Part II

Relators

In this part the focus of attention is narrowed to those cone complexes and cone categories in which the boundaries of the cones are homotopically equivalent to S^1. These 'circular' categories are introduced in Section 4. In Section 5, the key concept of a general relator is defined and the technical properties of these relators are investigated.

4 Circular Categories

In this section circular categories are introduced, and notations for words, cycles, paths and loops in labeled circular categories are established. The section concludes with the proof of a 1-dimensional version of the simplicial approximation theorem for circular categories. In particular, it is shown that in a circular category all topological paths which start and end at vertices are homotopic relative to their endpoints to a combinatorial path in the 1-skeleton.

4.1 Circular Complexes and Circular Categories

A circular complex is a special type of cone complex which is tailor-made to work as a building block for a general small cancellation theory. A cone complex whose 1-skeleton is a simplical complex and where the subspace ∂c is homotopically equivalent to a circle for all elements c of rank at least 2 is called a circular complex. By virtue of the fact that every circular complex C is a cone complex, many concepts have already been defined for circular complexes. The elements $c \in C$ are called open circular cones, and the principal ideals C/c are closed circular cones. A cone map between circular complexes is called a circular map.

A circular category, C, is a cone category in which every slice is the category of a closed circular cone. An object of the category is called an open circular cone, the slice category C/c with the attaching functor, ϕ_c, implicitly understood is called a closed circular cone and ∂c, which is equal to C/c minus its maximum element, is the boundary of c. A circular functor is a cone functor between circular categories. The definitions of rank, the n-skeleton, and geometric realization are the same as for cone categories. Notice that the 1-skeleton of a circular category is a graph so that it is possible to speak of edges and vertices in C. Circular categories, merging the specificity of circular complexes with the flexibility of cone categories, will be the constructions of primary interest in the remainder of the article. In the same way that cell categories are an economical way to construct and prove results about manifolds, circular categories are an economical way to construct and prove results about groups.

4.2 Inversions and Words

The language of inversions presented below is an extension of the usage in [5] and is used here to provide a unified way to describe the fact that most functions between invertible objects (such as letters, words, edges, paths, cycles, etc.) commute with inversion. The ease with which a labeled graph can be defined illustrates the advantages of this approach.

Let A be a set and let $\iota : A \to A$ be a function from A to itself such that $\iota(\iota(a)) = a$ for all $a \in A$. The function ι is called an inversion of A and, when ι is understood, A is called an invertible set. The element $\iota(a)$ is the formal inverse of a and is usually written a^{-1}. The set $\{a, \iota(a)\}$ is called an orbit. An inversion homomorphism between invertible sets A and B is a function $f : A \to B$ such that $f(\iota(a)) = \iota(f(a))$ for all $a \in A$. Invertible sets are closed under direct product in the obvious way. If ι_A is the inversion of A and ι_B is the inversion of B, then the function $\iota_{A\times B}$ defined by $\iota_{A\times B}(a, b) = (\iota_A(a), \iota_B(b))$ is an inversion on the product $A \times B$.

An orientation of an invertible set A is a distinguished subset $A^+ \subset A$ such that $\iota(A^+)$ is disjoint from A^+ and $A^+ \cup \iota(A^+) = A$. The elements of A^+ are called the positive elements of A, and those in $\iota(A^+)$ are negative elements. When A^+ is clear from context we say that A is oriented. A necessary and sufficient condition for an orientation to be possible is that the inversion, ι, must be fixed-point-free. If $A^+ \subset A$ is an orientation of A then $\iota(A^+) \subset A$ is another orientation of A called the opposite orientation. If f is an inversion homomorphism between oriented sets A and B, it is called orientation-preserving if $f(A^+) \subset B^+$ and orientation-reversing if $f(A^+) \subset \iota(B^+)$.

An alphabet is a set. All alphabets considered in this article will be, in addition, finite, oriented, and invertible. The free monoid over an alphabet A is denoted A^*, and $|W|$ represents the length of a word $W \in A^*$. If W, X, Y, and Z are possibly empty words such that $W = XYZ$ in A^*, then X is an initial segment of W, Y is a subword of W, and Z is a final segment of W. If $W = XY$ and $Z = YX$ then Z is a cyclic conjugate of W. A cyclic conjugate is proper if both X and Y are nonempty words. The equivalence class of all cyclic conjugates of a word W is called the cycle of W or simply the cycle W. The cycle of W can be thought of as a string of letters in A which are 'written in a circle'. The circle can be then be broken at any point to create a particular cyclic conjugate of W. The inverse of a word, denoted W^{-1}, is the formal inverse of each letter listed in the opposite order, and the inverse of the cycle of the word W is the cycle of W^{-1}. A word, W, is reduced when $\iota(a)a$ does not occur as a subword for any letter $a \in A$, and cyclically reduced if all cyclic conjugates of W are reduced. Define W^0 to be the empty word, $W^{n+1} = W^n W$ for all positive integers and $W^{-n} = (W^{-1})^n$ when $-n$ is a negative integer. If $W = V^n$ for some word V and integer $n > 1$, then W is a power of V. If W is not a power of any word V, then W is simple. Equivalently, W is simple iff it is not equal to any of its proper cyclic conjugates. If W is a subword of V^n then W is V-periodic and V is a period of W. The following lemma is easy to prove and well-known in the literature.

Lemma 4.1 *Let X and Y be simple non-conjugate words. If W is both X-periodic and Y-periodic, then $|W| < |X| + |Y|$. As a consequence, whenever X^i is Y-periodic for some integer $i > 1$, $|X| < \frac{1}{i-1}|Y|$.*

A word equivalent to the identity in a free group is called a Dyck word. These words have numerous well-known properties such as the fact that they can be reduced to the empty word by repeatedly removing subwords of the form aa^{-1} for some $a \in A$. Other easy properties are contained in the next two lemmas.

Lemma 4.2 *The set of Dyck words is closed under products, conjugates, and the removal of subwords which are also Dyck words. In addition, every Dyck word contains a subword of the form aa^{-1} for some $a \in A$, and the cycle of a Dyck word always contains at least two subwords of this form. All Dyck words have an even length. And finally, if W and W^{-1} are words in A^* which are conjugate to each other in the free group over A, then W is a Dyck word and thus equivalent to the identity element.*

4.3 Graphs and Cayley Graphs

A graph is a 1-dimensional simplicial category. Since 0-simplices and 1-simplices are indistinguishable from 0-cells and 1-cells, a graph could also be defined as a 1-dimensional cell category. The 0-simplices are called vertices, and the open 1-simplices are called edges. If a graph is also a simplicial complex then it is called a simplicial graph, in which case the attaching maps are injective and every pair of vertices determines at most 1 edge.

The slice category of an edge contains exactly 2 vertices called the endpoints of the edge, and the two possible orderings of these vertices yield two orientations of the edge. Although the endpoints are distinct in the slice category, they may be identified under the attaching map. An oriented edge is conventionally represented by an arrow going from one endpoint to the other. These vertices are called its initial and terminal vertices, respectively. Note that the arrow used to represent an oriented edge is very different from an arrow in the simplicial category of the graph. If Γ is a graph then the edge set of Γ, written Edges(Γ), is the set of oriented edges of Γ. The edge set of a graph is invertible and orientable since reversing the orientation of an edge generates a fixed-point-free involution of Edges(Γ). An oriented graph is a graph with an orientation of its edge set. To describe an oriented graph it is enough to describe the positive edges.

A graph morphism is a simplicial functor between graphs viewed as simplicial categories. Every graph morphism induces an inversion homomorphism between the edge sets. A graph morphism between oriented graphs is called orientation-preserving or -reversing according to the status of the inversion homomorphism between the edge sets. A morphism from a graph to itself is an automorphism and the set of all automorphisms forms a group called the automorphism group of the graph.

A graph Γ is labeled by an invertible alphabet A if there is an inversion homomorphism $f : \text{Edges}(\Gamma) \to A$. A graph which is labeled by A is called an A-graph or, if A is understood, simply a labeled graph. A labeled graph is deterministic if every pair of distinct oriented edges leaving the same initial vertex has distinct labels. A labeled circular category is one whose 1-skeleton is a labeled graph.

Let p be an inversion homomorphism from an alphabet A to a group G. When the unique extension of p to a monoid homomorphism $p^* : A^* \to G$ is onto, then it is said that A (via p) generates G, the letters in A are generators, and G is called an A-group. If the function p is injective, then the generators can be viewed as elements of G. The size of a generating set is the number of orbits in A. If G and H are A-groups via $p : A \to G$ and $q : A \to H$ respectively, and $f : G \to H$ is a group homomorphism such that $fp = q$, then f is called an A-group homomorphism. When X is a word in A^* sent to an element g in G, then X is called a representative of g.

