $I \to p(e)$. Lifting the homotopy to $SO(n+1)$ we see that γ is homotopic to a loop in $SO(n)$. It then follows from the induction hypothesis that $\pi_1(SO(n+1), e)$ is either $\mathbf{Z}/2$ or is the trivial group. In exercises 4.2.3 and 4.2.4 it is shown that $\pi_1(SO(n), e) \simeq \mathbf{Z}/2$ for $n \geq 3$. In the subsection on Clifford algebras we construct a double covering of $SO(n)$ which also shows that $\pi_1(SO(n+1), e) \simeq \mathbf{Z}/2$ for $n \geq 3$. Similarly by looking at the fibration $SU(n+1) \to S^{2n+1}$ with fibre $SU(n)$, and recalling that $SU(2) \simeq S^3$ one shows inductively that $SU(n)$ is simply connected for $n \geq 2$. ♠

Let $\pi : E \to M$ be a fibre bundle with fibres homeomorphic to a manifold Q of dimension q. We assume M and E are path-connected. Let $x \in E$, then $Q \simeq \pi^{-1}(\pi(x))$ (fibre over $\pi(x)$), and we have induced homomorphisms

$$\pi_1(Q, x) \xrightarrow{\imath_\sharp} \pi_1(E, x) \xrightarrow{\pi_\sharp} \pi_1(M, \pi(x)). \tag{4.2.1}$$

The generalization and properties of these homomorphisms are studied in the context of higher homotopy groups and long exact sequence for homotopy. For the time being, we make some observations about (4.2.1) which will become useful later in this chapter.

Lemma 4.2.1 *With the above notation and hypotheses, the sequence (4.2.1) is exact, i.e.,* $\mathrm{Im}(\imath_\sharp) = \mathrm{Ker}(\pi_\sharp)$.

Proof - It is clear that $\mathrm{Im}(\imath_\sharp) \subset \mathrm{Ker}(\pi_\sharp)$. Let $\gamma : I \to E$ be such that $\pi\gamma$ represents the identity in $\pi_1(M, \pi(x))$, then there is a homotopy $F : I \times I \to M$ between $\pi\gamma$ and the constant map. By the homotopy lifting property, the homotopy lifts to $\tilde{F}$ which is a homotopy between γ and a mapping $\tilde{F}(1, .) : I \to Q$ which shows that $[\gamma]$ lies in the image of $\imath_\sharp$ proving the required exactness. ♣

Example 4.2.7 Assume that that in the fibration $\pi : E \to M$, $M = S^1$ and the fibre Q is also path-connected. The the conclusion of lemma 4.2.1 can be strengthened to the exactness of the sequence

$$1 \longrightarrow \pi_1(Q, x) \xrightarrow{\imath_\sharp} \pi_1(E, x) \xrightarrow{\pi_\sharp} \mathbf{Z} \longrightarrow 0.$$

First we prove that $\imath_\sharp$ is injective. Let $\gamma : I \to Q$ and $F : I \times I \to E$ be a homotopy between γ and the constant map to E, i.e., $F(0, t) = \gamma(t)$, $F(s, \partial I) = x$ and $F(1, t) = x$. Then $\pi F : I^2 \to S^1$ with πF the constant map to $\pi(x)$. Example 4.2.5 is applicable to show that πF is homotopic to the constant map. This means we have a mapping $G : I \times I^2 \to S^1$ such that

1. $G(0, (t_1, t_2)) = \pi F(t_1, t_2)$;
2. $G(s, \partial I^2) = \pi(x)$;
3. $G(1, (t_1, t_2)) = \pi(x)$.

Let $\tilde{G}$ be the lift of this homotopy to E with $\tilde{G}(0, (t_1, t_2)) = F(t_1, t_2)$. Then the mapping $(s, t) \to \tilde{G}(s, (0, t))$ takes values in Q and gives the desired homotopy between γ and the constant map to $x \in Q$. The surjectivity of $\pi_\sharp$ only requires path-connectedness of Q. (Of course $\mathbf{Z}$ should be replaced with $\pi_1(M, \pi(x))$ if $M \neq S^1$.) Let $\gamma : I \to M$ and $\tilde{\gamma}$ be a lift of γ to $\tilde{\gamma} : I \to E$. Since Q is path-connected, there is $\delta : I \to Q = \pi^{-1}(\pi(x))$ such that $\delta(0) = \tilde{\gamma}(1)$ and $\delta(1) = \tilde{\gamma}(0)$. The composition of δ and γ define an element of $\pi_1(E, x)$ which maps to γ by $\pi_\sharp$. This completes the proof of the claimed exactness. For $M = S^1$, we have $\pi_1(M, \pi(x)) \simeq \mathbf{Z}$, and since $\mathbf{Z}$ is free, we have a splitting homomorphism $\rho : \mathbf{Z} \to \pi_1(E, x)$, i.e., $\pi_\sharp \rho = \mathrm{id}$. Therefore we have the semi-direct product decomposition

$$\pi_1(E) \simeq \pi_1(Q, x).\mathbf{Z}.$$

We will use this decomposition in connection with knots. ♠

Exercise 4.2.3 *Let $F : I^2 \to SO(3)$ be a continuous mapping such that $F(\partial I^2) = e$. Show that F lifts to a mapping $\tilde{F} : I^2 \to SU(2)$. Deduce that $\tilde{F}$ and therefore F is homotopic to the constant map to e. Using the fibration $SO(n+1) \to S^n$ with fibre $SO(n)$, show that for $n \geq 3$, every continuous map $F : I^2 \to SO(3)$ such that $F(\partial I^2) = e$ is homotopic to the constant map $g \to e$.*

Exercise 4.2.4 *By emulating the argument of example 4.2.7 show that the inclusion* $\imath : SO(n) \to SO(n+1)$, $n \geq 3$, *(where* $SO(n)$ *acts on the last* n *coordinates), induces an injection*

$$\imath_\sharp : \pi_1(SO(n), e) \to \pi_1(SO(n+1), e).$$

Deduce that $\pi_1(SO(n), e) \simeq \mathbf{Z}/2$ *for* $n \geq 3$.

4.2.2 Theorem of Van Kampen

The most important tool in computing fundamental groups is the theorem of van Kampen which may be regarded as the analogue of the Mayer-Vietoris sequence for the fundamental group.

Theorem 4.2.1 (van Kampen) *Let* $Z = X \cup Y$ *(union of connected simplicial complexes),* $A = X \cap Y$, $x \in A$, *and assume* A *is a path connected simplicial complex. Set* $\Gamma = \pi_1(X, x)$, $\Gamma' = \pi_1(Y, x)$ *and* $\Delta = \pi_1(A, x)$. *Let* $i_{X\sharp} : \Delta \to \Gamma$ *and* $i_{Y\sharp} : \Delta \to \Gamma'$ *be the homomorphisms induced by the inclusions* $i_X : A \to X$ *and* $i_Y : A \to Y$. *Then*

$$\pi_1(Z, x) \simeq \Gamma \star_\Delta \Gamma'.$$

Remark 4.2.1 This description of $\pi_1(Z, x)$ which is intuitively and geometrically reasonable, and we omit its formal proof. It generalizes the second assertion of proposition 4.1.1 above. The assumption that X, Y, Z and A are simplicial complexes is not necessary, although some mild restrictions on these spaces are necessary. A set of sufficient conditions is as follows:

1. X and Y are separable, regular topological spaces;
2. $X \setminus A$ and $Y \setminus A$ are open in in Z;
3. X, Y and A are locally contractible;
4. A is path-wise connected.

Necessary and sufficient conditions are given in [Ol]. ♡

Exercise 4.2.5 *Let* $L_1, \cdots, L_m$ *be lines in general position in* $\mathbf{R}^3$, $\Lambda = \cup L_j$, *and* $N = \mathbf{R}^3 \setminus \Lambda$. *Show that the fundamental group of* N *is isomorphic to the free group on* m *generators.*

Exercise 4.2.6 *Give an alternative proof of exercise 4.1.5 on the basis of van Kampen's theorem.*

Example 4.2.8 Let M be a manifold of dimension $m \geq 4$ and $\gamma : S^1 \rightarrow M$ be an embedding of a circle in M. Denote by N_γ a small (closed) tubular neighborhood of $\text{Im}(\gamma)$ and its interior by $\breve{N}_\gamma$. Let $M' = N \setminus \breve{N}_\gamma$. In this example we show that the fundamental group of M' is isomorphic to that of M. Let D^n denote the disc of dimension n, then $N_\gamma \simeq S^1 \times D^{m-1}$ and $\partial N_\gamma \simeq S^1 \times S^{m-2}$. Consequently, for $x \in \partial N_\gamma$, the inclusion $\partial N_\gamma \rightarrow N_\gamma$ induces an isomorphism of fundamental groups for $m \geq 4$:

$$\pi_1(\partial N_\gamma, x) \xrightarrow{\simeq} \pi_1(N_\gamma, x) \simeq \mathbf{Z}. \tag{4.2.2}$$

Since $M' \cap N_\gamma = \partial N_\gamma$, van Kampen's theorem and (4.2.2) imply

$$\pi_1(M, x) \simeq \pi_1(M', x) \star_{\mathbf{Z}} \mathbf{Z} \simeq \pi_1(M', x),$$

which is the desired result. ♠

A modification of the argument proving simple connectedness of $\mathbf{CP}(n)$ (example 4.2.2) has interesting consequences for the fundamental group of simplicial or regular cell complexes. Let X be a simplicial complex or a regular cell complex, $X^n = \cup_{j \leq n}(\text{j}-\text{cells})$ the n-skeleton of X and $x \in X^1$. Then the inclusion $X^k \subset X^n$, $k < n$, induces a homomorphism

$$\lambda_{kn} : \pi_1(X^k, x) \longrightarrow \pi_1(X^n, x). \tag{4.2.3}$$

Let $\gamma : I \rightarrow X^2$. For every 2-simplex or 2-cell D of X^2 we may assume $\text{Im}\gamma$ misses at least one point of D and by the argument in example 4.2.2, γ is homotopic, relative to ∂I to a map $\gamma' : I \rightarrow X^1$. This implies that the map λ_{12} is surjective. We can actually say more. In fact we have

Proposition 4.2.1 *With the above notation and hypotheses, the homomorphisms λ_{1n}, $n > 1$, are surjective, and the homomorphisms λ_{2n}, $n > 2$, are isomorphisms.*

Proof - Since we have already shown that λ_{12} is surjective, it suffices to prove the second assertion. Let D is an n-simplex or an n-cell, $n \geq 3$, $Y \subset Z$ simplicial or regular cell complexes with $Z = Y \cup D$ where $\partial D \subset Y$. Let $x \in \partial D$. By van Kampen's theorem

$$\pi_1(Z, x) = \pi_1(Y, x) \star_{\pi_1(\partial D, x)} \pi_1(D, x).$$

Since $\pi_1(D, x) = \pi_1(\partial D, x) = 0$ the required result follows. ♣

Example 4.2.9 In this example we use van Kampen's theorem to show that for every $n \geq 1$ and $m \geq 3$ there is a compact orientable manifold M of dimension m with fundamental group isomorphic $\mathbf{F}_n$. Let $M_1 = S^1 \times S^{m-1}$. The fundamental group of of M_1 is isomorphic to $\mathbf{Z}$ since $m \geq 3$. Let $M_j \simeq M_1$ for $j = 2, \cdots, n$. We apply the $\sharp$ construction to obtain $M_1 \sharp M_2$. In order to apply van Kampen's theorem we write

$$M_1 \sharp M_2 = M_1' \cup M_2', \quad \text{with} \quad M_1' \cap M_2' \simeq (0,1) \times S^{m-1}.$$

Since M_j' is obtained from M_j by removing a (small) disc with boundary diffeomorphic to S^{m-1},

$$\pi_1(M_j', x) \simeq \mathbf{Z}.$$

Now $\pi_1(M_1' \cap M_2', x) = 0$ in view of $M_1' \cap M_2' \simeq (0,1) \times S^{m-1}$, and consequently by van Kampen's theorem $\pi_1(M_1 \sharp M_2, x) \simeq \mathbf{F}_2$. Similarly, the fundamental group of $M_1 \sharp \cdots \sharp M_n$ is isomorphic to the free group $\mathbf{F}_n$ on n generators. *spadesuit*

Example 4.2.10 In example 4.2.9 we exhibited, for every $m \geq 3$, a compact orientable manifolds of dimension m whose fundamental group is the free group on n generators. In this example we refine that result by showing that for $m \geq 4$ and for every finitely presented group Γ there is a compact orientable manifold of dimension m whose fundamental group is Γ. We have an exact sequence

$$\{1\} \longrightarrow R \longrightarrow \mathbf{F}_n \longrightarrow \Gamma \longrightarrow \{1\},$$

with R the normal subgroup generated by the finitely many relations $R_1, \cdots, R_l \in \mathbf{F}_n$ and their conjugates. Let $M_\circ = S^1 \times S^{m-1}$ and $M = M_\circ \sharp \cdots \sharp M_\circ$ be the manifold constructed in example 4.2.9 with fundamental group isomorphic to $\mathbf{F}_n$. Let $x \in M$ and $\gamma : S^1 \rightarrow M$ be an embedding of the circle with $\gamma(1) = x$ and γ representing the relation $R_1 \in \pi_1(M, x)$. That this is possible follows from the transversality theorem and $m \geq 3$. Let N_γ be a small closed tubular neighborhood of $\mathrm{Im}(\gamma)$ as in example 4.2.8, and $M' = M \setminus N_\gamma$. Now note the key point that the distinct manifolds with boundary $D^2 \times S^{m-2}$ and $S^1 \times D^{m-1}$ have diffeomorphic boundaries:

$$\partial N_\gamma \simeq \partial(S^1 \times D^{m-1}) \simeq S^1 \times S^{m-2} \simeq \partial(D^2 \times S^{m-2}). \tag{4.2.4}$$

By the theorem of Cerf (see chapter 1) we obtain a manifold M_1 by attaching $D^2 \times S^{m-2}$ to M' via the diffeomorphism of the boundaries of N_γ and of $D^2 \times S^{m-2}$ as given by (4.2.4).

Since $\pi_1(M', x) = \pi_1(M, x)$ (example 4.2.8), and $D^2 \times S^{m-2}$ is simply connected for $m \geq 4$, van Kampen's theorem implies

$$\pi_1(M_1, x) \simeq \mathbf{F}_n/(R_1).$$

Repeating the construction for the remaining relations we obtain the desired manifold. ♠

Example 4.2.11 In chapter 3 we computed the homology of a finite connected graph, and showed that its first homology group is the free abelian on (1-number of vertices+number of edges) number of edges) generators. In the example below we show that its fundamental group is the free group on the same number of generators. Although this result is more elementary than van Kampen's theorem, we shall use the latter result in our computation to demonstrate the power of this theorem. The proof is by induction on the number of edges of the graph Z. If Z has only one edge then the conclusion is valid. If we can decompose Z into two disjoint connected graphs X and Y, and each having at least on edge, and joined at one vertex x, then from theorem 4.2.1 and the induction hypothesis we see that $\pi_1(Z, x)$ is the free group on $2-\chi(X)-\chi(Y)$ generators. Since the number of the edges of Z is the sum of those of X and Y, and the number of vertices of Z is the sum of those of X and Y minus 1, the claim follows in this case. Now assume that by disconnecting the edge λ at the vertex $y \in \lambda$, the graph does not become disconnected. Let X be the graph obtained by removing the edge λ. The the vertices of X and Z are identical, but X has one less edge, namely λ. Let x and y be the vertices of the edge λ. Since removing λ does not disconnect the graph Z, there is a path in X connecting x to y. Let A be a connected contractible subgraph of X joining x to y. Notice that the shortest path (i.e., the fewest number of edges) joining x to y is necessarily contractible since by eliminating loops we shorten the length of a path. Let Y be the graph obtained from A by the addition of the single edge λ joining x to y. $|Y|$ is clearly homeomorphic to the circle. We now have the decomposition $Z = X \cup Y$ with $X \cap Y = A$. Applying the induction hypothesis and invoking van Kampen's theorem we get the required result. ♠

Example 4.2.12 We apply van Kampen's theorem to compute the fundamental group of a surface of genus g. We know that for $g = 1$, $\pi_1(M_1, x) = \mathbf{Z}^2$. We can decompose $M_2 = X \cup Y$ in such a way X and Y are homeomorphic to a torus with a small disc removed, and $A = X \cap Y$ is a circle. We claim that $\pi_1(X, x) \simeq \pi_1(Y, x) \simeq \mathbf{F}_2$, the free group on two generators. In fact, we can represent X as a square with a disc in the middle removed and with the obvious identification of the edges as described in the computation of homology of surfaces. Then expanding the disc to the entire interior of the square, we see that the boundary of the square, with the identification of the edges, which is figure **8**, is a deformation retract of X. This proves the claim. Let $\{a_1, b_1\}$ (resp. $\{a_2, b_2\}$) be the generators of $\pi_1(X, x)$ (resp. $\pi_1(Y, x)$) corresponding to the loops of figure **8**. Then the generator of $\pi_1(A, x)$ as an element of $\pi_1(X, x)$ (resp. $\pi_1(Y, x)$) is $a_1 b_1 a_1^{-1} b_1^{-1}$ (resp. $a_2 b_2 a_2^{-1} b_2^{-1}$, after possibly renaming the generators). Therefore $\pi_1(M_2, x)$ is the quotient of $\mathbf{F}_4$, with generators $\{a_1, b_1, a_2, b_2\}$ and the single relation

$$a_1 b_1 a_1^{-1} b_1^{-1} a_2 b_2 a_2^{-1} b_2^{-1} = e.$$

In Figure 3.2 planar representations of surfaces of genus 2 and 3 are given. The boundary of the region consists of g connected components and it is immediate that the resulting surface is M_g after proper identification of sides. If we remove a small disc from M_g then the resulting surface M'_g has the homotopy type of a bouquet of $2g$ circles which we denote by $\mathcal{B}_{2g}$. This is easily seen by expanding the small discs so that their complement becomes a one dimensional figure. The fundamental group of $\mathcal{B}_{2g}$ is $\mathbf{F}_{2g}$ (exercise 4.1.4.) Furthermore, from the representation of M'_g we see that if A denotes the boundary of M'_g, then the image of a generator of $\pi_1(A, x)$, $(x \in A)$ in $\pi_1(M'_g)$ induced by the inclusion of A in M'_g, is

$$a_1 b_1 a_1^{-1} b_1^{-1} \cdots a_g b_g a_g^{-1} b_g^{-1}.$$

We this observation we can easily continue the process of computing the fundamental group of M_g, and show that $\pi_1(M_g, x)$ is isomorphic to the quotient of the free group $\mathbf{F}_{2g}$, with generators $\{a_1, b_1, \cdots, a_g, b_g\}$ by the single relation

$$a_1 b_1 a_1^{-1} b_1^{-1} \cdots a_g b_g a_g^{-1} b_g^{-1} = e. \tag{4.2.5}$$

This completes the computation of the fundamental group of M_g. ♠

Exercise 4.2.7 *Determine the conjugacy classes of the subgroups of $\pi_1(M_2, x)$ corresponding to the non-regular coverings of example 4.1.7.*

Exercise 4.2.8 *Let N_g be the compact nonorientable surface defined in chapter 3. Show that the fundamental group of $N_1 = \mathbf{RP}(2)$ is $\mathbf{Z}/2$. By mimmicking the argument of given for the $\pi_1(M_2, x)$, show that the fundamental group of $\pi_1(N_2, x)$ is the quotient of the free group on two generators, c_1 and c_2, by the single relation*

$$c_1^2 c_2^2 = e.$$

The extension of the computation in exercise 4.2.8 to $\pi_1(N_g, x)$ requires knowledge of $\pi_1(N'_g, x)$, where N'_g is obtained from N_g by removing a small disc. In Figure 3.3 we have represented N_g and N'_g as planar domains for $g = 2$ and 3. We have chosen these representations since their validity is immediate from the definition of N_g. By expanding the disc D we see that N'_g has the homotopy type of a bouquet of g circles. Consequently, $\pi_1(N'_g, x)$ is isomorphic to the free group on g generators. With this observation it is not difficult to compute $\pi_1(N_g, x)$. In fact,

Exercise 4.2.9 *Show that $\pi_1(N_g, x)$ is the quotient on g generators, $\{c_1, \cdots, c_g\}$, by the single relation*

$$c_1^2 \cdots c_g^2 = e.$$

Exercise 4.2.10 *Show that the fundamental group of the surface obtained by removing $n > 0$ distinct points from M_g (resp. N_g) is the free group on $2g+n-1$ (resp. $g+n-1$) generators.*

Exercise 4.2.11 *Show that there are 2^g double covers of N_g, and realize them geometrically for $g = 1$* and 2.

Exercise 4.2.12 *Which double cover(s) of N_g are orientable, and identify the surface.*

Example 4.2.11, proposition 4.2.1 and theorem 4.2.1, in principle, allow us to compute the fundamental group of any simplicial or regular cell complex. In fact, given any simplicial or cell complex, the fundamental group of its 1-skeleton is computable by example 4.2.11. Proposition 4.2.1 implies that the fundamental group of a finite simplicial or regular cell complex X is a quotient of that of its 1-skeleton. Adding 2-simplices creates relations which are obtained by invoking theorem 4.2.1. Furthermore, proposition 4.2.1 also implies that attaching simplices of dimension greater than two to the 2-skeleton of X does not affect the fundamental group.

4.2.3 Knots and Links

By a *knot* we mean a smooth embedding of the circle S^1 in $\mathbf{R}^3$ or S^3. For practical purposes, it is often more convenient to look at the knot as a compact piece-wise linear one dimensional manifold necessarily homeomorphic to the circle. Often it is the image of the embedding which is of geometric interest rather than the mapping itself. Therefore we use the word knot to mean both the embedding of the circle and its image. This will cause no confusion. One should note that if we only assume continuity of the embedding, then it is possible to construct *wild knots* with infinitely many crossings as shown in Figure 3.4. We would like to avoid such pathologies. To distinguish between different knots, one looks at the complement of a knot in the ambient space. Recall, however, from chapter 3 exercise (ZZZ) that the homology groups of the complement of a knot do not distinguish between the different embeddings. The fundamental group of the complement of a knot, on the other hand, is good invariant. We conveniently refer to the fundamental group of the complement of a knot K as the *group of the knot* K or a *knot group*. Let L_n be the disjoint union of n circles which is a one dimensional manifold with n connected components. A C^∞ or piece-wise linear embedding of L_n in $\mathbf{R}^3$ or S^3 is called a *link*. It is the image of the embedding and/or its complement which are of interest. Referring to the image as a link will not cause confusion and will be denoted by L_n again. The fundamental group of the complement of a link L_n is called the *group of the link* L_n or a *link group*. The ordinary embedding of the circle in $\mathbf{R}^3$ or S^3, or any embedding that differs from it by a homeomorphism of the ambient space, is called the *trivial* knot or an *unknotted circle*. It is a simple matter to draw knots and links which look very non-trivial, at least intuitively. Extensive tables with pictures can be found in [Rol].

Exercise 4.2.13 *Consider the ordinary embedding* $S^1 = \{e^{i\theta}\} \subset \mathbf{R}^2 \subset \mathbf{R}^3$. *Show that* $\pi_1(\mathbf{R}^3 \setminus S^1; x) \simeq \mathbf{Z}$.

Exercise 4.2.14 *Identify* S^3 *with the one point compactification of* $\mathbf{R}^3$ *(e.g., via the stereographic projection), and let* $K \subset \mathbf{R}^3$. *Show that* $\pi_1(S^3 \setminus K) \simeq \pi_1(\mathbf{R}^3 \setminus K)$.

Let $L_{n,1} = \{e^{i\theta}\} \subset \mathbf{R}^2 \subset \mathbf{R}^3$ be the standard embedding of the circle. Let $D_1 = \{(x_1, x_2, 0) | x_1^2 + x_2^2 \leq 1\}$, then $\partial D_1 = L_{n,1}$. Let $T_2, \cdots, T_n$ be diffeomorphisms of $\mathbf{R}^3$ and set $L_{n,j} = T_j(L_{n,1})$, $D_j = T_j(D_1)$.

Exercise 4.2.15 *With the above notation let* T_j*'s be such that for every* $i \neq j$, $D_j \cap L_{n,i} = \emptyset$. *Show that the link group of* L_n *is isomorphic to the free group on* n *generators.* (Geometrically, the hypothesis $D_j \cap L_{n,i} = \emptyset$ means that the circles $L_{n,j}$ are unlinked. This is the simplest case of a link which is not a knot.)

We want to introduce a class of knots, known as torus knots, which are of interest in algebraic geometry as well. It is convenient to begin with describing a decomposition of S^3 which plays an important role in the study of 3-manifolds.

Example 4.2.13 Consider S^3 with its standard embedding:

$$S^3 = \{x = (x_1, x_2, x_3, x_4) | \sum x_i^2 = 1\}.$$

Define the subsets U and V of S^3 by

$$U = \{x \in S^3 | x_1^2 + x_2^2 \geq 1/2\}; \quad V = \{x \in S^3 | x_1^2 + x_2^2 \leq 1/2\}.$$

Identify $\mathbf{R}^4$ with $\mathbf{C}^2$ with co-ordinates z_1 and z_2 given by $z_1 = x_1 + ix_2$ and $z_2 = x_3 + ix_4$. U and V are solid tori. To see this, note that U is the circle bundle with base the disc $D = \{z_2 \mid |z_2|^2 \leq 1/2\}$ and fibre (over the point z_2) the circle $\{z_1 \mid |z_1|^2 = 1 - |z_2|^2\}$. Therefore $U = S^1 \times D$ which is a solid torus. Furthermore, $U \cap V = \{(z_1, z_2) \mid |z_1|^2 = 1/2 = |z_2|^2\}$ is a torus. Therefore S^3 is the union of two solid tori with intersection a torus. ♠

Example 4.2.14 Let $m > 1$ and $n > 1$ be relatively prime integers, and consider the knot, called the *torus knot* with parameters (m, n) and denoted by $K_{m,n}$, given by the embedding

$$S^1 = \{e^{i\theta}\} \longrightarrow (\frac{e^{im\theta}}{\sqrt{2}}, \frac{e^{in\theta}}{\sqrt{2}}) \subset U \cap V \subset S^3. \tag{4.2.6}$$

To see what a torus knot looks like, it is convenient to start with the representation of the torus as a square, then identify two sides to get a cylinder, and finally identify the top and bottom of the cylinder to get a torus. Figure 3.5 shows how this is done for $K_{2,3}$. This knot is also called the *trefoil* knot. To see the relationship of torus knots to algebraic geometry in a special case, consider the complex locus $\Gamma_{m,n} \subset \mathbf{C}^2$ defined by the equation $f(z_1, z_2) \stackrel{\text{def}}{=} z_1^n - z_2^m = 0$, where m and n are relatively prime positive integers. For $m, n > 1$, $\Gamma_{m,n}$ is singular at the origin. Let S_ϵ be the sphere of radius $\epsilon > 0$ centered at the origin in $\mathbf{R}^4 \simeq \mathbf{C}^2$. For a nonsingular curve Γ, the intersection $S_\epsilon \cap \Gamma$ is an unknotted circle in S^3. However, $S_\epsilon \cap \Gamma_{m,n}$ is easily seen to be the torus knot $K_{m,n}$ (here we can take $\epsilon = 1$). Therefore one expects the structure of the knot to shed some light on the nature of the singularity of the curve at the origin. We shall return to the relationship between geometry and torus knots later, and the reader is referred to [EN] and [Mi3] for this area and its generalizations. ♠

Exercise 4.2.16 *Show that if m or n is unity, then the torus knot $K_{m,n}$ is an unknotted circle. What can you say about the locus $f = 0$ and the map in (4.2.6) if m and n are not relatively prime?*

Example 4.2.15 In this example we show that the group of the torus knot $K_{m,n}$ is isomorphic to the quotient of the free group on two generators ξ_1 and ξ_2 by the relation $\xi_1^m = \xi_2^n$. The proof is based on van Kampen's theorem and example 4.2.13. With the notation of example 4.2.13, let $U' = U \setminus K_{m,n}$, $V' = V \setminus K_{m,n}$, and $x \in U' \cap V'$. It is a simple matter to see that $A = U' \cap V'$ has the homotopy type of a circle, and therefore $\pi_1(A;x) \simeq \mathbf{Z}$. Also U' and V' are solid tori with the knot $K_{m,n}$, which lies on their common boundary, removed. Thus they have the homotopy type of the circle. Let η, ξ_1, and ξ_2 be generators for the fundamental groups of A, U' and V', all with base-point x respectively. Denote by ι_1 and ι_2 the inclusions of A into U' and V' respectively. Then, after possibly replacing one or more of the generators by their inverses, we obtain

$$\iota_{1\sharp}(\eta) = \xi_1^m; \quad \iota_{2\sharp}(\eta) = \xi_2^n. \tag{4.2.7}$$

To see this note that A winds around U' (resp. V') m (resp. n) times. The required result now follows from van Kampen's theorem. ♠

There are algorithmic ways of computing the fundamental group of a knot. These methods give presentations of the fundamental group in terms of generators and relations once the knot is given explicitly pictorially, i.e. as a planar diagram. One such method is the *Wirtinger presentation* of the fundamental group of a knot or a link. Example 4.1.1 plays a key role in the Wirtinger presentation. The algorithm is most easily explained by looking at an example, and it will become clear how to apply to other cases.

Example 4.2.16 Consider the trefoil knot for example, drawn as a piece-wise linear curve (see Figure 3.6). We assume that, except for the underpasses, the knot lies in the plane $z = 0$. At the underpasses the curve has the form $\sqcup$ with the bottom of the cup lying on the plane $z = -1$. Note that $K_{2,3} \cap \{z = -1\}$ consists of three disjoint small line segments which we denote by $A_1, A_2,$ and A_3. We give an orientation to the curve, and label the portions of the curves lying between consecutive underpasses. Thus the trefoil knot is labelled $x_1, x_2,$ and x_3. We may assume the segment x_i passes over A_i. To compute the group of the knot $K_{2,3}$ we decompose $\mathbf{R}^3 \setminus K_{2,3}$ into the union

$$\mathbf{R}^3 \setminus K_{2,3} = X \cup Y_1 \cup Y_2 \cup Y_3 \cup Z,$$

and apply van Kampen's theorem repeatedly. More precisely, let $X = \{(x, y, z)|z \geq -1\} \setminus K_{2,3}$, and $v = (v_1, v_2, v_3) \in X$ with v_3 large. It is clear that that we are in the situation described in example 4.1.1 and

$$\pi_1(X; v) \simeq \mathbf{F}_3.$$

Let Y_i''s be disjoint small solid cubical boxes in the half plane $z \leq -1$ attached to X such that $Y_i' \cap X$ is a rectangle in the plane $z = -1$ containing A_i in its interior. Let L_i be a line segment joining v to $Y_i \cap X$. We assume that $L_i \cap L_j = \{v\}$ for $i \neq j$. Set $Y_i = Y_i' \cup L_i$. $\pi_1(Y_i, v) = \{e\}$ and $\pi_1(Y_i \cap X, v) \simeq \mathbf{Z}$. Applying van Kampen's theorem to the space $X \cup Y_1$ and computing the representations for the loops as described in example 4.1.1, we obtain the relation

$$\xi_1\xi_2\xi_1^{-1}\xi_3^{-1} = e.$$

Proceeding along the curve and attaching Y_2 and Y_3 we obtain the following two relations:

$$\xi_3\xi_1\xi_3^{-1}\xi_2^{-1} = e, \quad \xi_2\xi_3\xi_2^{-1}\xi_1^{-1} = e. \tag{4.2.8}$$

The three relations are not independent, and, in fact, any two of them imply the third. Finally, let $Z' = \{(x, y, z) | z \leq -1\} \setminus (Y_1 \cup Y_2 \cup Y_3)$, and M be a string, lying in X, and connecting v to Z'. Set $Z = Z' \cup M$. Clearly, $\pi_1(Z, v) = e = \pi_1(Z \cap (X \cup Y_1 \cup Y_2 \cup Y_3), v)$. Therefore we can conclude that the fundamental group of the trefoil knot $K_{2,3}$ is the quotient of the free group on three generators by the relations (4.2.8). ♠

Remark 4.2.2 Note that the procedure described in example 4.2.16 can be applied to any knot or link in a mechanical way to obtain a presentation of the fundamental group of the complement of the knot or link in terms of generators and relations. In fact we draw a planar diagram of the knot or link as a piecewise linear manifold specifying the the underpasses. We orient the knot or the link and assign symbols $x_1, x_2, \cdots$ to each segment between two consecutive underpasses. To the $\sqcap$-shaped curve x_j we assign a generator ξ_j of the fundamental group. Around each underpass we draw an oriented rectangle. As one traverses a rectangle one encounters various oriented segments x_j. For every intersection of the sides of the rectangle and the line segments x_j we write (in cyclic order) ξ_j or ξ_j^{-1} according as the side of the rectangle and x_j form a positively or negatively oriented pair of vectors. These give the relations. ♡

The Wirtinger presentation exhibits the fundamental group by n generators ξ_i and n relations R_i. The following observation about the Wirtinger presentation shows that there is a little redundancy in the presentation:

Lemma 4.2.2 *Any one of the relations R_i can be deleted from the set of relations of a Wirtinger presentation without affecting the the group of the knot or link.*

Proof - To prove the redundancy of one relation, say R_n, we work with S^3 rather than R^3 by adjoining the point at infinity. The plane $z = -1$ is actually a copy of S^2. Attaching Y_n does not create a new relation since

$$Y_n \cap (X \cup Y_1 \cup \cdots \cup Y_{n-1})$$

is homeomorphic to the space obtained by deleting an arc from S^2. ♣

Exercise 4.2.17 *Prove algebraically that groups obtained as the fundamental group of the trefoil knot in examples 4.2.15 ($m = 2, n = 3$) and 4.2.16 are isomorphic.*

Exercise 4.2.18 *Show that the fundamental group of the complement of figure 8 knot shown in Figure 3.7 is the quotient of the free group on four generators ξ_j, $j = 1, 2, 3, 4$ by the relations*

$$\xi_3\xi_2 = \xi_1\xi_3, \quad \xi_4\xi_2 = \xi_3\xi_4, \quad \xi_3\xi_1 = \xi_1\xi_4.$$

Understanding or extracting information from a group defined by generators and relations is often a non-trivial matter. A useful general idea is to construct homomorphisms from the given group into other groups the structure of which is better understood. This idea will utilized in examples 4.2.20 and 4.5.3 later.

Exercise 4.2.19 *Consider the trefoil knot K with generators ξ_1, ξ_2, ξ_3 of $\pi_1(S_K, x)$ and subject to the relations (4.2.8). Show the following mappings define two homomorphisms of $\pi_1(S_K, x)$ into the alternating group $\mathcal{A}_5 \subset \mathcal{S}_5$:*

1. $\rho_1 : \xi_1 \longrightarrow (12345)$, $\xi_3 \longrightarrow (12345)$.
2. $\rho_2 : \xi_1 \longrightarrow (12345)$, $\xi_3 \longrightarrow (13542)$.

Prove furthermore that, up to an inner automorphism of $\mathcal{S}_5$, ρ_1 and ρ_2 are the only nontrivial homomorphisms of $\pi_1(S_K, x)$ into $\mathcal{S}_5$ mapping ξ_1 to a 5-cycle.

Example 4.2.17 In this example we study torus knots $K_{2,2n+1}$ for small n in some detail. It is not difficult to convince oneself that Figure 3.8 represents the torus knot $K_{2,9}$. Using Figure 3.8, it is straightforward to show that the Wirtinger presentation of the fundamental group of the complement of $K_{2,2n+1}$ is the group on $2n + 1$ generators $\xi_1, \cdots, \xi_{2n+1}$ subject to the relations

$$\xi_1\xi_{n+1} = \xi_{n+2}\xi_1, \quad \xi_{n+2}\xi_1 = \xi_2\xi_{n+2}, \quad \xi_2\xi_{n+2} = \xi_{n+3}\xi_2, \quad \cdots \tag{4.2.9}$$

The trefoil knot is the case $n = 1$. Notice that the left hand side of the k^{th} relation is the right hand side of the preceding one, and the right hand side of the k^{th} relation is obtained from the left hand side of the preceding one by augmenting the indices by 1. Addition of indices is mod $2n+1$ in the complete residue system $\{1, 2, \cdots, 2n+1\}$. Denote the group on generators $\xi_1, \cdots, \xi_{2n+1}$ subject to relations (4.2.9) by Γ_1. We know that the fundamental group $\pi_1(S_{K_{p,q}}, x)$ is isomorphic to the group Γ_2 on two generators ξ and η subject to the relation $\xi^p = \eta^q$. To understand the relationship of the two presentations of the fundamental group, let us fix $n = 4$ to simplify the argument. The indices (i, j) of $\xi_i\xi_j$ appearing in (4.2.9) are (for $n = 4$)

$$(1,5),\ (6,1),\ (2,6),\ (7,2),\ (3,7),\ (8,3),\ (4,8),\ (9,4),\ (5,9).$$

Thus every integer $k \in \{1, 2, \cdots, 9\}$ appears exactly once as the first coordinate and exactly once as the second coordinate of of the pairs (i, j). Now set

$$\eta = \xi_2\xi_6, \quad \xi = \xi_2\xi_6\xi_2\xi_6\xi_2\xi_6\xi_2\xi_6\xi_2.$$

Expanding η^9 we see that to prove $\eta^9 = \xi^2$ we have to show

$$\xi_2\xi_6\xi_2\xi_6\xi_2\xi_6\xi_2\xi_6\xi_2 = \xi_6\xi_2\xi_6\xi_2\xi_6\xi_2\xi_6\xi_2\xi_6. \tag{4.2.10}$$

Using the relations (4.2.9) we rewrite the left hand side of (4.2.10) as

$$\begin{aligned} \xi_2\xi_6\xi_2\xi_6\xi_2\xi_6\xi_2\xi_6\xi_2 &= \xi_6\xi_1\xi_2\xi_6\xi_2\xi_6\xi_2\xi_6\xi_2 &&= \xi_6\xi_1\xi_5\xi_9\xi_2\xi_6\xi_2\xi_6\xi_2 \\ &= \xi_6\xi_2\xi_6\xi_9\xi_2\xi_6\xi_2\xi_6\xi_2 &&= \xi_6\xi_2\xi_6\xi_9\xi_4\xi_8\xi_2\xi_6\xi_2 \\ &= \xi_6\xi_2\xi_6\xi_2\xi_6\xi_8\xi_2\xi_6\xi_2 &&= \xi_6\xi_2\xi_6\xi_2\xi_6\xi_8\xi_3\xi_7\xi_2 \\ &= \xi_6\xi_2\xi_6\xi_2\xi_6\xi_2\xi_6\xi_7\xi_2 &&= \xi_6\xi_2\xi_6\xi_2\xi_6\xi_2\xi_6\xi_2\xi_6 \end{aligned}$$

which proves (4.2.10). The idea of this derivation was to use the fact every index k appears once as first coordinate to successively convert the the left hand side to the right hand side. Therefore $\eta \to \xi_2\xi_6$, $\xi \to (\xi_2\xi_6)^4\xi_2$ gives a homomorphism $\psi : \Gamma_2 \to \Gamma_1$. Now $\eta^{-4}\xi = \xi_2 \in \Gamma_2$ and it follows easily from (4.2.9) that $\xi_k \in \Gamma_2$ for all k. Therefore ψ is surjective, and we in fact obtain explicit expressions for ξ_k's in terms of ξ and η. For example,

$$\xi_2 = \eta^{-4}\xi,\ \xi_6 = \xi\eta^{-4},\ \xi_7 = \eta^{-4}\xi^2\eta^{-4}\xi^{-1}\eta^4, \quad \text{etc.}$$

These expressions imply that homomorphism ψ has an inverse and therefore $\Gamma_1 \simeq \Gamma_2$. Of course, we knew this fact for geometric reasons and the point in this calculation was to show that the algebraic proof of the isomorphism of two groups given by generators and relations is often non-trivial and sometimes extremely difficult to prove. While our discussion was limited to the case $n = 4$, it can be extended to arbitrary n and the pattern of the argument is clear from the above analysis. The details of this extension are formal and therefore omitted. ♠

Exercise 4.2.20 *Prove algebraically the isomorphism between the group on two generators ξ and η subject to the relation $\xi^2 = \eta^{2n+1}$ and the Wirtinger presentation of $\pi_1(S_{K_{2,2n+1}}, x)$, for $n = 1, 2, 3$. Generalize to arbitrary n.*

Exercise 4.2.21 *The intersection of $S^3 \subset \mathbf{R}^4$ with the hyperplane $x_4 = 0$ is $S^2 \subset \mathbf{R}^3$ and separates S^3 into two closed hemispheres $S^3_{\pm}$ corresponding to $x_4 \geq 0$ and $x_4 \leq 0$. Let $\gamma : [0, 1] \to S^3_+$ be an embedding such that*

$$\mathrm{Im}(\gamma) \cap S^2 = \{\gamma(0), \gamma(1)\}, \quad \text{and} \quad \gamma(0) \neq \gamma(1).$$

Let K be the knot in S^3 obtained by joining $\gamma(0)$ to $\gamma(1)$ by a straight line segment, and $x \in S^3_+ \setminus K$. Prove that

$$\pi_1(S^3 \setminus K, x) \simeq \pi_1(S^3_+ \setminus \mathrm{Im}(\gamma), x),$$

by showing that the Wirtinger presentations for the two groups is the same set of generators and relations.