Lemma 4.3 *If G and H are A-groups, and $f : G \to H$ and $g : H \to G$ are A-group homomorphisms, then f and g are isomorphisms.*

Proof: If $p : A \to G$ and $q : A \to H$ are extended to the monoid homomorphisms $p^* : A^* \to G$ and $q^* : A^* \to H$, then $fp^* = q^*$ and $gq^* = p^*$. Since q^* is onto, so is f. Next, let $p^*(X)$ and $p^*(Y)$ be arbitrary elements of G. If $f(p^*(X)) = f(p^*(Y))$ then $q^*(X) = q^*(Y)$, so that $g(q^*(X)) = g(q^*(Y))$ and $p^*(X) = p^*(Y)$. Thus f is 1 to 1 as well as onto, so that f is an isomorphism. By symmetry, g is an isomorphism. □

The Cayley graph of an A-group G, written $\mathcal{C}(G, A)$, is an A-graph constructed as follows: the vertices correspond to the elements of G, and the edge set corresponds to $G \times A$. The oriented edge associated with (g, a) starts at the vertex g, ends at the vertex ga (or, to be more exact, the vertex labeled by g multiplied on the right by $p(a)$), and is labeled by the letter a. The inverse edge which starts at ga and ends at g corresponds to (ga, a^{-1}) and is labeled a^{-1}. The vertex corresponding to the identity of the group is considered a distinguished vertex. Since Cayley graphs are deterministic and every vertex is the initial vertex of an edge labeled by an arbitrary letter, given any word W and any vertex v in the Cayley graph, there is always a reading of a word W starting at v, and it is unique. In particular, for every word W there is a unique reading of W starting at the base point and the endpoint u of this reading is the vertex corresponding to the element g of G which is equal to the product of the letters in the word W in the group G. This endpoint in turn determines a unique automorphism of the Cayley graph, namely, the one which sends the distinguished vertex to u. Several important and elementary properties of Cayley graphs are recorded in the following lemmas without proof.

Lemma 4.4 *If G is an A-group then its Cayley graph is the unique, connected, deterministic A-graph with a distinguished vertex whose label-preserving automorphism group acts transitively on the vertex set and is canonically isomorphic to G in such a way that the automorphism corresponding to the generator a moves the base point to a vertex connected to the base point by an oriented*

$$1 \xrightarrow{1} 2 \xrightarrow{2} 3 \xrightarrow{3} 4 \quad \cdots \quad n \xrightarrow{n} n+1$$

Figure 10: The abstract path, P_n

edge labeled by a, for all $a \in A$. In particular, if Γ is a connected, deterministic A-graph on which the group of label-preserving automorphisms acts transitively on its vertex set, then Γ is the Cayley graph of its automorphism group.

Lemma 4.5 *Let G be an A-group and let C be its Cayley graph. Every word $W \in A^*$ determines a unique automorphism of C. Thus, the automorphism corresponding to W is the identity automorphism iff the word W is equivalent to the identity element in the group, and consequently a path in C is a loop iff the word read by the path is equivalent to 1 in the group.*

Lemma 4.6 *If G is an A-group and H is a normal subgroup of G, then there is a unique label-preserving graph morphism from $\mathcal{C}(G, A)$ to $\mathcal{C}(G/H, A)$ which sends the distinguished vertex to the distinguished vertex. Moreover, one way of constructing such a function is to quotient the graph $\mathcal{C}(G, A)$ by the action of the subgroup of the automorphism group corresponding to H under the canonical isomorphism.*

4.4 Paths, Lines, and Loops

The three examples listed below define types of graphs which will be needed later. To fix notation let $\mathbf{Z}$ be the integers, let $[n] = \{1, 2, \ldots, n\}$ with $[0] = \emptyset$, let $\mathbf{Z}_n = \mathbf{Z}/n\mathbf{Z}$ be the cyclic group of order n, let $\mathbf{D}_{2n}$ be the dihedral group of order $2n$, and let $\mathbf{D}_\infty$ be the infinite dihedral group, which is the free group generated by two involutions.

Example 1 An abstract path, P_n, $n \geq 0$, is an oriented graph with vertices $\cong [n+1]$ and positively oriented edges $\cong [n]$ such that the positive edge i goes from the vertex i to the vertex $i+1$ (see Figure 10). The automorphism group of P_n, $n > 0$, is the cyclic group $\mathbf{Z}_2$, since the only non-trivial automorphism is the morphism which sends the vertex i to $n + 2 - i$ for all $i \in [n+1]$.

Example 2 An abstract line, L, is an oriented graph with vertices $\cong \mathbf{Z}$ and positively oriented edges $\cong \mathbf{Z}$ such that the positive edge i goes from the vertex i to the vertex $i+1$ (see Figure 11). The automorphisms of L form the group $\mathbf{D}_\infty$, and they can be divided into several types: translations of the graph, reflections which fix a vertex, reflections which fix an edge, and the identity.

Example 3 An abstract loop, L_n, $n \geq 1$, is an oriented graph with vertices $\cong \mathbf{Z}_n$ and positively oriented edges $\cong \mathbf{Z}_n$ such that the positive edge i goes from the vertex i to the vertex $i+1$ (see Figure 12). The automorphisms of

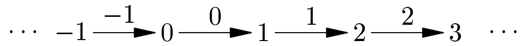

Figure 11: The abstract line, L

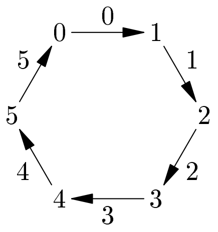

Figure 12: The abstract loop, L_6

L_n form the group $\mathbf{D}_{2n}$, and they can be divided into several types: rotations of the graph, reflections which fix a vertex, reflections which fix an edge, and the identity. If n is odd there is no difference between vertex- and edge-fixing automorphisms; if n is even then there is a difference. In particular notice that a reflection which fixes a vertex or an edge must also fix the graphical element which is directly opposite it. If n is even then for every vertex there is an opposite vertex and for every edge an opposite edge. If n is odd for every vertex there is an opposite edge and vice versa.

An abstract path has two endpoints, namely the vertices 1 and $n+1$, which are called the start and the end of the path respectively, since they are at the start and end of the sequence of positively oriented edges in P_n. The vertex 1 in L and L_n is called its base. All other vertices in P_n, L or L_n are called interior. Let C be a circular category. A finite path is a circular functor from P_n to C or equivalently a graph morphism into the 1-skeleton of C, a loop is a circular functor from L_n to C, and a line or infinite path is a circular functor from L to C. The start (end) of a path is simply the image of the start (end) of P_n and the base of a line or a loop is the image of the base of L or L_n.

The inverse of a path is the same functor from P_n to C but precomposed with the unique non-trivial automorphism of P_n. The inverse of a line or a loop is the same functor but precomposed with the unique non-trivial automorphism which fixes the base. A path, line, or loop is called reduced if the positively oriented edges i and $i+1$ are never sent to edges which are inverses of each other. The degree of a vertex v, written $\mathrm{d}(v)$, is the number of oriented edges with v as an initial vertex. A path, loop, or line is called simple iff the graph map into C is injective on vertices. An arc is a simple loop or a maximal simple path in which all interior vertices have degree 2. The length of a path or loop is the number of positive edges in its domain. A path from u to v of shortest length is called a geodesic. A line is called a geodesic if every path it contains is a geodesic. A graph is connected if every pair of vertices can be the endpoints of a path. Every connected graph Γ comes equipped with a metric called the graph metric in which the distance between the vertices u and v is the length of a geodesic from them. If C is a circular category, v is a vertex in C and n is a positive integer, then the ball $\mathrm{Ball}(v,n)$ is the connected subgraph of the 1-skeleton of C consisting of v, all the vertices u whose distance to v in the graph metric is less than n, and all of the edges of C whose endpoints are in $\mathrm{Ball}(v,n)$. The vertex v is called the center of the ball, and n is the radius. Notice that $\mathrm{Ball}(v,0) = \mathrm{Ball}(v,1)$ is the vertex v alone. Sometimes it will be convenient to include in $\mathrm{Ball}(v,n)$ all of the open cones whose entire 1-skeleton is already included under the above definition.

4.5 Labeled Paths and Loops

If the graph P_n is labeled by A then the labels of the positive edges of P_n form a word in A^*, and conversely, every word of length n corresponds to a unique labeling of the positive edges of P_n. Using this correspondence the word W will refer both to a particular finite string of letters and to the associated labeled abstract path. The form intended will be clear from context. The word W is reduced iff its graph is deterministic. If W is a reduced word then its graph is called $\mathrm{str}_1(W)$, the rank 1 straightline construction on W, and if W is not reduced then $\mathrm{str}_1(W)$ is the rank 1 straightline construction of its reduction in the free group.

Similarly, if the graph L_n is labeled by A then the labels of the positive edges form a cycle, and every cycle of length n corresponds to a labeling of the positive edges of L_n. If labelings of L_n which differ by an orientation-preserving automorphism are considered equivalent, then the equivalence classes of labelings of L_n correspond exactly to cycles of length n. Orientation-preserving automorphisms of L_n only serve to change the base and thus produce cyclic conjugates of the word read along the positive edges from the base. The cycle W will also refer to the associated labeled abstract loop. A word is cyclically reduced iff its cycle is reduced iff the corresponding labeling of L_n is deterministic. If the cycle W is reduced then the labeled graph L_n is called $\mathrm{cir}_1(W)$, the rank 1 circular construction on W, and if the cycle W is not reduced, then $\mathrm{cir}_1(W)$ is the rank 1 circular construction of its cyclic reduction in the free group.

Every path in a labeled graph Γ induces a labeling of P_n and thus corresponds to a word $W \in A^*$. The word W is called the word read by the path, the path reads W, and W is readable in Γ. Similarly, a loop in a labeled graph induces a labeling of the abstract loop which corresponds to a cycle, say W. The cycle W is called the cycle read by the loop, the loop reads W and the cycle W is readable in Γ. A path which reads the word W is called the path W, and a loop which reads the cycle W is called the loop W. At this point four distinct situations have been defined which can be indicated by the same finite string of letters, $W \in A^*$, namely, the word W, the cycle W, the path W, and the loop W. To summarize the distinctions: words and cycles refer to labeled abstract paths and loops, respectively, while paths and loops refer to implicitly understood, label-preserving functors from a word or cycle into a

$$0 \xrightarrow{a} 1 \xrightarrow{b} 2 \xrightarrow{a} 3 \xrightarrow{c} 4 \xrightarrow{d} 5$$

Figure 13: A rank 1 straightline construction

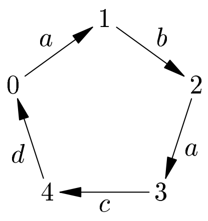

Figure 14: A rank 1 circular construction

labeled complex of some kind.

Example 4 In Figure 13 the rank 1 straightline construction str_1 is illustrated for the word $W = abacd$. Notice that it is also the rank 1 straightline construction of the word $abee^{-1}acd$.

Example 5 In Figure 14 the rank 1 circular construction cir_1 is illustrated for the word $W = abacd$. Notice that it is also the rank 1 circular construction of the word $eabacde^{-1}$.

An automaton is a labeled graph with a specified start state and one or more specified end states. In this article all automata considered will have exactly one end state. A word W is accepted by an automaton Γ if it is read by a path in Γ which starts at the start state and ends at one of the end states. The inverse of a word W, denoted W^{-1}, corresponds to the same labeled graph but with the opposite orientation. The inverse of a cycle, a path, or a loop is defined analogously. If a word is its own inverse then the non-trivial automorphism of P_n is label-preserving. If n is odd then the middle edge is fixed, and the alphabet is non-oriented, and if n is even then the middle two edges show that the word is not reduced. Thus no reduced word over an oriented alphabet is its own inverse. If W is a word read by path from u to v, then W^{-1} is read from v to u. Let X be a path from u to v and let Y be a loop based at v. The loop Y can be conjugated by X to produce a loop based at u. Specifically, the conjugation of Y by X is a loop which reads the word XYX^{-1}.

4.6 Approximations by Paths

A path, in the topological sense of the word, is simply a continuous function from the unit interval $[0, 1]$ to a topological space. The Simplicial Approximation Theorem on simplicial complexes can be used to show that every topological path between vertices of a circular category is homotopic relative to its endpoints to a path as defined above.

Lemma 4.7 *If C is a circular category and $f : [0, 1] \to C$ is a topological path which starts and ends at vertices in the space C, then f is homotopic relative to its endpoints to a simplicial path in the 1-skeleton of C.*

Proof: The crucial properties of circular cones are that, like simplices and cells, they are contractible spaces with path-connected boundaries when the rank is 2 or more. By Lemma 2.12 the circular category C can be subdivided so that the result is a simplicial complex. The traditional simplicial approximation theorem can then be applied to the subdivided complex. The finite number of vertices in the new path are each contained in a unique open cone of C by Lemma 1.16. Let k be the maximum rank of an open cone of C containing a vertex of the path. This implies that the new path is in the k-skeleton of the original category C. If $k \geq 2$, then the path can be shown to be homotopic to a path in the $(k-1)$-skeleton as follows.

Let c be an open cone of rank k which contains a vertex v of the new path. The edges on either side of v must also be contained in c. By looking far enough in each direction there is a unique vertex in the path which is farthest from v, in the boundary of c, and with all edges and vertices in the path between it and v in the open cone c. Let v_1 and v_2 be these vertices on either side of v. The portion of the path between v_1 and v_2 is in the closed cone C/c and it starts and ends in the boundary of c. Since the closed cone c has a path-connected boundary and is contractible and thus simply connected, there is a path between v_1 and v_2 which is contained in ∂c and is homotopic to the other path, keeping the endpoints fixed. By again applying the simplicial approximation theorem to the new path between v_1 and v_2 in the lower skeleton, the new path is also simplicial.

Technically speaking, the image of the closed cone C/c under the attaching functor may not be either simply connected or contractible. The homotopic path is actually to be found by lifting the portion between v_1 and v_2 into the closed cone C/c, choosing a path in the boundary, applying the Simplicial Approximation Theorem to the path in the boundary, finding a homotopy to the new path, and then pushing the homotopy, via the attaching map, back down into the construction. If this procedure is carried out for each of the open cones of rank k, then the maximum rank of a vertex in the path will be $k-1$, and the path is now in the simplicial subdivision of the $(k-1)$-skeleton of C.

Continuing in this way, the path will eventually be a path in the 1-skeleton. At this point, a similar procedure is followed. For all vertices v in the path which are not vertices of C, there is an open edge c containing v and vertices v_1 and v_2 in the boundary ∂c as defined above. If the lift of the path between v_1 and v_2 to the closed edge C/c lifts v_1 and v_2 to the same endpoint, then the path between v_1 and v_2 is homotopic to the constant path, and thus the original path is homotopic to the path obtained by removing the vertices and edges between v_1 and v_2 and connecting the two sections. If, on the other hand,

v_1 and v_2 are at opposite ends of the edge when lifted, then the portion of the path between v_1 and v_2 is homotopic to a nice linear function between the subinterval of $[0, 1]$ and the edge c. When this is done for all vertices in the path which are not vertices in the original category C, the result is a function from unit interval $[0, 1]$ to the 1-skeleton of C which by appropriately partitioning $[0, 1]$ is a circular functor into the 1-skeleton of C. □

Lemma 4.8 *A circular category is path-connected as a topological space iff its 1-skeleton is a connected graph.*

Proof: Lemma 4.7 shows that being topologically connected implies that the 1-skeleton is connected. To see the converse, notice that by Lemma 1.16 every point is contained in an open cone and thus contained in the image of a closed cone. Moreover, since a closed cone is path-connected and contains a vertex, every point in the space of the circular category is connected by a path to a vertex. These paths combined with the connectedness of the graph prove that the topological space is path-connected. □

A circular category which satisfies either condition of Lemma 4.8 will be called a connected circular category.

5 General Relators

In this section a generalization of the traditional notion of a relator is defined and some of the basic properties of these general relators are examined. More specifically, the lemmas in this section show that all paths in the boundary of a general relator beyond a certain minimum length can be assigned an orientation relative to the boundary, and that under relatively mild conditions, the shortest loops with winding number 1 must be simple. The section concludes with specialized results on cycles which are piecewise geodesic, relator metrics, and the structure of the automorphism groups of relators.

5.1 Length and Representatives

A generator is an edge labeled by A. A relator in the usual sense can be thought of variously as a cyclically reduced word $W \in A^*$, as a cyclically reduced word $W \in A^*$ together with all cyclic conjugates of W and its inverse, as a cycle, as a deterministically labeled loop, or as a 2-dimensional cell with a deterministic labeling by A. The last description is the most geometric and the one which is used as the basis of the generalization which follows.

A general relator over A is defined to be a circular cone of rank $n > 1$ with a deterministic labeling by A. Note that a general relator is, by definition, a circular complex; it cannot be a circular category which is not also a circular complex. This extends the usual notion of a relator in a group. The restriction that a general relator be a circular complex corresponds to the usual restriction that relators be cyclically reduced. If C is a labeled circular category, and c

is an object in C of rank at least 2, then $R = C/c$ is a general relator which inherits its labeling from the attaching map $\phi_c : R \to C$. The boundary of R, written ∂R, is defined to be ∂c, the boundary of c.