Example 4.2.18 Consider the link L_2 shown in Figure 3.9 Applying the algorithm of the Wirtinger presentation we see the group of this link is the quotient of the free group on two generators by the relation

$$\xi_1 \xi_2 \xi_1^{-1} \xi_2^{-1} = e.$$

In other words, the group of the link L_2 is the free abelian group on two generators. This could have been proven using example 4.2.13. In fact, by looking at a tubular neighborhood of a connected component of L_2 we see that removing one of the circles from S^3 is (homotopically) the same as removing a solid torus from S^3 leaving us with another solid torus. Removing a tubular neighborhood of the other circle is, up to homotopy, removing the interior of the remaining solid torus. Therefore we end up with a torus whose fundamental group is $\mathbf{Z}^2$. ♠

Exercise 4.2.22 *Compute the fundamental group of the complements of the Borromean and Whitehead links given in Figures 3.10 and 3.11.*

Let K_1 and K_2 be knots in S^3 (or $\mathbf{R}^3$) and P_k be a point on K_k. Cut K_k at P_k to obtain non-closed knots K_1' and K_2'. Denote the end points of K_k' by A_k and B_k. Now we can join K_1' to K_2' in two different ways to obtain a new knot, namely,

1. Join A_1 to A_2 and B_1 to B_2.
2. Join A_1 to B_2 and B_1 to A_2.

However, it is really immaterial how we join K_1 and K_2 and we denote them by the same symbol $K_1\sharp K_2$ in analogy with the construction $M_1\sharp M_2$. Figure 3.12 shows the granny knot and the square knot. The latter knot is obtained by joining a copy of the trefoil knot to its mirror image, while the former is constructed from two copies of the trefoil knot. It is clear that if K_2 is the trivial knot then $K_1\sharp K_2$ and K_1 are the same knot since we can map one to the other by a diffeomorphism of S^3 or $\mathbf{R}^3$. In view of this construction it is reasonable to call a knot K *prime* if for every decomposition $K = K_1\sharp K_2$ one of K_1 or K_2 is the trivial knot. An important implication of the concept of prime knot is

Proposition 4.2.2 *The complement of a prime knot K is determined by $\pi_1(S_K, x)$ (up to homeomorphism).*

The proof of proposition 4.2.2 requires techniques from the topology of manifolds of dimension three which we have not introduced. The interested reader is referred to [Whitt].

Example 4.2.19 One would like to obtain a description of $\pi_1(S_{K_1\sharp K_2}, x)$ in terms of $\Gamma_1 = \pi_1(S_{K_1}, x)$ and $\Gamma_2 = \pi_1(S_{K_2}, x)$. To obtain such a description we assume K_1 and K_2 lie in the half-spaces $U_\pm$ defined by $z > 0$ and $z < 0$ respectively, and "cutting" the knots produces two "knots" whose end points lie in the plane $z = 0$. Denote the "cut knots" by K_1' and K_2', and let

$$U_1' = U_+ \setminus K_1', \quad U_2' = U_- \setminus K_2'.$$

Adding the point at infinity ∞ to work in S^3 rather than $\mathbf{R}^3$, we obtain

$$S_{K_1\sharp K_2} = U_1 \cup U_2, \quad U_1 \cap U_2 = S^2 \setminus \{\text{two points}\}, \tag{4.2.11}$$

where $U_1 = U_1' \cup \infty$, and $U_2 = U_2' \cup \infty$. Let the Wirtinger presentation of the fundamental groups $\Gamma_1 = \pi_1(S_{K_1}, x)$ and $\Gamma_2 = \pi_1(S_{K_2}, x)$ be

$$\Gamma_k: \quad \text{Generators}: \ \xi_1^k, \cdots, \xi_{n_k}^k, \quad \text{Relations}: \ R_1^k, \cdots, R_{n_k-1}^k. \tag{4.2.12}$$

Applying van Kampen's theorem and exercise 4.2.21 to the decomposition (4.2.11), we obtain

$$\pi_1(S_{K_1\sharp K_2}, x) = \Gamma_1 \sharp_{\mathbf{Z}} \Gamma_2, \tag{4.2.13}$$

where the homomorphisms $\rho_k : \mathbf{Z} = \pi_1(U_1 \cap U_2, x) \to \Gamma_k$ are given by mapping a generator of $\pi_1(U_1 \cap U_2, x)$ to generators ξ_i^1 and ξ_j^2. It is no loss of generality to assume that the generators in the Wirtinger presentation are numbered so that $i = 1$ and $j = 1$. In this way we obtain the following presentation for $\pi_1(S_{K_1\sharp K_2}, x)$:

$$\text{Gen.}: \xi_1^1, \cdots, \xi_{n_1}^1, \xi_2^2, \cdots, \xi_{n_2}^2, \quad \text{Rel.}: R_1^1, \cdots, R_{n_1-1}^1, S_1^2, \cdots, S_{n_2-1}^2, \tag{4.2.14}$$

where the relation S_j^2 is obtained from R_j^2 by substituting ξ_1^1 for ξ_1^2. An immediate consequence of this description of the fundamental group of the complement of $K_1\sharp K_2$ is that the groups of the granny knot and the square knot isomorphic. The fact that their complements are not homeomorphic requires the introduction of peripheral systems given below. Thus the hypothesis of primeness of the knot is necessary for the validity of proposition 4.2.2. ♠

Let K be a knot, T_K be an open (small) tubular neighborhood of K and $S_K = S^3 \setminus T_K$. A closed curve on $\partial T_K = \partial S_K$ which bounds a disc in T_K but represents a generator of $H_1(S_K; \mathbf{Z})$ is called a *meridian* (see Figure 3.13 for $K = K_{2,3}$ the trefoil knot). A curve representing $0 \in H_1(S_K; \mathbf{Z})$ and a generator of $H_1(T_K; \mathbf{Z})$ is called a *longitude*. For instance, in the case of the trefoil knot, example 4.1.1 implies that the curve L in Figure 3.13 represents the element $\xi_3\xi_2\xi_1$ in $\pi_1(S_{K_{2,3}}, x)$ where the base point x is represented as a $\bullet$ in the Figure 3.13. Therefore a closed curve in the homotopy class of $\xi_3\xi_2\xi_1\xi_2^{-3}$ represents $0 \in H_1(S_{K_{2,3}}; \mathbf{Z})$ and a generator of $H_1(T_{K_{2,3}}; \mathbf{Z})$. It is convenient to assume $x \in \partial T_K$. Meridians and longitudes are not uniquely defined, however, the subgroup of $\pi_1(S_K, x)$ generated by a meridian and a longitude is the image $\imath_\sharp(\pi_1(T_K, x)) \subset \pi_1(S_K, x)$ and therefore is unambiguously defined. Here $\imath$ is the inclusion of ∂T_K in S_K. We denote $\imath_\sharp(\pi_1(T_K, x))$ by Π_K and refer to the pair $(\pi_1(T_K, x), \Pi_K)$ as a *peripheral system*.

Let $\mu = \mu_K$ and $\lambda = \lambda_K$ denote the homotopy classes of a (fixed) meridian and a (fixed) longitude in $\pi_1(S_K, x)$. It is clear that a homeomorphism of S_K onto $S_{K'}$ will necessarily map a peripheral system to a peripheral system and a meridian μ_K (respectively, a longitude λ_K) to a meridian $\mu_{K'}$ (respectively $\lambda_{K'}$). The notion of peripheral system can be used to distinguish between knot complements with isomorphic fundamental groups[2], and has applications to the topology of manifolds of dimension three.

Example 4.2.20 Let K and K' denote the granny knot and the square knot respectively. We noted in example 4.2.19 that $\Gamma = \pi_1(S_K, x)$ and $\Gamma' = \pi_1(S_{K'}, x)$ are isomorphic. We want to use example 4.2.19 and exercise 4.2.19 to distinguish between S_K and $S_{K'}$. Accordingly we denote the generators of Γ by $\xi_1^1, \xi_2^1, \xi_3^1, \xi_2^2, \xi_3^2$ and those of Γ' by $\eta_1^1, \eta_2^1, \eta_3^1, \eta_2^2, \eta_3^2$. Using exercise 4.2.19 and example 4.2.19, it is not difficult to show that, up to an inner automorphism of $\mathcal{S}_5$, there are only eight non-trivial homomorphisms of Γ into $\mathcal{A}_5$ mapping ξ_1^1 to a 5-cycle, namely,

[2]The issue of homeomorphisms of knot or link complements should not be confused with the stronger condition of a homeomorphism of S^3 mapping one knot/link to another. See for example [Whitt] and [Whi3]

1. $\rho_1 : \xi_1^1 \to (12345),\ \ \xi_3^1 \to (12345),\ \ \xi_3^2 \to (12345)$.
2. $\rho_2 : \xi_1^1 \to (12345),\ \ \xi_3^1 \to (12345),\ \ \xi_3^2 \to (13542)$.
3. $\rho_3 : \xi_1^1 \to (12345),\ \ \xi_3^1 \to (13542),\ \ \xi_3^2 \to (12345)$.
4. $\rho_4 : \xi_1^1 \to (12345),\ \ \xi_3^1 \to (13542),\ \ \xi_3^2 \to (13542)$.
5. $\rho_5 : \xi_1^1 \to (12345),\ \ \xi_3^1 \to (13542),\ \ \xi_3^2 \to (15324)$.
6. $\rho_6 : \xi_1^1 \to (12345),\ \ \xi_3^1 \to (13542),\ \ \xi_3^2 \to (14352)$.
7. $\rho_7 : \xi_1^1 \to (12345),\ \ \xi_3^1 \to (13542),\ \ \xi_3^2 \to (13254)$.
8. $\rho_8 : \xi_1^1 \to (12345),\ \ \xi_3^1 \to (13542),\ \ \xi_3^2 \to (15243)$.

Similarly, one shows that Γ' has eight non-trivial homomorphisms into $\mathcal{A}_5$ mapping η_1^1 to a 5-cycle. These homomorphisms are denoted by ρ'_j, $j = 1, \cdots, 8$ in accordance with the above description for ρ_j. To understand the peripheral systems associated to the granny and the square knots it is necessary to exhibit a longitude for each knot (a meridian is given by the generator ξ_1^1 for the granny knot and η_1^1 for the square knot). We have

1. Granny Knot: $\lambda = (\xi_2^1\xi_1^1\xi_3^1)(\xi_1^1)^{-3}(\xi_2^2\xi_1^1\xi_3^2)(\xi_1^1)^{-3}$.
2. Square Knot: $\lambda' = (\eta_2^1\eta_1^1\eta_3^1)(\eta_2^2\eta_1^1\eta_3^2)^{-1}$.

It is a simple calculation that under the homomorphisms ρ'_j, $j = 2, \cdots, 8$, the peripheral subgroup for the square knot is mapped onto a subgroup isomorphic to $\mathbf{Z}/5$ and $\rho_j(\Gamma')$ is non-abelian. On the other hand, computing the images of a peripheral subgroup for the granny knot under ρ_j we see that none of them has this property. Therefore there is no isomorphism $\Gamma \to \Gamma'$ preserving the peripheral systems and the knot complements S_K and $S_{K'}$ are not homeomorphic. ♠

4.2.4 Some Peculiar Observations

In this subsection we discuss certain phenomena some of which are plausible and some very implausible. Let Γ be a simple close curve in $\mathbf{R}^2$, i.e., the image of a continuous injective mapping of S^1 to $\mathbf{R}^2$. Then Γ decomposes the plane into two connected components which we naturally call the *interior* (denoted by Γ^i) and *exterior* (denoted by Γ^e) of Γ, where Γ^i is relatively compact. While this assertion is certainly very plausible, a formal proof is not trivial. Under the additional hypothesis that φ is smooth with nowhere vanishing derivative, the proof simplifies considerably. The same assertions are valid for the image of S^m in $\mathbf{R}^{m+1}$ and the smooth case is discussed in chapter 6. The fact that a homeomorphic image of S^m separates $\mathbf{R}^{m+1}$ into an exterior and an interior is discussed in [Sp]. We begin with the following plausible proposition:

Proposition 4.2.3 (Schönflies) *Let Γ_1 and Γ_2 be be simple closed curves in $\mathbf{R}^2$. Then any homeomorphism $\varphi : \Gamma_1 \rightarrow \Gamma_2$ extends to a homeomorphism $\overline{\Gamma_1^i} \rightarrow \overline{\Gamma_2^i}$. In particular, every simple closed curve in $\mathbf{R}^2$ bounds a disc.*

The proof of this proposition requires techniques of geometric topology which are different from those discussed in this text. For a proof see [Mo]. Under the stronger hypothesis that φ is a diffeomorphism, or the homeomorphism is piece-wise linear, it is not difficult to construct a proof. With the additional hypothesis of piece-wise linearity (or diffeomorphism) the proposition generalizes:

Proposition 4.2.4 *Let Δ^3 denote the convex closure of the the standard basis and the origin in $\mathbf{R}^3$, and $S \subset \mathbf{R}^3$ be a polyhedron homeomorphic to S^2 (or a set obtained from S be a homeomorphism of $\mathbf{R}^3$). Then there is an open convex set U containing S, and a piece-wise linear homeomorphism of $\mathbf{R}^3$ which is identity in $\mathbf{R}^3 \backslash U$ mapping S onto the boundary of the 3-simplex Δ^3.*

For a proof of this proposition see [Mo]. The surprising fact is that, contrary to dimension 2, this proposition is false without the assumption that S is polyhedral in the sense that it is homeomorphic to a polyhedron under a homeomorphism of $\mathbf{R}^3$ (see example 4.2.22 below). Even more elementary is the following counter intuitive phenomenon: The complement of a finite of discrete set of points in $\mathbf{R}^3$ is simply connected. However, it is possible to construct a compact totally disconnected set (generally called a *Cantor set*) in $\mathbf{R}^3$ whose complement is not simply connected. Anotoine's necklace is such a set and is described in the example below.

Example 4.2.21 Let $U_1 \subset \mathbf{R}^3$ be a standard (closed) solid torus. Inside U_1 consider k disjoint (closed) solid tori linked together as shown in Figure (XXXX). Denote the this set of k closed solid tori by U_2. Similarly inside each component of U_2 consider k (closed) solid tori linked together in the same way. Denote this set of k^2 (closed) solid tori by U_3. Continuing in the obvious manner we obtain sets U_j with U_j the disjoint union of k^{j-1} (closed) solid tori, and $U_j \subset U_{j-1}$. Let $U_\infty = \cap U_j$. Then U_∞ is a compact non-empty set which is totally disconnected since for every $x \in U_\infty$, the connected subset of U_∞ containing x is contained in a ball of radius ϵ for every $\epsilon > 0$. U_∞ is often referenced as *Antoine's necklace*. Let $\mathbf{x} \in \mathbf{R}^3 \backslash U_1$. It is geometrically clear that for every $j \geq 1$ the inclusion $\mathbf{R}^3 \backslash U_j \to \mathbf{R}^3 \backslash U_{j+1}$ induces an injective homomorphism

$$\rho_j : \pi_1(\mathbf{R}^3 \backslash U_j, \mathbf{x}) \to \pi_1(\mathbf{R}^3 \backslash U_{j+1}, \mathbf{x}).$$

Consequently, The induced map $\rho_{1j} : \pi_1(\mathbf{R}^3 \backslash U_1, \mathbf{x}) \to \pi_1(\mathbf{R}^3 \backslash U_j, \mathbf{x})$ is injective for all $j \geq 1$. This implies that the induced map

$$\rho_\infty : \pi_1(\mathbf{R}^3 \backslash U_1, \mathbf{x}) \to \pi_1(\mathbf{R}^3 \backslash U_\infty, \mathbf{x})$$

is also injective. In fact, let γ represent an element of $\pi_1(\mathbf{R}^3 \backslash U_1, \mathbf{x})$ and $F : I \times I \to \mathbf{R}^3 \backslash U_\infty$ a homotopy between γ and the constant map. If $F_n = \mathrm{Im}(F) \cap U_n =\neq \emptyset$ for every n, then $F_n \supset F_{n+1}$ implies $\emptyset \neq \cap F_n \subset U_\infty$ contrary to hypothesis. Therefore $\mathrm{Im}(F)$ is disjoint from U_n for n sufficiently large which implies that $\rho_{1n}(\gamma) = e$ from which injectivity of ρ_∞ follows. Thus Anotine's necklace $U_\infty \subset \mathbf{R}^3$ is a compact totally disconnected set whose complement is not simply connected. ♠

Example 4.2.22 The construction of Antoine's necklace in example 4.2.21 also yields a counter example to proposition 4.2.4 when the assumption that S is a polyhedron is removed. Let $V \supset U_\infty$ be an open set, then there is n such that $U_n \subset V$. We now construct a set $Y \subset V$, homeomorphic to S^2, which is the desired counter example. Let $p \in V \backslash U_n$. For every connected component $C_{n,k}$ of U_n, let $L_{n,k}$ be a broken line segment joining p to $C_{n,k}$. We assume that $C_{n,k} \cap C_{n,l} = \{p\}$ for all k, l and $L_{n,k} \cap U_n \in \partial C_{n,k}$. Denote this point by $p_{n,k}$. Let N_n be a thin closed neighborhood of $\cup_k L_{n,k}$ such that N_n is homeomorphic to the disc D^3 and the intersection of N_n with U_n is exactly the union (over k) of small discs $D^2_{n,k} \subset \partial C_{n,k}$ and $p_{n,k} \in D^2_{n,k}$. Now we can repeat the process with each $D_{n,k}$ (in place of p), thus extending N_n to $N_{n+1} \supset N_n$. N_{n+1} is still homeomorphic to the ball D^3, and for each connected component $C_{n,k,l}$ of U_{n+1}, the intersection $N_{n+1} \cap C_{n,k,l}$ is a small disc $D^2_{n,k,l} \subset \partial C_{n,k,l}$ as shown in Figure (XXX). Clearly the process can be continued indefinitely and we set

$$N' = \bigcup_{j=n}^{\infty} N_n, \quad N = N' \cup U_\infty.$$

The boundary ∂N (i.e., the set of points q such that every neighborhood of q intersects both N and its complement) is homeomorphic to S^2. It follows from example 4.2.21 that $\pi_1(\mathbf{R}^3 \backslash N, \mathbf{x})$ is non-trivial. Consequently ∂N is the required counter example since the fundamental group of the complement of the disc D^3 in $\mathbf{R}^3$ is trivial. This example may be regarded as an extention of the idea of a wild knot (see Figure (XXXX)). One can avoid such pathologies by requiring smoothness, and it is possible that even Hölder continuity would eliminate them. ♠

4.3 $SL(2,\mathbf{Z})$ and the Mapping Class Group

4.3.1 Free Groups and Subgroups of $SL(2,\mathbf{Z})$

In this subsection we use our knowledge of the fundamental group to establish some elementary facts about free groups and certain subgroups of $SL(2,\mathbf{Z})$. One generally refers to $SL(2,\mathbf{Z})$ as the *modular group* and we denote it by Γ in this subsection. We begin with the following proposition:

Proposition 4.3.1 *A subgroup H of the free group $\mathbf{F}_n$ on n generators is free (n can be infinite). Furthermore, if H has index m, then $H \simeq \mathbf{F}_{mn-m+1}$.*

Proof - Recall that $\mathbf{F}_n$ is the fundamental group of n circles joined at one point, $S^1 \vee ... \vee S^1$. The universal cover of $S^1 \vee ... \vee S^1$ is the infinite tree $\mathbf{T}_{2n}$ with $2n$ edges meeting at each vertex, and $\mathbf{F}_n$ acts freely on $\mathbf{T}_{2n}$. H is isomorphic to the fundamental group of the graph $H\backslash\mathbf{T}_{2n}$. The first assertion follows from the fact that the fundamental group of a graph is free. We have the covering projection

$$\pi_m : H\backslash\mathbf{T}_{2n} \longrightarrow S^1 \vee ... \vee S^1.$$

Since H has index m, this is an m-fold covering, and the Euler characterics are related by

$$\chi(H\backslash\mathbf{T}_{2n}) \;= \sharp(\text{vertices}) - \sharp(\text{edges}) \;= m.\chi(S^1 \vee ... \vee S^1) = m - mn.$$

Therefore $H \simeq \mathbf{F}_{mn-m+1}$ as desired. ♣

Next we investigate the structure of certain subgroups of the modular group Γ. Γ acts on the hyperbolic plane $\mathcal{H}$ by fractional linear transformations. Let Γ' be a subgroup of Γ. Then

$$\pi_{\Gamma'} : \mathcal{H} \rightarrow M_{\Gamma'} = \Gamma'\backslash\mathcal{H}$$

is a covering projection if Γ' acts *freely* (or without fixed points), i.e., $\gamma(z) = z$ for some $\gamma \in \Gamma'$ and $z \in \mathcal{H}$ implies $\gamma = \text{id}$. Γ does not act freely on $\mathcal{H}$. We briefly indicate the structure of $\Gamma\backslash\mathcal{H}$. The matrix $\tau_n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \in \Gamma$ acts by translation parallel to the x-axis. Therefore there is an element z in every orbit of Γ with $|\Re(z)| \leq 1/2$. Furthermore, for every $z = x + iy \in \mathcal{H}$, $\Gamma(z) \bigcap \mathsf{F}_\Gamma \neq \emptyset$, where $\mathsf{F}_\Gamma = \{z \in \mathcal{H} | x^2 + y^2 \geq 1 \text{ and } |x| \leq 1/2\}$. This is proven (geometrically) by repeated application of the inversion transformation, i.e., $w = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \Gamma$, and τ_n. We show that F_Γ is a *fundamental domain* for the action of Γ on $\mathcal{H}$ which means that

1. Every orbit of Γ intersects F_Γ;
2. Two points of F_Γ are not in the same Γ orbit unless they are both on the boundary of F_Γ.

This is proven by observing that for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $z \in \mathcal{H}$, $\Im(\gamma(z)) = \Im(z)/|cz+d|^2$, and $\Gamma(z) \cap \mathsf{F}_\Gamma$ is the point(s) in the orbit of z under Γ with maximal imaginary part(s). The equivalence of the boundary points of F_Γ under the action of Γ is easy to understand. In fact, the points $-1/2 + iy$ and $1/2 + iy$ are in the same orbit of Γ. Two points z and z' on the arc $\{z = x + iy \in \mathcal{H} || x|^2 + |y|^2 = 1 \text{ and } |x| \leq 1/2\} \subset \partial\mathsf{F}_\Gamma$ are in the same orbit of Γ if and only if $z' = w(z)$. The points i and $(\pm 1 + i\sqrt{3})/2$ are fixed by the matrices w and $\rho_\pm = \begin{pmatrix} \pm 1 & \mp 1 \\ 1 & 0 \end{pmatrix}$, and these are the only fixed points of Γ in F_Γ. It follows that $\Gamma\backslash\mathcal{H}$ is the sphere with the point at infinity removed. The projection $\mathcal{H} \to \Gamma\backslash\mathcal{H}$ is a covering map outside the (ramification) points $\Gamma(i)$ and $\Gamma((1+i\sqrt{3})/2)$. In particular, $e \neq \gamma \in \Gamma$ has a fixed point if and only if γ is conjugate, in Γ, to w or $\rho_\pm$. Note that this implies that a torsion-free subgroup of Γ acts freely on $\mathcal{H}$. Consequently

Corollary 4.3.1 *Torsion-free subgroups of $SL(2, \mathbf{Z})$ are free.*

Proof - Since a torsion free subgroup $\Gamma' \subset \Gamma$ acts freely on $\mathcal{H}$, Γ' is isomorphic to the fundamental group of $M_{\Gamma'}$. Clearly $M_{\Gamma'}$ is homeomorphic to the manifold obtained by removing several points from a compact orientable surface, and we showed in example **??** that the fundamental group of such a surface is free. ♣

It seems difficult to directly establish freeness of a torsion-free subgroup of Γ without realizing it as the fundamental group of a surface as described above.

Example 4.3.1 The *principal congruence subgroup of level N*, $\Gamma(N)$, is the kernel of the homomorphism

$$R_N : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \to \begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix} ,$$

where $\bar{a}$ denotes the reduction of a mod$N$ and $N \geq 2$. Since $\Gamma(N)$ is a normal subgroup and $w, \rho_\pm \notin \Gamma(N)$, it is free by corollary 4.3.1. Consider the covering projection $\pi_{\Gamma(N)} : \mathcal{H} \to \Gamma(N)\backslash\mathcal{H} = M_{\Gamma(N)}$. It is clear that $M_{\Gamma(N)}$ is diffeomorphic to the surface obtained by removing several points from a compact orientable surface $M'_{\Gamma(N)}$. Hence its fundamental group, $\Gamma(N)$, is the free group on $n_N + 2g_N - 1$ generators, where n_N is the number of the points that are removed, and g_N is the genus of $M'_{\Gamma(N)}$. One can also compute the numbers n_N and g_N in terms of N for the congruence subgroup $\Gamma(N)$. In fact,

$$n_2 = 3;\ n_N = \frac{N^2}{2}\prod_{p|N}(1-\frac{1}{p^2}) \text{ for } N \geq 3;$$

and

$$g_N = 0 \text{ for } N \leq 5;\quad g_N = 1 + \frac{N^2(N-6)}{24}\prod_{p|N}(1-\frac{1}{p^2}) \text{ for } N \geq 6.$$

The index of $\Gamma(N)$ in Γ is

$$[\Gamma : \Gamma(N)] = N^3\prod_{p|N}(1-\frac{1}{p^2}).$$

The proofs of these formulae are not relevant to the material of this text. The interested reader is referred to [Ran] for their proofs and more. ♠

Exercise 4.3.1 *Show that the six matrices*

$$\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad \begin{pmatrix}1&1\\0&1\end{pmatrix},\quad \begin{pmatrix}1&0\\1&1\end{pmatrix},$$
$$\begin{pmatrix}0&-1\\1&0\end{pmatrix},\quad \begin{pmatrix}0&-1\\1&1\end{pmatrix},\quad \begin{pmatrix}-1&-1\\1&0\end{pmatrix},$$

form a set of coset representatives for $\Gamma(2)$ *in* Γ.

Exercise 4.3.2 *For positive integers* N *and* N' *let* $\{N, N'\}$ *denote their* lcm. *Show that*

$$\Gamma(N) \cap \Gamma(N') = \Gamma(\{N, N'\}).$$

Exercise 4.3.3 *For positive integers* N *and* N' *let* (N, N') *denote their* gcd. *Show that the group generated by* $\Gamma(N)$ *and* $\Gamma(N')$ *is* $\Gamma((N, N'))$.

Example 4.3.2 With a little bit of number theory one can exhibit other torsion-free (and therefore free) subgroups of Γ. The subgroups

$$\Gamma_\circ(N) = \{\begin{pmatrix}a&b\\c&d\end{pmatrix} \mid c \equiv 0 \bmod N\}.$$

have important applications in number theory. Let $N = p$ be a prime. From our analysis of torsion elements in $SL(2,\mathbf{Z})$ we deduce that if $\gamma \in \Gamma_\circ(p)$ is a torsion element, then its characteristic polynomial is either $\lambda^2 + 1 = 0$ or $\lambda^2 - \lambda + 1 = 0$. Substituting from $\gamma = \begin{pmatrix} a & b \\ pc' & d \end{pmatrix}$ in $\det(\gamma) = 1$ and reducing mod p we obtain the relations

$$a^2 + 1 \equiv 0 \mod p, \quad \text{or} \quad a^2 - a + 1 \equiv 0 \mod p.$$

From elementary number theory of quadratic residues we know that if $p = 12k - 1$, then neither equation has a solution, and therefore for such p the subgroup $\Gamma_\circ(p)$ is free. ♠

Exercise 4.3.4 *Show that the index of $\Gamma_\circ(N)$ in Γ is*

$$[\Gamma : \Gamma_\circ(N)] = n \prod_{p|n} (1 + \frac{1}{p}).$$

Exercise 4.3.5 *Assume N is such that $\Gamma_\circ(N)$ is torsion free. Use the covering space structure $M_{\Gamma_\circ(N)} \rightarrow M_{\Gamma(N)}$, the preceding exercise and example 4.3.1, to determine the structure of the surface $M_{\Gamma_\circ(N)}$, i.e., determine the genus of its compactification $M'_{\Gamma_\circ(N)}$ and the number of points that are removed.*

Example 4.3.3 In this example we give another geometric interpretation to $SL(2,\mathbf{Z})$. All tori of a given dimension are diffeomorphic, however, they are not necessarily isometric or even conformally isometric relative to the standard flat metric. To understand this point we analyze conformal isometry classes of flat tori in dimension two. Let L be a lattice in $\mathbf{R}^2$. The standard metric $ds^2 = dx^2 + dy^2$ is invariant under translations and therefore induces a flat metric on $T = \mathbf{R}^2/L$. Multiplication by a scalar $\lambda > 0$ is a conformal isometry, and rotations leave the metric ds^2 invariant. Therefore we may assume that the lattices L has a basis

$$L: \ (\alpha, \beta), \ (\gamma, \delta) \quad \text{with} \quad \det \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = 1.$$

It is easily verified that $ds^2 = e^\rho(dx^2 + dy^2)$ is flat if and only if ρ is a constsant. Therefore the set $\mathcal{T}_1$ of conformally equivalent flat tori in dimension two are in natural bijection with the set of lattices in $\mathbf{R}^2$ of determinant 1 where two lattices differing by a rotation of $\mathbf{R}^2$ are identified. Thus $\mathcal{T}_1$ is in natural bijection with $SL(2,\mathbf{Z})\backslash SL(2,\mathbf{R})/SO(2)$. ♠

Exercise 4.3.6 *By introducing complex notation and considering lattices with basis $1, \omega = x + iy$, $y > 0$, show that $\mathcal{T}_1$ is the orbit space of the upper half plane under the action of $SL(2,\mathbf{Z})$ by fractional linear transformations.*

Since the group $SL(2,\mathbf{Z})$ is generated by the matrices

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad W = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

it is natural to try to determine the relations defining the group. It is more convenient to work with $SL(2,\mathbf{Z})/\pm I$. The images of W and T in $SL(2,\mathbf{Z})/\pm I$ will be denoted by the same letters. Set $S = TW$, then S is an element of order 3 in $SL(2,\mathbf{Z})/\pm I$. Our goal is to prove the free product decomposition

$$SL(2,\mathbf{Z})/\pm I \simeq \mathbf{Z}/2 \star \mathbf{Z}/3 \tag{4.3.1}$$

with W and S generators for the groups on the right hand side. We need a criterion for establishing that a group G is the free product of two subgroups A and B. Obviously, it is necessary for A and B to generate G. To prove that $G \simeq A \star B$, it is sufficient (also a necessary condition) to show that a word of the form

$$w = a_1 b_1 a_2 b_2 \cdots a_n b_n, \quad \text{with } a_i \in A,\ b_i \in B, \tag{4.3.2}$$

where $a_i \neq e$ except possibly for $i = 1$ and $b_i \neq e$ except possibly for $i = n$, is not the identity. The fact that these conditions imply $G \simeq A \star B$ is almost immediate. To establish validity of (4.3.1) by invoking this criterion, we look at the action of $SL(2,\mathbf{Z})/\pm I$ on the set of irrational real numbers by fractional linear transformations. Let $\mathbf{R}'_\pm$ denote the sets of positive and negative irrational real numbers. Clearly $\mathbf{R}'_+ \cup \mathbf{R}'_-$ is invariant under the action of $SL(2,\mathbf{Z})/\pm I$. Furthermore

$$W(\mathbf{R}'_\pm) = \mathbf{R}'_\mp, \quad S^{\pm 1}(\mathbf{R}_-) \subset \mathbf{R}_+. \tag{4.3.3}$$

Now consider a word of the form (4.3.2) with $A = \mathbf{Z}/2$ and $B = \mathbf{Z}/3$. If w has odd length then either $a_1 = e$ or $b_n = e$ (but not both). From (4.3.3) it follows that

$$WS^{\pm 1}W \cdots S^{\pm 1}W(\mathbf{R}'_+) \subset \mathbf{R}'_-, \quad S^{\pm 1}W \cdots WS^{\pm 1}(\mathbf{R}'_-) \subset \mathbf{R}'_+.$$

Therefore a product of odd length > 1 is never equal to the identity. Let $\mathbf{R}'_{>1}$ (respectively $\mathbf{R}'_{(0,1)}$) denote the set of irrational real numbers > 1 (respectively in $(0,1)$). For a product of even length of the form

$$w = S^{\pm 1}W \cdots S^{\pm 1}W, \tag{4.3.4}$$

it is easily verified that (note $S = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$)

$$w(\mathbf{R}'_+) \subset \mathbf{R}'_{>1}, \quad \text{or} \quad w(\mathbf{R}'_+) \subset \mathbf{R}'_{(0,1)},$$

according as S or S^{-1} is the left most element in the expansion (4.3.4) of w. In either case such product cannot be equal to e. For an word of the form $w = WS^{\pm 1} \cdots WS^{\pm 1}$, by conjugation with $S^{\mp 1}$ we reduce it to one of the form (4.3.4). Therefore we have shown

Corollary 4.3.2 *The group $SL(2, \mathbf{Z})/\pm I$ is isomorphic to the free product of $\mathbf{Z}/2$ and $\mathbf{Z}/3$ with generators of the latter groups being W and S as defined above.*

4.3.2 Dehn Twist and the Mapping Class Group

In example 4.3.3 we noted that the group $SL(2, \mathbf{Z})$ appears in the classification of flat two dimensional flat tori under conformal equivalence. There is an entirely different way in which $SL(2, \mathbf{Z})$ appears in the study of diffeomorphisms of tori. Let G be the group of orientation preseving diffeomorphisms (or homemophisms) of the torus T^2, and $\tilde{G}$ be the subgroup consisting of those diffeomorphisms which preserve the standard volume element $dx \wedge dy$ on T^2. G and $\tilde{G}$ are topologized by compact open topology and are topological groups. In chapter 1, example **??** we showed that $\tilde{G}$ is an infinite dimensional group by constructing a one parameter family of diffeomorphisms, preserving the volume element, for every smooth function on T^2. All these diffeomorphisms are in the connected component of the identity G° or $\tilde{G}^\circ$. The quotient group G/G° is called the *mapping class group* of T^2 and is denoted by M_1. The structure of M_1 is described by the following proposition:

Proposition 4.3.2 M_1 *is isomrphic to $SL(2, \mathbf{Z})$.*

This proposition is a special of a more general theorem about the mapping class group of compact orientable surfaces. To understand the case of a surface of genus g, it is useful to elaborate on proposition 4.3.2. It is clear that every element of $SL(2, \mathbf{Z})$ acting as a linear transformation of $\mathbf{R}^2$ induces an orientation preserving diffeomorphism of $T^2 = \mathbf{R}^2/\mathbf{Z}^2$. A linear transformation $g \in SL(2, \mathbf{Z})$ induces an automorphism of $\mathbf{Z}^2 = H_1(T^2; \mathbf{Z})$ which is given by the same matrix g relative to the appropriate basis. Therefore unless $g = e$, the transformation of T^2 induced by g is not homotopic to the identity and therefore does not lie in the connected component G° of identity. The transformations of T^2 induced by the matrices

$$U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad V = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

which we still denote by U and V, are special cases of diffeomorphisms called Dehn twists which are defined below. Proposition 4.3.2 states that Dehn twists U and V generate the mapping class group of T^2 since U and V generate $SL(2,\mathbf{Z})$. It is this form of the proposition which generalizes to arbitrary genus $g \geq 1$. To give the general definition of a Dehn twist, let γ be a simple closed curve on a surface representing a homology class. Denote the image of γ by C_γ and let $T_\gamma \simeq [-1,1] \times C_\gamma$ be a tubular neighborhood of C_γ. Denote the variable along $[-1,1]$ by x and that along γ by $\theta \in [0,1]$ with 1 identified with 0. Consider the diffeomorphism ϑ_γ of T_γ defined by

$$\vartheta_\gamma(x,\theta) = \begin{cases} (x, \theta + x), & \text{if } x \in [0,1]; \\ (x, \theta - x), & \text{if } x \in [-1,0]. \end{cases} \tag{4.3.5}$$

It is immediate that the diffeomorphisms of T^2 induced by U and V are of the form ϑ_γ for the appropriate choice of γ. We refer to ϑ_γ as the *Dehn twist* relative to γ. It is more appropriate to consider the isotopy class of diffeomorphisms or homeomorphisms of which a representative is ϑ_γ and call the isotopy class a Dehn twist. This abuse of language will cause no confusion in the sequel. For a surface M_g of genus g let $\gamma_1, \cdots, \gamma_{3g-1}$ be the simple closed curves as shown in Figure XXXX for the case $g = 3$. Then the generalization of proposition 4.3.2 to arbitrary genus is

Proposition 4.3.3 *Let $\gamma_1, \cdots, \gamma_{3g-1}$ be the simple closed curves as described above and ϑ_j be the Dehn twist relative to γ_j. Then $\vartheta_1, \cdots, \vartheta_{3g-1}$ generate the mapping class group M_g.*

(THIS SUBSECTION IS INCOMPLETE)

4.4 Relations with Geometry and Group Theory

4.4.1 Spaces of Constant Curvature

In this subsection we discuss some important classes of simply connected Riemannian manifolds. Our knowledge of local differential geometry, the fundamental groups and covering spaces enables us to understand the structure of simply connected complete Riemannian manifolds of constant sectional curvature. The following observation relates local isometries to covering projections:

Lemma 4.4.1 *Let $\pi : M \rightarrow N$ be a mapping of connected Riemannian manifolds of the same dimension such that $\pi^\star(ds^2_N) = ds^2_M$, i.e., π is a local isometry. If M is complete then N is also complete and π is onto.* (By taking a point out of M, we see that the hypothesis of completeness is necessary.)