Since ∂R is homotopically equivalent to a circle, this implies, among other things, that $\pi_1(\partial R, v) \cong \mathbf{Z}$ for every vertex v in ∂R. Since there are two possible isomorphisms of $\mathbf{Z}$ with itself, there are two possible isomorphisms between $\pi_1(\partial R, v)$ and $\mathbf{Z}$ which are called the two orientations of the relator, R. If a particular isomorphism from $\pi_1(\partial R, v)$ to $\mathbf{Z}$ is chosen then an integer can be assigned to every oriented loop based at v and by path conjugation to every oriented loop in ∂R. Since every loop in R is contained in its 1-skeleton and thus in ∂R, all loops in R are assigned an integer. If the other isomorphism with $\mathbf{Z}$ is chosen then the integer assigned to a particular loop will be opposite in sign. Thus the absolute value of this integer is a characteristic of the unoriented loop, which will be called the winding number of the loop. From algebraic topology it is known that a loop in a circle is contractible or homotopic to a point iff it has a winding number of 0. If both the loop and the relator are oriented then a plus or a minus sign is added to the winding number according to whether it is orientation-preserving or -reversing. A loop in R is called a representative of R if it has a winding number of 1, regardless of sign.

Lemma 5.1 *Every general relator has a representative. Specifically, if R is a general relator, then R contains a loop in its 1-skeleton which has winding number 1.*

Proof: Let $f : [0,1] \to \partial R$ be a topological path which starts and ends at the same point and has winding number 1 as a topological loop. Since the boundary of R is topologically connected f is conjugate to a loop which starts and ends at a vertex in ∂R. Thus, without loss of generality, assume this is true of f. By Lemma 4.7, f is homotopic relative to its endpoints to a path in the 1-skeleton of ∂R. Since the endpoints remain fixed, the loop formed in the 1-skeleton is homotopic relative to its basepoint to the loop formed by f. Because winding numbers are invariant under such homotopies, the proof is complete. □

Lemma 5.2 *For every general relator R there exists a simple loop U with some winding number, say n, such that for all loops V with winding number m,*

$$\frac{|U|}{n} \leq \frac{|V|}{m}$$

Proof: First of all, by Lemma 5.1 there exists at least one loop with a non-trivial winding number. If this loop is not simple then it can be decomposed into two loops of shorter length, one of which again has a non-zero winding number. Continuing in this way a simple loop with non-zero winding number. The process must stop since the loops are getting shorter. Since the relator R is finite, there are only a finite number of simple loops and thus there exists a simple loop U such that the length of U divided by its winding number is the smallest possible among all simple loops with non-zero winding number in R.

The ratio of the length of the simple loop U to its winding number is actually minimal among all loops with non-zero winding number in R. The argument proceeds by contradiction.

Suppose that there exists a loop V such that the length of V divided by its winding number is strictly less than the length of U over n. Suppose further that V is selected so that the length of V is the shortest possible among all counterexamples. Since U yields the smallest value among all simple loops, V must be non-simple. Let V_1 and V_2 be loops read in R so that the loop V_1 followed by the loop V_2 is the loop V. If the orientation of the loops V_1 and V_2 are identical then since $\frac{a+c}{b+d}$ is between $\frac{a}{b}$ and $\frac{c}{d}$ for all positive numbers a, b, c, and d, the ratio of length to winding number for at least one of the two smaller loops must be at least as small as that of V, contradicting the minimality of the counterexample V. If, on the other hand, the orientations of the two loops are opposite in sign, then one of the loops has a shorter length and a winding number at least as large, so that its length divided by its winding number is smaller than the ratio for V, again contradicting the minimality of V as a counterexample. Since all possibilities lead to contradictions, it must be that no counterexample exists. □

The length of the general relator R, written $|R|$, is defined to be $\frac{|U|}{n}$, where U and n are as described in Lemma 5.2. A simple loop U with winding number n for which $|U| = n|R|$ is called a standard geodesic of R.

Corollary 5.3 *If V is a loop in a general relator R and m is its winding number then $|V| \geq m|R|$.*

The universal cover of an abstract labeled loop W is called W^∞ since it is an abstract labeled line which reads the bi-infinite word

$$\dots WWWW \dots$$

By analogy, if R is a general relator, then the universal cover of ∂R is called R^∞. A path W in ∂R which lifts to a geodesic path in R^∞ will, by an abuse of notation, be called a geodesic. This is equivalent to requiring that the path W be the shortest path in its homotopy class relative to its endpoints. Similarly, a loop W in ∂R is called a geodesic in ∂R if there does not exist a shorter loop in ∂R which is homotopic to W. Since by Corollary 5.3 U is the shortest loop of winding number n, and since loops in circles are homotopic iff they have the same winding number, U certainly deserves to be called a geodesic loop. It is important to realize that geodesic loops in the boundary of a general relator R do not necessarily lift to geodesic loops or lines in the regular covers of ∂R. See Example 1 in Section 6. The standard geodesic U, however, does have this property.

Lemma 5.4 *If U is a standard geodesic of a general relator R, then the natural map from U^∞ to ∂R lifts a line U^∞ in R^∞ which is a geodesic.*

Proof: A path between points in the line U^∞ whose length is strictly less than the portion of U^∞ demarcated by these points can be used to create a loop in ∂R which violates the minimality of the loop U as guaranteed by Lemma 5.2. $\square$

5.2 Ends and Orders

Let R be a general relator, let R^∞ be the universal cover of its boundary, and let $f : R^\infty \to \partial R$ be the covering map. Since $\pi_1(\partial R, v) \cong \mathbf{Z}$ by definition, there is a 1 to 1 correspondence between the vertices of $f^{-1}(v)$ and the integers $\mathbf{Z}$. In particular, if W is an oriented representative of R based at v and v_0 is a vertex with $f(v_0) = v$, then define $v_n \in f^{-1}(v)$ to be the vertex which is the endpoint of the lift of W^n to R^∞ starting at v_0. The following lemma is immediate.

Lemma 5.5 *If R is a general relator, $f : R^\infty \to \partial R$ is the covering map, W is an oriented representative of R based at a vertex v, and v_0 is a vertex in R^∞ with $f(v_0) = v$, then for all integers n the lift of the path W starting at v_n ends at v_{n+1}.*

In order to facilitate the proofs in this subsection it is convenient to assign a unique subscript to every vertex (and open cone) in R^∞. This can be done efficiently by choosing a spanning tree in R. Let T be a fixed spanning tree of the 1-skeleton of R. More generally, T can be extended to a spanning tree T' of the poset which defines the cone complex ∂R. Since T' is a tree and thus simply connected, it follows by Lemma 1.13 and Lemma 1.14 that there are uniquely defined lifts $g_i : T' \to R^\infty$ which send v to v_i. The same lemmas can be used to show that the images of these lifts partition the open cones in R^∞ in the following sense.

Lemma 5.6 *If R is a general relator, and f, g_i, and T' are as described above, then the open cones in the images $g_i(T')$ are disjoint in R^∞ and their union contains every open cone in R^∞.*

For each open cone u in R, define $u_i = g_i(u)$. Since the images $g_i(T')$ partition the open cones in R^∞, this definition assigns a unique subscript to each open cone in R^∞.

Lemma 5.7 *Let R be a general relator, and let f, v, W and v_i, $i \in \mathbf{Z}$ be as described above. If a finite portion of the 1-skeleton of R^∞ is removed, then there is a large positive integer N, based on n and R, such that the open cones u_i are in the same connected component of the result for all $i \geq N$ and for all open cones u in R. Similarly, there is a large negative integer N' such that the open cones u_i are also in a connected component of the result for all $i \leq N'$ and for all u in R.*

Proof: Let j be the largest subscript assigned to a open cone in the ball $\mathrm{Ball}(v_0, n)$. Next, consider the path which starts at v_0 and reads the word W, and let k be the absolute value of the smallest subscript assigned to a vertex or edge in this path. The statement of the lemma will be true for all N strictly bigger than $j+k$. To see this, notice that $g_i(T')$ is connected and all of the open cones involved has the same subscript so that for all $i > j$ and for all $u \in R$, u_i is in the same connected component as v_i. And finally, v_i is connected to v_{i+1} by a path reading W and all of the vertices and edges involved have subscripts of at least $i-k$. An analogous argument works for large negative subscripts. $\Box$

When a finite portion of the 1-skeleton of R^∞ is removed, then the connected component of the result which contains the vertices v_i for arbitrarily large positive integers i will be called the positive end, and it will be denoted $R^{+\infty}$. Similarly, the connected component of the result which contains the vertices v_i for arbitrarily large negative integers i will be called the negative end, and it will be denoted $R^{-\infty}$. Notice that these do not necessarily describe distinct connected components. When these components are distinct the ends are said to be disconnected. The labeling of these components as ends conforms to the use of this term by John Stallings ([]). The following Corolllary is recorded for later use. It states that the positive and the negative end are the only possible ends.

Corollary 5.8 *If R is general relator and a finite number of open cones are removed from R^∞ then there are at most two connected components which contain an infinite number of open cones. In particular the only connected components with an infinite number of open cones are the (possibly identical) ends $R^{+\infty}$ and $R^{-\infty}$.*

Proof: This follows immediately from Lemma 5.7 once it is noticed that there are only a finite number of open cones which are assigned to each subscript. $\Box$

Now that the ends of R^∞ have been defined, the next several lemmas will investigate the ordering induced on disjoint, connected pieces of R^∞ which disconnect the ends. Each finite B_i under consideration will be the finite union of open cones. It may not be a subcomplex since the boundary of an open cone in B_i need not be included in B_i. Also, notice that B_i is required to be connected in the topological sense, but that the portion of the 1-skeleton in B_i need not be connected. Such a B_i is said to be connected to $R^{+\infty}$ if there are (topological) paths which start at open cones in B_i and end at vertices of arbitrarily large positive subscript. Similarly, B_i is connected to $R^{-\infty}$ if there are paths which start at open cones in B_i and end at vertices of arbitrarily large negative subscript. When no restrictions are placed on the connecting paths it is clear that all B_i are connected to both $R^{+\infty}$ and $R^{-\infty}$, but once restrictions are introduced, this is no longer the case.

Lemma 5.9 *For $i = 1, 2$, let B_i be a finite union of open cones in R^∞ which is topologically connected and whose removal disconnects the ends $R^{+\infty}$ and $R^{-\infty}$. If B_1 and B_2 are disjoint, then the following statements are equivalent:*

(1) B_1 can be connected to $R^{-\infty}$ by paths which is disjoint from B_2.

(2) B_1 cannot be connected to $R^{+\infty}$ by paths which is disjoint from B_2.
(3) B_2 can be connected to $R^{+\infty}$ by paths which is disjoint from B_1.
(4) B_2 cannot be connected to $R^{-\infty}$ by paths which is disjoint from B_1.

Proof: If (1) is true, and (2) is false, then the B_2 does not disconnect the ends, contradicting the hypothesis of the lemma. Thus $(1 \Rightarrow 2)$. A similar argument shows that $(3 \Rightarrow 4)$. Next, let U^∞ be the lift to R^∞ of the standard geodesic of R, and consider $B_i \cap U^\infty$. The points in the intersection are at most a bounded distance apart, bounded by the finite diameter of B_i. Thus there is a last vertex or edge in U^∞ which is contained in B_i. Notice that there must be at least one vertex or edge in the intersection since B_i disconnects the ends and U^∞ connects them. Moreover, the last vertices/edges in U^∞ of B_1 and B_2 are distinct since B_1 and B_2 are disjoint. Finally the B_i which contains the later vertex/edge is connected to the end $R^{+\infty}$ by the remaining portion of U^∞. This shows that at least one of the B_i is connected to R^∞ by paths which are disjoint from the other, and as a consequence $(2 \Rightarrow 3)$. An analogous argument shows $(4 \Rightarrow 1)$. $\square$

If any of the four equivalent conditions are satisfied then write $B_1 < B_2$. Notice that $B_1 < B_2$ implies (assumes) that they are disjoint.

Corollary 5.10 *For $i = 1, 2$, let B_i be a finite union of open cones in R^∞ which is topologically connected and whose removal disconnects the ends $R^{+\infty}$ and $R^{-\infty}$. If $B_1 < B_2$, then $B_2 \not< B_1$. Moreover, if B_1 and B_2 are disjoint, then either $B_1 < B_2$ or $B_2 < B_1$.*

Proof: The first statement is immediate since condition (2) for the inequality $B_1 < B_2$ and condition (3) for the ineqality $B_2 < B_1$ directly contradict each other. The second statement is immediate from the proof of Lemma 5.9. $\square$

Lemma 5.11 *For $i = 1, 2$, let B_i be a finite union of open cones in R^∞ which is topologically connected and whose removal disconnects the ends $R^{+\infty}$ and $R^{-\infty}$. If $B_1 < B_2$ and $B_2 < B_3$, then B_2 is 'between' B_1 and B_3 in the sense that any path connecting a point in B_1 to a point in B_3 must contain a point in B_2. In particular B_1 and B_3 are disjoint, and moreover, $B_1 < B_3$.*

Proof: By the definition of $<$, B_1 is connected to $R^{-\infty}$, and B_3 is connected to $R^{+\infty}$ by paths which are disjoint from B_2. If B_1 were connected to B_3 by a path disjoint from B_2, then B_2 would not disconnect the ends. This contradiction proves the first statement. If B_1 and B_3 are not disjoint, the the empty path connects the common point to itself. Finally, notice that by the first part of the lemma, the paths connecting B_1 to $R^{-\infty}$ which are disjoint from B_2, must also be disjoint from B_3. $\square$

For $i = 1, 2$, let B_i For $i = 1, 2$, let B_i be a finite union of open cones in R^∞ which is topologically connected and whose removal disconnects the ends $R^{+\infty}$ and $R^{-\infty}$.

Corollary 5.12 *Let $F = \{B_i | i \in I\}$ be a collection where each B_i is a finite union of open cones in R^∞ which is topologically connected and whose removal disconnects the ends $R^{+\infty}$ and $R^{-\infty}$. If the elements of F are pairwise disjoint, then $<$ defines a discrete total linear order on F which is order isomorphic to the integers, the positive integers, the negative integers, or the integers from* 1 *to* n.

Proof: The ordering is total and linear as an immediate consequence of Corollary 5.10 and Lemma 5.11. To see that it is discrete, let B_1 and B_2 be arbitrary members of F with $B_1 < B_2$. Since $B_1 \cup B_2$ contains only a finite number of open cones, by Corollary 5.8 there are only a finite number of open cones which are not in either B_1, B_2, $R^{+\infty}$ or $R^{-\infty}$. In particular, there are only a finite number of members of F which contain one of these open cones, and thus only a finite number of C in F with $B_1 < C < B_2$. Since F is totally ordered, there is a well-defined member of F which comes after B_1 and a well-defined member of F which comes before B_2. The final comment simply lists the possible discrete total orders. □

5.3 Width and Orientation

The width of a general relator R, written ω_R, can be described as the smallest non-negative integer such that the removal of a ball of radius ω_R centered at any vertex in R^∞ will disconnect the ends. Let u be a vertex in R, u_0 a lift of u to R^∞, W an oriented representative, and n a positive integer. If the removal of a ball of radius n centered at u_0 disconnects the ends of R^∞, then the ends will also be disconnected by the removal of a ball of radius n centered at any other lift of u. Also notice that the orientation of W and even the choice of the representative W is irrelevant since the reversal of the orientation merely switches the labels on the ends, and another representative would leave the lifts of u identical. By the above reasoning, it is sufficient to check that the ends are disconnected for a single lift of each of the finite vertices in R instead of the infinite number of vertices in R^∞. The following lemma shows that such a number always exists.

Lemma 5.13 *Let R be a general relator, and let f, v, W and v_i, $i \in \mathbf{Z}$ be as described above. There is a unique smallest integer ω_R such that for all vertices u in R^∞, the removal of a ball of radius ω_R centered at u will always disconnect the ends.*

Proof: The existence of such a unique minimum non-negative integer will be immediate once it is shown that there is at least one which satisfies the conditions. To begin let k be the maximum distance between vertices in the tree T, so that u_i and w_i are within k units of each other for all u and w in R, even when the possible paths are restricted to those which only use vertices with subscript i. Next, consider the edges of R. If a particular edge is lifted to R^∞ then its endpoints have subscripts, and the absolute value of the difference between these subscripts is a value which is independent of the lift of the edge. This is because any two lifts differ by an automorphism of R^∞ and the automorphism which sends v_0 to v_i also sends u_j to u_{i+j} for all $u \in R$ and for all $j \in \mathbf{Z}$. Since R is finite, there are only a finite number of edges in R and thus a maximum absolute change in the value of a subscript when traversing a single edge. Call this largest value m.

Finally, if n is chosen so that n is greater than $2k + m|W|$, then the ball $\mathrm{Ball}(u_i, n)$ contains all vertices with subscripts in the range $[i-m, i+m]$. This can be seen by tracing a path from u_i to v_i to the v_j with the correct subscript, and then to the final vertex with subscript j. These can be chosen so that their lengths are at most k, $|i-j| \cdot |W|$, and k, respectively. This must disconnect the vertices with large negative subscripts from those with large positive subscripts. For suppose otherwise: let X be a path from any vertex with a subscript less than $i - m$ with a vertex with subscript greater than $i + m$. Since it cannot contain any vertices with a subscript in $[i - m, i + m]$, it must contain a single edge which starts at a vertex with subscript less than $i-m$ and ends at a vertex with subscript greater than $i+m$, but this contradicts the definition of m. Since u and i are arbitrary, this shows that there is at least one n which works, and this completes the proof. □

If $4\omega_R \leq |R|$ then the general relator R is called thin. Let U be a path in a general relator R from v to u and let W be an oriented representative of R based at v, and let $v_i, i \in \mathbf{Z}$ be defined as above using the word W. If the path U in R^∞ which starts at v_0 ends at a vertex in the same connected component as the vertices v_i with large positive subscripts once the ball $\mathrm{Ball}(v_0, \omega_R)$ has been removed from the 1-skeleton of R^∞, then U is said to have the same orientation as W, or rather that U is positively oriented relative to W. If U ends at a vertex in the same connected component as the vertices v_i with large negative subscripts then U is said to have the opposite orientation as W, or that U is negatively oriented relative to W. If neither is the case, and this possibility does occur, then U is called unoriented relative to W.