Proof - It is clear from the local isometry property that $\mathrm{Im}\pi$ is open in N. Let $x \in \mathrm{Im}\pi$ and γ be a geodesic in N with $\gamma(0) = x$. Let $y \in M$ be such that $\pi(y) = x$. The condition $\pi^\star(ds^2_N) = ds^2_M$ implies that that $\pi_\star : \mathcal{T}_yM \rightarrow \mathcal{T}_xN$ is injective and consequently there is a geodesic δ in M with $\delta(0) = y$ such that $\pi\delta = \gamma$. This implies completeness of N which also implies that the map is onto since N is connected. ♣

Lemma 4.4.2 *Let $\pi : M \rightarrow N$ be a mapping of connected Riemannian manifolds of the same dimension such that $\pi^\star(ds^2_N) = ds^2_M$. If M is complete then (M, π, N) is a covering projection.*

Proof - Let $\rho > 0$ be sufficiently small so that the conditions of remark (XXX) of chapter 2 are fulfilled for $\mathcal{B}^N_\rho \subset \mathcal{T}_xN$. Let $B^N_\rho(x) = \mathrm{Exp}_x(\bar{\mathcal{B}}_\rho)$. It is clear from the above cited remark that π maps $B^M_\rho(y)$ isometrically onto $B^N_\rho(x)$ for every $y \in \pi^{-1}(x)$. We show that for $y \neq y' \in \pi^{-1}(x)$ we have $B^M_\rho(y) \cap B^M_\rho(y') = \emptyset$. If $z \in v$ then the geodesics joining y and y' to z yield either two geodesics joining x to $\pi(z) \in B^N_\rho(x)$ or a geodesic joining x to itself. Neither alternative is possible. To complete the proof that (M, π, N) is a covering projection we have to show

$$\pi^{-1}(B^N_\rho(x)) = \bigcup_{y\in\pi^{-1}(x)} B^M_\rho(y). \tag{4.4.1}$$

Let $w \in B^N_\rho(x)$, $\pi(z) = w$ and γ be the geodesic in $B^N_\rho(x)$ joining w to x. Then there is δ with $\delta(0) = z$ such that $\pi\delta = \gamma$ and by completeness we obtain $y' \in M$ such that $z \in B^M_\rho(y')$. This proves (4.4.1). ♣

To give a classification of complete, simply connected Riemannian manifolds of constant curvature we need the following simple lemma:

Lemma 4.4.3 *Let P and Q be Riemannian manifolds of constant curvature κ and of the same dimension. Let $\varphi : B_\epsilon^P(y) \rightarrow B_\epsilon^Q(\varphi(y))$ be an isometry, and $z \in \partial B_\epsilon^P(y)$. Then there is $\rho > 0$ such that the restriction of φ to $B_\rho^P(z) \cap B_\epsilon^P(y)$ extends to an isometry of $B_\rho^P(z)$ onto its image. The extension is unique.*

Proof - The lemma is a simple consequence of the classification of local isometries of spaces of constant curavture given in chapter 2 proposition (XXX). ♣

Proposition 4.4.1 *All simply connected complete Riemannian manifolds of constant sectional curvature $\kappa \leq 0$ of dimension n are isometric and consequently contractible.*

Proof - Let $\mathcal{H}_n$ be the hyperbolic space of dimension n with sectional curvature $\kappa < 0$, and N a Riemannian manifold of dimension n of constant sectional curvature κ. Let $z \in \mathcal{H}_n$ and $x \in N$. Since a Riemannian metric is locally determined by its sectional curvatures, there is $\epsilon > 0$ and an isometry π mapping $B_\epsilon^{\mathcal{H}_n}(z)$ onto $B_\epsilon^N(x)$ and $\pi(z) = x$. Since geodesics are mapped to geodesics by an isometry, the map π is

$$\pi : \mathrm{Exp}_z(\xi) \longrightarrow \mathrm{Exp}_x(\pi_\star(\xi)), \quad \text{for } \xi \in \mathcal{B}_\epsilon^{\mathcal{H}_n} \subset \mathcal{T}_z\mathcal{H}_n. \tag{4.4.2}$$

Exp is a diffeomorphism of $\mathcal{T}_z\mathcal{H}_n$ onto $\mathcal{H}_n$ since geodesics in $\mathcal{H}_n$ diverge. If $w \in \mathcal{H}_n$ then there is a unique geodesic joining z to w, i.e., there is a unique $\zeta \in \mathcal{T}_z\mathcal{H}_n$ such that $\mathrm{Exp}_z(\xi) = w$, and by defining $\pi(w) = \mathrm{Exp}_x(\pi_\star(\xi))$ we extend π to a mapping of $\mathcal{H}_n \rightarrow N$. To see that π is a local isometry, let $\tau > 0$ be the largest number such that for all $t < \tau$ and all $\xi \in \mathcal{T}_z\mathcal{H}_n$ of norm 1, the map π is an isometry at $\mathrm{Exp}_z(t\xi)$. By lemma 4.4.3 the isometry π extends beyond τ and is necessarily given by expression in (4.4.2) since an isometry maps geodesics to geodesics. Hence π is a local isometry and by lemma 4.4.2 and uniqueness of the universal cover, π is an isometry. Exactly the same argument works for $\kappa = 0$ with $\mathbf{R}^n$ replacing $\mathcal{H}_n$. This completes the proof of the proposition. ♣

An important consequence of lemma 4.4.1 is

Corollary 4.4.1 (Cartan-Hadamard) *A simply connected Riemannian manifold M with sectional curvatures ≤ 0 is diffeomorphic to $\mathbf{R}^m$.*

Proof - It follows from divergence property of geodesics in a non-positively curved Riemannian manifold that the mapping $\pi = \mathrm{Exp}_x : \mathcal{T}_xM \rightarrow M$ is a submersion. π is a local isometry relative to the metric $\pi^\star(ds_M^2)$ on $\mathbf{R}^m = \mathcal{T}_xM$ and by lemma 4.4.1, a covering projection. By simple connectedness it is a diffeomorphism. ♣

For spaces of constant positive curvature we have:

Proposition 4.4.2 *A complete simply connected Riemannian manifold of dimension m and of constant positive sectional curvature $\kappa > 0$ is isometric to the sphere of radius $\frac{1}{\sqrt{\kappa}}$ in $\mathbf{R}^{m+1}$.*

Proof - By scaling we may assume $\kappa = 1$. Let $z \in S^m$ be, for example, the "north pole", and $x \in M$. Since curvature locally determines the metric, there is $\epsilon > 0$ and an isometry $\pi : B^{S^m}_\epsilon(z)$ onto $B^M_\epsilon(x)$ with $\pi(z) = x$. Just as in the case of constant non-positive curvature this isometry is given by (4.4.2) with S^m replacing $\mathcal{H}_n$. From chapter 2 subsection (XXX) we know that the map Exp_z is a diffeomorphism from the ball of radius π in $\mathcal{T}_zS^m$ onto its image. Therefore the map π is in fact defined on $B^{S^m}_\pi(z)$. To extend π to S^m we have to show that $\mathrm{Exp}_x(\xi)$ is independent of the choice of unit tangent vector $\xi \in \mathcal{T}_xM$. Let $x_1 = \mathrm{Exp}_x(\xi) \neq \mathrm{Exp}_x(\eta) = x_2$ for unit tangent vectors $\xi, \eta \in \mathcal{T}_xM$, and consider a curve $\delta : I \to M$ joining x_1 to x_2. Then the set of geodesics joining x to the points $\delta(s)$ gives a variation of the geodesic $\gamma(t) = \mathrm{Exp}_x(t\xi)$ and a consequently a Jacobi field along γ which vanishes at x but not at $\mathrm{Exp}_x(\pi\xi)$ which is not possible. Therefore the desired extension of π exists. To see that π is necessarily smooth and a local isometry, let $0 < t_\circ < \pi$ and let $y = \mathrm{Exp}_z(t_\circ\xi)$. By looking at π as defined in a neighborhood of y we easily see that the extension of π is smooth and a local isometry. By lemma 4.4.2, $\pi : S^m \to M$ is a covering projection and in view of simple connected of M, it is an isometry. ♣

4.4.2 Growth of Fundamental Group

Let M be a Riemannian manifold with sectional curvatures bounded above by $\rho < 0$, and $\pi : \tilde{M} \to M$ be its universal covering space with the induced metric $\pi^\star(ds^2_M)$. The fundamental group $\Gamma = \pi_1(M, x)$ acts properly discontinuously a a group of isometries of $\tilde{M}$. Let us assume M is compact which implies that Γ has a finite set of generators $\{\gamma_1, \cdots, \gamma_n\}$. (For convenience we assume that the set of generators is invariant under taking inverses.) Thus every $\gamma = \Gamma$ is a word

$$\gamma = \gamma_{i_1}\gamma_{i_2}\cdots\gamma_{i_s}. \tag{4.4.3}$$

The *length of* γ, denoted by $L(\gamma)$, is the minimum of s over all representations of γ as a product of the generators. Let $N(s)$ denote the number of elements $\gamma \in \Gamma$ with $L(\gamma) \leq s$. One may pose the question of how $N(s)$ grows as $s \to \infty$. For example, does it have polynomial or exponential growth? We can gain some understanding of this question by invoking a little geometry. The following exercise shows that exponential growth is the best/worst (depending on the individual's attitude) that one can expect:

Exercise 4.4.1 *Let $\mathbf{F}_n$ be the free group on n generators. Show that the growth function N satisfies*

$$c_1 e^{\alpha s} \leq N(s) \leq c_2 e^{\beta s},$$

for some positive numbers c_1, c_2, α, β (In particular, one cannot expect super-exponential growth like $e^{\alpha e^{\beta s}}$ for some positive numbers α and β.)

We noted in example (XXX) of chapter 2 that the volume of a ball in a Riemannian manifold with sectional curvatures bounded above by a negative constant grows exponentially with the radius. Let $x \in \tilde{M}$ and $r > 0$ be sufficietly large so that $\pi(B_r(x)) = M$. Let $d(x, y)$ denote the distance of $x, y \in \tilde{M}$ and

$$\beta = \inf d(B_r(x), \gamma(B_r(x))),$$

where inf is taken with respect to all $\gamma \in \Gamma$ such that $B_r(x) \cap \gamma(B_r(x)) = \emptyset$. Since $\cup_\gamma \gamma(B_r(x)) = \tilde{M}$, the finite set $\Delta = \{\gamma_1, \cdots, \gamma_n\}$ consisting of those γ's for which $B_r(x) \cap \gamma(B_r(x)) \neq \emptyset$, is a set of generators for Γ. First we look at the growth function for this set of generators. Let $y \in \tilde{M}$ and $\delta : I \to \tilde{M}$ be a geodesic joining x to y. Let $0 = t_\circ < t_1 < \cdots t_{s-1} < t_s = 1$ such that

$$d(\delta(t_j), \delta(t_{j+1}) < \beta,$$

and set $x_j = \delta(t_j)$. Let $\delta_j \in \Gamma$ (possibly identity) be such that $x_j \in \delta_j(B_r(x))$ Since $d(x_j, x_{j+1}) < \beta$, $\delta_j^{-1}\delta_{j+1} \in \Delta$. Now

$$\delta_s = \delta_1 \cdot (\delta_1^{-1}\delta_2) \cdots (\delta_{s-2}^{-1}\delta_{s-1})(\delta_{s-1}^{-1}\delta_s)$$

which is a word of length s in the generators Δ. Therefore we have shown

Lemma 4.4.4 *With the above notation and hypotheses, if $d(x, \gamma(x)) < s\beta$, then $L_\Delta(\gamma) \leq s$.*

Recall from chapter 2, example (XXXX) that

$$\text{liminf}_{R\to\infty} \frac{\text{vol}(B_R(x))}{e^{\epsilon R}} > 0 \tag{4.4.4}$$

for some sufficiently small $\epsilon > 0$. Lemma 4.4.4 and (4.4.4) imply the existence of $\eta > 0$ such that

$$\text{liminf}_{s\to\infty} \frac{N_\Delta(s)}{e^{\eta s}} > 0; \tag{4.4.5}$$

in other words, the number of words of length s, relative to the generating set Δ, grows exponentially with s. The role of the particular generating set Δ is not essential. In fact, we have

Proposition 4.4.3 *Let M be a compact Riemannian manifold with sectional curvatures bounded above by a negative constant. Then the function $N(s)$ has exponential growth relative to any generating set.*

Proof - It only remains to show that the property of having exponential growth is independent of the choice of set of generators. If Δ_1 and Δ_2 are two sets of generators, by expressing elements of each in terms of the other, we see that there is an integer k such that

$$N_{\Delta_1}(s) \leq N_{\Delta_2}(ks), \quad N_{\Delta_2}(s) \leq N_{\Delta_1}(ks),$$

which implies the required result. ♣

Exercise 4.4.2 *Let $T^n = \mathbf{R}^n/\mathbf{Z}^n$ be the n-dimensional torus. Show that for its fundamnental group, $\mathbf{Z}^n$, the growth function N satisfies*

$$\alpha s^n \leq N(s) \leq \beta s^n$$

for some positive numbers α and β.

Exercise 4.4.3 *Let T be the group of real 3×3 upper triangular matrices with 1's along the diagonal, $\Gamma \subset T$ be the subgroup consisting of integral matrices. Show that the growth function of the fundamental group Γ of the quotient manifold $\Gamma\backslash T$ satisfies*

$$\alpha s^4 \leq N(s) \leq \beta s^4$$

for some positive numbers α and β.

It is reasonable to inquire whether there are finitely generated (abstract) groups whose growth function satisfies the inequalities

$$c_1 e^{s^\alpha} \leq N(s) \leq c_2 e^{s^\beta},$$

where $0 < \alpha < \beta < 1$, and $c_j > 0$. For a discussion of such groups see [GrKu].

4.4.3 Flat Riemannian Manifolds

One of the consequences of Bieberbach solution to Hilbert's eighteenth problem is

Theorem 4.4.1 *Every discrete subgroup $\Gamma \subset E(m)$ with compact quotient $\Gamma \backslash \mathbf{R}^m$ fits into an exact sequence*

$$\mathbf{0} \longrightarrow L \longrightarrow \Gamma \longrightarrow G \longrightarrow \mathbf{1}, \tag{4.4.6}$$

where the normal subgroup $L = \Gamma \cap \mathbf{R}^m$ is a finitely generated free abelian group of rank m, is a maximal abelian subgroup, and G is a finite group. Conversely, a group Γ containing a free abelian group L of rank m such that

1. *$L \subset \Gamma$ is a normal subgroup;*
2. *Γ / L is a finite group;*
3. *L is a maximal abelian subgroup Γ;*

can be realized as a discrete subgroup of $E(m)$ with compact orbit space $\Gamma \backslash \mathbf{R}^m$. (We assume $m \geq 2$; the case $m = 1$ is easily treated separately.)

An immediate consequence of theorem 4.4.1 is that a group Γ satisfying the hypotheses of the theorem can be realized as a subgroup of $\mathbf{Z}^m.GL(m, \mathbf{Z})$ (semi-direct product) whose image under the projection $\mathbf{Z}^m.GL(m, \mathbf{Z}) \rightarrow GL(m, \mathbf{Z})$ is finite.

Theorem 4.4.1 is not valid without the assumption of compactness of the orbit space $\Gamma \backslash \mathbf{R}^m$. In fact, let $\mathbf{R}^m = V_1 \oplus V_2$, $0 \neq u \in V_1 \simeq \mathbf{R}^n$, $A \in O(m - n)$ (acting on V_2) be an element of infinite order, and Γ be the subgroup generated by the Euclidean motion (u, A). It is clear that Γ is a discrete subgroup of $E(m)$ (isomorphic to $\mathbf{Z}$) and $\Gamma \cap \mathbf{R}^m = e$ and acts on V_1 freely as a group of translations. The following proposition, which incorporates this example, shows that a modified version of theorem 4.4.1 is still valid for the non-compact case.

Proposition 4.4.4 *Let $\Gamma \subset E(m)$ be a discrete subgroup, then Γ contains an abelian normal subgroup $\Gamma_\circ$ of finite index containing all the translations in in Γ. Given an abelian subgroup $\Gamma \subset E(m)$, then there is a subspace $V \subset \mathbf{R}^m$ such that Γ admits of the decomposition $\Gamma = \Gamma_1 \times \Gamma_2$ with Γ_2 finite and acting trivially on V and Γ_1 a free abelian group acting freely as a group of translations on V.*

The solution to Hilbert's eighteenth also contains the following finiteness result:

Proposition 4.4.5 *For each m, there are only finitely many crystallographic groups of the form (4.4.6). There are only finitely many diffeomorphism classes of compact flat Riemannian manifolds of a given dimension m.*

The proof of proposition 4.4.5 makes use of the following result (due to Minkowski) which is of independent interest and extends torsion freeness of congruence subgroups of $SL(2,\mathbf{Z})$ to $GL(m,\mathbf{Z})$.

Proposition 4.4.6 *Let $p \geq 3$ be an odd prime, and $R_p : GL(n,\mathbf{Z}) \rightarrow GL(n,\mathbf{Z}/p)$ be the reduction* mod p *map. Then* $\mathrm{Ker}(R_p)$ *is torsion free and every finite subgroup of $GL(n,\mathbf{Z})$ is mapped injectively into $GL(n,\mathbf{Z}/p)$.*

Proof - Let $\gamma \in GL(n,\mathbf{Z})$ be a torsion element so that $\gamma^l = I$, and assume $\gamma \in \mathrm{Ker}(R_p)$. If $l > 1$ then we may assume l is a prime. Define $T' = \gamma - I$, then $T' = p^r T$ where $p^r > 1$ is the highest power of p dividing the entries of the matrix T'. Therefore

$$0 = \gamma^l - I = lp^r T + \binom{l}{2} p^{2r} T^2 + \cdots \tag{4.4.7}$$

Since l is a prime, (4.4.7) implies $l = p$ and $r = 1$. Therefore

$$T + \binom{p}{2} T^2 + \binom{p}{3} pT^3 + \cdots = 0. \tag{4.4.8}$$

The hypothesis $p \geq 3$ implies $p | \binom{p}{2}$ which, in view of (4.4.8), contradicts the assumption that not every entry of T is divisible by p. Hence $\gamma = I$. ♣

Bieberbach's solution to Hilbert's eighteenth problem extends earlier results of Jordan and has been significantly simplified by Frobenius and others (see [Oli]) and will be discussed in the final subsection of this section. Here we mainly concentrate on some examples especially in connection with the theory of covering spaces and fundamental groups of compact flat Riemannian manifolds.

By a *crystallographic group* we mean a discrete subgroup $\Gamma \subset E(m)$ with compact quotient space $\Gamma\backslash\mathbf{R}^m$. Note that in theorem 4.4.1 we did not require the orbit space $\Gamma\backslash\mathbf{R}^m$ to be a manifold. Since $G \subset E(m) = \mathbf{R}^m.O(m)$, in case $M = \Gamma\backslash\mathbf{R}^m$ is a manifold, the standard Euclidean metric on $\mathbf{R}^m$ is invariant under Γ and induces a flat Riemannian metric on M. A discrete subgroup of $E(m)$ is necessarily closed, and it is straightforward to verify that a subgroup $\Gamma \subset E(m)$ acts discontinuously on $\mathbf{R}^m$ if and only is it is *discrete*, i.e., there is neighborhood U of $e \in E(m)$ such that the sets $\gamma(U)$ and U are disjoint except possibly for finitely many γ's. The first issue we address is when a discrete subgroup $\Gamma \subset E(m) = \mathbf{R}^m.O(m)$ determines a covering projection $p : \mathbf{R}^m \rightarrow \Gamma\backslash\mathbf{R}^m$, i.e., when is the action properly discontinuous. If the action of every $\gamma \neq e$ on $\mathbf{R}^m = E(m)/O(m)$ is free from fixed points (i.e. the action is free), then M is a manifold and $\pi : \mathbf{R}^m \rightarrow M$ is a covering projection. Conversely, if $\gamma \neq e$ has a fixed point in $\mathbf{R}^m$ then $p : \mathbf{R}^m \rightarrow \Gamma\backslash\mathbf{R}^m$ is not a covering projection.

Lemma 4.4.5 *Let $\Gamma \subset E(m)$ be a discrete subgroup. Then $M = \Gamma\backslash\mathbf{R}^m = \Gamma\backslash E(m)/O(m)$ is a manifold and the canonical mapping $p : \mathbf{R}^m \rightarrow M$ is a covering projection if and only if Γ is torsion free.*

Proof - If $\gamma \in \Gamma$ is torsion, then the point $\sum_j \gamma^j(\mathbf{x})$ is fixed by γ where $\mathbf{x} \in \mathbf{R}^m$. Conversely assume $\mathbf{x} \in \mathbf{R}^m$ is fixed by $\gamma \in \Gamma$. Let $\tau_{\mathbf{x}}$ be translation by $\mathbf{x}$. Then $\gamma \in \tau_{\mathbf{x}} O(m)\tau_{-\mathbf{x}} \cap \Gamma$. Since Γ is closed, $\tau_{\mathbf{x}} O(m)\tau_{-\mathbf{x}} \cap \Gamma$ contains torsion elements unless $\tau_{\mathbf{x}} O(m)\tau_{-\mathbf{x}} \cap \Gamma = e$. ♣

Example 4.4.1 For a lattice $L \subset \mathbf{R}^m \subset E(m)$ acting on (left) on $\mathbf{R}^m$ by translations, the quotient space $L\backslash\mathbf{R}^m$ is a compact manifold, in fact an m-dimensional torus. In this example, we construct a noncommutative discrete subgroup $\Gamma \subset E(m)$ such that the quotient space $\Gamma\backslash\mathbf{R}^m$ is a compact manifold and the canonical mapping $\pi : \mathbf{R}^m \rightarrow M = \Gamma\backslash\mathbf{R}^m$ is a covering projection. Now let $e_1, \cdots, e_m$ be the standard basis for $\mathbf{R}^m$ and A be the permutation matrix mapping e_i to e_{i+1} and e_m to e_1. Then $A^m = I$. Let $L = \mathbf{Z}^m$ be the lattice of vectors with integer coordinates. Let Γ be the subgroup of $E(m)$ generated Euclidean motions of the form (e_j, A). Thus $G \simeq \mathbf{Z}/m$ in notation of theorem 4.4.1. We show that Γ has the required properties. Discreteness and noncommutativity of Γ are clear and it suffices to show that it is torsion free. Every $\gamma \in \Gamma$ is of the form $\gamma = (v, A^k)$ for some $v \in L$. Let $V \subset \mathbf{R}^m$ be the orthogonal complement of the vector $w = \sum e_i$, $(Aw = w)$, i.e., the subspace consisting of vectors of the form $\sum a_i e_i$ with $\sum a_i = 0$. If γ has finite order then $v \in L \cap V$ in view of the above analysis. Since $(u, A)^{-1} = (-A^{-1}u, A^{-1})$ and (v, A^k) is a product of terms of the form (e_i, A) and $(-e_k, A^{-1})$, we have the expression

$$\gamma = (v, A^k) = (\epsilon_1 e_{i_1} + \cdots + \epsilon_l e_{i_l}, A^{\epsilon_1} \cdots A^{\epsilon_l}),$$

where $\epsilon_j = \pm 1$. Thus $v \in V$ if and only if $\sum \epsilon_j = 0$ in which case

$$A^k = A^{\epsilon_1} \cdots A^{\epsilon_l} = A^{\sum \epsilon_j} = I.$$

Therefore Γ is torsion free. Notice that Γ contains the abelian normal subgroup L' generated by $(e_i - e_j, I), (e_1 + \cdots + e_m, I)$, $1 \leq i, j \leq m$, and Γ/L' is isomorphic to the cyclic group $\mathbf{Z}/(m)$. However Γ is not isomorphic to the semi-direct product of L' and $\mathbf{Z}/(m)$ since it is torsion free. ♠

Exercise 4.4.4 *Show that the discrete subgroup Γ of example 4.4.1 fits into an exact sequence*

$$\mathbf{0} \longrightarrow \mathbf{Z}^{m-1} \longrightarrow \Gamma \longrightarrow \mathbf{Z} \longrightarrow 0.$$

If we remove the requirement of torsion freeness, then we can easily construct many groups $\Gamma \subset E(m)$ containing an abelian normal subgroup L of finite index. Of course for a finite subgroup $G \subset O(m)$ leaving a lattice $L \subset \mathbf{R}^m$ invariant, one can consider the semi-direct product $L.G$. However, there are many examples of discrete subgroups $\Gamma \subset E(m)$ which are not semi-direct products. One refers to an exact sequence of the form (4.4.6) as an *extension of* L *by* G or simply a *group extension*. The case of the semi-direct product is called the *trivial extension*. In view of the abelian assumption on L, the conjugation action of Γ on L induces a homomorphism $G \rightarrow \mathrm{Aut}(L)$ which enables one to define semi-direct product. A necessary and sufficient condition for Γ to be isomorphic to the semi-direct product of G and L is the existence of a homomorphism $\chi : G \rightarrow \Gamma$ whose composition with the homomorphism $\pi : \Gamma \rightarrow G$ is the identity map of G. In fact, the semi-direct product structure in such a case is given by

$$\gamma \longrightarrow (\gamma(\chi\pi(\gamma))^{-1}, \pi(\gamma)) \in L \cdot G.$$

The following exercise shows that there are non-trivial group extensions which nevertheless contain torsion elements.

Exercise 4.4.5 *Let $G = \mathcal{S}_m$ be the symmetric group on m letters represented by $m \times m$ permutation matrices, i.e., a permutation is regarded as the orthogonal transformation effecting the same permutation of the standard basis $e_1, \cdots, e_m$ of $\mathbf{R}^m$. Let*

$$\gamma_1 = (e_1, \sigma_{12}), \ \gamma_2 = (e_2, \sigma_{23}), \ \ \cdots, \ \gamma_{m-1} = (e_m, \sigma_{m1}),$$

where σ_{ij} denotes the permutation transposing the indices i and j. Let $\Gamma \subset E(m)$ be the group generated by $\gamma_1, \cdots, \gamma_m$, and $\rho : \Gamma \rightarrow \mathcal{S}_m$ be the homomorphism $\rho((u, \sigma)) = \sigma$. Show that for $\sigma \in \mathcal{S}_m \setminus \mathcal{A}_m$, every $(u, \sigma) \in \rho^{-1}(\sigma)$ has infinite order, and deduce that

$$0 \longrightarrow L \longrightarrow \Gamma \longrightarrow G \longrightarrow \mathbf{1}$$

is a non-trivial group extension. Exhibit torsion elements in Γ.

Our immediate goal is to show that example 4.4.1 can be generalized in the sense that for every finite group G one can construct an exact sequence (4.4.6) with Γ torsion free. It is convenient to introduce a definition. If the orbit space $M = \Gamma \backslash \mathbf{R}^m$ is a manifold, one then refers to $G = \Gamma / L$ as the *linear holonomy group* of the Riemannian manifold M. Implicit in this definition is that the Riemannian metric on M is the flat metric induced from $\mathbf{R}^m$. The key observation in proving that every finite group may be realized as the linear holonomy group of a flat Riemannian manifold is

Lemma 4.4.6 *Let $\mathbf{F}_N$ be the free group on N generators and $G = \mathbf{F}_N/R$ for some normal subgroup R (of relations). Then the commutator subgroup $[R, R]$ is a normal subgroup of $\mathbf{F}_N$, and $\mathbf{F}_N/[R, R]$ is torsion free.*

Proof - The first assertion is immediate. To prove the second assertion let $e \neq \gamma \in \mathbf{F}_N/[R, R]$. If $\gamma \in R/[R, R]$ then clearly γ has infinite order. So assume $\gamma \notin R$ and let Δ be the subgroup of $\mathbf{F}_N$ generated by R and γ. $\Delta/[\Delta, \Delta]$ is a free abelian group and since $[\Delta, \Delta] \supset [R, R]$, if γ is a torsion element of $\mathbf{F}_N/[R, R]$ then $\gamma \in [\Delta, \Delta]$. Now $[\Delta, \Delta] \subset R$ since

$$(\gamma^a r)(\gamma^b r')(r^{-1}\gamma^{-a})(r'^{-1}\gamma^{-b}) = (\gamma^a r \gamma^{-a})(\gamma^{a+b} r' r \gamma^{-a-b})(\gamma^b r'^{-1}\gamma^{-b})$$

which clearly lies in R. Therefore $\gamma \in R$ contrary to hypothesis. ♣

Lemma 4.4.6 suggests that the obvious choice for a crystallographic group with linear holonomy G is $\Gamma = \mathbf{F}_N/[R, R]$, where $G \simeq \mathbf{F}_N/R$. In order for theorem 4.4.1 to be applicable we still need to establish that $R/[R, R]$ is a maximal abelian subgroup of $\mathbf{F}_N/[R, R]$. The following lemma implies the required maximality:

Lemma 4.4.7 *Let $R \subset \mathbf{F}_N$ be a normal subgroup of finite index m of the free group $\mathbf{F}_N$ and $N \geq 2$. Then the conjugation action of the finite group $G = \mathbf{F}_N/R$ on $R/[R, R]$ is effective, i.e., if $g \in G$, and for all $r \in R$, $grg^{-1}[R, R] = r[R, R]$, then $g = e$*

Proof - Assume the contrary and let $g' \in \mathbf{F}_N \setminus R$ be an element representing $g \in \mathbf{F}_N/R$, $F \subset \mathbf{F}_N$ the subgroup generated by R and g'. Let q be the order of $g \in G$, then $\{e, g, \cdots, g^{q-1}\}$ is a complete set of coset representatives for F/R. By proposition 4.3.1 F is a free group on s generators and since $R \simeq \mathbf{F}_{Nm-m+1}$ we have

$$s = 1 + \frac{(N-1)m}{q}. \tag{4.4.9}$$

Clearly $[F, F] = [R, R]$ and therefore $F/[R, R]$ is a free abelian group containing a subgroup (namely, $R/[R, R]$) isomorphic to $\mathbf{Z}^{Nm-m+1}$. Therefore

$$s \geq Nm - m + 1. \tag{4.4.10}$$

Relations (4.4.9) and (4.4.10) show that $s = N = 1$. ♣

It is clear the lemma 4.4.7 implies that R is a maximal abelian subgroup of $\mathbf{F}_N/[R, R]$ and consequently theorem 4.4.1 is applicable to show

Corollary 4.4.2 *For every finite group G there is a lattice of rank m and a torsion free crytallographic group $\Gamma \subset E(m)$ containing L as the subgroup of translations and $G \simeq \Gamma/L$. Equivalently, there is a compact flat Riemannian manifold with linear holonomy group G.*

Another consequence of theorem 4.4.1 is the fact the fundamental group of a compact flat Riemannian manifold determines the manifold up to diffeomorphism.

Corollary 4.4.3 *Let M and M' be compact flat Riemannian manifolds, of dimensions m and m', with fundamental groups $\Gamma = \pi_1(M, x)$ and $\Gamma' = \pi_1(M', x')$. If Γ and Γ' are isomorphic, then $m = m'$ and there is a linear isomorphism of $\mathbf{R}^m$ onto itself inducing a linear diffeomorphism (but not necessarily an isometry) of M and M'.*

Proof - Let $L = \Gamma \cap \mathbf{R}^m$ and $L' = \Gamma' \cap \mathbf{R}^{m'}$. Then L and L' are the maximal abelian normal subgroups of their respective groups and any isomorphism $\Phi : \Gamma \to \Gamma'$ necessarily maps L isomorphically onto L'. Denote this isomorphism by T. In particular $m = m'$ and we obtain the row exact commutative diagram with vertical arrows isomorphisms:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & L & \longrightarrow & \Gamma & \longrightarrow & G & \longrightarrow & 0 \\
 & & T \downarrow & & \Phi \downarrow & & \phi \downarrow & & \\
0 & \longrightarrow & L' & \longrightarrow & \Gamma' & \longrightarrow & G' & \longrightarrow & 0
\end{array}$$

Here G and G' are finite subgroups of $O(m)$. From the semi-direct product decomposition $E(m) = \mathbf{R}^m.O(m)$ it follows that for every $(u, h), (v, g) \in \Gamma$ we have $(v, g)(u, h) = (v + g \cdot u, gh)$ and consequently

$$(T(v) + \phi(g) \cdot T(u), \phi(gh)) = (T(v) + T(g \cdot u), \phi(gh)).$$

In other words, we have

$$\phi(g) = TgT^{-1}.$$

It follows that T induces a diffeomorphism of the of the flat manifolds M and M'. Since T may not be a Euclidean motion, the induced diffeomorphism is not necessarily an isometry. ♣

A geometric method for the construction of discrete subgroups of $E(m)$, which extends to the constant curvature as well, will be discussed in connection with Riemannian manifolds of constant negative curvature.

Example 4.4.2 In this example we show that the assumption that Γ is a subgroup of $E(m)$ is essential and theorem 4.4.1 is not valid if we replace Euclidean motions with affine transformations. We use the notation of example (XXXX) of chapter 2. Let U be the group of 3×3 upper triangular matrices with 1's along the diagonal. Recall that relative to a left invariant metric ds^2, U which is diffeomorphic to $\mathbf{R}^3$, is not flat. The left action of U on itself yields affine transformations of $\mathbf{R}^3$ which are isometries relative to ds^2, but are not Euclidean motions. Let $U_{\mathbf{Z}} \subset U$ be the subgroup consisting of integral matrices. Clearly, $U_{\mathbf{Z}}$ acts by left translations as a properly discontinuous group of isometries of U relative to ds^2, and $M = U_{\mathbf{Z}}\backslash U$ is a compact manifold. However $U_{\mathbf{Z}}$ does not contain an abelian normal subgroup of finite index. Notice that the action of $U_{\mathbf{Z}}$ on U is by affine transformations of $\mathbf{R}^3 = U$ which are not Euclidean isometries. It is a simple matter to generalize this example to obtain many properly discontinuous groups of affine transformations of $\mathbf{R}^m$, $m \geq 3$, with compact quotient. ♠

It appears that the theory of torsion free discrete subgroups of $E(m)$ is not adequately developed. For example, given a representation $\rho : G \rightarrow O(m)$, where G is a finite group, and $\rho(G)$ leaves a lattice L invariant, it is interesting to know when there is a torsion free discrete subgroup $\Gamma \subset E(m)$ containing a lattice $L' \subset L$ as a maximal abelian subgroup and $\Gamma/L' \simeq G$. For some results about torsion free discrete subgroups of $E(m)$ see [FH]. The following lengthy example demonstrates the difficulty in establishing torsion freeness.

Example 4.4.3 Let D_n denote the dihedral group of order $2n$. In this example we construct, for every odd integer n, an exact sequence of the form (4.4.6) with $G = D_n$ and Γ torsion free. Recall that D_n is the semi-direct product $C_n.\mathbf{Z}/2$ of the of the cyclic group $C_n \simeq \mathbf{Z}/n$ of order n and $\mathbf{Z}/2$ with the latter group acting on the former by mapping an element to its inverse. We write t for the non-identity element of $\mathbf{Z}/2$ and fix a generator α for C_n. Consider the representation ρ of C_n on $\mathbf{R}^n$ with the standard basis $e_1, \cdots, e_n$ given by

$$\rho(\alpha)(e_j) = e_{j+1}$$

where the indices are computed mod n. Let $\mathbf{R}^n \oplus \mathbf{R}^n = \mathrm{Ind}_\rho^{D_n}$ in the notation of representation theory. To describe this explicitly consider $\mathbf{R}^{2n}$ with basis $e_1, \cdots, e_{2n}$ and let C_n act on the span of $e_1, \cdots, e_n$ according to the representation ρ. We let $\mathbf{Z}/2$ act on $\mathbf{R}^{2n}$ by defining action of $t \in \mathbf{Z}/2$ by $t : e_i \rightarrow e_{n+i}$ for $i \leq n$. Then C_n leaves the span of $e_{n+1}, \cdots, e_{2n}$ invariant and the action of α on this subspace is given by

$$\alpha : e_{n+i} \longrightarrow \alpha.t.(e_i) = t(t\alpha t).(e_i) = t\alpha^{-1}.(e_i) = e_{n+i-1}, \quad \text{for } \ i > 1$$

and $\alpha(e_{n+1}) = e_{2n}$. n being odd, we let $2k = n + 1$ and set $t' = \alpha^{-k}t\alpha^k$. The group D_n can be generated by two elements t and t'. In fact note that $t\alpha^{-k}t\alpha^k = \alpha^{2k} = \alpha$. Let $\Gamma \subset E(2n)$ be the subgroup generated by transformations of the form

$$(e_i, t),\ (e_a, t') \quad \text{where} \quad i \leq n \quad \text{and} \quad a \geq n + 1, \tag{4.4.11}$$

and their inverses. Here t and t' are regarded as orthogonal transformation of $\mathbf{R}^{2n}$. Γ is a discrete subgroup of $E(2n)$ and $\Gamma\backslash\mathbf{R}^{2n}$ is compact. It is clear that any expansion of $(v, \sigma) \in \Gamma$ as a product of generators given in (4.4.11) and their inverses contains either an even number or an odd number of factors according as $\sigma \in C_n$ or $\sigma \notin C_n$. Since the vector $z = e_1 + \cdots + e_{2n}$ is fixed by the action of D_{2n}, if v is not orthogonal to z, then $(v, \sigma) \in \Gamma$ has infinite order. Therefore

1. If (v, σ) is a product of odd length when as expressed in terms of generators (4.4.11) and their inverses, then (v, σ) is not a torsion element.

To facilitate understanding the argument for torsion freeness of Γ we introduce a mapping $\lambda : \Gamma \longrightarrow \mathbf{Z} \oplus \mathbf{Z}$. An element of Γ has a unique expression of the form

$$(v, \sigma) = (a_1e_1 + \cdots + a_{2n}e_{2n}, \sigma), \quad \text{with} \quad a_j \in \mathbf{Z}, \sigma \in D_{2n}.$$

Define

$$\lambda(v, \sigma) = (a_1 + \cdots + a_n, a_{n+1} + \cdots + a_{2n}).$$

Each of the n dimensional subspaces spanned by $e_1, \cdots, e_n$ and $e_{n+1}, \cdots, e_{2n}$ is invariant under C_n and the vectors $e_1 + \cdots + e_n$ and $e_{n+1} + \cdots + e_{2n}$ are fixed by this subgroup. The second observation in proving torsion freeness of Γ is:

2. An element $(v, \sigma) \in \Gamma$ is torsion if and only if $\lambda(v, \sigma) = (0, 0)$ and $e \neq \sigma \in C_n$.

This statement follows immediately from example (XXX) of chapter 1 or by an elementary calculation.

If A is an element of the form (e_i, t), $(i \leq n)$, and B an element of the form (e_a, t'), $(a \geq n + 1)$, or their inverses, then for the value of λ on products of length 2 we have the following table:

Type 1	λ	*Type 2*	λ	*Type 3*	λ	*Type 4*	λ
AA	(1,1)	AB	(2,0)	BB	(1,1)	BA	(0,2)
AA^{-1}	(0,0)	AB^{-1}	(1,-1)	BB^{-1}	(0,0)	$B^{-1}A$	(-1,1)
$A^{-1}A$	(0,0)	$A^{-1}B$	(1,-1)	$B^{-1}B$	(0,0)	BA^{-1}	(-1,1)
$A^{-1}A^{-1}$	(-1,-1)	$A^{-1}B^{-1}$	(0,-2)	$B^{-1}B^{-1}$	(-1,-1)	$B^{-1}A^{-1}$	(-2,0)

If $(v, \sigma) \in \Gamma$ is product of length two and $(u, \tau) \in \Gamma$, then

$$\lambda((v,\sigma)(u,\tau)) = \lambda(v,\sigma) + \lambda(u,\tau). \tag{4.4.12}$$

It follows that (4.4.12) is valid for all $(v,\sigma), (u,\tau) \in \Gamma^e$ where Γ^e is the subgroup of all products of even length, i.e., all (v,σ)'s with $\sigma \in C_n$. We write $\lambda = (\lambda_1, \lambda_2)$ and let $\bar{\lambda} = (\bar{\lambda_1}, \bar{\lambda_2})$ where $\bar{\lambda}_j$ denotes the reduction of λ_j mod n. Now we can show

3. Let (v, α^r) be a product of r elements of type 2, and (u, α^{-s}) be a product of s elements of type 4 in the table. Then
$$\bar{\lambda}(v, \alpha^r) = \bar{\lambda}(u, \alpha^{-s}) \tag{4.4.13}$$
if and only if $r \equiv s \bmod n$.