Lemma 5.14 *Let R be a general relator and let U be a path in R whose lift to R^∞ has endpoints which are a distance of at least $2\omega_R$ apart. If U starts at a vertex v in R, and W is an oriented representative of R which is based at the same vertex v then U is oriented either positively or negatively relative to W.*

Proof: Let the vertices v_i, $i \in \mathbf{Z}$ be defined as usual from W and a particular lift v_0 of v. By the definition of ω_R it is known that when the ball $\mathrm{Ball}(v_0, \omega_R)$ is removed from the 1-skeleton of R^∞, the v_i with large negative and the v_j with large positive subscripts are contained in distinct connected components, but it is not known whether there are other connected components. The proof will show that all of the vertices in the other connected components must be very close to v_0. Let u be the endpoint of the path U lifted to R^∞ so that it starts at v_0, and assume that u is in a connected component other than the two expected ones.

Consider a path in R^∞ from a vertex v_i with a large negative subscript to a vertex v_j with a large positive subscript. If this path contains a vertex in a component other than the two expected ones and the ball $\mathrm{Ball}(v_0, \omega_R)$, then there must be a subpath which starts and ends in $\mathrm{Ball}(v_0, \omega_R)$ which contains the vertex in the other component. Since $\mathrm{Ball}(v_0, \omega_R)$ is connected (through paths to and from v_0), this subpath can be removed and replaced by a subpath completely in $\mathrm{Ball}(v_0, \omega_R)$. Continuing in this way, the path from v_i to v_j will eventually contain only vertices in the components of v_i and v_j, and the ball $\mathrm{Ball}(v_0, \omega_R)$.

The assumption about the path U implies that the ball $\mathrm{Ball}(u, \omega_R)$ does not contain any of the vertices in $\mathrm{Ball}(v_0, \omega_R)$. But this in turn means that $\mathrm{Ball}(u, \omega_R)$ does not contain a vertex of the path described above, and so does not disconnect the component of v_i from the component of v_j, contradicting the definition of ω_R. Thus U must end in one of the two standard components. $\square$

Up until this point, the orientation of a path U has only been defined relative to an oriented representative of R. The orientation of a path U relative to an unoriented representative W will now be defined, provided that U is a subword of the cycle W. The definition goes as follows. Let W' be the cyclic conjugate of the cycle W or W^{-1} such that U is an initial segment of the word W', that is, so that $W' = UV$ for some word V. The orientation of U relative to the unoriented representative W is then the orientation of U relative to the oriented representative W'. Using this definition, it is possible to have a representative loop $W = XUYU^{-1}$ and to have both instances of the word U be positively oriented relative to W. Intuitively, the orientation of U is positively oriented relative to W if the subword U, as it is situated in W, helps to complete the loop, and it is negative if, as it is situated in W, goes in the opposite direction.

5.4 Simple Representatives

Although by Lemma 5.1 every general relator contains a representative loop, it is not true that this representative must be simple or even that a simple representative must exist at all. A presentation of $\mathbf{Z}$ is given in Section 6 which will provide a counterexample. The lemmas below show that if, however, a general relator is thin, then there does have to exist a simple representative, which can then be used to create a deformation retraction.

Lemma 5.15 *For all geodesic paths X read in R, there exist paths Y, P and Q in R such that the path Y is a geodesic read in a standard geodesic loop U, the lengths of P and Q are each less than ω_R, and the cycle $XPYQ$ is readable in R as a contractible loop. Since both X and Y are geodesic, $||X| - |Y|| \leq |P| + |Q| < 2\omega_R$.*

Proof: Lift the path X to R^∞ and consider its relation to the line U^∞, the lift of a standard geodesic loop U. For convenience the lift of X will also be called X. Since the balls of radius ω_R disconnect R^∞, there is a vertex or edge of U^∞ contained in the ball centered at the start of X. And since edges are included in an open ball only if its endpoints are, there must be a vertex. If Q is the geodesic path from the vertex in U^∞ to the start of X, then the length of Q is strictly less than ω_R. Similarly, there is a path P from the endpoint of X to a vertex of U^∞ whose length is strictly less than ω_R. By Lemma 5.4 the portion of U^∞ marked off by the end of P and the start of Q is a geodesic path which will be called Y. Since $XPYQ$ is a loop in the universal cover, it is contractible. The homotopy which proves the contractibility of $XPYQ$ in R^∞, when composed with the projection down into the boundary of R, provides a contraction of the image of the loop $XPYQ$ in the boundary of R. Finally, X geodesic implies that $|X| \leq |Y| + |P| + |Q|$ and Y geodesic implies that $|Y| \leq |X| + |P| + |Q|$. These inequalities can be rearranged to show that $||X| - |Y|| \leq |P| + |Q| < 2\omega_R$. □

The next lemma is a technical result which will be used to prove Lemma 5.17.

Lemma 5.16 *Let $X = X_1X_2$ be a geodesic loop in a general relator R with $|X_i| \geq 2\omega_R$ for $i = 1, 2$. Let x_0 be a lift of x to R^∞, let x_i be the endpoint of the path reading X^i which starts at x_0, let x'_i be the endpoint of the path reading X_1 starting at x_i, and let B_i and B'_i be the open balls of radius ω_R centered at x_i and x'_i respectively. The following four implications are true:*

(1) If $B_i < B'_i$ then $B'_{i-1} < B_i < B'_i < B_{i+1}$

(2) If $B'_i < B_i$ then $B_{i+1} < B'_i < B_i < B'_{i-1}$

(3) If $B'_i < B_{i+1}$ then $B_i < B'_i < B_{i+1} < B'_{i+1}$

(4) If $B_{i+1} < B'_i$ then $B'_{i+1} < B_{i+1} < B'_i < B_i$

In particular, $B_0 < B'_0$ implies that $B_i < B_j$ for all integers $i < j$, and $B'_0 < B_0$ implies that $B_j < B_i$ for all integers $i < j$.

Proof: Since all of these implications have essentially identical proofs, it is sufficient to show that $B_i < B'_i$ implies $B'_i < B_{i+1}$. First of all, each of these three balls is a finite subgraph of the 1-skeleton of R^∞ whose removal disconnects the infinite ends by the definition of ω_R and by Lemma 5.13. Next, B'_i and B_{i+1} are disjoint since X_2 is the shortest path between their centers and it was to have a length of at least $2\omega_R$. Similarly B_i and B'_i are disjoint and B_i and B_{i+1} are disjoint, since the shortest distances between their centers are $|X| \geq m|R| \geq |R| \geq 4\omega_R$ and $|X_1| \geq 2\omega_R$, respectively. Therefore, by Corollary 5.12, B_i, B'_i and B_{i+1} are linearly ordered. Without loss of generality, assume that $B_i < B_{i+1} < B'_i$. By Lemma 5.11, the path X_1 from x_i to x'_i must contain a point in B_{i+1}. Let $X_1 = X_{11}X_{12}$ be a partition of X_1 so that the endpoint of X_{11} lies in B_{i+1}. The fact that $X_{12}X2$ is a subword of X guarantees that it is a geodesic, its least is at least that of X_2 which in turn is at least $2\omega_R$, but since it starts and ends in B_{i+1}, there is a path connecting the startpoint to the endpoint of length strictly less than $2\omega_R$, contradiction. Thus the assumption that $B_{i+1} < B'_i$ must have been false. The final statements follow immediately by iterating the other implications. □

Lemma 5.17 *For all geodesic loops X in a thin relator R, $||X| - m|R|| < 2\omega_R$ where m is the winding number of the loop X, and moreover the winding number of X is uniquely determined by this inequality. In particular, $\frac{|X|}{|R|}$ rounded to the nearest integer is the winding number of X.*

Proof: Let U be a standard geodesic loop read in R and let n be its winding number. Orient the loops U and X so that they induce the same orientation on R. Next, let x be a point in the loop X, let P be a path of minimal length from x to any point u in U, and without loss of generality assume that X and U are the words determined by the vertices x and u. Since the loops X and U are geodesic loops, by definition X and U are geodesic words. Since the loops X^n and U^m both have the same winding number, namely, nm, it follows that $X^nPU^{-m}P^{-1}$ is a loop of winding number 0.

Let x_0 be any point in universal cover R^∞ which is sent to x by the covering map f. Next, let u_0 be the endpoint of the path P, lifted to R^∞ so that it starts at x_0. Then define x_i and u_i as the endpoints of the paths X^i and U^i lifted to R^∞ to start at x_0 and u_0, respectively. Since the loop $X^nPU^{-m}P^{-1}$ has winding number 0 in ∂R, it lifts to a loop based at x_0 in R^∞. Notice that this shows that the path P lifts so that it starts at x_n and ends at u_m. If the loop $X^nPU^{-m}P^{-1}$ is lifted so that it starts at x_n then we can conclude that the path P lifted to start at x_{2n} will end at u_{2m}. In general, the vertex x_{jn} is connected to the vertex u_{jm} by a path whose length is at most the length of P, for all integers j. The above steps could be repeated with a loop Y in ∂R of winding number 1, based at a vertex y and with a path Q starting at y and ending at x. The same argument would show that y_{jmn} is connected to x_{jn} by a path whose length is at most the length of Q, and it is connected to the vertex u_{jm} by a path whose length is at most the length of QP. This will now be used to show that the vertices x_i for large positive i are contained in the positive end $R^{+\infty}$.

Consider the removal of the ball of radius ω_R centered at x_0. By the definition of ω_R this ball disconnects the ends of R^∞. Since u_{jm} for large positive values of j is within $|QP|$ of y_{jmn}, and since by Lemma 5.4 U^∞ is an infinite geodesic, it follows that j can be chosen large enough so that the path QP from y_{jmn} to u_{jm} is disjoint from the ball around x_0. If this were false, then $|QP| + \omega_R$ would bound $|U^{jm}|$ for all j, contradiction. Similarly, there is a large negative k such that the path QP from y_{kmn} to u_{km} is disjoint from the ball centered at x_0. This shows specifically that the path $U^{(j-k)m}$ connecting u_{km} to u_{jm} in R^∞ is a path connecting the two disconnected ends $R^{+\infty}$ and $R^{-\infty}$, and must therefore pass through the ball centered at x_0. In particular, there is a point in U^∞ which is connected to x_0 by a path of length less than ω_R. Since the path P was chosen to have minimal length, this shows that $|P| < \omega$.

Next, by repeating the above arguments with x_i replacing x_0 as the center of the ball to be removed, there always exists some vertex in U^∞ which is connected to x_i by a path of length less than ω_R. Define v_i to be the last vertex or edge in $\mathrm{Ball}(x_i, \omega_R) \cap U^\infty$. Since edges are contained in open balls only when their endpoints are as well, the last vertex or edge is in fact a vertex. Let P_i denote the shortest path from x_i to v_i. Notice that $|P_i| < \omega_R$ for all i. Also note that by symmetry v_0 and v_n are connected by a cyclic conjugate of the word U^m whose length is exactly $|U^m| = m|U| = mn|R|$. The second equals sign follows from the definition of $|R|$ as the length of its standard geodesic divided by its winding number. Since U^∞ is a geodesic in R^∞, it follows that $mn|R|$ is the minimum distance in R^∞ from v_0 to v_n.

Because R is thin and X is a geodesic loop, $|X| \geq m|R| \geq |R| \geq 4\omega_R$. In particular, there is a partition of $X = X_1X_2$ such that $|X_i| \geq 2omega_R$ for $i = 1, 2$. Let B_i be the open ball $\mathrm{Ball}(x_i, \omega_R)$. Since it is known that $B_0 < B_m$, it follows by Lemma 5.16 that $B_i < B_j$ for all $i < j$. In particular, the vertices v_i occur in order within U^∞, and since U^∞ is a geodesic the length of the minimal path from v_0 to v_n is exactly the minimal distance from v_0 to v_1 plus the minimal distance from v_1 to v_2 plus ... plus the minimal distance from v_{n-1} to v_n. This in turn means that there must be some subscript i $(0 \leq i < n)$ such that the distance from v_i to v_{i+1} is at most $m|R|$. For this i identified above, there is a loop $XP_{i+1}VP_i^{-1}$ which goes from x_i to x_{i+1} to v_{i+1} to v_i and back to x_i. The word V (or its inverse) is a subword of U^∞. Since X is a geodesic loop in the general relator R, the path X lifts to a geodesic in R^∞ (since any shorter path in R^∞ is mapped under the covering map f to a shorter path in R, contradiction). Thus, $|X| \leq |P_i| + |P_{i+1}| + |V| \leq 2\omega_R + m|R|$ by the above estimates. On the other hand $|X| \geq m|R|$ by Corollary 5.3. Finally, $m|R| \leq |X| < m|R| + 2\omega_R$, or $||X| - m|R|| \leq 2\omega_R$, as was to be shown. If in addition R is a thin relator, then by definition $|R| \geq 4\omega_R$, and in particular the above inequalities become $m|R| \leq |X| < m|R| + 2\omega_R \leq m|R| + \frac{1}{2}|R|$, or $m \leq \frac{|X|}{|R|} < m + \frac{1}{2}$, which completes the proof. □

Lemma 5.18 *If X and Y are loops in a thin relator R which are based at the same vertex, and X followed by Y is a geodesic loop in R, then X and Y are identically oriented. In particular, the winding number of XY is that of X plus that of Y.*

Proof: By Lemma 5.17 the length of the geodesic loop XY is strictly within $2\omega_R$ of the winding number of XY times $|R|$. Since XY is a geodesic, Y does not have winding number 0, or else it could be removed without changing the homotopy class of the loop. Consequently, $|Y| \geq |R|$ and $|X| \leq |XY| - |R|$. Using the fact that R is thin in combination with Lemma 5.17, the winding number of X is seen to be strictly less than that of XY. A similar argument can be used to show the winding number of Y is less than that of XY. Since the winding number of XY is a combination of the winding number of X and the winding number of Y with the appropriate signs attached, the appropriate signs must both be positive in this case. □

Although the content of Lemma 5.18 has great intuitive appeal, the assumption that the relator is thin is necessary for the conclusion to follow, as Example 1 in Section 6 will show. The same counterexample will also show that the conclusion of the following lemma is unwarranted when the word ‘thin’ is removed.

Lemma 5.19 *If X is a geodesic representative of a thin general relator R, then X is simple.*

Proof: Let X be a geodesic representative which is not simple. In this case X can be written as the sum of two distinct loops, X_1 and X_2, based at the same vertex. By Lemma 5.18, the sum of the winding numbers of X_1 and X_2 equals the winding number of X, which is 1. Thus one of the loops is of winding number 0 and can be removed without changing the homotopy class of the loop. This would imply that the origin loop X is not a geodesic, contradiction, so the assumption that X is not simple must be false. □

Lemma 5.20 *If X is a geodesic representative of a thin general relator R, then there is a deformation retraction from ∂R to X.*

Proof: By Lemma 5.19, the representative X is simple, and since X is simple the functor f from L_n to R is injective and can be viewed as an inclusion map. Because the winding number of X is 1, it induces an isomorphism on fundamental groups. All of the higher homotopy groups are trivial in both spaces which means that f is actually a weak homotopy equivalence, and by Lemma 1.6 there is a deformation retraction from ∂R to X. □

5.5 Geodesic n-gons

If a loop $V = V_1 V_2 \dots V_n$ is read in a labeled circular category C and each V_i is a geodesic in the homotopy class relative its endpoints in C, then the loop V is called a geodesic n-gon in C. Let V be a geodesic n-gon in the boundary of a general relator R and let the endpoints of V_i be called v_{i-1} and v_i respectively. If the loop V is a representative of R and for each i, $\mathrm{Ball}(v_i, \omega_R)$ does not contain any of the other endpoints or any of the edges of the geodesic paths other than the ones which contain v_i as an endpoint, then V is called a geodesic n-gon representing R. This condition is precisely what is needed to prove the following lemma.

Lemma 5.21 *If $V = V_1 V_2 \dots V_n$ is a geodesic n-gon representing a general relator R, then $|V| < |R| + 2n\omega_R$.*

Proof: Let U be the standard geodesic and suppose that it has a winding number of k. Lift V^k to R^∞ and consider a particular reading of the infinite geodesic U^∞. Let v_0 be the initial vertex of V^k and let v_i be the terminal vertex of the initial segment V^i for all $i = 1, \dots, k$. By the definition of ω_R there is a vertex u_i in U^∞ which passes strictly within ω_R units of each v_i. Moreover, since U has winding number k the vertex u_k can be chosen so that there are paths from v_0 to u_0 and from v_k to u_k which have the same label. In particular, pushing the path from v_0 to u_0 into R and then lifting the result so that it starts at v_k will select the vertex u_k having this property. One consequence of this choice is that the length of the subword of U^∞ marked off by u_0 and u_k is exactly $k|R|$, and thus there is an integer i such that u_i and u_{i+1} mark off a subword of U^∞ of length at most $|R|$. Without loss of generality, assume that $i = 0$ is an integer with this property. Attention now shifts to the word V read from v_0 to v_1.