To prove this assertion notice that in view of validity of 4.4.12 on Γ^e and the possible values of λ as described in the table, we have partitions

$$r = r_1 + r_2 + r_3, \quad s = s_1 + s_2 + s_3,$$

and

$$\bar{\lambda}(v, \alpha^r) = (2r_1 + r_2, -r_2 - 2r_3), \quad \bar{\lambda}(u, \alpha^{-s}) = (-2s_1 - s_2, s_2 + 2s_3).$$

Therefore the equation (4.4.13) is equivalent to $2r \equiv 2s \bmod n$ or $r \equiv s \bmod n$ since n is odd. Having proven validity of (3) we can establish torsion freeness of Γ. In view of (1) the only possible torsion elements are products of even length in the generators of Γ and their inverses. Every $(v,\sigma) \in \Gamma^e$ is a product of elements of the four types described in the table. Let r be the number elements of type 2 and s the number of elements of type 4 in the table. It follows easily from the table, (2) and (4.4.13) that in order for (v,σ) to be torsion it is necessary that $\bar{\lambda}(v, \alpha^r) = \bar{\lambda}(u, \alpha^{-s})$. From (3) we see easily that for such (v,σ) we have $\sigma = e$. This completes the proof of torsion freeness of Γ. ♠

4.4.4 2 and 3-D Crystals

Understanding the structure of crystals was an important factor in the development of the theory of discrete groups of Euclidean motions. Mathematically, a *crystal structure* (denoted by (L, G)) consists of a lattice $L \subset \mathbf{R}^m$, together with a finite subgroup $G \subset GL(m, \mathbf{R})$ (called *point group* in chemistry literature) leaving the lattice L invariant. By a general theorem of algebra, given G and L there are finitely many group extensions of the form

(4.4.6). The question arises given a finite group $G \subset GL(m, \mathbf{R})$, whether there are G-invariant lattices $L \subset \mathbf{R}^m$. Clearly if such L exists then $g(L)$, where g lies in the normalizer $\mathcal{N}_G$ of G in $GL(m, \mathbf{R})$, is also invariant under G. Therefore it is reasonable to designate two G-invariant lattices L and L' as G-*equivalent* if they differ by an element of $\mathcal{N}_G$. For example, if $G = \{e\}$, then all lattices in $\mathbf{R}^m$ are G-equivalent. We should also define an equivalence relation on the set of crystal structures. Given a crystal structure (L, G) and $g \in GL(m, \mathbf{R})$, then $(g(L), gGg^{-1})$ is also a crystal structure. Thus one defines two crystal structures as equivalent if they differ by an element $g \in GL(m, \mathbf{R})$. In this subsection we consider crystal structures in dimensions 2 and 3.

Example 4.4.4 Let $G = \mathbf{Z}/n$ be the cyclic group of order n acting on $\mathbf{R}^2$ by rotations $R_{\frac{2k\pi}{n}}$ through angles $\frac{2k\pi}{n}$. We want to determine lattices in $\mathbf{R}^2$ invariant under G. If $n = 2$ then $G = \{\pm I\}$ and every lattice is invariant under G. Furthermore, G lies in the center of $GL(2, \mathbf{R})$ and and therefore all lattices are G-equivalent. If $G = \mathbf{Z}/3$, then the lattice L is invariant under rotation by $\frac{2\pi}{3}$. Let $0 \neq v_1 \in L$ be a non-zero vector of minimal length, $v_2 = R_{\frac{2\pi}{3}}(v_1)$ and $L' \subset L$ be the sublattice generated by v_1, v_2. Let w a vector of minimal length in $L \setminus L'$. In view of invariance of L and L' under $G = \mathbf{Z}/3$ and the minimality of length assumption, we may assume that the angle between w and v_1 lies in the open interval $(\frac{2\pi}{3}, \pi)$. But then the vector $w + v_1$ has length less than that of w contrary to hypothesis. Therefore $L = L'$. The normalizer of $\mathbf{Z}/3$ in $GL(2, \mathbf{R})$ is $\mathbf{R}^\times \times O(2)$. Therefore all $G = \mathbf{Z}/3$-invariant lattices in $\mathbf{R}^2$ are G-equivalent to one generated by the vectors $v_1, v_2 = R_{\frac{2\pi}{3}}(v_1)$. This lattice is also invariant under $G = \mathbf{Z}/6$. It follows that $\mathbf{Z}/6$-invariant lattices are the same as $\mathbf{Z}/3$-invariant lattices. For $G = \mathbf{Z}/4$ we proceed similarly. Let $v_1 \in L$ be a non-zero vector of minimal length, $v_2 = R_{\frac{\pi}{2}}(v_1)$, and L' the lattice generated by v_1, v_2. Let $w \in L \setminus L'$ be a vector of minimal length. From minimality and $\mathbf{Z}/4$-invariance assumptions, we deduce w can be chosen so that the angle between w and v_1 lies in the interval $(\frac{2\pi}{3}, \pi)$ and therefore $w + v_1$ is a vector of length less than that of w. Therefore $L' = L$, and all $\mathbf{Z}/4$-invariant lattices in $\mathbf{R}^2$ are $\mathbf{Z}/4$-equivalent to the standard lattice spanned by $e_1 = (1, 0)$ and $e_2 = (0, 1)$. Finally assume $n = 5$ or $n \geq 7$, and let v_1 be a non-zero vector of minimal length in L. Let

$$w = \begin{cases} R_{\frac{\pi(k-1)}{k}}(v_1), & \text{if } n = 2k; \\ R_{\frac{2\pi(k-1)}{n}}(v_1), & \text{if } n = 2k - 1. \end{cases}$$

Then the vector $w + v_1$ has length less than that of v_1. Consequently for $n = 5$ or $n \geq 7$ there are no $\mathbf{Z}/n$-invariant lattices in $\mathbf{R}^2$. ♠

For $n \geq 3$, an injective homomorphism $\rho : \mathbf{Z}/n \rightarrow GL(2,\mathbf{R})$ is necessarily conjugate to one into $SO(2)$. In fact, the eigenvalues of $\rho(1)$ are necessarily primitive n^{th} roots of unity, i.e., they are $e^{\pm\frac{2\pi ik}{n}}$ for some k relatively prime to n, which implies that $\rho(1)$ is conjugate to a 2×2 rotation matrix. For $n = 2$, either both eigenvalues of $\rho(1)$ are -1 in which case $\rho(1)$ is rotation through π, or there is a basis v_1, v_2 such that

$$\rho(1)(v_1) = v_1, \quad \rho(1)(v_2) = -1.$$

The two cases are distinguished according as $\det(\rho(1)) = \pm 1$.

Let $G \simeq \mathbf{Z}/2\times\mathbf{Z}/2$. To distinguish two copies of $\mathbf{Z}/2$ we write $G \simeq \{e, \epsilon\}\times\{e', \epsilon'\}$. Then there is a basis v_1, v_2 for $\mathbf{R}^2$ such that the matrices (ϵ, e') and (e, ϵ') are represented as

$$(\epsilon, e') \leftrightarrow \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (e, \epsilon') \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

After possibly replacing $G \simeq \mathbf{Z}/2\times\mathbf{Z}/2$ we may assume that v_1, v_2 is the standard orthonormal basis e_1, e_2.

Exercise 4.4.6 *Let $G \simeq \mathbf{Z}/2 \times \mathbf{Z}/2$ acting on $\mathbf{R}^2$ as diagonal matrices $\begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$. Show that there are two equivalence classes of of crystal structures (L, G), namely,*

1. *L is generated by e_1, e_2.*
2. *L is generated by e_1 and $w = \frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2$.*

Exercise 4.4.7 *For the group G of order 8 of symmetries of the square, show that there are two inequivalent crystal structures (L, G).* (Assume the standard basis e_1, e_2 are eigenvectors for $\mathbf{Z}/2 \times \mathbf{Z}/2$, and consider the lattices generated e_1, e_2 and $2e_1, e_1 + e_2$.)

Exercise 4.4.8 *Let $G = \mathcal{S}_3$ and L be the lattice spanned by $e_1, -\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2$ with the even permutation (123) acting as rotation by $\frac{2\pi}{3}$. Show that*

1. $\rho((12))(e_1) = -e_1, \quad \rho((12))(-\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2) = \frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2.$
2. $\rho'((12))(e_1) = e_1, \quad \rho((12))(-\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2) = -\frac{1}{2}e_1 - \frac{\sqrt{3}}{2}e_2.$

extend the action of $\mathbf{Z}/3$ to $\mathcal{S}_3$. Denote the representations of $\mathcal{S}_3$ by ρ and ρ'. Show that ρ and ρ' are equivalent representations and $\rho(\sigma)$, $\rho'(\sigma')$, $\sigma \in \mathcal{S}_3$, are integral matrices relative to the basis $e_1, -\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2$. Prove that there is no $g \in GL(2, \mathbf{Z})$ implementing the equivalence of ρ and ρ'.

In chapter 1, §XXXX, we determined finite subgroups of $SO(3)$. The first step in the determination of 3-D crystal structures is to determine which of these finite groups leave a lattice in $\mathbf{R}^3$ invariant. The following simple lemma eliminates most finite subgroups and reduces the problem to a finite number of cases:

Lemma 4.4.8 *Let $e \neq g \in SO(3)$ leave a lattice $L \subset \mathbf{R}^3$ invariant. Then g is a rotation by π, $\frac{2\pi}{3}$, $\frac{\pi}{2}$ or $\frac{\pi}{3}$.*

Proof - Assume g is not a rotation through angle π, and $v \in L \subset \mathbf{R}^3$ be linearly independent from the axis of rotation of g. Let Π_g denote the plane of rotation of g. Then the vectors $v - g(v)$ and $g(v) - g^2(v)$ are linearly independent and lie in $L \cap \Pi_g$. Therefore by the two dimensional case (example 4.4.4) g cannot be rotation through $\frac{2\pi}{n}$ where $n = 5$ or $n \geq 7$. ♣

Note in particular that lemma 4.4.8 eliminates the icosahedral group ($\simeq \mathcal{A}_5$) as a point group in dimension 3. Using this lemma, one can determine all the possible point groups in dimension three. The book-keeping is facilitated by splitting up the set of point groups into three classes, viz.,

1. Groups constructed from rotation group $\mathbf{Z}/n$ in the plane.
2. Groups constructed from the dihedral group $D_n = \mathbf{Z}/n.\mathbf{Z}/2$ acting on the plane.
3. Groups constructed from the group of proper symmetries of the (regular) tetrahedron or the cube.

Each case is subdivided according as the group is contained in $SO(3)$, or contains the inversion $-I$ or a reflection. Since the arguments are straightforward, lengthy and hardly enlightening, we tabulate the complete list of thirty two possible point groups. It is convenient to introduce some notation. We make use of the representations

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \leftrightarrow [\bar{x}, y, \bar{z}], \quad \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \leftrightarrow [\bar{x}, \bar{z}, y] \quad \text{etc.}$$

Recall from chapter 1 §XXXX that the group of proper symmetries of the regular tetrahedron is isomorphic to the alternating group $\mathcal{A}_4$. B_3 is the group of order 48 of symmetries of the cube and contains the group B'_3 of proper symmetries of the cube as a subgroup of index two. The following table is more or less self-explanatory except for the numbers in the last column which will explained shortly:

Class	$G \subset SO(3)$	$-I \in G$	$-I \notin G$	L
1	I	$I \cup (-I)$	$-$	1
1	$\mathbf{Z}/2 = I \cup [\bar{x}, \bar{y}, z]$	$\mathbf{Z}/2 \cup (-I)\mathbf{Z}/2$	$I \cup [x, y, \bar{z}]$	2
1	$\mathbf{Z}/3$	$\mathbf{Z}/3 \cup (-I)\mathbf{Z}/3$	$-$	5
1	$\mathbf{Z}/4 = \mathbf{Z}/2 \cup [y, \bar{x}, z]\mathbf{Z}/2$	$\mathbf{Z}/4 \cup (-I)\mathbf{Z}/4$	$\mathbf{Z}/2 \cup [\bar{y}, x, \bar{z}]\mathbf{Z}/2$	4
1	$\mathbf{Z}/6 = \mathbf{Z}/3 \cup [\bar{x}, \bar{y}, z]\mathbf{Z}/3$	$\mathbf{Z}/6 \cup (-I)\mathbf{Z}/6$	$\mathbf{Z}/3 \cup [\bar{y}, x, \bar{z}]\mathbf{Z}/3$	5
2	$D_2 = \mathbf{Z}/2 \cup [x, \bar{y}, \bar{z}]\mathbf{Z}/2$	$D_2 \cup (-I)D_2$	$\mathbf{Z}/2 \cup [\bar{x}, y, z]\mathbf{Z}/2$	3
2	$D_3 = \mathbf{Z}_3 \cup [y, x, \bar{z}]\mathbf{Z}/3$	$D_3 \cup (-I)D_3$	$\mathbf{Z}/3 \cup [\bar{y}, \bar{x}, z]\mathbf{Z}/3$	5
2	$D_4 = \mathbf{Z}/4 \cup [x, \bar{y}, \bar{z}]\mathbf{Z}/4$	$D_4 \cup (-I)D_4$	$\mathbf{Z}/4 \cup [\bar{x}, y, z]\mathbf{Z}/4$	4
2	$-$	$-$	$D_2 \cup [\bar{y}, x, \bar{z}]D_2$	4
2	$D_6 = \mathbf{Z}/6 \cup [y, x, \bar{z}]\mathbf{Z}/6$	$D_6 \cup (-I)D_6$	$\mathbf{Z}_6 \cup [\bar{y}, \bar{x}, z]\mathbf{Z}/6$	5
2	$-$	$-$	$D_3 \cup [x, y, \bar{z}]D_3$	5
3	$\mathcal{A}_4$	$\mathcal{A}_4 \cup (-I)\mathcal{A}_4$	$-$	7
3	B_3'	$B_3 = B_3' \cup (-I)B_3'$	$\mathcal{A}_4 \cup [y, x, z]\mathcal{A}_4$	7

The action of the cyclic groups $\mathbf{Z}/n$, $n = 2, 3, 4, 6$, on $\mathbf{R}^3$ is defined in the first column of the table. So is the action of the dihedral groups D_n, $n = 2, 3, 4, 6$. Note that $D_2 \cup (-I)D_2 \simeq (\mathbf{Z}/2)^3$. The isomorphisms

$$\mathbf{Z}/4 \cup [x, \bar{y}, \bar{z}]\mathbf{Z}/4 \simeq D_2 \cup [y, \bar{x}, z]D_2, \quad \mathbf{Z}/6 \cup [y, x, \bar{z}]\mathbf{Z}/6 \simeq D_3 \cup [\bar{x}, \bar{y}, z]D_3,$$

explain the blanks in the first two columns of the table. The thirty two point groups enumerated in the above table leave a lattice invariant. Therefore it is necessary to exhibit a lattice invariant under each of the given point groups. We exhibit below seven lattice which prove that the groups in the above table do indeed leave a lattice invariant. The numbers in the last column of the table refer to these lattices as enumerated below. The lattices given below are representatives from G-equivalence classes and we also have indicated the prevalent terminology from crystallography. Throughout, v_1, v_2, v_3 denotes a basis for L, the length of v_i is denoted by l_i, and A_{ij} denotes the angle between between v_i and v_j. It is convenient to interpret a statement such as $l_1 = l_2$, l_3 *arbitrary* as $l_3 \neq l_1 = l_2$ in spite of the fact in the equivalence class there is a lattice with $l_3 = l_1 = l_2$. The latter lattice will have more symmetries and maybe regarded as a degenerate case of the former.

1. (*Triclinic*) - v_1, v_2, v_3 are any three linearly independent vectors.

2. (*Monoclinic*) - l_i's are arbitrary, but $A_{23} = A_{13} = \frac{\pi}{2}$ and A_{12} is arbitrary.

3. (*Orthorhombic*) - l_i's are arbitrary, but $A_{ij} = \frac{\pi}{2}$.

4. (*Tetragonal*) - $l_1 = l_2$, l_3 arbitrary, and $A_{ij} = \frac{\pi}{2}$.
5. (*Hexagonal*) - $l_1 = l_2$, l_3 arbitrary, $A_{23} = A_{13} = \frac{\pi}{2}$, and $A_{12} = \frac{2\pi}{3}$.
6. (*Trigonal*) - $l_1 = l_2 = l_3$, and $A_{12} = A_{23} = A_{13} < \frac{2\pi}{3}$.
7. (*Cubic*) - $l_1 = l_2 = l_3$ and $A_{ij} = \frac{\pi}{2}$.

The seven lattices enumerated above are not the only ones admitting of the given finite groups of symmetry (see for example, exercise 4.4.7). In crystallography literature the above seven are known as *primitive Bravais classes.* There are also seven *imprimitive* lattices which are derived from the above as indicated below. The resulting fourteen lattices are called *Bravais lattices.*

1. Lattice generated by the monoclinic lattice together with the vector $\frac{1}{2}(v_1 + v_2)$.
2. Three lattices derived from the orthorhombic lattice:
   (a) By addition of the vector $\frac{1}{2}(v_1 + v_2)$;
   (b) By addition of the vector $\frac{1}{2}(v_1 + v_2) + v_3$;
   (c) By addition of the vectors $\frac{1}{2}(v_1 + v_2)$, $\frac{1}{2}(v_2 + v_3)$ and $\frac{1}{2}(v_3 + v_1)$.
3. Lattice obtained by adding the vector $\frac{1}{2}(v_1 + v_2 + v_3)$ to the tetragonal lattice.
4. Two lattices derived from the Cubic lattice:
   (a) By addition of the vector $\frac{1}{2}(v_1 + v_2 + v_3)$;
   (b) By addition of the vectors $\frac{1}{2}(v_1 + v_2)$, $\frac{1}{2}(v_2 + v_3)$ and $\frac{1}{2}(v_3 + v_1)$.

Having determined the equivalence clases of crystal structures, we still have to determine all goup extensions of the form (4.4.6). This is not a trivial problem and the solution involves lengthy case by case constructions. In dimension two one obtains 17 such groups; there are 230 such groups in 3-D and 4783 in dimension four. We have already indicated the explicit construction of some of these groups in the preceding subsection. For a complete construction of all such groups in dimensions two and three see [Bur]. In dimension four a complete enumeration is given in [B...Z].

Only a small number of the crystallographic groups are torsion free. In fact in dimension two there are only two such groups namely $\mathbf{Z}$ and the torsion free group constructed in example 4.4.1 for $m = 2$. The latter group is the fundamental group of the Klein bottle. In dimension three there are ten torsion free crytallographic groups of which six are contained in $SE(3)$ so that the orbit space $\Gamma\backslash\mathbf{R}^3$ is an orientable manifold of dimension three. An enumeration of these groups is given in the following examples:

Example 4.4.5 It is not difficult to describe the fundamental groups of compact flat orientable Riemannian manifolds M of dimension three. The matrices in the following enumeration are relative to the standard basis e_1, e_2, e_3 of $\mathbf{R}^3$:

1. $\pi_1 \simeq \mathbf{Z}^3$ in which case the M is a torus of dimension three.

2. The point group $G = \mathbf{Z}/2$ which acts on $\mathbf{R}^3$ via the matrix $A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. A set of generators for Γ is
$$\gamma = (e_3, A),\ \tau_1 = (e_1, I),\ \tau_2 = (e_2, I).$$

3. The point group is $\mathbf{Z}/3$ which acts on $\mathbf{R}^3$ via the matrix
$$A = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
Γ is generated by
$$\gamma_1 = (e_3, A),\ \tau_1 = (e_1, I),\ \tau_2 = (-\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2, I).$$
This is group is isomorphic to the group described in example 4.4.1 for $m = 3$.

4. The point group is $\mathbf{Z}/4$ generated by rotation by $\frac{\pi}{2}$ given by the matrix
$$A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
The group Γ is generated by
$$\gamma_1 = (e_3, A),\ \tau_1 = (e_1, I),\ \tau_2 = (e_2, I).$$

5. The point group is $\mathbf{Z}/2 \times \mathbf{Z}/2$ with the generators acting by the matrices
$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$
The subgroup Γ is generated by
$$\gamma_1 = (e_3, A),\ \gamma_2 = (e_1, B), \tau_1 = (e_2, I).$$

6. The point group is $\mathbf{Z}/6$ acting by rotation by $\frac{k\pi}{3}$ generated by:
$$A = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
The group Γ is generated by
$$\gamma_1 = (e_3, A),\ \tau_1 = (e_1, I),\ \tau_2 = (\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2, I).$$

The non-orientable case is given below. ♠

Exercise 4.4.9 *Show that the groups exhibited in example 4.4.5 are torsion free.*

Example 4.4.6 Continuing with example 4.4.5, we enumerate compact flat non-orientable manifolds.

1. The point group is $\mathbf{Z}/2$ acting on $\mathbf{R}^3$ by the matrix $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. The group $\Gamma = \Gamma_1 \times \mathbf{Z}$ where Γ_1 is the group constructed in example 4.4.1 for $m = 2$. The orbit space $\Gamma \backslash \mathbf{R}^3$ is $K \times S^1$ where K is the Klein bottle.

2. The point group is $\mathbf{Z}/2$ acting on $\mathbf{R}^3$ via the matrix $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. The group Γ is generated by
$$\gamma_1 = (e_1, A),\ \tau_1 = (2e_2, I),\ \tau_2 = (e_1 + e_2 - e_3, I).$$

3. The point group is $\mathbf{Z}/2 \times \mathbf{Z}/2$ acting diagonally on $\mathbf{R}^3$ via matrices

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

   The group Γ is generated by

$$\gamma_1 = (e_2, A), \ \gamma_2 = (e_1, AB), \ \tau_1 = (e_3, I).$$

4. The point group is $\mathbf{Z}/2 \times \mathbf{Z}/2$ acting diagonally on $\mathbf{R}^3$ via matrices

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

   The group Γ is generated by

$$\gamma_1 = (e_1, C), \ \gamma_2 = (e_2 + e_3, A), \ \tau_1 = (2e_3, I).$$

   The verification of torsion free-ness of these groups is left as an exercise. ♠

Exercise 4.4.10 *Show that there is one and only one compact flat three dimensional manifold with vanishing first Betti number, and its linear holonomy group is* $\mathbf{Z}/2 \times \mathbf{Z}/2$.(This manifold is called a *Hantsche-Wendt* manifold.)

4.4.5 Proofs of Bieberbach's theorems

In this subsection we prove theorem 4.4.1 and propositions 4.4.4 and 4.4.5. We represent elements $\gamma, \delta \in \Gamma$ in the form $\gamma = (u, A)$ and $\delta = (v, B)$ following the semi-direct product decomposition $E(m) = \mathbf{R}^m.O(m)$. We recall that $E(m) \subset GL(m+1, \mathbf{R})$ where (u, A) is represented as the matrix $\begin{pmatrix} A & u \\ 0 & 1 \end{pmatrix}$. The standard Euclidean metric on $\mathbf{R}^{n^2}$ restricts to the metric on $GL(n, \mathbf{R})$ given by $||C||^2 = \mathrm{Tr}(CC')$ and called the *Hilbert-Schmidt* norm. For $a > 0$ we let Γ_a be the subgroup generated by elements $(u, A) \in \Gamma$ such that $||A - I|| < a$. The following properties of the Hilbert-Schmidt norms of matrices A, B and $u \in \mathbf{R}^m$ are easy to verify:

1. $||A|| = m$ for $A \in O(m)$;

2. $||AB|| \leq ||A||.||B||$, and $||Au|| \leq ||A||.||u||$;

3. $||AB|| = ||B||$ for $A \in O(m)$

The idea of the proof of theorem 4.4.1 is to show that for all $a > 0$, Γ_a is a normal subgroup of finite index in Γ and if a is sufficiently small (in fact, $a \leq \frac{1}{2}$), then Γ_a is abelian. Then one shows that Γ_a is contained in $\mathbf{R}^m$. To carry out the proof we begin with

Lemma 4.4.9 *$\Gamma_a \subset \Gamma$ is a normal subgroup of finite index.*

Proof - Normality follows from the identity $||UAU^{-1}-I|| = ||A-I||$ where $U \in O(m)$. Since $O(m)$ is compact, for every $a > 0$ there is N such that for every subset $\{A_1, \cdots, A_N\} \subset \Gamma$ there are indices i, j with $||A_i - A_j|| < a$. Given $a > 0$ let $\{(u_1, A_1), \cdots, (u_n, A_n)\} \subset \Gamma$ be a maximal subset with the property $||A_i - A_j|| \geq a$ for all $1 \leq i, j \leq n$. For $(v, B) \in \Gamma$, $||A_i - B|| < a$ for some i by the maximality assumption. Now $(u_i, A_i)^{-1}(v, B) = (v - A_i^{-1}u_i, A_i^{-1}B)$ and $||A_i^{-1}B - I|| = ||A_i - B|| < a$ proving $(u_i, A_i)\Gamma_a = (v, B)\Gamma_a$, i.e., Γ_a has finite index. ♣

The fact that Γ_a is abelian for $a > 0$ sufficiently small is proven in several steps. The first lemma is a simple exercise in linear algebra:

Lemma 4.4.10 *Let $A, B \in O(m)$ and $||B - I|| < 2$. If A and BAB^{-1} (or $BA^{-1}B^{-1}$) commute, then so do A and B.*

Proof -Consider the decomposition of the complexification $\mathbf{C}^m = V_1 \oplus \cdots \oplus V_r$ into eigenspaces corresponding to distinct eigenvalues of A. The subspaces V_j are orthogonal. Since A and BAB^{-1} commute, the subspaces V_j are invariant under BAB^{-1}, and consequently $AB^{-1}(V_j) = B^{-1}(V_j)$, i.e., $B^{-1}(V_j)$ is invariant under A. Therefore $B^{-1}(V_j) = \sum_i (B^{-1}(V_j) \cap V_i)$. Now if $B^{-1}(v)$ is orthogonal to v, $0 \neq v \in \mathbf{R}^m$, then necessarily $||B^{-1}-I|| = ||B-I|| \geq 2$ contradicting $||B - I|| < 2$. Therefore $B^{-1}(V_j) = V_j = B(V_j)$ for all j which implies that A and B commute. The proof in case of commutativity of A and $BA^{-1}B^{-1}$ is similar. ♣

Lemma 4.4.11 *Let $\gamma = (u, A), \delta = (v, B)$ be in Γ_1. If $AB = BA$ then $\gamma\delta = \delta\gamma$.*

Proof - Let $\gamma_1 = \gamma\delta\gamma^{-1}\delta^{-1}$ and inductively define $\gamma_n = \gamma\gamma_{n-1}\gamma^{-1}\gamma_{n-1}^{-1}$. It is a simple calculation that

$$\gamma_1 = (c, I), \quad \text{and} \quad \gamma_n = ((A - I)^{n-1}w, I),$$

where $w = (I - A)u - (I - B)v$. Now $||(I - A)^n w|| \leq ||I - A||^n ||w|| \to 0$ as $n \to \infty$, and by discreteness of Γ, $\gamma_n = (0, I)$ for all $n \geq N$, i.e., $(I - A)^{N-1}(w) = 0$. Let $\mathbf{R}^m = V_1 \oplus V_2$

where V_1 is eigenspace corresponding to eigenvalue 1, and V_2 is its orthogonal complement. Since $I-A$ is bijective on V_2, $(I-A)^{N-1}(w)=0$ implies $\gamma_2=(0,I)$. Now $\gamma_2=\gamma\gamma_1\gamma^{-1}\gamma_1^{-1}$ and so γ_1 commutes with γ, and consequently γ commutes with $\delta\gamma^{-1}\delta^{-1}$. By lemma 4.4.10 γ and δ commute. ♣

Lemma 4.4.12 *Let $\gamma=(u,A)\in\Gamma_{1/2}$ and $\delta=(v,B)\in\Gamma_2$. Then $AB=BA$.*

Proof - Let $\delta_1=\delta\gamma\delta^{-1}$ and inductively define $\delta_n=\delta_{n-1}\gamma\delta_{n-1}^{-1}$. Set $\delta_n=(w_n,C_n)$. Then $C_n=C_{n-1}AC_{n-1}^{-1}$ and

$$w_n=(I-C_{n-1}AC_{n-1}^{-1})(w_{n-1})+C_{n-1}(u). \tag{4.4.14}$$

Therefore

$$||w_n||\leq||u||+a||w_{n-1}||\leq\frac{1}{1-a}||u||,$$

where $a\leq\frac{1}{2}$. Thus the sequence $\{w_n\}\subset\mathbf{R}^m$ remains bounded and consequently $\{\delta_n\}$ has a convergent subsequence. Now, unless $A=C_n$ for all n sufficiently large, we have

$$\begin{aligned}||A-C_{n+1}|| &= ||AC_n-C_nA||\\ &= ||(A-C_n)(A-I)-(A-I)(A-C_n)||\\ &\leq 2||I-A||.||A-C_n||.\end{aligned}$$

This means that $C_n\rightarrow A$ as $n\rightarrow\infty$ since $2||I-A||<1$, and by discreteness $A=C_n$ for all n sufficiently large. In particular, A commutes with $C_n=C_{n-1}AC_{n-1}^{-1}$ and since $||C_j-I||<2$ for all j, C_{n-1} commutes with A by lemma 4.4.10. Proceeding inductively, we see that $C_1=BAB^{-1}$ commutes with A and consequently A and B commute. ♣

Lemmas 4.4.11 and 4.4.12 imply

Corollary 4.4.4 *$\Gamma_{1/2}$ is abelian.*

Next we show that if $\Gamma\backslash\mathbf{R}^m$ is compact and Γ is abelian, then Γ is necessarily a lattice in the normal subgroup $\mathbf{R}^m\subset E(m)$. The first observation in this direction is

Lemma 4.4.13 *Let $\Gamma\subset E(m)$ be an abelian group and $(u,A)\in\Gamma$. There is $(v,I)\in E(m)$ such that $Au'=u'$ where $(u',A)=(-v,I)(u,A)(v,I)$.*

Proof - Consider the orthogonal direct sum decomposition $\mathbf{R}^m=V_1\oplus V_2$ where V_1 is the eigenspace for eigenvalue 1 of A and V_2 its orthogonal complement. Clearly $V_2=\mathrm{Im}(A-I)$ and $u'=u+(A-I)v$. The required result follows immediately. ♣

Continuing with the assumption that Γ is an abelian discrete subgroup of $E(m)$ (not necessarily with compact quotient) we let $(u, A) \in \Gamma$ be such that the eigenvalue 1 has minimal multiplicity. If 1 were not an eigenvalue of A, then in view of lemma 4.4.13, we may assume $u = 0$. Therefore if $(v, B) \in \Gamma$, then commutativity of Γ implies $v = 0$ and $AB = BA$. Consequently $\Gamma \subset O(m)$ and it is a finite group by discreteness. We have shown

Lemma 4.4.14 *Let $\Gamma \subset E(m)$ be a discrete abelian group of infinite order. Then*

$$\min_{(u,A)\in\Gamma}(\text{Multiplicity of Eigenvalue 1 of } A) \geq 1.$$

Let $(\mathbf{R}^m)^A$ be the eigenspace for eigenvalue 1, and Γ be as in lemma 4.4.14. Choose $(u, A) \in \Gamma$ such that $\dim(\mathbf{R}^m)^A = k \geq 1$ is minimal. Writing (u, A) as an $(m+1) \times (m+1)$ matrix we have, after possibly replacing Γ by a conjugate subgroup in $E(m)$, we have

$$(u, A) = \begin{pmatrix} I_k & 0 & u_1 \\ 0 & A_1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where I_k is $k \times k$ identity matrix, u_1 a column vector in $\mathbf{R}^k$ and A_1 an $(m-k) \times (m-k)$ orthogonal matrix all whose eigenvalues are $\neq 1$. Representing $(v, B) \in \Gamma$ in the similar block form

$$(v, B) = \begin{pmatrix} B_2 & 0 & v_1 \\ 0 & B_1 & v_2 \\ 0 & 0 & 1 \end{pmatrix},$$

we immediately see that commutativity of Γ implies $v_2 = 0$. It follows that we have orthogonal direct sum decomposition

$$\mathbf{R}^m \simeq V_1 \oplus V_2, \quad \text{with } \ V_1 \simeq \mathbf{R}^k, \text{ and } V_2 \simeq \mathbf{R}^{m-k},$$

with V_1 invariant under Γ. Summarizing, we have,

Lemma 4.4.15 *With the above notation, V_1 is invariant under Γ.*

Proof of proposition 4.4.4 - IN view of lemma 4.4.9 and corollary 4.4.4 it remains to prove the proposition in the special case where Γ is abelian. We proceed by induction on m. Let $V_1 \simeq \mathbf{R}^k$ be as in lemma 4.4.15. Then V_1 is invariant under Γ and so we have homomorphism $\rho : \Gamma \to \Gamma' \subset E(k)$ by restriction to V_1. The kernel of ρ is a discrete subgroup of $O(m-k)$

and is therefore finite. If $\dim V_1 < m$ then we can apply the induction hypothesis to V_1 to obtain the decomposition $\Gamma' = \Gamma'_1 \times \Gamma'_2$ as stated in the proposition. We obtain the exact sequence

$$0 \longrightarrow \Delta \longrightarrow \Gamma \longrightarrow \Gamma'_1 \longrightarrow 0,$$

with Δ finite since Kerρ is finite. This sequence splits in view of the fact that Γ'_1 is a free abelian group and therefore $\Gamma = \Gamma' \times \Delta$. The required result follows easily assuming that $\dim V_1 < m$. If $V_1 = \mathbf{R}^m$, then Γ is a group of translations and the required result is clearly valid. ♣

Corollary 4.4.5 *An abelian discrete subgroup of $\Gamma \subset E(m)$ with compact quotient is the free abelian group of translations generated by a basis for $\mathbf{R}^m$.*

Proof of theorem 4.4.1 - Since maximality of the normal abelian subgroup $L \subset \Gamma$ of theorem 4.4.1 is clear, the proof of the direct implication of that theorem is complete. We prove the converse statement. Let $f_1, \cdots, f_m \in \Gamma$ be a basis for the free abelian group L and for $\gamma \in \Gamma$ define the matrix $A^\gamma = (A^\gamma_{jk})$ by

$$\gamma f_j \gamma^{-1} = \sum_{k=1}^{m} A^\gamma_{kj} f_k.$$

Since L is maximal, the mapping $\gamma \rightarrow A^\gamma$ is an embedding of Γ/L into $GL(m, \mathbf{Z})$, and Γ is realized as a subgroup of $\mathbf{Z}^m.GL(m, \mathbf{Z})$ (semi-direct product). Since Γ/L is finite we may assume its image lies in $O(m)$. This completes the proof of theorem 4.4.1. ♣

The finiteness statement 4.4.5 is a purely algebraic corollary of the theory we have developed. We consider the realization of the fundamental group Γ of a compact flat manifold of dimension m as a subgroup of $\mathbf{Z}^m.GL(m, \mathbf{Z})$.

Lemma 4.4.16 *$GL(m, \mathbf{Z})$ contains only finitely many conjugacy classes of finite subgroups.*

Proof - Let $\rho_3 : GL(m, \mathbf{Z}) \rightarrow GL(m, \mathbf{Z}/3)$ denote the reduction mod 3 map, and $\Gamma'(3)$ denote its kernel. For every finite subgroup $G' \subset GL(m, \mathbf{Z})$ let $G = \rho_3(G')$ and $\Gamma' = \rho_3^{-1}(G)$. Consider the the exact sequence

$$(1) \longrightarrow \Gamma'(3) \longrightarrow \Gamma' \longrightarrow G \longrightarrow (1).$$

Since ρ_3 is injective on finite subgroups (proposition 4.4.6), we have a splitting homomorphism $G \rightarrow \Gamma'$ and therefore $\Gamma' = \Gamma'(3).G$ (semi-direct product). It follows that the subgroups of $GL(m, \mathbf{Z})$ which map isomorphically onto G under ρ_3 form a single conjugacy class in Γ' by a theorem of algebra[3]. The finiteness assertion of the lemma follows from finiteness of

[3]This follows from the vanishing of the cohomology set $H^1(G; \Gamma(3))$ or that every crossed homomorphism is inner.

$GL(m, \mathbf{Z}/3)$. ♣

Proof of proposition 4.4.5 - It is a standard theorem in algebra that there are only only finitely many group extensions of the form (**??**) for a given finite group and action of G on L. In view of lemma 4.4.16 and corollary 4.4.3 imply the required finiteness. The second assertion follows from the first and corollary 4.4.3. ♣

4.4.6 The Laplacian and the Fundamental group

(THIS SUBSECTION IS NOT INCLUDED)

4.4.7 Braids and Configuration Spaces

(THIS SUBSECTION IS NOT INCLUDED)

4.5 Theorem of Hurewicz and Knot Groups

4.5.1 Fundamental Group and First Homology

We noted earlier that van Kampen's theorem is the analogue of the Mayer-Vietoris sequence for the fundamental group. To emphasize this similarity, we derive the basic relation between the fundamental group and the first homology group by relating van Kampen's theorem to the Mayer-Vietoris sequence. Let Z be a simplicial or cell complex and $\gamma \in \pi_1(Z, z)$. Then γ defines an element of $H_1(Z; \mathbf{Z})$ since γ has empty boundary and if γ and γ' are homotopic relative to ∂I, then the homotopy shows that γ and γ' define the same element in $H_1(Z; \mathbf{Z})$. The resulting map

$$\mathsf{h} : \pi_1(Z, z) \to H_1(Z; \mathbf{Z})$$

is called the *Hurewicz homomorphism*. Since $H_1(Z; \mathbf{Z})$ is abelian, $\mathrm{Ker}\mathsf{h}$ contains the commutator subgroup of $\pi_1(Z, z)$. As another application of van Kampen's theorem we show the kernel is in fact equal to the commutator subgroup of $\pi_1(X, x)$.

We have already noted the fundamental group and first homology groups of a finite graph are the free group and the free abelian group on the same number of generators. Since in this case a cycle is precisely a loop, it is easy to see that the Hurewicz map h is surjective, and the kernel is precisely the commutator subgroup. Clearly the same conclusion is valid for cell complexes of dimension one.

For a general simplicial or cell complex Z, a comparison of van Kampen's theorem and the Mayer-Vietoris sequence shows that the kernel of the Hurewicz map of $\pi_1(Z, z) \to H_1(Z; \mathbf{Z})$ is indeed the commutator subgroup. Although the theorem is true under very general conditions, we limit our proof to the case of finite simplicial or cell complexes and we prove it by induction on the number of simplices or cells. We already know the theorem to be true for cell complexes of dimension one. Therefore it suffices to investigate the relationship between fundamental groups and first homology groups of X, Y, Z and A where $Z = X \cup Y$ and $A = X \cap Y$ and all spaces in consideration are path connected. Path-connectedness is necessary to make van Kampen's theorem applicable. The main point in the proof is a simple algebraic observation. Let Γ, Γ' and Δ be groups, and $\imath : \Delta \to \Gamma$ and $\imath' : \Delta \to \Gamma'$ be homomorphisms. Let $\Gamma_1 = \Gamma/[\Gamma, \Gamma]$, Γ'_1 and Δ'_1 denote the abelianizations of Γ, Γ' and Δ respectively. Motivated by van Kampen's theorem, we define $\Gamma \star_\Delta \Gamma'$ as the quotient of the free product $\Gamma \star \Gamma'$ by the normal subgroup R generated by the elements

$$\imath(\delta)\imath'(\delta)^{-1}$$

as δ ranges over Δ. For $\gamma \in \Gamma$, $\delta \in \Delta$ etc. let $\bar{\gamma}$, $\bar{\delta}$ etc. denote their images in Γ_1, Δ_1 etc. The map $\imath$ induces a homomorphism $\bar{\imath} : \Delta_1 \to \Gamma_1$; and $\bar{\imath}'$ is similarly defined. The proof of the following lemma is straighforward:

Lemma 4.5.1 *With the above notation we have*

$$[\Gamma \star_\Delta \Gamma']_1 \simeq (\Gamma_1 \oplus \Gamma'_1)/N,$$

where $N \subset \Gamma_1 \oplus \Gamma'_1$ is generated by elements of the form

$$(\bar{\imath}(\bar{\delta}), \bar{\imath}'(\bar{\delta}))$$

Pictorially one can exhibit the lemma as the following commutative row-exact diagram:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & R & \longrightarrow & \Gamma \star \Gamma' & \longrightarrow & \Gamma \star_\Delta \Gamma' & \longrightarrow & 1 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & N & \longrightarrow & \Gamma_1 \oplus \Gamma'_1 & \xrightarrow{\rho} & [\Gamma \star_\Delta \Gamma']_1 & \longrightarrow & 0 \end{array}$$

It is clear that the homomorphism ρ is onto and the proof of the lemma is the verification of $\mathrm{Ker}\rho = N$ which is straightforward. Applying the lemma to $Z = X \cup Y$, with $A = X \cap Y$, $\Gamma = \pi_1(X, x)$, etc. one obtains

Theorem 4.5.1 (Hurewicz) *Let Z be a cell or simplicial complex. Then the kernel of the Hurewicz homomorphism $\mathsf{h} : \pi_1(Z, z) \to H_1(Z; \mathbf{Z})$ is the commutator subgroup $[\pi_1(Z, z), \pi_1(Z, z)]$ of $\pi_1(Z, z)$.*

As a first application of the theorem of Hurewicz we consider a knot $K \subset S^3$. Set $S_K = S^3 \setminus K$, and $\pi_1 = \pi_1(S_K; x)$. Since $H_1(S_K; \mathbf{Z}) \simeq \mathbf{Z}$, the theorem of Hurewicz gives the exact sequence

$$1 \longrightarrow \hat{\pi}_1 \longrightarrow \pi_1 \longrightarrow \mathbf{Z} \longrightarrow 0, \tag{4.5.1}$$

where $\hat{\pi}_1$ is the commutator subgroup of π_1.

Corollary 4.5.1 *The fundamental group $\pi_1 = \pi_1(S_K; x)$ of a knot complement admits of the semi-direct product decomposition*

$$\pi_1 \simeq \hat{\pi}_1.\mathbf{Z}.$$

Proof - Since $\mathbf{Z}$ is free, the exact sequence (4.5.1) admits of a splitting homomorphism $\chi : \mathbf{Z} \to \pi_1$. ♣

Example 4.5.1 We discuss Poincaré's example of a compact manifold with the same homology as S^3 but with a different fundamental group. Identify S^3 with the group of unit quaternions or equivalently with $SU(2)$ and recall that we have a surjective homomorphism $\rho : SU(2) \to SO(3)$ with $\mathrm{Ker}\rho = \pm I$. From example **??** of chapter 1 we know that $\mathcal{A}_5$, the alternating group on five letters, can be realized as a subgroup of $SO(3)$ by identifying it with the group of rotational symmetries of the icosahedron or dodecahedron. Let $\Gamma = \rho^{-1}(\mathcal{A}_5) \subset SU(2)$. The left translation action of Γ on $SU(2)$ is orientation preserving and the quotient space $M = \Gamma\backslash SU(2)$ is an orientable manifold. In particular, $H_3(M; \mathbf{Z}) \simeq \mathbf{Z}$. For a group G let $\hat{G}$ denote the its commutator subgroup. Then

$$\rho(\hat{\Gamma}) = \hat{\mathcal{A}}_5 = \mathcal{A}_5, \tag{4.5.2}$$

and consequently either $\hat{\Gamma} = \Gamma$ or $\hat{\Gamma}$ has index two in Γ. We show that the latter alternative is not possible. In fact, if it were, then ρ would map $\hat{\Gamma}$ bijectively onto $\mathcal{A}_5$ and consequently ρ^{-1} would be a splitting homomorphism $\mathcal{A}_5 \to \Gamma$. Since $\pm I$ is the center of $SU(2)$ this gives the direct product decomposition $\Gamma \simeq \mathcal{A}_5 \times (\mathbf{Z}/2)$. Consider the subgroup

$$\{e, (12)(34), (13)(24), (14)(23)\}$$

of $\mathcal{A}_5$. Every non-identity element of this subgroup is realized as a rotation by π in $SO(3)$ and their pre-images in $SU(2)$ are conjugate to the quaternion $\mathbf{i}$ which has order 4. Therefore the decomposition $\Gamma \simeq \mathcal{A}_5 \times (\mathbf{Z}/2)$ is not possible and we necessarily have $\hat{\Gamma} = \Gamma$. By the theorem of Hurewicz

$$H_1(M; \mathbf{Z}) = 0. \tag{4.5.3}$$

It is an immediate consequence of Poincaré duality (which will be discussed in chapter 6) that $H_2(M; \mathbf{Z}) = 0$. Anticipating this result we have exhibited an orientable manifold with the same homology as S^3 but not the same homotopy type. A compact manifold with the same homology as S^n is called a *homology n-sphere*. The group $\Gamma = \rho^{-1}(\mathcal{A}_5) \subset SU(2)$ is called the *binary icosahedral* group. This example has a very special character. For no other n one can find a compact homogeneous space with the same homology as S^n but with a different homotopy type (see [Br]). The binary icosahedral group is unique in the sense that it is the only finite perfect group which admits of an irreducible complex representation which acts without fixed points on the unit sphere. The representation is also unique and is the restriction of the natural representation of $SU(2)$ as above. The proof of these uniqueness assertions is rather lengthy, and the interested reader is referred to [Wolf]. In the following examples we construct many other homology spheres based on a different approach. ♠

Example 4.5.2 We can use our knowledge of torus knots to construct examples of homology 3-spheres (other than S^3.) Let $K \subset S^3$ be a knot and T_K be a small open tubular neighborhood of K. Consider two copies M_1, M_2 of $S_K = S^3 \setminus T_K$ and let $\varphi : \partial M_1 \to \partial M_2$ be a diffeomorphism. Let us compute the homology of $M = M_1 \sharp_\varphi M_2$ which is a compact manifold (see remark (XXX) of chapter 1). The Mayer-Vietoris sequence yields (coefficient group $\mathbf{Z}$ is omitted)

$$\mathbf{Z} \oplus \mathbf{Z} = H_1(\partial S_K) \xrightarrow{A} H_1(M_1) \oplus H_1(M_2) \longrightarrow H_1(M) \longrightarrow 0.$$

Here $H_1(M_1) \oplus H_1(M_2) \simeq \mathbf{Z} \oplus \mathbf{Z}$ and the integral 2×2 matrix A depends on the map φ. If $\det A = \pm 1$, then $H_1(M; \mathbf{Z}) = 0$. Recall that a meridian and a longitude represent generators of $H_1(\partial T_K; \mathbf{Z})$. There is a diffeomorphism ψ of ∂T_K interchanging a given meridian and a given longitude. In fact, we can identify ∂T_K with $\mathbf{R}^2/\mathbf{Z}^2$ such that the given meridian and longitude are represented by the line segments joining the origin to $(1, 0)$ and $(0, 1)$ respectively. Then the diffeomorphism ψ is induced by the orientation reversing linear map $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Now let the diffeomorphism φ be the map ψ. Then $A = \psi$ and $H_1(M; \mathbf{Z}) = 0$. Since ψ is orientation reversing, the $M_1 \sharp_\varphi M_2$ is an orientable manifold M. From vanishing of $H_2(S_K; \mathbf{Z})$, we obtain $H_2(M; \mathbf{Z}) = 0$ by an application of the Mayer-Vietoris sequence. Consequently, M is a homology 3-sphere. In the next example we discuss the fundamental group of M. ♠

Example 4.5.3 We continue with example 4.5.2 and specialize to the case $K = K_{2,3}$ is the trefoil knot. We want to apply van Kampen's theorem to obtain a presentation of the fundamental group of M. The trefoil knot corresponds to the case $n = 1$ in example 4.2.17. In terms of the Wirtinger presentation we have $\xi = \xi_2\xi_3\xi_2$ and $\eta = \xi_2\xi_3$. Then, by a straightforward calculation, a meridian and a longitude are represented by

$$\mu = \xi_2 = \eta^{-1}\xi, \quad \lambda = \xi_3\xi_2\xi_1\xi_2^{-3} = \xi^2(\xi^{-1}\eta)^6$$

respectively. Therefore $\pi_1(M, x)$ is is the group on four generators ξ, η, α, β subject to the relations

$$\xi^2 = \eta^3, \ \alpha^2 = \beta^3, \ \xi^2(\xi^{-1}\eta)^6 = \beta^{-1}\alpha, \ \alpha^2(\alpha^{-1}\beta)^6 = \eta^{-1}\xi.$$

We show that $\pi_1(M, x) \neq 0$ which implies that M is (homotopically) distinct from S^3. Let $\mathcal{A}_m$ denote the alternating group on m letters. It is straightforward to check that the mapping

$$\xi, \alpha \longrightarrow (34)(67), \quad \eta, \beta \longrightarrow (123)(456)$$

defines a homomorphism of $\pi_1(M, x)$ into $\mathcal{A}_7$ and therefore $\pi_1(M, x) \neq 0$. ♠

Remark 4.5.1 The construction of the preceding example generalizes to give infinitely many distinct homology 3-spheres. In fact using the knot $K_{2,2n+1}$ (see example 4.2.17) in exactly the same way we obtain a manifold which we denote by M_n. Its fundamental group is defined by four generators ξ, η, α, β subject to the relations

$$\xi^2 = \eta^{2n+1},\ \alpha^2 = \beta^{2n+1},\ \xi^2(\xi^{-1}\eta^n)^{4n+2} = \beta^{-n}\alpha,\ \alpha^2(\alpha^{-1}\beta^n)^{4n+2} = \eta^{-n}\xi.$$

Then the assignment

$$\begin{aligned} \xi, \alpha &\longrightarrow (2n+1\ 2n+2)(4n+2\ 4n+3), \\ \eta, \beta &\longrightarrow (12\cdots 2n+1)(2n+2\ 2n+3\cdots 4n+2). \end{aligned}$$

defines a homomorphism of $\pi_1(M_n, x)$ into $\mathcal{A}_{4n+3}$. Furthermore, one shows that if $(4n+3)! < 2m+1$, and $2m+1$ is a prime, then any homomorphism of $\pi_1(M_m, x)$ into $\mathcal{S}_{4n+3} \supset \mathcal{A}_{4n+3}$ is necessarily trivial and therefore M_n and M_m are homotopically distinct. The reader is referred to [Gro] for a detailed treatment. ♡

4.5.2 Theorem of Hurewicz and Alexander Polynomials

In this subsection we use the the theorem of Hurewicz and the theory of covering spaces to obtain an invariant of a knot which is less precise but more manageable than the fundamental group of the knot complement. Let $K \subset \mathbf{R}^3$, $S_K = S^3 \setminus K$ and $\tilde{S}_K = \tilde{S}$ be its universal cover. Fix a base point $x \in S_K$ and let $\pi_1 = \pi_1(S_K, x)$ and $\hat{\pi}_1$ denote the commutator subgroup of π_1. According to the theorem of Hurewicz $\pi_1/\hat{\pi}_1 \simeq \mathbf{Z}$ since $H_1(S_K;\mathbf{Z}) \simeq \mathbf{Z}$ for every knot K. Regular covering spaces of S_K are in one to one correspondence with normal subgroups of $\pi' \subset \pi_1$. We denote by $p_{ab} : \tilde{S}_{K,ab} \to S_K$ the regular covering space corresponding to the normal subgroup $\hat{\pi}_1$. The group of covering transformations of the regular covering space corresponding to the normal subsgroup π' is π/π'. The maximal regular covering space of S_K with abelian group of covering transformations is $p_{ab} : \tilde{S}_{K,ab} \to S_K$. Any other regular covering of S_K with abelian group of covering transformations is a quotients $\tilde{S}_{K,ab}$ by a subgroup of $\mathbf{Z}$. Therefore all such coverings have cyclic groups of covering transformations, namely, $\mathbf{Z}$ or $\mathbf{Z}/(m)$. We denote the covering space corresponding to the latter group by $\tilde{S}_{K,m}$. The above can be summarized in the following diagram:

$$\begin{array}{ccccc} \tilde{S}_K & \to & \tilde{S}_{K,ab} & \to & \tilde{S}_{K,m} \\ & \searrow & \downarrow & \swarrow & \\ & & S_K & & \end{array}$$

One can invoke the theorem of Hurewicz to gain some understanding of $\tilde{S}_{K,ab}$ for a presentation of the fundamental group of S_K by generators and relations. Since $\tilde{S}_{K,ab}$ is the quotient of $\tilde{S}_K$ by $\hat{\pi}_1$, the first homology of $\tilde{S}_{K,ab}$ is $H_1(\tilde{S}_{K,ab};\mathbf{Z}) \simeq \hat{\pi}_1/[\hat{\pi}_1,\hat{\pi}_1]$. It is convenient to look at $H_1(\tilde{S}_{K,ab};\mathbf{Z})$ as a module over the ring of finite Laurent series in one variable and with integer coefficients, namely, $\mathcal{R} \stackrel{\text{def}}{=} \mathbf{Z}[t,t^{-1}]$. Note that $\mathcal{R}$ is simply the group algebra of $\pi_1/\hat{\pi}_1 \simeq \mathbf{Z}$ with $\mathbf{Z}$ coefficients. It may be judicious to distinguish between $\mathbf{Z}$ as a coefficient group and as the group $\pi_1/\hat{\pi}_1$. For this reason we write J for the integers regarded as $\pi_1/\hat{\pi}_1$, so that $\mathcal{R} = \mathbf{Z}[J]$. We represent J multiplicatively as $\{t^n\}_{n\in\mathbf{Z}}$. Let us first describe the $\mathcal{R}$-module structure of $H_1(\tilde{S}_{K,ab})$.

Let $\epsilon \in \pi_1(S_K)$ be such that it maps to $t = 1 \in H_1(S_K;\mathbf{Z}) = J$ under the Hurewicz map. (The equality $t = 1$ means we have fixed one of the two isomorphisms $J \simeq \mathbf{Z}$.) For $\alpha \in \hat{\pi}_1$ define $t.\alpha = \epsilon^{-1}\alpha\epsilon \in \hat{\pi}_1$. $t.\alpha$, thus defined, depends on the choice of ϵ, however, the image of $t.\alpha$ in $\hat{\pi}_1/[\hat{\pi}_1,\hat{\pi}_1]$ is independent of the choice ϵ since two such choices differ by an element of $\hat{\pi}_1$. Hence we have endowed $H_1(\tilde{S}_{K,ab};\mathbf{Z}) \simeq \hat{\pi}_1/[\hat{\pi}_1,\hat{\pi}_1]$ with the structure of an $\mathcal{R}$-module. The following simple lemma is useful in understanding the structure of $H_1(\tilde{S}_{K,ab};\mathbf{Z})$:

Lemma 4.5.2 *There is a set of generators $\{\gamma;\beta_1,...,\beta_k\}$ for $\pi_1(S_K)$ with $\beta_j \in \hat{\pi}_1$, and $\gamma \in \pi_1$.*

Proof - The assertion follows from the fact that $\pi_1/\hat{\pi}_1 \simeq \mathbf{Z}$. ♣

Remark 4.5.2 Lemma 4.5.2 does not imply that $\hat{\pi}_1$ is generated by $\{\beta_1,...,\beta_k\}$. As we shall see later $\hat{\pi}_1$ is generally not finitely generated, however, lemma 4.5.2 says that as an $\mathcal{R}$-module, $\hat{\pi}_1$ is finitely generated. ♡

The key idea is to choose a set of generators for π_1 as specified by Lemma 4.5.2, and represent $H_1(\tilde{S}_{K,ab};\mathbf{Z})$ as the cokernel of an $\mathcal{R}$-linear operator A on $\mathcal{R}\oplus...\oplus\mathcal{R}$. Notice that in this description of $H_1(\tilde{S}_{K,ab};\mathbf{Z})$, it is realized as a module over $\mathbf{Z}[t,t^{-1}]$ which is a richer structure than just an abelian group. We refer to A as an *Alexander matrix* of the knot K. Of course A is not uniquely determined by K. The determinant of A (if non-zero) is called the *Alexander polynomial* of the knot K and will be denoted by $\Delta_K(t)$. More generally, if A is an $n\times n$ matrix then the ideal generated by $(n-k)\times(n-k)$ minors of A is called the k^{th} Alexander ideal of the knot and is denoted by $\mathcal{I}_k^K$. For $\Delta_K(t)$ and $\mathcal{I}_k^K$ to be a geometrically meaningful it is necessary to show that they only depend on the cokernel of A and not A itself since the cokernel is $H_1(\tilde{S}_{K,ab};\mathbf{Z})$. For the proof of this basic fact see corollary 4.7.1 is the final subsection of this chapter.

The Alexander polynomial is defined up to multiplication by a unit of $\mathbf{Z}[t,t^{-1}]$. The units of this ring are $\{\pm t^n\}$. Therefore the *degree of the Alexander polynomial* is defined as the difference of the degrees of the highest and lowest powers of t that appear in $\Delta_K(t)$. Similarly, the *constant term* of the Alexander polynomial is the coefficient of the lowest power of t that appears in $\Delta_K(t)$. *Unless stated to the contrary, we normalize the Alexander polynomial so that the lowest power of t appearing in $\Delta_K(t)$ is zero.* This makes $\Delta_K(t)$ a polynomial and uniquely determines it up to multiplication by ± 1. The fact that the Alexander polynomial depends only on K and not the matrix A has the following important consequence:

Corollary 4.5.2 *Let N be the degree of the Alexander polynomial, then*

$$t^N \Delta_K(t^{-1}) = \Delta_K(t).$$

Proof - In the group algebra $\mathcal{R} = \mathbf{Z}[J]$ we have the freedom of a choice of the generator for $J \simeq \mathbf{Z}$. Thus replacing the generator t by t^{-1} induces an automorphism of the $\mathbf{R}$-module Coker(A), and det(A) remains unchanged up to multiplication by a unit of $\mathcal{R}$, i.e., $t^N \Delta_K(t^{-1}) = \pm\Delta_K(t)$. Substituting $t = 1$ we see that the sign $\pm$ is necessarily $+$. ♣

Example 4.5.4 Consider the torus knot $K = K_{m,n}$. The fundamental group of S_K is isomorphic to the quotient of the free group on two generator x and y by the relation $x^m = y^n$. Let us compute a set of generators as specified by lemma 4.5.2. The images of x and y in $\pi_1/\hat{\pi}_1$ are nt and mt where t is ± 1. Let r and s be integers such that $rm + sn = 1$. (Recall that in the definition of the torus knot $K_{m,n}$ we assumed that m and n are relatively prime positive integers.) Then $x^s y^r$ maps to $t \in \mathbf{Z}$. We set $\gamma = x^s y^r \in \pi_1$. Let $\beta_1 = x\gamma^{-n}$ and $\beta_2 = y\gamma^{-m}$, then $\{\gamma; \beta_1, \beta_2\}$ is a set of generators of the required form for $\pi_1(S_K)$. The relations are

$$(\beta_1\gamma^n)^m = (\beta_2\gamma^m)^n, \qquad \gamma = (\beta_1\gamma^n)^s(\beta_2\gamma^m)^r.$$

Note that a word in β_i's and γ lies in $\hat{\pi}_1$ if and only if the sum of the exponents of γ is zero. Therefore we multiply the first relation by γ^{-mn} to obtain a relation in $\hat{\pi}_1$. Assuming $r < 0$ (one of r and s is negative) we rewrite the second relation in the form $\gamma(\beta\gamma^m)^{-r} = (\beta_1\gamma^n)^s$. Multiplying both sides of this equation by γ^{-ns} we obtain an equation where both sides are in $\hat{\pi}_1$. Abelianizing the relations in their new form yields (recall $r < 0$):

$$\begin{array}{rcl} (t^n + t^{2n} + ... + t^{mn})\beta_1 - (t^m + t^{2m} + ... + t^{mn})\beta_2 & = & 0; \\ (t^n + t^{2n} + ... + t^{ns})\beta_1 - (t^m + t^{2m} + ... + t^{-mr})\beta_2 & = & 0. \end{array}$$

This means that $\hat{\pi}_1/[\hat{\pi}_1, \hat{\pi}_1]$ is the cokernel of the $\mathcal{R}$-linear mapping of $\mathcal{R}\oplus\mathcal{R}$ into itself given by the matrix

$$A = \begin{pmatrix} t^n + t^{2n} + ... + t^{mn} & -t^m - t^{2m} - ... - t^{mn} \\ t^n + t^{2n} + ... + t^{ns} & -t^m - t^{2m} - ... - t^{-mr} \end{pmatrix}$$

To determine the structure of this module we try to diagonalize A by transformations of the form $A \rightarrow PAQ$ where P and Q are matrices with entries from $\mathcal{R}$ and determinant a unit in $\mathcal{R}$. Since $\mathcal{R}$ is not a principal ideal domain, there is no guarantee of success and in fact there are knots for which this is not possible. However, in this particular example, after a simple calculation we obtain the following canonical form for A:

$$\begin{pmatrix} 1 & 0 \\ 0 & \Delta_K(t) \end{pmatrix}, \quad \text{where} \ \ \Delta_K(t) = \frac{(1-t)(1-t^{mn})}{(1-t^m)(1-t^n)}.$$

Of course, $\Delta_K(t)$ is, up to a unit in $\mathcal{R}$, the determinant of the matrix A. The fact that $\Delta_K(t)$ is a polynomial with integer coefficients in t and constant term 1 is easily proved by looking at the roots of the denominator. It is the Alexander polynomial of the torus knot $K = K_{m,n}$. From the canonical form of the matrix A and the expansion $\Delta_K(t) = t^{(m-1)(n-1)} + \cdots + 1$, it follows that as a module over $\mathbf{Z}[t, t^{-1}]$, $H_1(\tilde{S}_{K,ab}; \mathbf{Z})$ is isomorphic to $\mathbf{Z}^N$ where $N = (m-1)(n-1)$, and $1, t, \cdots, t^{N-1}$ a $\mathbf{Z}$ basis exhibiting the action of t. ♠

Exercise 4.5.1 *With the presentation of the group of the figure 8 knot K given by four generators $\{\xi_1, \xi_2, \xi_3, \xi_4\}$ subject to the relations*

$$\xi_3\xi_2 = \xi_1\xi_3, \quad \xi_4\xi_2 = \xi_3\xi_4, \quad \xi_3\xi_1 = \xi_1\xi_4,$$

show that we can set $\tau = \xi_3$, $\beta_1 = \xi_1^{-1}\xi_3$, $\beta_2 = \xi_3\xi_2^{-1}$ and $\beta_3 = \xi_4^{-1}\xi_3$ in the notation of lemma 4.5.2. Rewriting the relations in terms of β's and t, show that the Alexander matrix of K is the $\mathcal{R}$-linear map of $\mathcal{R} \oplus \mathcal{R} \oplus \mathcal{R}$ into itself given by the matrix

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1-(1/t) \\ -1+(1/t) & 0 & -1/t \end{pmatrix}.$$

Deduce that the Alexander polynomial of the figure 8 knot is $t^2 - 3t + 1$, and $H_1(\tilde{S}_{K,ab}; \mathbf{Z}) \simeq \mathbf{Z}^2$. What is the action of t?

The examples of Alexander polynomial given above (torus knots and figure 8 knot) are particularly simple. While obtaining the Wirtinger presentation is essentially a mechanical procedure, experimenting with the computation of the Alexander polynomial of more complex knots will convince the reader that obtaining the matrix A may require some degree of ingenuity. As such the calculation of the matrix A is not sufficiently simplified so that that the computation of the Alexander polynomial can be implemented by a straightforward algorithm. The problem of an easily implementable algorithm for computing the Alexander polynomial of a knot is solved by invoking our knowledge of covering spaces which reduces the problem to a straightforward calculation. This method was abstracted into an algebraic formalism called *Free Differential Calculus* (see e.g. [CrF]). It is both more illuminating and expedient for our immediate goal to simply explain the geometric content of this algebraic formalism and understand its application to the calculation of Alexander polynomials rather than develop the algebraic machinery formally.

It is convenient to remove a tubular neighborhood of K from S^3 so that S_K is a compact manifold with boundary and therefore a finite cell complex. It is no loss of generality to assume that S_K has only one cell of dimension zero, namely the base point $x \in S_K$. As usual we denote by S_K^j the j-skeleton of S_K. The group $J \simeq \mathbf{Z}$ acts on $\tilde{S}_{K,ab}$. We endow $\tilde{S}_{K,ab}$ with the structure of regular a cell complex so that the cell structure is invariant under the action of the group of covering transformations J. This is done by defining the zero cells of $\tilde{S}_{K,ab}$ to be $p_{ab}^{-1}(x)$. Then lifting the 1-cells to all points $y \in p_{ab}^{-1}(x)$ to obtain $\tilde{S}^1_{K,ab}$. Using the homotopy lifting property, we then lift 2-cells etc. In the computation of π_1 or H_1, cells of dimension three play no role.

A loop γ representing an element of π_1 also represents an element of $\hat{\pi}_1$ if and only if its lift to $\tilde{S}_{K,ab}$ (with base point y) is a loop, and this requirement is independent of the choice of the base point $y \in \tilde{S}_{K,ab}$ with $p_{ab}(y) = x$. Fix $y_\circ \in p_{ab}^{-1}(x)$. For $\gamma \in \pi_1$ we denote by $\tilde{\gamma}$ its lift to $\tilde{S}_{K,ab}$ with base point $y_\circ$. For $t \in \mathbf{Z}$, we denote by $t.\tilde{\xi}$ the lift of ξ with base point $t.y_\circ$. Denote the canonical homomorphism $\pi_1 \rightarrow \mathbf{Z}$ by τ, and the image of ξ in $\mathbf{Z}$ by τ_ξ. It is clear that for $\xi, \delta \in \pi_1$ we have

$$\widetilde{\eta\xi} = \tau_\xi.\tilde{\eta} \bullet \tilde{\xi}, \tag{4.5.4}$$

where we have used $\bullet$ to denote the composition of paths in $\tilde{S}_{K,ab}$. Equation (4.5.4) is the fundamental geometric fact which reduces the calculation of the Alexander polynomial(s) to a straightforward algorithm. It is also the basis for the algebraic development of free differential calculus.

The Wirtinger presentation of the fundamental group is often convenient for computing the Alexander polynomial or Alexander ideals. The following lemma plays an important role in the development of the algorithm for this purpose and its proof is almost immediate:

Lemma 4.5.3 *The Wirtinger presentation of the fundamental group S_K has the following properties:*

1. *Every generator ξ_j of the Wirtinger presentation is mapped to a generator of J by τ;*

2. *Every relation of the Wirtinger presentation lifts to a loop in $\tilde{S}^1_{K,ab}$.*

Recall that $H_1(\tilde{S}^1_{K,ab}, \tilde{S}^\circ_{K,ab}; \mathbf{Z})$ is the free abelian group on all the lifts $t.\tilde{\xi}_j$, $t \in J \simeq \mathbf{Z}$. Therefore it is the free $\mathcal{R}$-module on generators $\xi_1, \cdots, \xi_n$ (in the Wirtinger presentation). Let γ_j denote the image of ξ_j in $H_1(\tilde{S}^1_{K,ab}, \tilde{S}^\circ_{K,ab}; \mathbf{Z})$. Note that $t = \tau_{\xi_j} \in H_1(S_K; \mathbf{Z}) = J$ independently of of j, and for all j, k,

$$\tau_{\xi_j}(\xi_k) = t.\gamma_k. \tag{4.5.5}$$

Connectedness of $\tilde{S}^1_{K,ab}$ implies that the long exact sequence reduces to

$$\begin{array}{ccccccc} 0 & \longrightarrow & H_1(\tilde{S}^1_{K,ab}; \mathbf{Z}) & \longrightarrow & H_1(\tilde{S}^1_{K,ab}, \tilde{S}^\circ_{K,ab}; \mathbf{Z}) & \xrightarrow{\delta_1} & \\ & & H_\circ(\tilde{S}^\circ_{K,ab}; \mathbf{Z}) & \longrightarrow & H_\circ(\tilde{S}^1_{K,ab}; \mathbf{Z}) & \longrightarrow & 0. \end{array} \tag{4.5.6}$$

Kerδ_1 is precisely $H_1(\tilde{S}^1_{K,ab}; \mathbf{Z})$ (identified with its image in $H_1(\tilde{S}^1_{K,ab}, \tilde{S}^\circ_{K,ab}; \mathbf{Z})$). Observe that the lifts of ξ_i and ξ_j with base point $y_\circ$ are two distinct curves joining $y_\circ$ to $t.y_\circ$. With this observation one proves easily that a basis for Ker$\delta_1 = H_1(\tilde{S}^1_{K,ab}; \mathbf{Z})$ as an $\mathcal{R}$-module is given by

$$\gamma_1 - \gamma_2,\ \gamma_2 - \gamma_3, \cdots, \gamma_{n-1} - \gamma_n. \tag{4.5.7}$$

Note that Kerδ_1 is precisely the abelian group generated by the loops in $H_1(\tilde{S}^1_{K,ab}, \tilde{S}^\circ_{K,ab}; \mathbf{Z})$. Therefore by lemma 4.5.3 the relations in the Wirtinger presentation are represented by homology classes in Kerδ_1. From the calculation of homology of a space given as a cell complex we know that $H_1(\tilde{S}_{K,ab}; \mathbf{Z})$ is the quotient

$$H_1(\tilde{S}^1_{K,ab}; \mathbf{Z})/N,$$

where $N = \delta_2(H_2(\tilde{S}^2_{K,ab}, \tilde{S}^1_{K,ab}; \mathbf{Z}))$ is the set of relations in homology. Now N is easily computable from (4.5.4) ($\bullet$ becomes $+$ in homology) and the Wirtinger presentation, and therefore we can effectively calculate the matrix A. To see this point clearly we work out an example.

Example 4.5.5 Consider the knot in Figure 5.1 (generally known as 6_1). In the Wirtinger presentation the generators are $\xi_1, \cdots, \xi_6$ and the relations are

$$\xi_1\xi_5 = \xi_6\xi_1, \quad \xi_2\xi_5 = \xi_5\xi_1, \quad \xi_2\xi_4 = \xi_5\xi_2, \quad \xi_4\xi_2 = \xi_3\xi_4, \quad \xi_6\xi_3 = \xi_3\xi_1.$$

Lifting the first relation to $\tilde{S}^1_{K,ab}$ and using (4.5.4) we obtain

$$\tau_{\xi_5}\tilde{\xi}_1 \bullet \tilde{\xi}_5 = \tau_{\xi_1}\tilde{\xi}_6 \bullet \tilde{\xi}_1.$$

Abelianizing this relation and expressing it in terms of the basis $\gamma_i - \gamma_{i+1}$, we get

$$0 = (\gamma_1 - \gamma_5) - t \cdot (\gamma_1 - \gamma_6) = \sum_{i=1}^{4} (1-t) \cdot (\gamma_i - \gamma_{i+1}) - t \cdot (\gamma_5 - \gamma_6).$$

Repeating the process for the remaining relations we obtain the matrix

$$A = \begin{pmatrix} 1-t & 1-t & 1-t & 1-t & -t \\ 1 & 1-t & 1-t & 1-t & 0 \\ 0 & 1-t & 1-t & -t & 0 \\ 0 & 1 & 1-t & 0 & 0 \\ 1 & 1 & t & t & t \end{pmatrix}$$

Computing the determinant of A and normalizing by dividing by $-t^2$ we obtain the Alexander polynomial $2t^2 - 5t + 2$ of the knot in question. ♠

Exercise 4.5.2 *Compute the Alexander matrices of the following knots as shown in Figure 5.2 and show the Alexander polynomials are as indicated:*

1. *(5_2 Knot):* $\Delta_K(t) = 2t^2 - 3t + 2$;
2. *(7_3 Knot):* $\Delta_K(t) = 2t^4 - 3t^3 + 3t^2 - 3t + 2$;
3. *(7_7 Knot):* $\Delta_K(t) = t^4 - 5t^3 + 9t^2 - 5t + 1$.

Exercise 4.5.3 *Show that the Alexander polynomial of the knot shown in Figure 5.3 is* $2t^2 - 5t + 2$ *(just as for the knot in example 4.5.5). However, the ideal generated by* $2t - 1$ *and* $t-2$ *is an Alexander ideal for the knot in this exercise, while the Alexander ideals, except for* $(\Delta_K) = (\det(A))$, *for the knot in example 4.5.5 are the unit ideal.*

Remark 4.5.3 It is not difficult to construct nontrivial knots K with Alexander polynomial $\Delta_K(t) = 1$. A discussion of this fact is postponed to chapter 6 since it is most easily understood in the context of Seifert surfaces which are introduced in that chapter. ♡

4.5.3 Torus Knots

In our development of knot theory so far, we relied mainly on algebraic techniques related to the theorem of Hurewicz. Deeper insight into the structure of knots requires the introduction geometric ideas and especially the notion of a Seifert surface. The application of the concept of a Seifert surface to knot theory requires the notion of linking number which will be discussed in the context of cohomology in chapter 6. In this subsection we study some geometric properties of torus knots which allow us to introduce Seifert surfaces in an analytical manner in the context of torus knots. These knots are very special and considerably simpler than most knots. Nevertheless, it is useful to understand certain concepts in this simpler context.

Let $m > 1$ and $n > 1$ be relatively prime positive integers and $K_{m,n}$ be the torus knot which may be defined as the intersection of the locus $z_1^n - z_2^m = 0$ in $\mathbf{C}^2$ with the the sphere $|z_1|^2 + |z_2|^2 = 1$. We set $f(z) = z_1^n - z_2^m$ and notice that we have the mapping

$$\varphi = \varphi_{m,n} : S_K \longrightarrow S^1, \qquad \varphi_{m,n}(z_1, z_2) = \frac{f(z)}{|f(z)|}.$$

Our goal is to prove that for the torus knot $K = K_{m,n}$, $\varphi_{m,n} : S_K \to S^1$ is actually a fibre bundle with a typical fibre M_K a connected orientable surface and determine the structure of M_K. Let $M_{K,\theta} = \varphi^{-1}(e^{i\theta})$ denote the fibre over the point $e^{i\theta}$. We refer to a fibre $M_{K,\theta} = \varphi^{-1}(e^{i\theta})$ as a *Seifert surface.*

First we introduce some notation. Let $<,>$ denote the standard Hermitian inner product on $\mathbf{C}^2$ so that its real part $\Re <,>$ is the standard inner product on $\mathbf{R}^4$. To relate the the derivative of the map φ to that of the polynomial f, we let ∇ denote the complex gradient operator, that is

$$\nabla = (\frac{\partial}{\partial x_1} - i\frac{\partial}{\partial y_1}, \frac{\partial}{\partial x_2} - i\frac{\partial}{\partial y_2}),$$

where $z_j = x_j + iy_j$. Set $\varphi(z) = e^{i\theta(z)}$ where θ is a many-valued but real function of z, and let $D = (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial y_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial y_2})$. Then it is a simple calculation that

$$iD\theta = \frac{Df}{f} - \frac{D|f|}{|f|} = \frac{1}{2}(\frac{Df}{f} - \frac{D\bar{f}}{\bar{f}}).$$

Therefore

$$D\theta = (\frac{\partial v}{\partial x_1}, \frac{\partial v}{\partial y_1}, \frac{\partial v}{\partial x_2}, \frac{\partial v}{\partial y_2}), \tag{4.5.8}$$

where $\log f = u + iv$. In view of the Cauchy Riemann equations we have

$$\nabla \log f = 2i(\frac{\partial v}{\partial x_1} + i\frac{\partial v}{\partial y_1}, \frac{\partial v}{\partial x_2} + i\frac{\partial v}{\partial y_2}), \tag{4.5.9}$$

Now we can prove

Lemma 4.5.4 *Let $K = K_{m,n}$ be the torus knot of type (m, n) as described above. Then for every $e^{i\theta} \in S^1$, $M_{K,\theta}$ is a smooth surface.*

Proof - It suffices to show that $\varphi_{m,n}$ has no critical point in the complement of the knot $K_{m,n}$. Critical points of φ are the points where $D\theta(z) = \lambda(x_1, y_1, x_2, y_2)$, for some real λ. In view of (4.5.8) and (4.5.9) this condition is equivalent to

$$i\overline{\nabla \log f(z_1, z_2)} = \lambda(z_1, z_2), \tag{4.5.10}$$

for some real λ. To prove that φ has no critical point, we show that (4.5.10) can be fulfilled only for λ lying on the positive purely imaginary axis. Let $z_j = \rho_j e^{i\theta_j}$ and $z_1^n - z_2^m = Re^{i\Delta}$. Then, in view of (4.5.10), at a critical point we have

$$\Delta + \frac{\pi}{2} - (n-1)\theta_1 = \arg(\lambda) + \theta_1, \text{ and } \Delta + \frac{3\pi}{2} - (m-1)\theta_2 = \arg(\lambda) + \theta_2.$$

Therefore $n\theta_1$ and $m\theta_2$ differ by π implying that z_1^n and $-z_2^m$ have identical arguments. Consequently $\Delta = n\theta_1 = -m\theta_2$. Thus $\arg(\lambda) = \frac{\pi}{2}$, i.e., λ lies on the positive purely imaginary axis as desired. ♣

Let ζ be the vector field on S_K defined by $\zeta = (\frac{iz_1}{n}, \frac{iz_2}{m})$, then

$$\Re < \zeta, (z_1, z_2) >= 0,$$

and consequently ζ is tangent to the sphere S^3. Let $p(t)$ denote a solution curve to ζ. The differential equation $\frac{dp(t)}{dt} = \zeta$ being linear, is easily integrated to yield the solution

$$p(t) = (Ae^{\frac{it}{n}}, Be^{\frac{it}{m}}),$$

where A and B are arbitrary complex numbers. For $|A|^2 + |B|^2 = 1$, the solution curve $p(t)$ lies on S^3 since ζ is tangent to it. If in addition $A^n = B^m$ then $p(t)$ moves along the knot $K_{m,n}$ and otherwise it is disjoint from it.

In view of (4.5.8) and (4.5.9) we have

$$\frac{d\theta(p(t))}{dt} = \frac{1}{2}\Im(\nabla \log f \cdot \zeta) = 1. \tag{4.5.11}$$

The solution to (4.5.11) is

$$\theta(p(t)) = \theta_\circ + t, \tag{4.5.12}$$

with $\theta_\circ$ referring the initial surface. The solution (4.5.12) shows that the one parameter family $p(t)$ preserves the fibres of $\varphi_{m,n} : S_K \to S^1$ by mapping each fibre to another. We have

Lemma 4.5.5 *$\varphi_{m,n} : S_K \to S^1$ is a fibration, and its fibres $M_{K,\theta}$ are orientable.*

Proof - The existence of fibre preserving one parameter family $p(t)$ implies local triviality of $\varphi_{m,n} : S_K \to S^1$ and that it is a fibration. It follows from (4.5.11) that ζ is everywhere transverse to the fibres of $\varphi_{m,n} : S_K \to S^1$ and therefore the fibres are orientable. ♣

Lemma 4.5.6 *Every fibre $M_{K,\theta}$ is connected.*

Proof - Since $\varphi : S_K \to S^1$ is a fibre bundle it suffices to show that one fibre is connected. We show $M_{K,\circ}$ (the fibre over $1 \in S^1$) is connected. Let $z = (z_1, z_2) = (r_1 e^{i\alpha_1}, r_2 e^{i\alpha_2})$, $w = (w_1, w_2) = (R_1 e^{i\beta_1}, R_2 e^{i\beta_2})$ be points in $M_{K,\circ}$. Then $r_1^2 + r_2^2 = 1$ and $R_1^2 + R_2^2 = 1$ and so we set

$$r_1 = \cos\phi,\ r_2 = \sin\phi, \quad R_1 = \cos\psi,\ R_2 = \sin\psi, \quad 0 \leq \phi, \psi \leq \frac{\pi}{2}$$

The assumption $z, w \in M_{K,\circ}$ means

$$e^{in\alpha_1} \cos^n\phi - e^{im\alpha_2} \sin^m\phi = \rho_1, \quad e^{in\beta_1} \cos^n\psi - e^{im\beta_2} \sin^m\psi = \rho_2, \tag{4.5.13}$$

with $\rho_j > 0$. To construct a path joining z to w, represent $\alpha = (\alpha_1, \alpha_2)$ and $\beta = (\beta_1, \beta_2)$ as pairs of unit vectors $(e^{in\alpha_1}, e^{i(\pi+m\alpha_2)})$ and $(e^{in\beta_1}, e^{i(\pi+m\beta_2)})$ in the complex plane. For the pair $(e^{in\xi}, e^{i(\pi+m\eta)})$ solvability of

$$e^{in\xi} \cos^n\phi - e^{im\eta} \sin^m\phi \in \mathbf{R}_+, \quad 0 \leq \phi, \psi \leq \frac{\pi}{2}$$

is equivalent to the statement that $1 \in \mathbf{C}$ lies in the cone[4] formed by the vectors $e^{in\xi}, e^{i(\pi+m\eta)}$. It is now easy to see how to construct a path in $M_{K,\circ}$ joining z to w (see Figure 5.4). ♣

An important consequence of the realization of S_K as a fibre bundle over S^1 is that we can gain better understanding of the structure of $\pi_1(S_K, x)$. Example 4.2.7 and lemmas 4.5.4, 4.5.5 and 4.5.6 imply

Corollary 4.5.3 *The fundamental group $\pi_1(S_K, x)$, where K is the torus knot $K_{m,n}$, admits of the semi-direct product decomposition*

$$\pi_1(S_K, x) \simeq \mathbf{F}_s \cdot \mathbf{Z}.$$

The free group $\mathbf{F}_s$ in this decomposition is the commutator subgroup $\hat{\pi}$ and $s = (m-1)(n-1)$.

[4]The *cone* formed by vectors v, w means the set $\{av + bw\} \subset \mathbf{C} = \mathbf{R}^2$ as a, b range over positive real numbers.

Proof - Since the fundamental group of a compact surface with boundary is free, it only remains to prove the last assertion about the commutator subgroup. Denote $\pi_1(S_K, x)$ by π_1 and its commutator subgroup by $\hat{\pi}_1$. Since the commutator subgroup is the smallest normal subgroup with abelian quotient, $\hat{\pi}_1 \subseteq \mathbf{F}_s$. This together with $\pi_1/\hat{\pi}_1 \simeq \mathbf{Z}$ imply that we have the exact sequence

$$0 \longrightarrow \mathbf{F}_s/\hat{\pi}_1 \longrightarrow \mathbf{Z} \longrightarrow \mathbf{Z} \longrightarrow 0,$$

which splits and is therefore impossible unless $\mathbf{F}_s = \hat{\pi}_1$ as desired. By example 4.5.4 $H_1(\tilde{S}_{K,ab}; \mathbf{Z}) \simeq \mathbf{Z}^{(m-1)(n-1)}$. Therefore $s = (m-1)(n-1)$ as asserted. ♣

The following lemma helps one visualize how the fibres $M_{K,\theta}$ are situated in S_K:

Lemma 4.5.7 *Every fibre $M_{K,\theta}$ has points arbitrarily close to every point of $K_{m,n}$.*

Proof - Let $w_1^n - w_2^m = 0$ and set

$$z_1 = w_1 + \epsilon_1 e^{i\beta_1}, \quad z_2 = w_2 + \epsilon_1 e^{i\beta_2}.$$

Then

$$z_1^n - z_2^m = n\epsilon_1 e^{i\beta_1} w_1^{n-1} - m\epsilon_2 e^{i\beta_2} w_1^{m-1} + \cdots \text{(higher order terms in } \epsilon_j\text{'s)}$$

Let $\epsilon_j > 0$ be such that $|n\epsilon_1 w_1^{n-1}| = |m\epsilon_2 w_2^{m-1}|$. The required result follows by taking ϵ_j's sufficiently small and varying β_j's (or using the implicit function theorem to be precise). ♣

Lemma 4.5.7 and the above analysis imply that $S_K = \cup_\theta M_{K,\theta}$ with each fibre $M_{K,\theta}$ a compact smooth surface with boundary $K_{m,n}$.

Using the notion of a Seifert surface, we construct the universal abelian cover $\tilde{S}_{K,ab}$ by cutting and pasting. Make a cut in S_K along a Seifert surface, for example $M_{K,\circ}$, and open up the incision to obtain a new manifold S'_K with boundary. Since $S_K \to S^1$ is a fibre bundle, $S'_K \simeq [0, 2\pi] \times M_{K,\circ}$. The boundary $\partial S'_K$ is connected and has a natural decomposition into

$$\partial S'_K = M^+_{K,\circ} \cup M^-_{K,\circ}. \tag{4.5.14}$$

where $M^+_{K,\circ}$ and $M^-_{K,\circ}$ are disjoint copies of $M_{K,\circ}$. Let $S_{K,j}$, for $j \in \mathbf{Z}$, be a copy of S'_K and set

$$\partial S_{K,j} = M^+_{K,j} \cup M^-_{K,j},$$

with the obvious meaning for the sets on the right. Identify $M^+_{K,j}$ with $M^-_{K,j+1}$ for all $j \in \mathbf{Z}$ to obtain a connected manifold of dimension 3. This manifold is $\tilde{S}_{K,ab}$.

The identification of $M^+_{K,j}$ and $M^-_{K,j+1}$ is described by the the solution $p(t)$ of the vector field ζ. In view of (4.5.12), $p(2k\pi)$'s lie on the same surface $M_{K,\theta_\circ}$ for $k \in \mathbf{Z}$. We can assume the labeling of $M^{\pm}_{K,j}$ is such that $p(-2\pi)$ is a diffeomorphism of $M^+_{K,j}$ onto $M^-_{K,j}$. Composing this diffeomorphism with the identification of $S_{K,j}$ with $S_{K,j+1}$ we obtain the map which identifies $M^+_{K,j}$ and $M^-_{K,j+1}$.

The action of the (monodromy) group $J \simeq \mathbf{Z}$ of covering transformations of $\tilde{S}_{K,ab} \rightarrow S_K$ is given by

$$t : S_{K,j} \longrightarrow S_{K,j+1}.$$

The diffeomorphism $p(-2\pi) : M^+_{K,j} \simeq M^-_{K,j}$ induces an isomorphism of $H_1(M^+_{K,j};\mathbf{Z})$ onto $H_1(M^-_{K,j};\mathbf{Z})$, which yields an integral matrix

$$A : \mathbf{Z}^{(m-1)(n-1)} \longrightarrow \mathbf{Z}^{(m-1)(n-1)}$$

of determinant ± 1. Using the Mayer-Vietoris sequence it is straightforward to see that the first homology group of $\tilde{S}_{K_{m,n},ab}$ is the cokernel of the $\mathcal{R}$-module homomorphism ($\mathcal{R} = \mathbf{Z}[t, t^{-1}]$)

$$tI - A : \mathcal{R}^{(m-1)(n-1)} \longrightarrow \mathcal{R}^{(m-1)(n-1)}. \tag{4.5.15}$$

The matrix $tI - A$ is an Alexander matrix for the torus knot $K_{m,n}$ and its determinant is the Alexander polynomial. The practical calculation of the matrix A, valid for an arbitrary knot, requires the notion of linking number which is introduced in chapter 6 in the context of cohomology.

Let $\tilde{S}_{K,l}$ denote the cyclic l-sheeted covering of S_K where $K = K_{m,n}$. With this picture in mind, it is (in principle) straightforward to calculate $H_1(\tilde{S}_{K,l};\mathbf{Z})$. Geometrically, l-sheeted covering means we identify $M^+_{K,l-1}$ with $M^-_{K,\circ}$. Therefore to calculate $H_1(\tilde{S}_{K,l};\mathbf{Z})$ we add the relation $t^l - 1 = 0$ to the Alexander matrix and then compute the cokernel. The resulting module is not the entire $H_1(\tilde{S}_{K,l};\mathbf{Z})$. One should take the direct sum of this module with $\mathbf{Z}$ to obtain $H_1(\tilde{S}_{K,l};\mathbf{Z})$. The reason is that there is a cycle in $H_1(\tilde{S}_{K,l};\mathbf{Z})$ which is mapped to l times a generator of $H_1(S_K;\mathbf{Z})$ and is not represented in $\mathrm{Coker}(tI - A)$. The pre-image under the covering map $\tilde{S}_{K,l} \rightarrow S_K$ of a meridian of the torus ∂S_K represents this cycle in $H_1(\tilde{S}_{K,l};\mathbf{Z})$. The case of 1-sheeted covering gives interesting information. Here we have to set $t = 1$ in the presentation for the Alexander module. Since a 1-sheeted cover is the manifold S_K itself, the matrix $I - A$ should be invertible which implies

$$\Delta_K(1) = \pm 1. \tag{4.5.16}$$

The above considerations about l-sheeted coverings and in particular (4.5.16) are valid for all knots. We will return to subject later.

Exercise 4.5.4 *Verify the following statements about l-sheeted coverings of the complement of the trefoil knot K:*

$$H_1(\tilde{S}_{K,3};\mathbf{Z}) \simeq \mathbf{Z}\oplus\mathbf{Z}/3\oplus\mathbf{Z}/3, \quad H_1(\tilde{S}_{K,4};\mathbf{Z}) \simeq \mathbf{Z}\oplus\mathbf{Z}/3,$$
$$H_1(\tilde{S}_{K,5};\mathbf{Z}) \simeq \mathbf{Z}, \quad H_1(\tilde{S}_{K,6};\mathbf{Z}) \simeq \mathbf{Z}\oplus\mathbf{Z}\oplus\mathbf{Z}.$$

The geometric picture of the torus knot described above is what one tries to generalize to arbitrary knots. While the notion of Seifert surface generalizes and one can do the cutting and pasting to construct $\tilde{S}_{K,ab}$, there is no fibre bundle structure $S_K \rightarrow S^1$ and no vector field ζ whose integral curves give a diffeomorphism of $M^-_{K,j}$ and $M^+_{K,j}$. In chapter 6 we show that we still have a mapping $\varphi : S_K \rightarrow S^1$ which when composed with any branch of logarithm on $S^1 \subset \mathbf{C}$ gives a (local) Morse function on S_K, i.e., a circle valued Morse function. For certain (but not all) regular values $e^{i\theta}$, the submanifold $\varphi^{-1}(e^{i\theta})$ is a Seifert surface for the knot. This Morse function necessarily has critical points for a general knot and the structure of the critical points and a linear mapping that it induces determine an Alexander module, and therefore the Alexander polynomial and ideals.

4.6 Discrete Subgroups of $SL(2,\mathbf{R})$ and $SL(2,\mathbf{C})$

4.6.1 Hyperbolic and Loxodromic Transformations

In this subsection we study some elementary properties of discrete subgroups Γ of $SL(2,\mathbf{R})$ and $SL(2,\mathbf{C})$ such that the canonical map $\mathcal{H}_i \rightarrow M = \Gamma\backslash\mathcal{H}_i$, $i = 2,3$, is a covering projection and M is a compact manifold. Although we limit ourselves to hyperbolic spaces in dimensions two and three, there are far reaching generalizations of these results which we will not go into. Recall from example (XXXX) of chapter 1 that the hyperbolic space $\mathcal{H}_3$ can be realized as the upper half space and the action of $SL(2,\mathbf{C})$ is by fractional linear transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : \; w \longrightarrow (aw+b)(cw+d)^{-1},$$

where $w = z + t\mathbf{j}$, $t > 0$, $z \in \mathbf{C}$, and the quantities in consideration are computed according to the algebra of quaternions. In particular, the isotropy subgroup of the point $\mathbf{j} \in \mathcal{H}_3$ is $SU(2)$. We have

Lemma 4.6.1 *$M = \Gamma\backslash\mathcal{H}_i$ is a manifold and the canonical map $\pi : \mathcal{H}_i \rightarrow M = \Gamma\backslash\mathcal{H}_i$ a covering projection if and only if Γ is torsion free.*

Proof - We consider the case $i = 3$; the case $i = 2$ is similar. It is clear that $\mathcal{H}_3 \rightarrow M = \Gamma\backslash\mathcal{H}_3$ is a covering projection (and consequently M a manifold) if and only if Γ acts properly discontinuously on $\mathcal{H}_3$. It is easy to see that latter condition is equivalent to the action Γ being free (i.e., $e \neq \gamma \in \Gamma$ then $\gamma(w) \neq w$ for all $w \in \mathcal{H}_3$). Since the isotropy subgroup of $\mathbf{j}$ is $SU(2)$ and $SL(2,\mathbf{C})$ acts transitively on $\mathcal{H}_3$, the isotropy subgroup of a point $w = g(\mathbf{j})$ is $gSU(2)g^{-1}$. Therefore the action of Γ is free if and only if

$$\Gamma \cap gSU(2)g^{-1} = e.$$

Since Γ is discrete this is equivalent to Γ containing no torsion element other than the identity. ♣

It is customary to refer to the discrete subgroup Γ is such that $M = \Gamma\backslash\mathcal{H}_i$ is compact as a *cocompact* subgroup. By a *unipotent* matrix we mean a matrix all whose eigenvalues are 1. By a *loxodromic* transformation[5] we mean an element $g \in SL(2,\mathbf{C})$ with distinct eigenvalues λ_i and $|\lambda_i| \neq 1$. It is easy to see that loxodromic tranformations are characterized by the property of having two distinct fixed points both lying on the boundary $\mathbf{C} \cup \infty$ of the upper half space. A loxodromic transformation $A \in SL(2,\mathbf{R})$ is called *hyperbolic*.

[5]There is some discrepancy in the literature about the terms *loxodromic* and *hyperbolic* transformation.

Lemma 4.6.2 *The fixed points of a hyperbolic transformation $A \in SL(2, \mathbf{R})$ necessarily lie in $\mathbf{R} \cup \infty$.*

Proof - The characteristic polynomial of a loxodromic transformation is $\lambda^2 - (a + d)\lambda + 1$ with discriminant $\Delta = (a + d)^2 - 4$. Now $a + d$ is real and if $(a + d)^2 > 4$ then the equation $\frac{az+b}{cz+d} = z$ has distinct real roots unless $b = c = 0$ in which case 0 and ∞ are fixed points of A. If $(a+d)^2 < 4$ then the characteristic polynomial can be written as $\lambda^2 - (2\cos\theta)\lambda + 1$. It follows that A is conjugate to a rotation matrix and its eigenvalues have norm 1 contradicting the hypothesis. ♣

Lemma 4.6.3 *Let $A \in SL(2, \mathbf{R})$. The following condition are equivalent:*

1. *A is hyperbolic.*
2. *A has distinct real eigenvalues.*
3. $\mathrm{Tr}(g)^2 > 4$.

Proof - The lemma follows easily fom the observation that for $A \in SL(2, \mathbf{C})$, the quadratic equations $\det(\lambda I - A) = 0$ and $\frac{a\lambda+b}{c\lambda+d} = \lambda$ have identical discriminant $\Delta = (a + d)^2 - 4$. ♣

An important property of loxodromic (or hyperbolic) transformations is that every such matrix γ leaves a unique geodesic $\tilde{\gamma}$ invariant, namely the geodesic joining its fixed points which lie on the boundary. One refers to the geodesic $\tilde{\gamma}$ as the *axis* of the loxodromic (hyperbolic) transformation γ. The concept of axis plays an essential role in the subsection on closed geodesics on $\Gamma\backslash\mathcal{H}_j$ and line geometry.

Exercise 4.6.1 *Given a pair of distinct points $\zeta_1, \zeta_2 \in \mathbf{C}$ (resp. $\mathbf{R}$) and a complex number τ (resp. a real number τ with $\tau^2 > 4$) there is a pair of loxodromic (resp. hyperbolic) transformations of trace τ whose axis is the unique geodesic with end points ζ_1, ζ_2.*

The following lemma is a consequence of the Lebesgue covering lemma:

Lemma 4.6.4 *Let M be a compact Riemannian manifold. Then there is $\delta > 0$ such that for all $x \in M$ the ball of radius δ centered at x is a contractible neighborhood of x.*

Now we can relate compactness of M to the group theory of Γ.

Lemma 4.6.5 *Let Γ be a torsion free cocompact discrete subgroup of $SL(2, \mathbf{R})$ or $SL(2, \mathbf{C})$. Then Γ contains no unipotent matrix other than the identity.*

Proof - Let $\gamma \in \Gamma$ be a unipotent element. After replacing Γ with a conjugate $g\Gamma g^{-1}$ we may assume $\gamma = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$. Let $z = it$ in case of $\Gamma \subset SL(2, \mathbf{R})$ and $z = t\mathbf{j}$ otherwise. Then the length of the straight line segment joining z to $\gamma(z)$ tends to zero as $t \to \infty$. This line segment represents $\gamma \neq e$ in the fundamental group of M. On the other hand, lemma 4.6.4 implies that by taking t sufficiently large we can make this loop contractible to a point and therefore $\gamma = e$. ♣

Corollary 4.6.1 *Let Γ be a torsion free cocompact discrete subgroup of $SL(2, \mathbf{R})$ Then every $\gamma \in \Gamma$, $\gamma \neq \pm e$, is hyperbolic.*

Proof - In view of lemmas 4.6.5 and 4.6.3 we need to show $\mathrm{Tr}(\gamma)^2 < 4$ is not possible. Since $\Gamma \subset SL(2, \mathbf{R})$ and the discriminant of the quadratic equation $\frac{az+b}{cz+d} = z$ is $\mathrm{Tr}(\gamma)^2 - 4$, the condition $\mathrm{Tr}(\gamma)^2 < 4$ implies that γ has a fixed point in $\mathcal{H}_2$. Then $\mathcal{H}_2 \to M$ is not a covering projection contrary to lemma 4.6.1. ♣

Lemma 4.6.6 *Let $\Gamma \subset SL(2, \mathbf{C})$ be a torsion free cocompact discrete subgroup. If $\pm e \neq \gamma \in \Gamma$ then the eigenvalues of γ do not lie on the unit circle, or equivalently, γ is loxodromic.*

Proof - Assume the contrary. After replacing Γ with a conjugate we may assume $\gamma = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$ or is a unipotent matrix. Then γ leaves the point $\mathbf{j} \in \mathcal{H}_3$ fixed and $\mathcal{H}_3 \to M$ is not a covering projection. ♣

Corollary 4.6.2 *Let Γ be a cocompact torsion free discrete subgroup of $SL(2, \mathbf{R})$ or $SL(2, \mathbf{C})$. Then every abelian subgroup of Γ is cyclic.*

Proof - Let $e \neq \gamma \in \Gamma \subset SL(2, \mathbf{R})$. Then γ is hyperbolic and therefore has distinct real eigenvalues. Therefore after replacing Γ by a conjugate subgroup we may assume $\gamma = \begin{pmatrix} e^s & 0 \\ 0 & e^{-s} \end{pmatrix}$ with $s \neq 0$. If $e \neq \delta \in \Gamma$ commutes with γ then $\delta = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$ with $t \neq 0$. If s and t are linearly dependent over the rationals then there is $q \in \mathbf{Q}$ such that $s = aq$ and $t = bq$ with $a, b \in \mathbf{Z}$ which shows that the subgroup generated γ and δ is cyclic. If s and t are linearly independent over $\mathbf{Q}$ then from elementary number theory we know that we can approximate 0 arbitrarily closely by sums of the form $as + bt$ with $a, b \in \mathbf{Z}$. This implies that the subgroup generated by γ and δ has elements arbitrarily close to the identity and therefore is not discrete. The case of $SL(2, \mathbf{C})$ is similar. Here we can assume

$$\gamma = \begin{pmatrix} e^{s+i\sigma} & 0 \\ 0 & e^{-s-i\sigma} \end{pmatrix}, \quad \delta = \begin{pmatrix} e^{t+i\tau} & 0 \\ 0 & e^{-t-i\tau} \end{pmatrix}$$

with $s \neq 0$ and $t \neq 0$ in view of lemma 4.6.6. If s and t are linearly independent over the rational numbers, then let a_j, b_j be a sequence of integers such that $a_j s + b_j t \to 0$ as $j \to \infty$. Then the sequence $\gamma^{a_j}\delta^{b_j}$ gets arbitrarily close to the unit circle $\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$ contradicting discreteness of Γ. So assume $as + bt = 0$ for some integers a and b. Then $e^{i(a\sigma + b\tau)} = 1$ for otherwise Γ either contains torsion elements or is not discrete. It follows that there is $r + i\rho$ such that $e^s = e^{n(r+i\rho)}$ and $e^t = e^{m(r+i\rho)}$ for some integers m, n. ♣

In order to understand the structure of the quotient space $\Gamma\backslash\mathcal{H}_j$, it is important to construct a fundamental domain for the action of Γ. There is a standard procedure for constructing a fundamental domain for a group Γ acting properly discontinuously and by isometries on a complete Riemannian manifold M. Fix $x \in M$ and for every $\gamma \in \Gamma$ let

$$U_\gamma = \{y \in M \mid d(x,y) \leq d(\gamma.x, y)\}.$$

It is not difficult to show that $\bigcap_\gamma U_\gamma$ is a fundamental domain for the action of Γ. The proof of this fact can be found in many texts, e.g. [Ra]. For certain subgroups $\Gamma \subset SL(2,\mathbf{R})$ acting by fractional linear transformations on $\mathcal{H}_2$ there is a different method for constructing fundamental domains which is computationally more tractable than the method described above. This method makes use of the concept of isometric circles which is described below. For this method to be applicable to subgroups $\Gamma_{a,p}$, it is necessary to conformally map the upper half plane onto the unit disc $\mathcal{D}$ by a fractional linear transformation C, and accordingly replace $\Gamma_{a,p}$ with $C\Gamma_{a,p}C^{-1}$. We simply write Γ for $C\Gamma_{a,p}C^{-1}$, and note that $CSL(2,\mathbf{R})C^{-1} = SU(1,1)$. The following lemma isolates the advantage of working on $\mathcal{D}$ rather than $\mathcal{H}_2$ as will become clear shortly:

Lemma 4.6.7 *Let $\gamma \in \Gamma_{a,p}$ be a hyperbolic element. Then $C\gamma C^{-1}$ is not diagonal.*

Proof - Since $\gamma \neq \pm I$ is a hyperbolic element, its eigenvalues are real. A 2×2 diagonal matrix of determinant 1 with real entries does not leave the unit disc $\mathcal{D}$ invariant unless it is $\pm I$. ♣

Given $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SU(1,1)$,with $c \neq 0$, the equation $|cz + d| = 1$ defines a circle C_γ. The reason for transforming the problem from $\mathcal{H}_2$ to $\mathcal{D}$ is the requirement $c \neq 0$ which is fulfilled by lemma 4.6.7. The center and radius of C_γ are

$$C_\gamma: \quad \text{Center} = (-\frac{d}{c}, 0), \quad \text{Radius} = \frac{1}{|c|}. \tag{4.6.1}$$

C_γ is called the *isometric circle* of γ. Geometrically, C_γ is the subset of $\mathbf{C}$ where the fractional linear transformation defined by γ is an isometry relative to the Euclidean metric

$dx^2 + dy^2 = dzd\bar{z}$. The transformationγ is length decreasing in the exterior of the isometric circle and length increasing in its interior. Denoting the exterior and interior of C_γ by C_γ^e and C_γ^i respectively we can restate this fact as

$$\gamma^\star(dzd\bar{z}) < dzd\bar{z}, \text{ for } z \in C_\gamma^e, \text{ and } \gamma^\star(dzd\bar{z}) > dzd\bar{z}, \text{ for } z \in C_\gamma^i. \tag{4.6.2}$$

Note that γ maps C_γ^e onto $C_{\gamma^{-1}}^i$ and vice versa. The following lemma plays the key role in the application of the concept of isometric circle:

Lemma 4.6.8 *Let $\gamma, \delta \in SU(1,1)$ and assume that the isometric circles $C_{\gamma^{-1}}$ and C_δ are exterior to each other. Then $C_{\gamma\delta}$ lies in the interior of C_γ.*

Proof - Let z lie in the exterior of C_γ, then $\gamma(z)$ lies in the interior $C_{\gamma^{-1}}$ which is contained in the exterior of C_δ. Therefore

$$(\gamma\delta)^\star(dzd\bar{z}) = \delta^\star\gamma^\star(dzd\bar{z}) < dzd\bar{z}.$$

Therefore $C_{\gamma\delta}^e \subset C_\gamma^e$ and the required result follows. ♣

Corollary 4.6.3 *Let $\gamma \in SU(1,1)$ and assume $C^{-1}\gamma C \in SL(2,\mathbf{R})$ is a hyperbolic transformation. Then C_γ and $C_{\gamma^{-1}}$ are exterior to each other and for $n \geq 1$, $C_{\gamma^{n+1}}^i \subset C_{\gamma^n}^i$.*

Proof - Since for a hyperbolic transformation $|a + d| > 2$, the first assertion is a simple computation using (4.6.1). The second assertion is an inductive argument where lemma 4.6.8 is applied with $\delta = \gamma^n$. Q E D

The following simple corollary relates the concept of isometric circle to a fundamental domain for Γ:

Corollary 4.6.4 *Let $\Gamma_1 \subset \Gamma = C\Gamma_{a,p}C^{-1}$ be a subgroup such that $C^{-1}\gamma C \in SL(2,\mathbf{R})$ is a hyperbolic transformation for all $\pm e \neq \gamma \in \Gamma_1$. A fundamental domain for Γ is*

$$\mathsf{F}_\Gamma = \mathcal{D}\bigcap_\gamma C_\gamma^e.$$

In practical applications of corollary 4.6.4 the intersection is only over a finite set $\{\gamma_1, \cdots, \gamma_N\}$ generating the group Γ. This point is clarified in lemma 4.6.13 below. Some examples of fundamental domains for explicitly given groups are given in the next section.

4.6.2 Quaternion Groups

In this subsection we give an arithmetical construction for discrete groups Γ acting on the hyperbolic plane or space $\mathcal{H}_j$, $j = 2,3$, such that $\Gamma \setminus \mathcal{H}_j$ is compact. There are several reasons for doing this. As a consequence of our construction we obtain a proof of a version of the uniformization theorem according to which every compact orientable surface of genus ≥ 2 admits of a metric of constant negative curvature. In particular, the upper half plane is the universal covering space of such a surface. The construction is explicitly given by 2×2 matrices and lends itself to calculation more readily than the geometric proof. The explicit nature of the arithmetical construction makes it possible to study the behavior of other geometric quantities. In the next section we study the behavior of closed geodesics on compact manifolds $\Gamma\backslash\mathcal{H}_j$, $j = 2,3$.

First we consider the two dimensional case. Let $p > 2$ be a prime and $a \in \mathbf{Z}$ a *quadratic non-residue* mod p, i.e., the equation $x^2 \equiv a \bmod p$ has no solution. Let $\Gamma = \Gamma_{a,p}$ be the subgroup of $SL(2,\mathbf{R})$ consisting of matrices of the form

$$\Gamma_{a,p}: \quad \begin{pmatrix} x_\circ + x_1\sqrt{a} & x_2\sqrt{p} + x_3\sqrt{ap} \\ x_2\sqrt{p} - x_3\sqrt{ap} & x_\circ - x_1\sqrt{a} \end{pmatrix}, \qquad x_j \in \mathbf{Z}. \tag{4.6.3}$$

The requirement $\Gamma \subset SL(2,\mathbf{R})$ means

$$x_\circ^2 - ax_1^2 - px_2^2 + apx_3^2 = 1. \tag{4.6.4}$$

Note that matrices of the above form in $SL(2,\mathbf{R})$ for a group. The fact that $\Gamma_{a,p}$ is an infinite discrete group is elementary number theory. For example, it is classical that the so-called Pell's equation $x^2 - ay^2 = 1$ has an infinity of solutions (in integers) generated by a *least* solution. Given a solution (x, y), then the matrices

$$\begin{pmatrix} x + y\sqrt{a} & 0 \\ 0 & x - y\sqrt{a} \end{pmatrix}^n$$

yield an infinity of solutions to Pell's equation and an infinite abelian subgroup of Γ.

Lemma 4.6.9 *Let p be of the form $4m + 1$. Then subgroup $\Gamma_{a,p}$ is torsion free and consequently $\mathcal{H}_2 \to M_\Gamma = \Gamma \setminus \mathcal{H}_2$ is a covering projection.*

Proof - It suffices to prove the first assertion. The law of quadratic reciprocity and the assumption on p imply that -1 is a quadratic residue. Let $\pm e \neq \gamma \in \Gamma_{a,p}$ be an element of finite order, then its eigenvalues are $e^{\pm i\theta}$ and its characteristic equation is $\lambda^2 - (2\cos\theta)\lambda + 1 =$

0. Therefore with the notation of (4.6.3) $x_\circ = 0, \pm 1$, and γ is conjugate to a rotation matrix. In the latter case, $x_\circ = \pm 1$, γ is conjugate to $\pm I$ and therefore $\gamma = \pm I$. In the former case, $x_\circ = 0$ and (4.6.4) becomes $1 + ax_1^2 + px_2^2 = apx_3^2$. Reducing mod p we obtain

$$1 + ax_1^2 \equiv 0 \bmod p,$$

which is not possible since -1 is a quadratic residue and a is not. ♣

Remark 4.6.1 If the prime p is of the form $4m+3$ then $\Gamma_{a,p}$ may contain elements of finite order. For example, let $a = 2$, $p = 3$, then the matrix

$$\begin{pmatrix} \sqrt{2} & \sqrt{3}+\sqrt{6} \\ \sqrt{3}-\sqrt{6} & -\sqrt{2} \end{pmatrix}$$

has order 4. ♡

Example 4.6.1 The trace of an element of $\Gamma_{a,p}$ is an even integer. Since an element of finite order is conjugate to a rotation matrix, every element ($\neq \pm e$) of finite order in $\Gamma_{a,p}$ necessarily has trace zero. Let $\Gamma^\circ_{a,p}$ be the subgroup of $\Gamma_{a,p}$ consisting of elements with $x_2 \equiv 0 \bmod a$ in the notation of (4.6.3). It is trivial that such matrices form a subgroup of finite index in $\Gamma_{a,p}$. For $\gamma \in \Gamma^\circ_{a,p}$ we have $x_\circ \equiv 1 \bmod a$ and therefore $\Gamma^\circ_{a,p}$ is torsion free. ♠

Example 4.6.2 The subgroups $\Gamma_{a,p}$ are not neceessarily maximal among discrete subgroups of $SL(2,\mathbf{R})$. For a positive integer $\delta \equiv 1 \bmod 4$ and y_1, y_2 integers of the same parity, the quantities

$$\xi_1, \xi_2 = \frac{y_1 \pm y_2\sqrt{\delta}}{2}$$

are integers in $\mathbf{Q}(\sqrt{\delta})$. In fact, ξ_j's are the roots of an equation of the form $x^2 + bx + c = 0$ with $\delta = b^2 - 4c$. Therefore matrices of the form

$$\gamma = \begin{pmatrix} \frac{x_\circ + x_1\sqrt{\delta}}{2} & \frac{(x_2+x_3\sqrt{\delta})\sqrt{p}}{2} \\ \frac{(x_2-x_3\sqrt{\delta})\sqrt{p}}{2} & \frac{x_\circ - x_1\sqrt{\delta}}{2} \end{pmatrix}$$

where the integers x_j satisfy $x_\circ \equiv x_1 \bmod 2$ and $x_2 \equiv x_3 \bmod 2$, form a discrete subgroup $\Gamma'_{\delta,p} \subset SL(2,\mathbf{R})$. Clearly $\Gamma_{\delta,p} \subset \Gamma'_{\delta,p}$. Let $\gamma' \in \Gamma'_{\delta,p}$ be of the same form as γ with the integers x_j denoted by y_j's. If $x_j \equiv y_j \bmod 2$, then it is readily verified that $\gamma'^{-1}\gamma \in \Gamma_{\delta,p}$. It follows that $\Gamma_{\delta,p}$ has index 4 in $\Gamma'_{\delta,p}$. ♠

HERE WE ADD EXPLICIT EXAMPLES OF FUNDAMENTAL DOMAINS

We have so far constructed some examples of compact quotients of $\mathcal{H}_2$ by explicitly exhibiting torsion free discrete subgroups of $SL(2,\mathbf{R})$. An immediate consequence of the above example is a version of the *uniformization theorem*, namely,

Corollary 4.6.5 *A compact orientable surface of genus ≥ 2 admits of a metric of constant negative curvature. In particular, The universal covering space of a compact orientable surface of genus $g \geq 2$ is the upper half plane.*

Proof - Example **??** shows that a surface of genus 2 admits of a metric of constant negative curvature. Since an n-sheeted covering of a compact orientable surface of genus $g = 2$ is a surface of genus of genus $n+1$ (example **??** of chapter4)the required result follows. ♣

Exercise 4.6.2 *Let $g > 1$ and N_g be the compact non-orientable surface constructed in chapter 3 subsection 5 (i.e., a compact non-orientable surface other than the projective plane or the Klein bottle). Show that the universal covering space of N_g is the upper half plane.*

To extend the arithmetical construction of discrete subgroups of $SL(2,\mathbf{R})$ to that of $SL(2,\mathbf{C})$ we recall some facts from elementary number theory. Let $\mathcal{O} = \mathbf{Z}[i]$ be the ring of Gaussian integers. We have

Proposition 4.6.1 *$\mathcal{O}$ is a principal ideal domain. A prime $p \in \mathbf{Z}$ remains a prime in $\mathcal{O}$ if and only if $p \equiv 3 \bmod 4$. For such p the field $\mathcal{O}/(p)$ is isomorphic to the finite field of p^2 elements.*

Example 4.6.3 For $p \not\equiv 3 \bmod 4$, p splits in $\mathcal{O}$. For example,

$$2 = (1+i)(1-i), \quad 5 = (1+2i)(1-2i), \quad 13 = (2+3i)(2-3i),$$
$$17 = (1+4i)(1-4i), \quad 29 = (5+2i)(5-2i), \quad 37 = (6+i)(6-i)$$
$$41 = (4+5i)(4-5i), \quad 53 = (7+2i)(7-2i), \quad 61 = (5+6i)(5-6i)$$

For primes $p \equiv 3 \bmod 4$ the finite field $\mathcal{O}/(p) \simeq \mathcal{F}_{p^2}$ is isomorphic to and is identified with $\mathcal{F}_{p^2}$ can be realized as the set $\{a+bi\}$ with a, b ranging over $\mathbf{Z}/p$ and field operations defined accordingly. Let $p = 3$. The quadratic residues in $\mathcal{F}_9^\times$ are $\{\pm 1, \pm i\}$, and the quadratic non-residues are $1 \pm i$ and $2 \pm i$. For $p = 7$ the quadratic residues in $\mathcal{F}_{49}^\times$ are

$$\pm 1, \quad \pm 2, \quad \pm 3, \quad \pm i, \quad \pm 2i, \quad \pm 3i,$$
$$1 \pm i, \quad 2 \pm i, \quad 3 \pm 2i, \quad 4 \pm i, \quad 5 \pm 4i, \quad 6 \pm 5i.$$

The remaining elements of $\mathcal{F}_{49}^\times$ are quadratic non-residues. ♠

Let $\tilde{\Gamma}_{a,p} \subset SL(2,\mathbf{C})$ be the subgroup consisting of matrices of the form

$$\gamma = \begin{pmatrix} z_\circ + z_1\sqrt{a} & z_2\sqrt{p} + z_3\sqrt{ap} \\ z_2\sqrt{p} - z_3\sqrt{ap} & z_\circ - z_1\sqrt{a} \end{pmatrix}$$

where $z_j \in \mathcal{O}$, and $\det(\gamma) = z_\circ^2 - az_1^2 - pz_2^2 + apz_3^2 = 1$. Let $\Lambda_\circ \subset \mathbf{C}^3 \simeq \mathbf{R}^6$ be the lattice which is $\mathcal{O}$-linear combinations of the matrices

$$f_1 = \begin{pmatrix} \sqrt{a} & 0 \\ 0 & -\sqrt{a} \end{pmatrix}, \quad f_2 = \begin{pmatrix} 0 & \sqrt{p} \\ \sqrt{p} & 0 \end{pmatrix}, \quad f_3 = \begin{pmatrix} 0 & \sqrt{ap} \\ -\sqrt{ap} & 0 \end{pmatrix},$$

It is immediately verified that $\Lambda_\circ$ is invariant under the conjugation action $X \to gXg^{-1}$ of $\tilde{\Gamma}_{a,p}$. Consequently, $\tilde{\Gamma}_{a,p}$ is a discrete subgroup of $SL(2,\mathbf{C})$.

Lemma 4.6.10 *For a prime integer $p \equiv 3 \bmod 4$ and $a \in \mathcal{O}$ a quadratic non-residue in $\mathcal{F}_{p^2}$, then the only solutions in $\mathcal{O}$ of the equation*

$$az_1^2 + pz_2^2 - apz_3^2 = 0$$

are $z_j = 0$.

Proof - Reducing the equation $az_1^2 + pz_2^2 - apz_3^2 = 0 \bmod p$ we obtain the equation $z_1^2 a = 0$ in $\mathcal{F}_{p^2}$. It follows that $z_1 = v_1 p$ for some $v_1 \in \mathcal{O}$. Substituting we obtain

$$apv_1^2 + z_2^2 - az_3^2 = 0. \tag{4.6.5}$$

Reduction mod$p$ shows that (4.6.5) has no solution in $\mathcal{O}$ since a is a quadratic non-residue. ♣

Lemma 4.6.11 *For a prime integer $p \equiv 3 \bmod 4$ and $a \in \mathcal{O}$ a quadratic non-residue in $\mathcal{F}_{p^2}$, $\tilde{\Gamma}_{a,p}/\pm I$ is a torsion free discrete subgroup of $SL(2,\mathbf{C})$.*

Proof - Since a torsion element $\gamma \in \tilde{\Gamma}_{a,p}$ is conjugate to a unitary matrix, we necessarily have $z_\circ = 0$ for $\gamma \neq \pm I$. Therefore $az_1^2 + pz_2^2 = apz_3^2 + 1$. Reducing mod$p$ we obtain $az_1^2 = -1$ in $\mathcal{F}_{p^2}$. Now $-1 = i^2$ and therefore a is a quadratic residue in $\mathcal{F}_{p^2}$ contrary to hypothesis. ♣

We now prove compactness of the orbit spaces $\Gamma_{a,p}\backslash\mathcal{H}_2$ and $\tilde{\Gamma}_{a,p}\backslash\mathcal{H}_3$. The proof, follows an idea of Mostow and Tamagawa (see [Bo]) and is based on Mahler's compactness criterion (see chapter 1, example (XXXX)). It is applicable to general arithmetic groups, a subject that we shall not go into. We identify $\mathbf{C}^3$ (resp. $\mathbf{R}^3$) with the space of 2×2 complex (res. real) trace zero matrices and let $SL(2,\mathbf{C})$ (resp. $SL(2,\mathbf{R})$) act on it via the adjoint representation $X \to gXg^{-1}$. Fix the basis

$$f_1 = \begin{pmatrix} \sqrt{a} & 0 \\ 0 & -\sqrt{a} \end{pmatrix},\ f_2 = \begin{pmatrix} 0 & \sqrt{p} \\ \sqrt{p} & 0 \end{pmatrix},\ f_3 = \begin{pmatrix} 0 & \sqrt{ap} \\ -\sqrt{ap} & 0 \end{pmatrix},$$

for $\mathbf{C}^3$ (resp. $\mathbf{R}^3$), and let $\Lambda_\circ$ (resp. $L_\circ$) be the lattice which is $\mathcal{O} = \mathbf{Z}[i]$ (resp. $\mathbf{Z}$) linear combinations of f_1, f_2, f_3. Then $\tilde{\Gamma}_{a,p}$ (resp. $\Gamma_{a,p}$) is the subgroup of $SL(2,\mathbf{C})$ (resp. $SL(2,\mathbf{R})$) leaving the lattice $\Lambda_\circ$ (resp. $L_\circ$) invariant. Therefore the orbit space $\tilde{\Gamma}_{a,p} \setminus SL(2,\mathbf{C})$ (resp. $\Gamma_{a,p} \setminus SL(2,\mathbf{R})$) can be identified with the set $\mathcal{R}$ of lattices of the form $gL_\circ g^{-1}$ with the obvious topology. Every lattice $L \subset \mathbf{C}^3$ (resp. $L \subset \mathbf{R}^3$) is of the form $A(\Lambda_\circ)$ (resp.$A(L_\circ)$) where $A \in GL(3,\mathbf{C})$ (resp. $A \in GL(3,\mathbf{R})$). For any basis $\{v_j\}$ for L, the volume of the parallelpiped spanned by the vectors $\{v_j\}$ is independent of the choice of the basis and is equal to $|\det(A)|^2 V_\circ$ (resp. $|\det(A)|V_\circ$) where $V_\circ$ is the volume of the parallelepiped spanned by $f_1, f_2, f_3, if_1, if_2, if_3$ (resp. f_1, f_2, f_3). One refers to $|\det(A)|V_\circ$ as the *volume* of the lattice L and denotes it by $\mathrm{vol}(L)$. By Mahler's compactness criterion (see chapter 1, example(XXX)) compactness of $\mathcal{R}$ is equivalent to the following conditions:

1. There is a neighborhood U of $\mathbf{0} \in \mathbf{C}^3$ (resp. $\mathbf{R}^3$) such that for every $X \in L_\circ$ and $g \in SL(2,\mathbf{C})$ (resp. $g \in SL(2,\mathbf{R})$) we have $gXg^{-1} \notin U$;

2. $\mathrm{vol}(L)$ is bounded on $\mathcal{R}$.

Verification of (1) - The verification of (1) depends on the following lemma (which also generalizes to $n \times n$ matrices in the appropriate form):

Lemma 4.6.12 *Let X be 2×2 trace zero complex matrix with distinct eigenvalues $\pm\beta$. Then the orbit $\{gXg^{-1}\}_{g\in GL(2,\mathbf{C})}$ is a closed subset of $\mathbf{C}^3$ disjoint from the ball of radius $|\beta|$ centered at the origin.*

Proof - Since the orbit space $\{gXg^{-1}\}_{g\in GL(2,\mathbf{C})}$ is defined by equation $Y^2 + \det(X)I = 0$ (Y is a variable 2×2 complex traceless matrix), it is a closed subset of $\mathbf{C}^3$. For $X = \begin{pmatrix} \beta & 0 \\ 0 & -\beta \end{pmatrix}$ and $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $\det g = 1$ we have

$$gXg^{-1} = \begin{pmatrix} (ad+bc)\beta & -2ab\beta \\ 2cd\beta & -(ad+bc)\beta \end{pmatrix}.$$

Therefore for $Y = (y_{ij})$ and relative to the Euclidean norm $||Y|| = \sqrt{\sum_{i,j} |y_{ij}|^2}$, we have

$$||gXg^{-1}||^2 = [4|ab|^2 + 4|cd|^2 + 2|ad + bc|^2]|\beta|^2 \geq 2|ad - bc|^2|\beta|^2.$$

Since $\det(g) = 1$ the required result follows. ♣

In view of the lemma, the validity of condition (1) follows once we show that the eigenvalues of a matrix $X \in L_\circ$ are uniformly bounded away from zero. The characteristic polynomial of

$$X = \begin{pmatrix} z_1\sqrt{a} & z_2\sqrt{p} + z_3\sqrt{ap} \\ z_2\sqrt{p} - z_3\sqrt{ap} & -z_1\sqrt{a} \end{pmatrix} \tag{4.6.6}$$

is $\lambda^2 + \delta = 0$ where

$$\delta = -az_1^2 - pz_2^2 + apz_3^2. \tag{4.6.7}$$

Since $\delta \in \mathcal{O}$, the eigenvalues $\pm\sqrt{-\delta}$ are uniformly bounded away from zero unless δ vanishes. For a solution $(z_1, z_2, z_3) \in \mathcal{O}^3$ of the diophantine equation $\delta = 0$ we have $z_1 \in p\mathcal{O}$ and setting $z_1 = py_1$ we obtain the equation

$$az_3^2 = apy_1^2 + z_2^2.$$

Reducing mod p we see that the equation (4.6.7) has no solution in $\mathcal{F}_{p^2}$ in view of the assumption that a is a quadratic non-residue in $\mathcal{F}_{p^2}$. This completes the verification of condition (1).

Verification of (2) - The symmetric bilinear pairing

$$(X, Y) \longrightarrow \mathrm{Tr}(XY)$$

is nondegenerate on the space of traceless matrices and the image of $SL(2, \mathbf{C})$ lies in the orthogonal group of this pairing. Since det is a bounded function on the orthogonal group (in fact, ± 1), the validity of (2) follows and we have established

Proposition 4.6.2 *1. Let a be a quadratic non-residue* mod p*, then the orbit space $\Gamma_{a,p}\backslash\mathcal{H}_2$ is compact.*

2. For $p \equiv 3$ mod 4 and $a \in \mathcal{O} = \mathbf{Z}[i]$ a quadratic non-residue in $\mathcal{F}_{p^2} = \mathcal{O}/(p)$, the orbit space $\tilde{\Gamma}_{a,p}\backslash\mathcal{H}_3$ is a compact manifold.

Remark 4.6.2 The essential point in the above proof was that there are no *unipotent* matrices (i.e., both eigenvalues are 1) in $\Gamma_{a,p}$ or $\tilde{\Gamma}_{a,p}$ (other than $\pm I$) or the lattice $L_\circ$ contained no non-zero nilpotent matrix. Note that the orbit of $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ approaches the origin under conjugation. On the other hand, the orbit $\{gXg^{-1}\}$, for X in (4.6.6, is closed. This is characteristic property of diagonalizable as opposed to nilpotent matrices. In the language of algebraic groups, the choice of the lattice $L_\circ$ defines a rational structure for the algebraic group $G = SL(2,\mathbf{C})$ or $G = SL(2,\mathbf{C}) \times SL(2,\mathbf{C})$ for which $\Gamma_{a,p}$ or $\tilde{\Gamma}_{a,p}$ is an arithmetic subgroup. The condition that a is a quadratic non-residue mod p or as an element of $\mathcal{F}_{p^2}$ implies that every element of $G_{\mathbf{Q}}$ is diagonalizable (over $\mathbf{C}$) which is a necessary and sufficient condition for the quotient of the real points of a semi-simple algebraic group defined over $\mathbf{Q}$ by an arithmetic subgroup to be compact (see [Bo]). ♡

Let F'_Γ be a relatively compact set in $\mathcal{H}_j$, containing a fundamental domain for the Γ. Then there are only finitely many $\gamma's$ such that

$$\gamma(\mathsf{F}'_\Gamma) \cap \mathsf{F}'_\Gamma \neq \emptyset.$$

Let Θ denote this finite set.

Lemma 4.6.13 *With the above notation and hypothesis, Θ is a set of generators for Γ.*

Proof - Let Γ' be the subgroup generated by Θ, $z \in \mathsf{F}'_\Gamma$, and $\gamma \in \Gamma$. Then there is $\delta \in \Gamma'$ such that $\delta\gamma(z) \in \mathsf{F}'_\Gamma$. Now this implies that $\delta\gamma \in \Gamma'$ and so $\Gamma' = \Gamma$. ♣

An immediate corollary is

Corollary 4.6.6 *The groups $\Gamma_{a,p}$ and $\tilde{\Gamma}_{a,p}$ are finitely generated.*

The assumption of compactness is not necessary for lemma 4.6.13 and corollary 4.6.6. The same conclusion is valid for general arithmetic groups. The above arithmetical construction of cocompact discrete subgroups of $SL(2,\mathbf{R})$ and $SL(2,\mathbf{C})$ can be generalized to give many other such subgroups. Our purpose here was merely to show that cocompact torsion free discrete subgroups can be easily constructed. One refers to the groups $\Gamma'_{\delta,p}$, $\Gamma_{a,p}$, $\tilde{\Gamma}_{a,p}$ and their subgroups of finite index as *quaternion groups.*

4.6.3 Poincaré's Construction

(THIS SUBSECTION IS NOT INCLUDED)

4.6.4 Closed Geodesics

There is extensive literature on the subject of closed geodesics on Riemannian manifolds. Since geodesics are solutions of systems of (non-linear) second order ordinary differential equations, in fact of Hamiltonian systems (see chapter 2, §Symplectic Manifolds), it is not surprising that this subject relates to the theory of dynamical systems. The problem of finding closed geodesics is that of finding periodic solutions of the corresponding system of differential equations. In this subsection we concentrate on compact manifolds of the form $M = \Gamma\backslash\mathcal{H}_j$, $j = 2, 3$, and Γ a discrete subgroup of the type explicitly constructed in the subsection on quaternion groups. Then the generally difficult problem of detecting periodic solutions is easily solved by elementary algebra.

Let $\Gamma \subset SL(2, \mathbf{R})$ or $SL(2, \mathbf{C})$ be a discrete subgroup such that $\Gamma\backslash\mathcal{H}_j$ is a compact manifold. Every $\gamma \in \Gamma$, $\gamma \neq \pm e$, is necessarily hyperbolic or loxodromic (corollary **??**). Then γ has two distinct fixed points (z_1, z_2) on $\mathbf{R}$ or $\mathbf{C}$ according as γ is hyperbolic or loxodromic. Let $\tilde{\gamma}$ be the axis of γ, i.e., the geodesic with end-points z_1 and z_2. Since the end-points of a geodesic invariant under γ remain fixed under γ, $\tilde{\gamma}$ is the unique geodesic in $\mathcal{H}_j$ invariant under γ. Thus for every $z \in \mathrm{Im}(\tilde{\gamma})$, $\gamma(z)$ also lies on $\mathrm{Im}(\tilde{\gamma})$, and the tangent vector to $\tilde{\gamma}$ at z is mapped to the tangent vector to $\tilde{\gamma}$ at $\gamma(z)$. Since a geodesic is completely determined by a point and the tangent vector to it at the point (i.e., the solution to a system of second order ordinary differential equations with given initial conditions), the image $\bar{\gamma}$ of $\tilde{\gamma}$ in $\Gamma\backslash\mathcal{H}_j$ is a closed geodesic. This is the key observation that makes it possible to explicitly exhibit closed geodesics on $\Gamma\backslash\mathcal{H}_j$.

Lemma 4.6.14 *To every non-identity element γ of the fundamental group of M corresponds a closed geodesic $\bar{\gamma}$ in M. Every closed geodesic on M corresponds to a non-identity element of the fundamental group.*

Remark 4.6.3 Note that the closed geodesics corresponding to γ and γ^n are the same except that the latter winds around n times the former. ♡

Proof - The fundamental group of M is isomorphic to $\Gamma/\pm e$ (if $-e \notin \Gamma$, the fundamental group is isomorphic to Γ). Therefore the above analysis assigns a unique closed geodesic to every non-identity element of the fundamental group. Conversely, given a closed geodesic $\bar{\gamma}$ in M, we lift it to a (complete) geodesic $\tilde{\gamma}$ in $\mathcal{H}_j$. Since $\bar{\gamma}$ is a closed geodesic, there is $\gamma \in \Gamma$ mapping $z \in \mathrm{Im}\tilde{\gamma}$ to $\gamma(z) \in \mathrm{Im}\tilde{\gamma}$ and mapping the tangent vector at z to $\tilde{\gamma}$ to the tangent vector at $\gamma(z)$ to $\tilde{\gamma}$. It follows that $\tilde{\gamma}$ is invariant under γ or equivalently $\tilde{\gamma}$ is the axis of γ. ♣

Having reduced the problem of detecting closed geodesics in M to an algebraic one, we address the issue of the length (or period) l_γ of the closed geodesic $\tilde{\gamma}$. Naturally l_γ is the distance traveled along image of $\tilde{\gamma}$ in M, starting at any point $z_\circ$, until the first return to $z_\circ$ with the tangents vectors also being identical. The geodesic in M may have many self-intersections. Let $\gamma \in \Gamma$ and $z \in \mathrm{Im}(\tilde{\gamma})$, then the length of the closed geodesic in $M = \Gamma\backslash\mathcal{H}_j$ corresponding to γ is

$$l_\gamma = \mathrm{dist}(z, \gamma(z)). \tag{4.6.8}$$

We have

Lemma 4.6.15 *Let β, β^{-1} be the eigenvalues of $\gamma \in \Gamma$, $\gamma \neq \pm e$. The length of the closed geodesic in M corresponding to γ is*

$$l_\gamma = \pm 2 \log |\beta|.$$

proof - We prove the assertion for the upper half space; the two dimensional case being a special case. If γ is diagonal (with diagonal entries β, β^{-1}) then the line $e^t\mathbf{j}$ is the geodesic invariant under γ. Consequently

$$l_\gamma = \mathrm{dist}(z, \gamma(z)) = \pm 2 \log |\beta|$$

as desired. The general case is reduced to the diagonal case by replacing γ with a digonal matrix $A\gamma A^{-1}$, $A \in SL(2,\mathbf{C})$, and noting that the quantities l_γ and β remain unchanged. ♣

Let us specialize to $M = \Gamma\backslash\mathcal{H}_2$. The eigenvalues $e^\alpha, e^{-\alpha}$ of a hyperbolic transformation $\gamma \in \Gamma \subset SL(2,\mathbf{R})$ are real, and after possibly replacing γ by $-\gamma$ we have

$$\mathrm{Tr}(\gamma) = e^\alpha + e^{-\alpha}.$$

It follows that

$$l_\gamma = 2\log(\mathrm{Tr}(\gamma)) + O(\frac{1}{\mathrm{Tr}(\gamma)^2}). \tag{4.6.9}$$

For a given positive number L, let $\vartheta_{a,p}(L)$ denote the number of possible lengths $< L$ of closed geodesics on $M = \Gamma_{a,p}\backslash\mathcal{H}_2$. Then $\vartheta_{a,p}(L)$ is approximately equal to the number of possible values of $\mathrm{Tr}(\gamma) < e^{L/2}$. Now $\gamma \in \Gamma_{a,p}$ is of the form

$$\gamma = \begin{pmatrix} x_\circ + x_1\sqrt{a} & x_2\sqrt{p} + x_3\sqrt{ap} \\ x_2\sqrt{p} - x_3\sqrt{ap} & x_\circ - x_1\sqrt{a} \end{pmatrix}$$

with $x_j \in \mathbf{Z}$. Let $\mathcal{N}(m)$ denote integers $n < m$ such that the diophantine equation

$$ax_1^2 + px_2^2 - apx_3^2 = n^2 - 1$$

has a solution $x_j \in \mathbf{Z}$. Since $\mathrm{Tr}(\gamma) = 2x_\circ$ our analysis proves

Proposition 4.6.3 *Let $p \equiv 1 \bmod 4$ and $a \in \mathbf{Z}$ be a quadratic non-residue mod p. Then*

$$\vartheta_{a,p}(L) = \mathcal{N}(\frac{1}{2}e^{L/2}) + O(1).$$

Exercise 4.6.3 *Let $p \equiv 3 \bmod 4$ and a be a quadratic non-residue* mod$p$. *Let $\Gamma^{\circ}_{a,p}$ be the subgroup of $\Gamma_{a,p}$ consisting matrices with $x_2 \equiv 0 \bmod a$ (see example* **??***). Formulate and prove the analogue of proposition 4.6.3 for $M = \Gamma^{\circ}_{a,p}\backslash\mathcal{H}_2$.*

Next we consider the three dimensional case $M = \Gamma\backslash\mathcal{H}_3$. Let $\beta = a + ib$ and without loss of generality assume $|\beta| > 1$ in lemma 4.6.15. Then

$$e^{l_\gamma} = a^2 + b^2 \sim |\mathrm{Tr}\gamma|^2. \tag{4.6.10}$$

Let $\tilde{\vartheta}(L)$ denote the number of lengths $< L$ of closed geodesics on $M = \tilde{\Gamma}_{a,p}\backslash\mathcal{H}_3$. Similarly, let $\tilde{\mathcal{N}}(m)$ denote the number of $n < m$ for which the system of diophantine equations

$$|z_\circ|^2 = n, \quad z_\circ^2 - az_1^2 - pz_2^2 + apz_3^2 = 1$$

have a solution $z_j \in \mathcal{O}$. Then the above analysis implies

Proposition 4.6.4 *Let $p \equiv 3 \bmod 4$ and $a \in \mathcal{O} = \mathbf{Z}[i]$ be a quadratic non-residue in $\mathcal{F}_{p^2} = \mathcal{O}/(p)$. Then*

$$\tilde{\vartheta}(L) = \tilde{\mathcal{N}}(\frac{1}{4}e^{L}) + O(1).$$

Propositions 4.6.3 and 4.6.4 (and exercise 4.6.3) relate the lengths of closed geodesics to number theory, and make it feasible to numerically investigate the distribution of lengths of closed geodesics.

4.6.5 Line Geometry of Hyperbolic Plane/Space

For a better understanding of geometry in the hyperbolic plane/space it is convenient to introduce some algebraic machinery. An application of this algebraic development is the Fenchel-Nielsen description of the moduli space of surfaces which is given in the next subsection.

Given a (complete) geodesic γ in $\mathcal{H}_i$, $i = 2, 3$, there is a unique one parameter subgroup $\{\exp(t\xi)\} \subset SL(2, \mathbf{C})$ leaving the geodesic invariant. In fact, for the positive y-axis, $\xi =$

$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and for other geodesics it is given by $g^{-1}\xi g$ for some $g \in SL(2,\mathbf{C})$ (or $g \in SL(2,\mathbf{R})$ if the geodesic is in $\mathcal{H}_2$.) Note also that for every $t \neq 0$, $\exp t\xi$ is loxodromic (or hyperbolic) and $\gamma = \gamma_{\exp t\xi}$. If the geodesic is of the form γ_A for a loxodromic or hyperbolic matrix A, then the corresponding one parameter group is

$$\exp(t\xi) \quad \text{with} \quad \xi = A - A^{-1}, \quad t \in \mathbf{R}. \tag{4.6.11}$$

Thus we have assigned to a geodesic γ in $\mathcal{H}_i$, $i = 2, 3$, a matrix $\xi \in \mathcal{SL}(2,\mathbf{R})$ or $\mathcal{SL}(2,\mathbf{C})$ which is unique up to multiplication by a nonzero scalar. This matrix is often called the *line matrix* of the geodesic γ. Clearly the end points of the geodesic γ_A are the fixed points of the line matrix $\xi = A - A^{-1}$. The line matrix of a loxodromic transformation is necessarily non-singular since

$$\det(A - A^{-1}) = 4 - (a + d)^2.$$

In particular, for a hyperbolic transformation $A \in SL(2,\mathbf{R})$ we have $\det(A - A^{-1}) < 0$. Thus by multiplying the line matrix ξ by a scalar matrix we can normalize ξ to have determinant -1. To make the line matrices unique we consider (complete) geodesics with orientations which is equivalent to assigning an order to their end points (i.e., fixed points). Given a line matrix $U = (u_{ij})$, traceless and of determinant -1, its fixed points are

$$z = \frac{u_{11} - 1}{u_{21}}, \quad z' = \frac{u_{11} + 1}{u_{21}}, \quad \text{if} \quad u_{21} \neq 0.$$

If $u_{21} = 0$, then $u_{11} = \pm 1$ and its fixed points are

$$z = \frac{-u_{12}}{2}, \; z' = \infty, \; \text{if } u_{11} = 1; \quad z = \infty, \; z' = \frac{u_{12}}{2}, \; \text{if } u_{11} = -1.$$

Thus we have represented the end points as an ordered pair (z, z'). Notice that the assigned orientation is compatible with the action of $SL(2,\mathbf{C})$ (or $SL(2,\mathbf{R})$). This means that if the oriented geodesic γ has line matrix U and ordered set of end points (z, z') according to the above rule, then the line matrix of the oriented geodesic $g(\gamma)$ is gUg^{-1} and its ordered set of end points is $(g(z), g(z'))$. With this normalization we have a bijection between the set of oriented complete geodesics and normalized line matrices. The normalized line matrix of a loxodromic (or hyperbolic) transformation A will be denoted by L_A. This matrix representation of geodesics is quiet useful in understanding the geometry of polygons in the hyperbolic plane and space. Let us prove some of the properties of this correspondence:

Lemma 4.6.16 *Let $A \in SL(2, \mathbf{C})$ be a loxodromic transformation and $\eta \in \mathcal{SL}(2, \mathbf{C})$ be the line matrix of a geodesic δ. The line matrix $A - A^{-1}$ is traceless and nonsingular, and δ is orthogonal to γ_A if and only if* $\mathrm{Tr}(A\eta) = 0$.

Proof - We may assume A is diagonal with diagonal entries a and a^{-1}. Then γ_A is the line $t\mathbf{j}$ for the hyperbolic space and ti for the hyperbolic plane where $t > 0$. Writing $\eta = (\eta_{jk})$, the condition $\mathrm{Tr}(A\eta) = 0$ becomes $(a - \frac{1}{a})\eta_{11} = 0$. Therefore $\eta_{11} = 0 = \eta_{22}$. The geodesic corresponding to the line matrix η is the half circle joining the points $\pm\sqrt{\frac{\eta_{12}}{\eta_{21}}}$, and therefore intersects γ_A orthogonally. ♣

Let γ_A and γ_B be the axes of loxodromic transformations A and B. We say γ_A and γ_B are *disjoint* if their closures in $\mathcal{H}_3 \cup \mathbf{C} \cup \infty$ are disjoint. An immediate consequence of the lemma is

Corollary 4.6.7 *Let A and B be loxodromic transformations and assume that the corresponding axes γ_A and γ_B are disjoint. Then γ_A and γ_B have a unique common normal and $AB - BA$ is a line matrix for it.*

Note that two intersecting geodesics cannot have a common normal since otherwise we obtain a triangle with angle sum $> \pi$. By the corollary the common normal of two disjoint geodesics is unique, and we denote by $\lambda = \lambda(A, B)$ the signed length of the segment of the common normal of γ_A and γ_B which connects them. The sign $\pm$ depends on the orientation as explained below.

Corollary 4.6.8 *Let A and B be loxodromic (or hyperbolic) transformations with disjoint axes γ_A and γ_B. Let L_C be the normalized line matrix of their common normal. Then*

$$\cosh\lambda = \frac{1}{2}\mathrm{Tr}(L_A L_B), \qquad \sinh\lambda = \frac{1}{2}\mathrm{Tr}(L_A L_B L_C),$$

where $\lambda = \lambda(A, B)$ and in the second equation orientations are incorporated (see below for explanation).

Proof - We may assume that the line matrices are

$$L_A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad L_B = \begin{pmatrix} 0 & e^s \\ e^{-s} & 0 \end{pmatrix}, \quad L_C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and the required result follows by direct calculation. ♣

The quantity $\lambda = \lambda(A,B)$ appearing in the statement of the corollary 4.6.8 is the length of the geodesic segment together with a sign. Let L_A, L_B and L_C be as in the proof of corollary 4.6.8. From the description of the orientation attached to a geodesic described by its line matrix it is clear that the ordered pair of fixed points for the line matrix L_A is $(-1,1)$, for L_B is $(-e^s,e^s)$, and for L_C is $(0,\infty)$. For $s>0$ if we move along the three geodesics according to the directions determined by the ordered pairs of end-points, then the sign is negative. Reversing the orientation of a geodesic which is expressed by replacing a line matrix by its negative, will multiply λ by -1.

Exercise 4.6.4 *Let L_i be the normalized line matrices of oriented geodesics γ_i, $i=1,2,3$, passing through a given point $p \in \mathcal{H}_3$. Show that the unit tangents to γ_1, γ_2 and γ_3 form a positively oriented orthonormal frame if and only if*

$$L_3L_2L_1 = I.$$

Exercise 4.6.5 *Let ξ and η be the line matrices of two (complete) geodesics on $\mathbf{H}_2$. Show that the geodesics γ_ξ and γ_η intersect in $\mathcal{H}_2$ (resp. do not intersect) if and only if* $\det([\xi,\eta]) > 0$ *(resp.* $\det([\xi,\eta]) < 0$*). Prove also that* $\mathrm{rank}([\xi,\eta]) = 1$ *is a necessary and sufficient condition for γ_ξ and γ_η to have exactly one common end-point.*

Exercise 4.6.6 *Let ξ and η be the line matrices of two (complete) geodesics on $\mathbf{H}_3$. Show that the geodesics γ_ξ and γ_η intersect transversally if and only if* $\det([\xi,\eta])$ *is real and positive. Prove also that* $\mathrm{rank}([\xi,\eta]) = 1$ *is a necessary and sufficient condition for γ_ξ and γ_η to have exactly one common end-point.*

By a *polygon* or *n-gon* in $\mathcal{H}_2$ we mean a compact region P bounded by n geodesic arcs. In general these geodesic arcs may intersect and P does not necessarily lie on one side of the every bounding (complete) geodesic (see Figure XXXX). An n-gon P which lies on one side of every bounding (complete) geodesic is called *convex*. *Hexagons* (i.e., regions bounded by six geodesic arcs) in the hyperbolic plane are especially significant in studying geometry of surfaces. By a *right n-gon* we mean one all whose angles are $\frac{\pi}{2}$. The following algebraic lemma is useful:

Lemma 4.6.17 *Let L_i, $i=1,\cdots,6$ denote the normalized line matrices of the sides S_i of the right hexagon $S \subset \mathcal{H}_2$. We assume S_i is adjacent to $S_{i\pm 1}$ (i mod 6). Then*

$$L_i^2 = I, \quad L_iL_{i+1} = -L_{i+1}L_i, \quad \mathrm{Tr}(L_4L_3L_2)\mathrm{Tr}(L_3L_2L_1) = -2\mathrm{Tr}(L_4L_1).$$

Proof - The first two relations follow by direct calculation since we may assume L_i and L_{i+1} are of the form L_B and L_C given in the proof of corollary 4.6.8. For 2×2 matrices A and B we have

$$\mathrm{Tr}(AB) + \mathrm{Tr}(\check{A}B) = \mathrm{Tr}(A)\mathrm{Tr}(B), \tag{4.6.12}$$

where $\check{A}$ denotes the matrix of cofactors of A. For a traceless 2×2 matrix A of determinant -1 we have $\check{A} = A$, and the matrix of cofactors of AB is $\check{B}\check{A}$. Therefore

$$\mathrm{Tr}(L_4L_3L_2)\mathrm{Tr}(L_3L_2L_1) = \mathrm{Tr}(L_4L_3L_2L_3L_2L_1) + \mathrm{Tr}(L_2L_3L_4L_3L_2L_1).$$

Applying the first two relations we obtain the desired result. ♣

Corollary 4.6.9 (Law of sines of a right hexagon) *Let S be a right hexagon with sides S_i and corresponding normalized line matrices L_i. We assume S_i is adjacent to $S_{i\pm1}$*(mod6), *and denote the signed length of S_i by λ_i. Then*

$$\frac{\sinh\lambda_1}{\sinh\lambda_4} = \frac{\sinh\lambda_3}{\sinh\lambda_6} = \frac{\sinh\lambda_5}{\sinh\lambda_2}.$$

Proof - Let λ_{14} denote the signed length of the common normal connecting the sides S_1 and S_4. Then in view of corollary 4.6.8 and the third relation of lemma 4.6.17 we have

$$\cosh\lambda_{14} = -\sinh\lambda_2\sinh\lambda_3.$$

Similarly,

$$\cosh\lambda_{14} = -\sinh\lambda_5\sinh\lambda_6.$$

The required identity follows immediately. ♣

Corollary 4.6.10 (Law of cosines of a right hexagon) *Let S be a right hexagon with sides S_i and corresponding normalized line matrices L_i. We assume S_i is adjacent to $S_{i\pm1}$*(mod6), *and denote the signed length of S_i by λ_i. Then*

$$\cosh\lambda_n = -\cosh\lambda_{n-2}\cosh\lambda_{n+2} + \sinh\lambda_{n-2}\sinh\lambda_{n+2}\cosh\lambda_{n+3},$$

Proof - Using lemma 4.6.17, equation (4.6.12) and with the same notation and reasoning similar to the proof of corollary 4.6.9 we have

$$\begin{aligned}
\mathrm{Tr}(L_5L_4L_3)\mathrm{Tr}(L_4L_2)\mathrm{Tr}(L_3L_2L_1) &= \mathrm{Tr}(L_5L_4L_3)[\mathrm{Tr}(L_4L_2L_3L_2L_1) + \mathrm{Tr}(L_2L_4L_3L_2L_1)] \\
&= -2\mathrm{Tr}(L_5L_4L_3)\mathrm{Tr}(L_4L_3L_1) \\
&= -2\mathrm{Tr}(L_5L_4L_3L_4L_3L_1) - 2\mathrm{Tr}(L_3L_4L_5L_4L_3L_1) \\
&= 2\mathrm{Tr}(L_5L_1) + 2\mathrm{Tr}(L_3L_5L_3L_1) \\
&= 4\mathrm{Tr}(L_5L_1) + 2\mathrm{Tr}(L_3L_5)\mathrm{Tr}(L_3L_1).
\end{aligned}$$

In view of corollary 4.6.8 this identity implies the required result for $n = 6$. Cyclic permutation of indices $1, 2, \cdots, 6$ yields the desired result. ♣

Example 4.6.4 Let $a < 0 < b$ be real numbers and γ be the unique geodesic with end points (a,b) (as ordered pair). It is a simple calculation that the normalized line matrix of γ is

$$L_{a,b} = \begin{pmatrix} \frac{a+b}{b-a} & \frac{2ab}{a-b} \\ \frac{2}{b-a} & \frac{a+b}{a-b} \end{pmatrix}$$

The line matrix of the imaginary axis is $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and therefore

$$\mathrm{Tr}(L_{a,b}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}) = 2\frac{b+a}{b-a}. \tag{4.6.13}$$

Computing the angle of intersection θ of γ and the imaginary axis we obtain

$$\cot\theta = \frac{a+b}{2\sqrt{-ab}}.$$

Comparing with (4.6.13) we deduce that if two geodesics γ and δ intersect at an angle θ then

$$\cos\theta = \frac{1}{2}\mathrm{Tr}(L_\gamma L_\delta) \tag{4.6.14}$$

where L_γ and L_δ are the corresponding (normalized) line matrices. It is clear that the conclusion is valid for $\mathcal{H}_3$ as well as for the hyperbolic plane. ♠

Example 4.6.5 While corollary 4.6.10 was about right hexagons, its proof together with example 4.6.4 imply a law of cosines for the pseudo-right pentagons and pseudo-right quadrilaterals as well. In this and the following example we make this point precise. By a *pseudo-right pentagon* we mean a pentagon (with sides geodesic segments) such that four of its five angles are $\frac{\pi}{2}$. We denote the non-right angle by θ. For definiteness let L_i denote the normalized line matrix of side S_i and assume θ is the angle between S_1 and S_5. An examination of the proofs of corollary 4.6.10 and lemma 4.6.17 shows that we have

$$\frac{1}{2}\mathrm{Tr}(L_5L_1) = -\cosh\lambda_4\cosh\lambda_2 + \sinh\lambda_4\sinh\lambda_2\cosh\lambda_3. \tag{4.6.15}$$

Now S_1 and S_5 intersect and therefore they do not have a common normal, however, example 4.6.4 is applicable to yield

$$\cos\theta = -\cosh\lambda_4\cosh\lambda_2 + \sinh\lambda_4\sinh\lambda_2\cosh\lambda_3. \tag{4.6.16}$$

This is the *law of cosines for pseudo-right pentagons.* ♠

Exercise 4.6.7 *Construct convex right pentagons in $\mathcal{H}_2$.*

Example 4.6.6 The case of a pseudo-right quadrilateral is simpler. By a pseudo-right quadrilateral we mean a 4-gon, with sides geodesic arcs, such that three of its four angles are $\frac{\pi}{2}$. Obviously there are no right quadrilaterals in $\mathcal{H}_2$ since the sum of the angles of a triangle in the hyperbolic plane is $< \pi$. Let $\theta < \frac{\pi}{2}$ be the angle between the sides S_1 and S_4. An examination of the proof of lemma 4.6.17 and corollary 4.6.8 show that the third identity of the lemma implies

$$\frac{1}{2}\mathrm{Tr}(L_4L_1) = -\sinh\lambda_2\sinh\lambda_3,$$

where as usual orientations are incorporated to obtain correct signs. Just as in example 4.6.5, the left hand side of the above eqaution is $\cos\theta$ and therefore we obtain

$$\cos\theta = -\sinh\lambda_2\sinh\lambda_3. \tag{4.6.17}$$

This is the *law of cosines for pseudo-right quadrilaterals.* ♠

Example 4.6.7 The law of cosines for a right hexagon in the hyperbolic plane (corollary 4.6.10) gives rise to an interesting algebraic identity which one can verify directly. Motivated by this law we let u_1, u_3 and u_5 be indeterminates and define v_2, v_4 and v_6 as

$$v_2 = \frac{u_5 + u_1u_3}{\sqrt{(u_1^2-1)(u_3^2-1)}},\quad v_4 = \frac{u_1 + u_3u_5}{\sqrt{(u_3^2-1)(u_5^2-1)}},\quad v_6 = \frac{u_3 + u_1u_5}{\sqrt{(u_1^2-1)(u_5^2-1)}}.$$

The remarkable fact is that u_1, u_3 and u_5 are defined by similar formulae in terms of v_2, v_4 and v_6, namely,

$$u_1 = \frac{v_4 + v_2v_6}{\sqrt{(v_2^2-1)(v_6^2-1)}},\quad u_3 = \frac{v_6 + v_2v_4}{\sqrt{(v_2^2-1)(v_4^2-1)}},\quad u_5 = \frac{v_2 + v_4v_6}{\sqrt{(v_4^2-1)(v_6^2-1)}}.$$

The validity of these formulae is established by direct substitution or by using a symbolic manipulation software. The algebraic validity of these relations is in fact useful in understanding the geometry of right hexagons in the hyperbolic plane. ♠

We can now prove

Proposition 4.6.5 *Given positive real numbers α, β and γ, there is a unique (up to an isometry of the hyperbolic plane) convex right hexagon, such that*

$$\lambda_1 = \alpha,\quad \lambda_3 = \beta,\quad \lambda_5 = \gamma.$$

Here λ_i denotes the length of the side S_i of the hexagon and S_i is adjacent to $S_{i\pm1}(\mathrm{mod} 6)$.

Proof - Uniqueness follows easily from the law of cosines for hexagons and the fact that all angles are $\frac{\pi}{2}$. Using the law of cosines for a right hexagon in the hyperbolic plane (corollary 4.6.10) we determine λ_2, λ_4 and λ_6. Let S_1 be a geodesic segment of length λ_1. Draw a geodesic segment S_2 of length λ_2 from an end point of S_1 and orthogonal to S_1. Similarly draw a segment of length λ_3 from the other end point of S_2 orthogonal to S_2. There is an ambiguity in drawing S_3 since the (complete) geodesic containing S_2 disconnects the hyperbolic plane into two components. The choice should be such that S_1 and S_3 lie on the same side of S_2. One proceeds in the obvious way to construct the sides $S_1, \cdots, S_6$. It is necessary to show that the construction can be carried out in such a way that for every i, the sides S_j, $j \neq i$, lie on the same side of the complete geodesic containing S_i and that we obtain a hexagon. Since S_1 and S_3 have a common normal, namely S_2, they do not intersect. Assume S_4 and S_1 intersect. Then example 4.6.6 is applicable and in particular (4.6.17) is valid. Substituting for $\sinh \lambda_2$ we obtain

$$\cos^2 \theta = \frac{(\cosh \lambda_5 + \cosh \lambda_1 \cosh \lambda_3)^2}{\sinh^2 \lambda_1} > 1,$$

which is impossible. Therefore the (complete) geodesics containing S_1 and S_4 do not intersect and S_1, S_2 and S_3 lie on the same side of the (complete) geodesic containing S_4. Now assume the (complete) geodesics containing S_1 and S_5 intersect. Then example 4.6.5 is applicable and in particular (4.6.16) is valid. In view of example 4.6.7 we can substitute in (4.6.16) for $\cosh \lambda_3$ in terms of $\cosh \lambda_{2i}$, $(i = 1, 2, 3)$, to obtain

$$\cos^2 \theta = \cosh^2 \lambda_6 > 1,$$

which is impossible. Therefore the (complete) geodesics containing S_1 and S_5 do not intersect and the condition regarding the sides S_j, $j \neq i$, lying in the same side is fulfilled. The common normal of S_1 and S_5 necessarily has length λ_6. ♣

Exercise 4.6.8 *Show that α, μ and β are the lengths of three consecutive sides of a right hexagon if and only if*

$$\cosh \mu > \frac{1 + \cosh \alpha \cosh \beta}{\sinh \alpha \sinh \beta}.$$

4.6.6 Decomposition of Surfaces

Let M be a compact orientable surface of genus $g > 1$. By cutting M along $3g - 3$ simple closed curves as indicated in Figure XXXX we can decompose M into $2g - 2$ surfaces with boundary each diffeomorphic to S^2 with three (small) discs removed. Such a surface will be called a *pantaloon*[6]. The choice of simple closed curves to obtain such a decomposition is not unique not even up to free homotopy. However, once a choice is made, say $\gamma'_1, \cdots, \gamma'_{3g-3}$, it is possible to uniquely choose simple closed geodesics $\gamma_1, \cdots, \gamma_{3g-3}$ such that γ_i is freely homotopic to γ'_i as shown in lemma 4.6.18 below.

Lemma 4.6.18 *Let M be a compact orientable surface of genus $g > 1$, and δ be a simple closed curve representing a nontrivial element of $\pi_1(M; p)$. Then there is a unique simple closed geodesic γ freely homotopic to δ.*

Proof - We have $\delta : I \rightarrow M$ with $\delta(0) = \delta(1) = p$. The universal cover of M is the hyperbolic plane $\mathcal{H}$ and $\pi_1(M; p)$ is identified with a discrete subgroup $\Gamma \subset SL(2, \mathbf{R})$. Therefore the loop δ is identifed with an element $A \in \Gamma$. Let $\tilde{p} \in \mathcal{H}$ be any point lying over $p \in M$. Then the image γ' of the unique geodesic segment in $\mathcal{H}$ joining $\tilde{p}$ to $A(\tilde{p})$ in M is homotopic to δ, however the end points may have distinct tangents and therefore γ' may not be smooth. In view of proposition **??**, $A \in \Gamma$ is necessarily hyperbolic and let γ_A be its axis. Let $q \in \gamma_A$ be any point, then the image of the segment of γ_A joining q to $A(q)$ is the desired simple closed curve. Q E D

From now on we assume that the decomposition of the surface M into pantaloons is accomplished by cutting along simple closed geodesics, and regard a pantaloon as a surface of constant curvature -1 with boundary consisting of three simple closed geodesics. As noted earlier, the universal cover of a compact orientable surface of genus $g > 1$ is the upper half plane H. However, the $4g$-gon in the proof of the topological uniformization theorem is far from unique and there may be many diffeomorphic surfaces which are not isometric relative to the metric induced from the upper half plane. Our goal now is to show how the decomposition into pantaloons enables one to gain some understanding of the space of compact surfaces of a given genus under isometric equivalence relative to the Poincaré metric.

Let P be a pantaloon and P' be its double. Then we can attach P and P' together along the corresponding boundaries to obtain a compact surface N of genus two which admits of a metric of constant negative curvature -1. The boundary curves represent non-trivial free homotopy classes in the fundamental group of N. The universal covering space of N is the upper half plane and its fundamental group $\pi_1(N; p)$ acts on $\mathcal{H}_2$ as a discontinuous group of fractional linear transformations. Let γ_i, $i = 1, 2, 3$, be simple closed geodesics in N representing the boundary curves of P as given in lemma 4.6.18, and $\tilde{\gamma}_i$'s be their respective lifts to complete geodesics in $\mathcal{H}$.

[6]In view of the discrepancy between the American and British usages of *pants* we use pantaloon rather than a pair of pants to describe such a surface.

Lemma 4.6.19 *Let M be a compact orientable surface of constant negative curvature -1, γ and δ two distinct non-intersecting closed geodesics on M. Let $\tilde{\gamma}$ and $\tilde{\delta}$ be their lifts to the upper half plane. Then $\tilde{\gamma}$ and $\tilde{\delta}$ are disjoint.*

Proof - It suffices to show that $\tilde{\gamma}$ and $\tilde{\delta}$ do not have a common end-point. Two geodesics in the upper half plane having one common end-point can be transformed to two vertical lines. Since the hyperbolic distance between two such geodesics tends to zero as one moves to infinity, they cannot project to two disjoint closed geodesics in M. ♣

By lemma 4.6.19 $\tilde{\gamma}_i$ and $\tilde{\gamma}_j$ are disjoint and therefore there is a unique geodesic $\tilde{\nu}_{ij}$ normal to both $\tilde{\gamma}_i$ and $\tilde{\gamma}_j$. The image ν_{ij} of $\tilde{\nu}_{ij}$ in $P \subset N$ is a geodesic orthogonal to the boundary geodesics γ_i and γ_j and it is the unique geodesic in P orthogonal to both boundary components. We refer to ν_{ij} as a *seam* of the pantaloon P. P has three seams.

Lemma 4.6.20 *A pantaloon is determined up to isometry by the lengths of its boundary closed geodesics.*

Proof - Let λ_i, $i = 1,2,3$ be the lengths of the boundary geodesics of the pantaloon P. Cutting the pantaloon P along its three seams, we obtain two right hexagons Q_1 and Q_2 whose alternate sides (corresponding to the seams along which we cut) are equal (see Figure XXXX). By proposition 4.6.5 Q_1 and Q_2 differ by an isometry of the upper half plane. Consequently the three remaining alternate sides of P_i have lengths $\frac{\lambda_1}{2}$, $\frac{\lambda_2}{2}$ and $\frac{\lambda_3}{2}$. Invoking proposition 4.6.5 we see that Q_1 and Q_2 and therefore P are determined uniquely up to an isometry by λ_1, λ_2 and λ_3. ♣

The proof of lemma 4.6.20 and lemma 4.6.18 imply

Corollary 4.6.11 *Every pantaloon can be give a metric of constant negative curvature -1 such that the boundary curves are geodesics.*

Given a decomposition of M into pantaloons, $M = \cup_{a=1}^{2g-2} P_a$, we assign to M, relative to this decomposition, $3g-3$ real numbers $\log \lambda_j$ where $j = 1, \cdots, 3g-3$ and λ_j's are lengths of the simple closed geodesics determining the decomposition of M into pantaloons. Now given $2g-2$ pantaloons P_a, $a = 1, \cdots 2g-2$ we endows the pantaloons with metrics of constant curvature -1 with the boundary of P_a consisting of three simple closed geodesics γ_{a1}, γ_{a2} and γ_{a3} of lengths $\lambda_{a1}, \lambda_{a2}$ and λ_{a3} respectively (see corollary 4.6.11). We assume that a pairing between the boundary curves of the pantaloons is given so that for every pair (a,i) there is a unique pair $\jmath(a,i) = (b,j)$ with the properties $a \neq b$ and $\lambda_{ai} = \lambda_{bj}$ is specified. This correspondence allows one to attach the pantaloons together to obtain a compact orientable surface of genus g. However, the attaching process is not unique. To understand this let the seams of the pantaloon P_a be denoted by σ_{ak} so that σ_{ai} is the normal to the boundary curves γ_{ai} and γ_{ai+1}, $i, i+1$ mod 3. In attaching the pantaloon P_a to pantaloon P_b along the curves γ_{ai} and γ_{bj} where $\jmath(a,i) = (b,j)$, the seams σ_{ai} and σ_{bj} do not necessarily line up. We denote the angle between them by θ_{ai} or θ_{bj} where $\jmath(a,i) = (b,j)$. In the definition of θ_{ai} there is a sign ambiguity which we can resolve by fixing orientations for the pantaloons. We assign an arbitrary orientation to one pantaloon, say to P_1, then assign orientations to P_a, P_b and P_c, where $\jmath(1,1) = (a,i)$, $\jmath(1,2) = (b,j)$ and $\jmath(1,3) = (c,k)$ so that the induced orientations on the corresponding boundaries cancel. It is clear that we can continue the process until we have assigned definite orientations on all the $2g-2$ pantaloons. Now let θ_{ai} be measured by moving in the direction of the induced orientation on the boundary geodesic γ_{ai} by P_a. This defines θ_{ai} unambiguously and $\theta_{ai} = \theta_{bj}$ if $\jmath(a,i) = (b,j)$.

Now note that in attaching P_a to P_b along the boundary geodesics γ_{ai} and γ_{bj} we can choose the angle $\theta_{ai} = \theta_{bj}$ arbitrarily and the resulting surface $P_a \cup P_b$ will still carry a metric of constant curvature -1. This is due to the fact that motion along the boundary geodesic γ_{ai} is given by a transformation $\exp(t\xi)$ where $\xi \in \mathcal{SL}(2,\mathbf{R})$ which leaves the Poincaré metric invariant. Therefore we have assigned to a compact orientable surface M of genus $g > 1$ and constant negative curvature -1, together with a fixed decomposition $M = \cup P_a$ into pantaloons, $6g-6$ real numbers $(\log\lambda_{ak}, \theta_{ak})$. The real numbers $(\log\lambda_{ak}, \theta_{ak})$ are called the *Fenchel-Nielsen coordinates* of the compact orientable surface M (with fixed decomposition into pantaloons). Two compact orientable surfaces of genus g together Riemannian metrics of constant negative curvature -1 are considered equivalent if there is an isometry of one onto the other. With this notion of equivalence of surfaces we can restate the above considerations as

Corollary 4.6.12 *The space of compact orientable surfaces of genus $g > 1$ together with a Riemannian metric of constant negative curvature -1 depends on $6g-6$ real parameters.*

We can also consider isometric equivalence classes of pataloons together with Riemannian metrics of constant negative curvature -1 such that the boundary curves are geodesics. To each such pantaloon we assign the triple $(\log\lambda_1, \log\lambda_2, \log\lambda_3) \in \mathbf{R}^3$. With this notion of equivalence the analysis of the geometry of pantaloons or right hexagons in $\mathcal{H}_2$ implies

Corollary 4.6.13 *The space of pantaloons together with Riemannian metrics of constant negative curvature such that the boundary curves are geodesics is the quotient of $\mathbf{R}^3$ under the permutation action of $\mathcal{S}_3$.*

4.7 Some Algebra and its Applications

4.7.1 Schanuel's Lemma and Alexander Polynomials

In the proof of well-definedness of Alexander polynomials, The principal issue is gaining some understanding of linear algebra over a ring $\mathcal{R}$ which is not necessarily a principal ideal domain. Lemmas 4.7.1 and 4.7.2 below, are valid under very general hypothesis on the ring $\mathcal{R}$ and it is not necessary to assume $\mathcal{R} = \mathbf{Z}[t, t^{-1}]$. The proof of the following critical observation, known as *Schanuel's Lemma*, is reminiscent of the construction of the pull-back of a fibre bundle.

Lemma 4.7.1 *Let A and B be $n \times n$ matrices with entries from $\mathcal{R}$. Assume that the $\mathcal{R}$-modules M and N, defined by the following exact sequences, are isomorphic:*

$$0 \rightarrow \mathcal{R}^n \xrightarrow{A} \mathcal{R}^n \xrightarrow{\pi_A} M \rightarrow 0, \quad 0 \rightarrow \mathcal{R}^n \xrightarrow{B} \mathcal{R}^n \xrightarrow{\pi_B} N \rightarrow 0.$$

Then $\mathrm{Im}(A) \oplus \mathcal{R}^n \simeq \mathrm{Im}(B) \oplus \mathcal{R}^n$.

Proof - Let $\phi : M \rightarrow N$ be an isomorphism, $L \subset \mathcal{R}^n \oplus \mathcal{R}^n$ be the submodule defined as

$$L = \{(x, y) \in \mathcal{R}^n \oplus \mathcal{R}^n \mid \phi\pi_A(x) = \pi_B(y)\}.$$

Since $\mathcal{R}^n$ is free, there is an $\mathcal{R}$-module homomorphism $h_A : \mathcal{R}^n \rightarrow \mathcal{R}^n$ such that $\pi_B \cdot h_A = \phi \cdot \pi_A$, or equivalently the following diagram commutes:

$$\begin{array}{ccc} \mathcal{R}^n & \xrightarrow{\pi_A} & M \\ h_A \downarrow & & \downarrow \phi \\ \mathcal{R}^n & \xrightarrow{\pi_B} & N \end{array}$$

Similarly one defines h_B. Consider the homomorphisms $\psi_A : \mathrm{Im}(A) \oplus \mathcal{R}^n \rightarrow L$ and $\psi_B : \mathrm{Im}(B) \oplus \mathcal{R}^n \rightarrow L$ defined by

$$\psi_A(A(u), y) = (A(u) + h_B(y), y), \quad \psi_B(B(v), x) = (x, B(v) + h_A(x)).$$

Injectivity of ψ_A and ψ_B is clear and surjectivity follows from the exactness hypotheses. ♣

Exercise 4.7.1 *Show, by means of an example, that the homomorphisms h_A and h_B in the above proof need not be isomorphisms.*

The proof of Schanuel's Lemma gives us $(2n) \times (2n)$ matrices

$$\Psi_A \leftrightarrow \begin{pmatrix} A & h_B \\ 0 & I \end{pmatrix}, \quad \Psi_B \leftrightarrow \begin{pmatrix} I & 0 \\ h_A & B \end{pmatrix}$$

Let $\mathcal{J}_k^A$ denote the ideal generated in $\mathcal{R}$ by the $(2n-k) \times (2n-k)$ minors of the matrix Ψ_A. Since $\mathrm{Im}(\psi_A) = L = \mathrm{Im}(\psi_B)$, the columns and rows of Ψ_A are linear combinations of those of Ψ_B and vice versa. It follows that

$$\mathcal{J}_k^A = \mathcal{J}_k^B. \tag{4.7.1}$$

Let $\mathcal{I}_k^A$ denote the ideal generated in $\mathcal{R}$ by the $(n-k) \times (n-k)$ minors of A. It is clear that

$$\mathcal{J}_k^A = \mathcal{I}_k^A, \qquad \mathcal{J}_k^B = \mathcal{I}_k^B. \tag{4.7.2}$$

Equations (4.7.1) and (4.7.2) imply

Lemma 4.7.2 *Let A and B be $n \times n$ matrices with entries from $\mathcal{R}$ such that* $\mathrm{Coker}(A) \simeq \mathrm{Coker}(B)$ *as $\mathcal{R}$-modules. Then $\mathcal{I}_k^A = \mathcal{I}_k^B$.*

An immediate consequence of lemma 4.7.2 is

Corollary 4.7.1 *The Alexander ideals and polynomial are well-defined.*

Proof - In order to deduce the required result from lemma 4.7.2 it suffices to remove the restriction that the matrices A and B have the same size. If A is $n \times n$ and B is $m \times m$ and $m < n$ then replace B by $\begin{pmatrix} I & 0 \\ 0 & B \end{pmatrix}$ to make it $n \times n$. The hypothesis on isomorphisms of cokernels is simply the fact that the cokernels are $H_1(\tilde{S}_{K,ab}; \mathbf{Z})$ as an $\mathcal{R}$-module. ♣

Let $A : \mathcal{R}^n \rightarrow \mathcal{R}^n$ be an $n \times n$ matrix with non-zero determinant $\det(A) = t^s(a_\circ + a_1 t + \cdots + a_l t^l)$. We assume $a_\circ a_l \neq 0$. The implication of the following algebraic lemma about $H_1(\tilde{S}_{K,ab}; \mathbf{Z})$ is given in the corollary 4.7.2 below:

Lemma 4.7.3 *With the above notation assume* $\gcd(a_\circ, a_1, \cdots, a_l) = 1$. *Then* $\mathrm{Coker}(A)$ *has no torsion relative to* $\mathbf{Z}$, *i.e., if $b \in \mathbf{Z}$, $z \in \mathcal{R}^n$ and $bz \in \mathrm{Im}(A)$, then $z \in \mathrm{Im}(A)$.*

Proof - Writing $b\zeta = A\eta$, where ζ and η are column vectors, we obtain $bA^\star\zeta = \det(A)\eta$ where $A^\star$ denotes the adjoint of the matrix A. Let p be a prime and $q = p^f$ the highest p^{th} power dividing b. Denote the components of the vector η by $\eta_i(t) = t^{r_i}(\eta_{i\circ} + \eta_{i1}t + \cdots + \eta_{is}t^s)$. By a familiar algebraic argument, $\gcd(a_\circ, a_1, \cdots, a_m) = 1$ implies all the coefficients η_{ij} are divisible by $q = p^f$. Therefore $\eta = b\xi$ with $\xi \in \mathcal{R}^n$, and consequently $b\zeta = bA\xi$ and $\xi \in \mathrm{Im}(A)$. ♣

Corollary 4.7.2 *The homology group* $H_1(\tilde{S}_{K,ab};\mathbf{Z})$ *has no torsion.*

Proof - Since the Alexander polynomial of $\Delta(t)$ satisfies $\Delta(1) = \pm 1$, the hypothesis of lemma 4.7.3 are fulfilled, and the required result follows. ♣

The structure of $H_1(\tilde{S}_{K,ab};\mathbf{Z})$ can be determined if the constant term $\Delta_K(0) = \pm 1$. We obtain a slightly stronger result. Let $c = \Delta_K(0)$ and $\mathbf{Z}[\frac{1}{c}]$ be the ring of rational numbers whose denominators are powers of c. In view of the symmetry property of the Alexander polynomial $\Delta_K(t)$, the coefficients of the highest order and lowest order terms are c. Recall $\Delta_K(t) = \det(A)$ where $A : \mathcal{R}^n \to \mathcal{R}^n$ and we may also regard A as an $\mathcal{R}_c$-module mapping of $\mathcal{R}_c^n$ to itself. Let l be the degree of Δ_K.

Lemma 4.7.4 *The monomials* $1, t, t^2, \cdots, t^{l-1}$ *form a set of generators form* $\mathrm{Coker}(A)$ *as a* $\mathbf{Z}[\frac{1}{c}]$*-module.*

Proof - In view of the identity

$$A^\star A = AA^\star = \det(A)I, \tag{4.7.3}$$

where $A^\star$ denotes the adjoint of the matrix A, $\det(A)$ annihilates $\mathrm{Coker}(A)$. We can write

$$ct^{-1} = \frac{1}{t}\Delta_K(t) - \sum_{j\circ}^{l-1} a_j t^j.$$

Since c and t are units in $\mathcal{R}_c$ and $\Delta_K(t)$ annihilates $\mathrm{Coker}(A)$, image of t^{-1} in $\mathrm{Coker}(A)$ is in the span of $1, t, t^2, \cdots, t^{l-1}$. Similarly, t^l lies in the span of $1, t, t^2, \cdots, t^{l-1}$. The required result follows easily. ♣

In view of lemma 4.7.4 and corollary 4.7.2, $\mathrm{Coker}(A)$ is a finitely generated free $\mathbf{Z}[\frac{1}{c}]$-module with the structure of an $\mathcal{R}$-module ($\mathcal{R} = \mathbf{Z}[t, t^{-1}]$). Action of $t \in \mathcal{R}$ on $\mathrm{Coker}(A)$ is given by an $N \times N$ matrix $T = (T_{jk})$ and therefore from standard linear algebra, $\mathrm{Coker}(A)$ is isomorphic to the $\mathrm{Coker}(tI - T)$ where $tI - T : \mathcal{R}^N \to \mathcal{R}^N$. By lemma 4.7.2 the ideal generated by $(N-k)\times(N-k)$ minors of $tI-T$ is the identical with the ideal $(n-k)\times(n-k)$ minors of A. In particular for $k = 0$ we obtain

$$\Delta_K(t) = \det(A) = \alpha \det(tI - T), \tag{4.7.4}$$

where α is a unit in $\mathcal{R}$. Therefore $N = l$, and we obtain

Corollary 4.7.3 *Let* $c \in \mathbf{Z}$ *be the constant term of the Alexander polynomial* $\Delta_K(t)$*, and* $l = \deg(\Delta_K(t))$*. Then*

$$H_1(\tilde{S}_{K,ab};\mathbf{Z}[\frac{1}{c}]) \simeq \mathbf{Z}[\frac{1}{c}]^l.$$

4.7.2 Clifford Algebras

In this section we give a generalization of Hamilton quaternions which enables us to construct the double covering of $SO(n)$ (see examples 4.2.1, 4.2.6 and exercise 4.2.4.) As a by-product of the algebraic construction, we exhibit a number (in fact, the maximal number possible) of linearly independent vector fields on odd dimensional spheres. Clifford algebras have many applications to geometry (see for example [LM] and [Mor]). Let k be the field of real or complex numbers, V be a vector space of dimension m over k, and q be a (non-degenerate) quadratic form on V. Denote the associated bilinear form by B_q. For $k = \mathbf{R}$ we assume q has signature $(r, m-r)$, i.e., q has r negative eigenvalues and $m-r$ positive ones. Let $\mathcal{T}(V)$ be the tensor algebra on V. The *Clifford algebra* on V relative to q is the quotient of $\mathcal{T}(V)$ by the ideal generated by

$$u \otimes u - q(u), \ u \in V.$$

For $k = \mathbf{R}$ we denote the Clifford algebra by $\mathcal{A}(r, m-r)$ or $\mathcal{A}(q)$ and for $k = \mathbf{C}$ we denote the Clifford algebra by $\mathcal{A}(m)$ or $\mathcal{A}$. Note that over $\mathbf{C}$ all non-degenerate quadratic forms are equivalent, so that we may assume that the matrix of q is the identity matrix. Over $\mathbf{R}$, q may be put in diagonal form with all eigenvalues ± 1. It is clear that the complexification of $\mathcal{A}(r, m-r)$ is $\mathcal{A}(m)$. Our objective in this section in to understand the structure of the Clifford algebras and give some applications. It is an immediate consequence of the definition of Clifford algebra that it satisfies the following:

- (**Universal Mapping Property**) Let W be a vector space over k and q a non-degenerate quadratic form on W. Given a linear map φ of W into an associative algebra A over k such that $\varphi(w)\varphi(w) = q(w)1$, then it extends uniquely to an algebra homomorphism $\varphi' : \mathcal{A}(q) \to A$ such the following diagram commutes:

$$\begin{tikzcd} W \arrow[rr, "\varphi"] \arrow[dr, "\iota"'] & & A \\ & \mathcal{A}(q) \arrow[ur, "\varphi'"'] & \end{tikzcd}$$

  where ι is the obvious inclusion of W in $\mathcal{A}(q)$.

It is not difficult to show that as a vector space $\mathcal{A}(r, m-r)$ or $\mathcal{A}(m)$ is isomorphic to the exterior algebra over V. In fact, if $\bullet$ denotes the multiplication in the Clifford algebra, then by considering $(u+v)\bullet(u+v)$ we obtain

$$u \bullet v + v \bullet u = 2B_q(u, v). \tag{4.7.5}$$

Consequently, if $\{e_1, ..., e_m\}$ is a basis for V, then $\{1 = e_0, e_{i_1} \bullet ... \bullet e_{i_k}\}$ for $i_1 < ... < i_k$ and $k \leq m$ is a basis for the Clifford algebra. The subspace spanned by products of the form $e_{i_1} \bullet ... \bullet e_{i_k}$, for $i_1 < ... < i_k$, will be denoted $\mathcal{L}(k)$, and $\mathcal{A}(r, m-r) \simeq \bigoplus \mathcal{L}(k)$ (direct sum of vector spaces).

Example 4.7.1 Consider the special case $k = \mathbf{R}$, $m = 2$ and $q(x) = -||x||^2$ where $|| \cdot ||$ denotes the Euclidean norm. Then $\mathcal{A}(q)$ has dimension 4. Let $\{e_1, e_2\}$ be an orthonormal basis for V; set $e_0 = 1$ and $e_3 = e_1 \bullet e_2$. Then the algebra structure of $\mathcal{A}(q)$ is described by the relations

$$e_0 = \text{identity},\ e_i^2 = -1,\ e_i \bullet e_j + e_j \bullet e_i = 0 \ \text{ for } i = 1, 2, 3 \text{ and } j > i.$$

Therefore $\mathcal{A}(q)$ is isomorphic to the division algebra of the Hamilton quaternions $\mathbf{H}$. ♠

Exercise 4.7.2 *Prove the following isomorphisms:*

1. $\mathcal{A}(1, 0) \simeq \mathbf{C}$;
2. $\mathcal{A}(0, 1) \simeq \mathbf{R} \oplus \mathbf{R}$;
3. $\mathcal{A}(0, 2) \simeq M_2(\mathbf{R}) \simeq \mathcal{A}(1, 1)$;

where $M_n(k)$ is the full matrix algebra of $n \times n$ matrices over k.

A useful tool for determining the structure of an algebra is the following simple observation:

Lemma 4.7.5 *Let A be an associative algebra with unit 1, and e and f be central elements in A such that $ef = 0$,and $e + f = 1$. Then Ae and Af are two-sided ideals in A and $A \simeq Ae \oplus Af$ as algebras.*

The proof of the lemma is straightforward and the isomorphism is given by $a \rightarrow ae + af$. ♣

Example 4.7.2 Consider the special case $k = \mathbf{R}$, $m = 3$ and $q(x) = -||x||^2$. Then $\mathcal{A}(q)$ has dimension 8. It is a simple matter to verify that the center of $\mathcal{A}(q)$ has dimension two and is spanned by 1 and $u = e_1 \bullet e_2 \bullet e_3$. Now set $e = (1 + u)/2$ and $f = (1 - u)/2$ and note that the conditions of lemma 4.7.5 are satisfied. Therefore $\mathcal{A}(q) \simeq \mathcal{A}(q)e \oplus \mathcal{A}(q)f$. A basis for $\mathcal{A}(q)e$ is $\{e, (e_2 \bullet e_3 - e_1)/2, (e_3 \bullet e_1 - e_2)/2, (e_1 \bullet e_2 - e_3)/2\}$. With this choice of basis it is trivial to verify that $\mathcal{A}(q)e$ is isomorphic to $\mathbf{H}$. Similarly, one shows $\mathcal{A}(q)f \simeq \mathbf{H}$, and consequently $\mathcal{A}(q) \simeq \mathbf{H} \oplus \mathbf{H}$. ♠

Exercise 4.7.3 *Prove the following isomorphisms:*

1. $\mathcal{A}(0,3) \simeq M_2(\mathbf{C}) \simeq \mathcal{A}(2,1)$;
2. $\mathcal{A}(1,2) \simeq \mathbf{H} \oplus \mathbf{H}$.

While it is possible to use lemma 4.7.5 together with some elaborate book-keeping to determine the structure of Clifford algebras, it is more enlightening to invoke the following ingenious device known as the *Periodicity theorem*:

Theorem 4.7.1 (Periodicity Theorem) *There are isomorphisms*

1. $\mathcal{A}(m+2) \simeq \mathcal{A}(m) \otimes M_2(\mathbf{C})$ *for* $m \geq 0$;
2. $\mathcal{A}(m+2,0) \simeq \mathcal{A}(0,m) \otimes \mathcal{A}(2,0)$ *for* $m \geq 0$;
3. $\mathcal{A}(0,m+2) \simeq \mathcal{A}(m,0) \otimes \mathcal{A}(0,2)$ *for* $m \geq 0$;
4. $\mathcal{A}(r+1,m-r+1) \simeq \mathcal{A}(r,m-r) \otimes \mathcal{A}(1,1)$ *for* $m, r \geq 0$.

Proof - The proof of these statements being similar, we only prove (4), and indicate the necessary changes for the other cases. Let $\{e_1, ..., e_{m+2}\}$ be a basis for $W = \mathbf{R}^{m+2}$ such that $q(e_i) = -1$ for $i \leq r+1$ and $q(e_i) = 1$ for $i \geq r+2$. We let $V \simeq \mathbf{R}^m$ and $V' \simeq \mathbf{R}^2$ be the subspaces of W spanned by the vectors $\{e_1, ..., e_r, e_{r+2}, ..., e_{m+1}\}$ and and $\{e_{r+1}, e_{m+2}\}$ respectively. Then the restrictions of q to V and V' have signatures $(r, m-r)$ and $(1,1)$ respectively. Consider the linear map φ of W into $\mathcal{A}(r,m-r) \otimes \mathcal{A}(1,1)$ defined by

$$\varphi(e_i) = e_i \otimes e_{r+1} \bullet e_{m+2}, \quad \text{for } i \leq r, \text{ and } r+2 \leq i \leq m+1,$$

and

$$\varphi(e_{r+1}) = 1 \otimes e_{r+1}, \quad \varphi(e_{m+2}) = 1 \otimes e_{m+2}.$$

Then φ extends to an algebra homomorphism φ' of $\mathcal{A}(r+1,m-r+1)$ into $\mathcal{A}(r,m-r)\otimes\mathcal{A}(1,1)$. Since φ' is surjective, it is an isomorphism for dimension reasons. This proves (4). To prove (2), we let $W \simeq \mathbf{R}^{m+2}$, q be negative definite, and $\{e_1, ..., e_{m+2}\}$ be a basis for W such that $B_q(e_i, e_j) = -\delta_{ij}$. Let V and V' be real vector spaces of dimensions m and 2 respectively. Let q_1 and q_2 be positive and negative definite quadratic forms on V and V' respectively. We choose bases $\{f_1, ..., f_m\}$ and $\{g_1, g_2\}$ for V and V' respectively such that $B_{q_1}(f_i, f_j) = \delta_{ij}$ and $B_{q_2}(g_i, g_j) = -\delta_{ij}$. Then set

$$\varphi(e_i) = f_i \otimes g_1 \bullet g_2, \quad \text{for } i \leq m,$$

and

$$\varphi(e_i) = 1 \otimes g_{i-m}, \quad \text{for } i = m+1, m+2.$$

The rest of the proof is just as before. ♣

With the aid of theorem 4.7.1, we can easily describe the structure of all the Clifford algebras. The complex case is given by

Corollary 4.7.4 *We have isomorphisms* $\mathcal{A}(2k) \simeq M_{2^k}(\mathbf{C})$, *and* $\mathcal{A}(2k+1) \simeq M_{2^k}(\mathbf{C}) \oplus M_{2^k}(\mathbf{C})$.

The real case is more complex, and is given by

Corollary 4.7.5 *We have isomorphisms* $\mathcal{A}(r+8, m-r) \simeq \mathcal{A}(r, m-r) \otimes M_{16}(\mathbf{R})$, *and* $\mathcal{A}(r, m-r+8) \simeq \mathcal{A}(r, m-r) \otimes M_{16}(\mathbf{R})$.

The low dimensional cases are given in the following table (the proof of the low dimensional cases is by repeated application of the periodicity theorem):

Real Clifford Algebras in Low Dimensions

$\mathcal{A}(r, m-r)$	$r=0$	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$	$r=7$
$m-r=0$	$\mathbf{R}$	$\mathbf{C}$	$\mathbf{H}$	$2\mathbf{H}$	$M_2(\mathbf{H})$	$M_4(\mathbf{C})$	$M_8(\mathbf{R})$	$2M_8(\mathbf{R})$
$m-r=1$	$2\mathbf{R}$	$M_2(\mathbf{R})$	$M_2(\mathbf{C})$	$M_2(\mathbf{H})$	$2M_2(\mathbf{H})$	$M_4(\mathbf{H})$	$M_8(\mathbf{C})$	$M_{16}(\mathbf{R})$
$m-r=2$	$M_2(\mathbf{R})$	$2M_2(\mathbf{R})$	$M_4(\mathbf{R})$	$M_4(\mathbf{C})$	$M_4(\mathbf{H})$	$2M_4(\mathbf{H})$	$M_8(\mathbf{H})$	$M_{16}(\mathbf{C})$
$m-r=3$	$M_2(\mathbf{C})$	$M_4(\mathbf{R})$	$2M_4(\mathbf{R})$	$M_8(\mathbf{R})$	$M_8(\mathbf{C})$	$M_8(\mathbf{H})$	$2M_8(\mathbf{H})$	$M_{16}(\mathbf{H})$
$m-r=4$	$M_2(\mathbf{H})$	$M_4(\mathbf{C})$	$M_8(\mathbf{R})$	$2M_8(\mathbf{R})$	$M_{16}(\mathbf{R})$	$M_{16}(\mathbf{C})$	$M_{16}(\mathbf{H})$	$2M_{16}(\mathbf{H})$
$m-r=5$	$2M_2(\mathbf{H})$	$M_4(\mathbf{H})$	$M_8(\mathbf{C})$	$M_{16}(\mathbf{R})$	$2M_{16}(\mathbf{R})$	$M_{32}(\mathbf{R})$	$M_{32}(\mathbf{C})$	$M_{32}(\mathbf{H})$
$m-r=6$	$M_4(\mathbf{H})$	$2M_4(\mathbf{H})$	$M_8(\mathbf{H})$	$M_{16}(\mathbf{C})$	$M_{32}(\mathbf{R})$	$2M_{32}(\mathbf{R})$	$M_{64}(\mathbf{R})$	$M_{64}(\mathbf{C})$
$m-r=7$	$M_8(\mathbf{C})$	$M_8(\mathbf{H})$	$2M_8(\mathbf{H})$	$M_{16}(\mathbf{H})$	$M_{32}(\mathbf{C})$	$M_{64}(\mathbf{R})$	$2M_{64}(\mathbf{R})$	$M_{128}(\mathbf{R})$

N.B. - $2A$ in the above table means $A \oplus A$.

Exercise 4.7.4 *Let* $\mathcal{Z}(r, m-r)$ *(or* $\mathcal{Z}(q)$*) denote the center of the Clifford algebra* $\mathcal{A}(r, m-r)$ *(or* $\mathcal{A}(q)$*). Show that*

1. *If* $m - 2r \equiv 1 \; mod(4)$, *then* $\mathcal{Z}(r, m-r) \simeq \mathbf{R} \oplus \mathbf{R}$;
2. *If* $m - 2r \equiv 3 \; mod(4)$, *then* $\mathcal{Z}(r, m-r) \simeq \mathbf{C}$;

3. *If $m-2r$ is even, then $\mathcal{Z}(r,m-r)\simeq \mathbf{R}$;*

4. *The center $\mathcal{Z}(m)$ of $\mathcal{A}(m)$ is $\mathbf{C}$ or $\mathbf{C}\oplus\mathbf{C}$ according as m is even or odd.*

The subspace $\mathcal{L}(2)$ is a Lie algebra under the bracket operation

$$[e_i \bullet e_j, e_k \bullet e_l] = e_i \bullet e_j \bullet e_k \bullet e_l - e_k \bullet e_l \bullet e_i \bullet e_j.$$

This follows easily from (4.7.5). Furthermore, by the same reasoning $V \simeq \mathbf{R}^m \subset \mathcal{A}(q)$ is also invariant under the linear transformation

$$T(\xi)(v) = \xi \bullet v - v \bullet \xi, \quad v \in V, \xi \in \mathcal{L}(2).$$

Furthermore, from (4.7.5) it follows easily that the linear transformation $T(\xi)$ satisfies

$$B_q(T(\xi)v, w) + B_q(v, T(\xi)w) = 0, \tag{4.7.6}$$

i.e., $T(\xi)$ is skew-symmetric relative to the symmetric bilinear form B_q. Thus T is a homomorphism of $\mathcal{L}(2)$ into the Lie algebra $\mathcal{SO}(r,m-r)$.

Exercise 4.7.5 *Show that $[V,\mathcal{L}(2k)]\subseteq \mathcal{L}(2k-1)$, and $[V,\mathcal{L}(2k-1)]\subseteq \mathcal{L}(2k)$. Hence, or otherwise, prove that if $\xi\in\mathcal{A}(r,m-r)$ is such that $[\xi,V]\subseteq V$, then $\xi\in\mathcal{Z}(q)\oplus\mathcal{L}(2)$. Show also that T is an isomorphism of Lie algebras.*

Example 4.7.3 We now construct the universal covering groups of $SO(m)$ for $m\geq 3$. Let Spin(m), called the *spin group*, be the analytic subgroup of the group $\mathcal{A}(m,0)^\times$ of invertible elements of $\mathcal{A}(m,0)$ corresponding to the Lie algebra $\mathcal{L}(2)$. The Lie algebra homomorphism T induces a homomorphism τ of Spin(m) into $SO(m)$ by

$$\tau(\exp\xi) = \exp(T(\xi)).$$

From exercise 4.7.5 it follows that τ is onto. Assume $m\geq 3$. Let $\{e_1,e_2\}$ be such that $B_q(e_i,e_j) = -\delta_{ij}$. Then by a simple calculation

$$\exp(te_1\bullet e_2) = (\cos t)e_0 + (\sin t)e_1\bullet e_2.$$

Consequently, $\pm e_0\in$ Spin(m), and therefore Ker$(\tau)\supseteq\{\pm e_0\}$. Since $\pi_1(SO(m),e)=\mathbf{Z}/2$, Spin$(m)$ is the universal covering group of $SO(m)$. ♠

Exercise 4.7.6 *By imitating the argument of example 4.7.3 construct the double covering of $SO^\circ(m-1,1)$ where $m\geq 4$ and $SO^\circ(m-1,1)$ denotes the connected component of $SO(m-1,1)$. Prove also that this is the universal covering group of $SO^\circ(m-1,1)$.*

Exercise 4.7.7 *Show that the universal covering group of $SO(m,2)$ is infinitely sheeted. Describe the universal covering group of $SO^\circ(m-r,r)$ where $m-r, r \geq 3$.*

To facilitate further applications of Clifford algebras, we have to make an observation. Since the algebra $\mathcal{A}(r, m-r)$ is semi-simple, the irreducible (left)-modules for $\mathcal{A}(r, m-r)$ are isomorphic to the minimal (left) ideals which are easily described. In fact, it is easy to verify that the minimal left ideals in $M_n(k)$ are isomorphic to the ideal consisting of matrices with zeroes everywhere except along one column where the entries are arbitrary. In particular, they are isomorphic to k^n.

Exercise 4.7.8 *Let $m = 8k+r$. Show that the dimension of a non-trivial irreducible module W for $\mathcal{A}(m,0)$ is*

$$dim_{\mathbf{R}}(W) = \begin{cases} 2^{4k} & \text{if } r = 0; \\ 2^{4k+1} & \text{if } r = 1; \\ 2^{4k+2} & \text{if } r = 2,3; \\ 2^{4k+3} & \text{if } r = 4,5,6,7. \end{cases}$$

Exercise 4.7.9 *Assume q is negative definite and let $\{e_1, ..., e_m\}$ be an orthonormal basis for $\mathbf{R}^m$, i.e., $q(e_i) = -1$. Show that with respect to the basis $\{1 = e_0, e_{i_1} \bullet ... \bullet e_{i_k}\}$ of $\mathcal{A}(m,0)$, the matrix of multiplication by e_i is skew-symmetric. Deduce that every irreducible $\mathcal{A}(m,0)$ module admits of an inner product such the action of $V \simeq \mathbf{R}^m \subset \mathcal{A}(m,0)$ is by skew symmetric linear transformations.*

Example 4.7.4 We use our knowledge of the Clifford algebras to construct linearly independent vector fields on odd dimensional spheres. By this we mean vector fields $\xi_1, ..., \xi_r$ which are linearly independent at every point of the sphere S^{2n-1}. Of course there are no such vectors fields on even dimensional spheres. Let $W \simeq \mathbf{R}^N$ be a module for $\mathcal{A}(m,0)$ and as noted in exercise 4.7.9 there is an inner product $<,>$ on W relative to which the action of $V \simeq \mathbf{R}^m \subset \mathcal{A}(m,0)$ is by skew symmetric transformations. Therefore $< e \bullet w, w >= 0$ for $e \in V$, $w \in W$, and $e \bullet w$ is perpendicular to w. Hence $e \bullet w$ may be regarded as a vector field ξ_e on S^{N-1}. Now $e \bullet w \neq 0$ for $e \neq 0$ and $w \neq 0$ since $< e \bullet w, e \bullet w >= - < e^2 \bullet w, w >= -q(e) < w, w >$, so that ξ_e is a nowhere vanishing vector field. Similarly, if $\{e_1, ..., e_m\}$ is basis for V, then the vector fields $\{\xi_{e_1}, ..., \xi_{e_m}\}$ are linearly independent at every point. This construction yields m linearly independent vector fields on S^{N-1} provided $W \simeq \mathbf{R}^N$ is a direct sum of (non-trivial) irreducible $\mathcal{A}(m,0)$ modules. Clearly all such modules are even dimensional which confirms the fact there are no nowhere vanishing vector fields on even dimensional spheres. We can also determine the maximal number of linearly independent vector fields on odd dimensional spheres which can be obtained in this fashion. To do so we have to determine, for a given $N = 2n$, the largest integer m such that $\mathbf{R}^N$ is a direct sum of non-trivial irreducible $\mathcal{A}(m, 0)$ modules. Writing $N = 2^{4a+b}(2l + 1)$ where $0 \leq b \leq 3$, and using exercise 4.7.8, it easy to show that the largest such m is $8a + 2^b - 1$. Therefore S^{2n-1} admits of (at least) $8a + 2^b - 1$ linearly independent vector fields. The fact that this number is exact is considerably deeper. ♠

Bibliography

[Bre] Bredon, G. - On Homogeneous Cohomology Spheres; *Annals of Math.*, **73**, no.3, (1961), pp. 556-565.

[CrF] Crowell, R. H. and R. H. Fox - *Introduction to Knot Theory*, Springer-Verlag, Berlin-Heidelberg-New York, (1977).

[EN] Eisenbud, D. and W. Neumann *Thee-Dimensional Link Theory and Invariants of Plane Curve Singularities*, Princeton University Press, Princeton, NJ, (1985).

[Gri] Griffiths, H. B. - Fundamental Group of Spaces with a Common Point; *Quart. J. of Math.*, **5** (1954), pp.175-190.

[GrKu] Grigorchuk, R. I. and P. F. Kurchanov - Some Questions of Group Theory Related to Geometry; in *Algebra VII; Encyclopedia of Mathematical Sciences* (A. N. Parshin and I. R. Shafarevich (Editors)), **58**, Springer-Verlag, Berlin-Heidelberg-New York (1990).

[Gro] Gross, J. L. - An Infinite Class of Irreducible Homology 3-Spheres; *Proc. Amer. Math. Soc.*, **25**, (1971), pp. 173-176.

[LM] Lawson, H. B. and M. L. Michelson - *Spin Geometry*, Princeton University Press, Princeton, NJ, (1989)

[Mi3] Milnor, J. W. - *Singular Points of Complex Hypersurfaces*, Princeton University Press, Princeton, NJ, (1968).

[Mor] Morgan, J. - *The Seiberg-Witten Equations and Applications to the Topology of Smooth Four Manifolds*, Princeton University Press, Princeton, NJ, (1996).

[Ol] Olum, P.- Non-abelian Cohomology and van Kampen's Theorem; *Annals of Math.*, **68**, no.3, (1958), pp.658-668.

[Rol] Rolfsen, D. - *Knots and Links*, Publish or Perish, Inc., Wilmington, Delaware, (1976).

[Whi3] Whitehead, J. H. C. - On Doubled Knots; *Jour. London Math. Soc.*, **12**, (1937), pp. 63-71.

[Whitt] Whitten, W. - Knot Complements and Groups; *Topology*, **26**, no.1, (1987), pp. 41-44.