The terminal vertex of each V_j is within ω_R units of a vertex in U^∞, for $j = 1, \ldots n-1$. Let U_j be the portion of U^∞ marked off by the vertices of U^∞ selected to correspond to the endpoints of V_j in V for all $j = 1, \ldots, n$. Since the definition of a geodesic n-gon representing R implies that the selected vertices in U^∞ occur between u_0 and u_1 in order, the sum of the lengths of the U_j, $j = 1, \ldots, n$ is equal to the length of the path in U^∞ between u_0 and u_1 and thus is at most $|R|$. Since each V_j is a geodesic, $|V_j| < |U_j| + 2\omega_R$ for all $j = 1, \ldots, n$. Combining these inequalities yields the result. □

By the diameter of a general relator R is meant the largest distance between two vertices in R, and it is denoted Diameter(R).

Lemma 5.22 *If R is a thin general relator then Diameter$(R) \leq \frac{1}{2}|R| + 2\omega_R$.*

Proof: Let u and v be vertices in R and let U be a geodesic path between them. If $|U| < 2\omega_R$ then the lemma holds for this pair of vertices, so suppose that $|U| \geq 2\omega_R$. Since U is a geodesic, the lift of the vertex u to R^∞ is strictly between two distinct lifts of the vertex v. The lift of the path U connects the lift of u to one of the lifts of v. Let V be a geodesic path in R^∞ from the other neighboring lift of v to the lift of u. The image of VU in R is now a geodesic 2-gon representing R. By Lemma 5.21 $|R| + 4\omega > |U| + |V| \geq 2|U|$ since U is a geodesic in R. Thus $|U|$ is less than $\frac{1}{2}|R| + 2\omega_R$, completing the proof. □

Suppose that $V = V_1 V_2 \ldots V_n$ is a geodesic n-gon in the boundary of a general relator R, but that it is not a geodesic n-gon representing R because there is an i such that Ball(v_i, ω_R) contains the vertex v_{i-1}. If V' is a geodesic path v_{i-1} to v_{i+1} then the path $V_i V_{i+1}$ can be replaced by the path V', in a process called the consolidation of a pair of vertices. The new loop is a geodesic $(n-1)$-gon in the boundary of R, and its length has been shortened by less than $2\omega_R$ units. To see this notice that since V_{i+1} is a geodesic, $|V_{i+1}| \leq |V'| + |V_i|$, and this can be rewritten as $|V_{i+1}| + |V_i| - |V'| \leq 2|V_i| < 2\omega_R$. When these observations are combined with Lemma 5.21, the result is the following.

Lemma 5.23 *If $V = V_1 V_2 \ldots V_n$ is a geodesic n-gon in the boundary of a general relator R, and there exists a sequence of consolidations of pairs of vertices which results in a geodesic k-gon which represents R, then $|V| < |R| + 2n\omega_R$.*

Proof: The comments above show that the $n-k$ consolidations shorten the length less than $2(n-k)\omega_R$, and by Lemma 5.21, the resulting k-gon has length less than $|R| + 2k\omega_R$. □

Lemma 5.24 *Let $V = V_1 V_2 \ldots V_n$ be a geodesic n-gon of winding number 1 in the boundary of a general relator R. If every subword of the form $V_i V_{i+1} \cdots V_j$ in the cycle V which has an orientation is positively oriented with respect to V, then $|V| < |R| + 2n\omega_R$.*

Proof: The properties of V listed in the lemma are inherited under consolidation of pairs of vertices. Moreover, once no more consolidations are possible, it is clear that the result is a geodesic n-gon representing R. The result then follows by Lemma 5.23. □

5.6 Relator Metrics

In order to generalize the small cancellation hypotheses to the context of general relators it is necessary to introduce appropriate functions which can be used to measure the extent to which a particular path wraps around the boundary of a general relator. These functions will be called relator metrics.

Let R be a general relator. A relator metric on R, denoted d_R, is any metric on R^∞, the universal cover of the boundary of R, which satisfies two additional properties: (1) it must be invariant under label-preserving automorphisms of the universal cover, and (2) the distance between the endpoints of a path which forms a loop when pushed into R must be at least the winding number of the loop. A relator metric on R can be used to measure the length of a path in R, by defining its length as the length of a lift of the path to the universal cover. By property (1) this definition is well-defined. By an abuse of notation, d_R will also represent the function which assigns a length to every path in R, so that the length of a path U will be written $d_R(U)$. Alternatively, the function assigning a length to every path in R can be defined directly without using the universal cover. Specifically, a function d_R which assigns a non-negative real number to every path in R is called a relator metric if it satisfies the following six properties:

1. $d_R(U) = d_R(V)$ whenever UV^{-1} is a contractible loop in ∂R
2. $d_R(U) \geq 0$ and $d_R(U) = 0$ iff U is a contractible loop in ∂R
3. $d_R(U) = d_R(U^{-1})$
4. $d_R(UV) \leq d_R(U) + d_R(V)$
5. if U is a path which forms a loop in ∂R then $d_R(U) \geq \text{wind.num.}(U)$
6. if U and V are paths in R which differ by an automorphism of R then $d_R(U) = d_R(V)$.

Property 1 guarantees that the function is only dependent on the homotopy class of the path, which implies that it defines a function on the endpoints of the lift of the path to the universal cover. Properties 2,3, and 4 are the usual axioms for a metric written in terms of paths pushed down to R. Finally, the fifth and sixth properties correspond to the extra conditions placed on relator metrics as described above. A general relator R together with a relator metric on R is called a measured relator. A set of general relators $\mathcal{R}$ together with the set of corresponding relator metrics d_R, one for each $R \in \mathcal{R}$, is called a set of measured relators.

In traditional small cancellation theory, the relators are cyclically reduced so that the loops formed by these cycles are deterministic and have a width of 0. The natural measure of distance in such a setting is the graph metric on R^∞. In particular if U is a path in R then the length of U can be defined as the geodesic distance between the endpoints of U when it is lifted to R^∞. For traditional relators this is equivalent to the length of the reduction of U in the free group. Another metric, which is a variation on the graph metric, is what will be called the normalized graph metric. The length of a path U in the normalized graph metric is given by dividing the length of U in the graph metric by the length of the relator in which it is read. The normalized graph metric of a path U in a relator R will be denoted $|U|_R$. Note that if U is a path in R which lifts to a geodesic in R^∞ then $|U|_R = \frac{|U|}{|R|}$. Also, property 5 is always an equality for the normalized graph metric. Finally, since a length $|R|$ was defined for all general relators R, the normalized graph metric is well-defined for any general relator.

Lemma 5.25 *The sum of the lengths of the arcs of a representative cell of a measured relator is at least 1. That is, if $W = U_1U_2\ldots U_j$ is a representative of a measured relator R then $\sum_{i=1}^{j} d_R(U_i) \geq 1$. Consequently, if $UXVY$ is a representative of R in which both X and Y measure at most $1-3\alpha$, and both U and V measure less than α, then both X and Y measure at least α.*

Proof: By properties 4 and 6, $\sum_{i=1}^{j} d_R(U_i) \geq d_R(W) \geq 1$. For the second statement, the measures of the four pieces add up to at least 1, but since $d_R(X) \leq 1-3\alpha$, $d_R(U) < \alpha$, and $d_R(V) < \alpha$, it follows that $d_R(Y) > \alpha$. The proof for X is analogous. □

General relators are more complicated than traditional relators because they typically have a non-zero width. As a consequence, distance can be measured in many different ways, using possibly distinct metrics. This gives rise to a variety of ways in which a word or a cycle may be 'reduced'. The different types of reductions are described below as a prelude to the general theory.

Let $\mathcal{R}$ be a set of measured relators, and let μ be a real number with $\frac{1}{2} \leq \mu < 1$. A word W is called μ-reduced if it is reduced in the free group and there do not exist words U and V and a general relator $R \in \mathcal{R}$ such that U is a subword of W, UV is a representative of R, and $|U| > \mu|UV|$. Alternatively, the inequality can be written as $|U| > \frac{\mu}{1-\mu}|V|$. In this form it is clear that over a set of standard relators, a word is $\frac{1}{2}$-reduced iff it is Dehn-reduced in the usual sense, that is, iff it does not contain strictly more than one-half of a representative of a relator. Over a set of general relators, the notion of being one-half of a relator loses its sense, but being μ-reduced is still well-defined. As a result, even in the general case, a word will be called Dehn-reduced iff it is $\frac{1}{2}$-reduced with respect to $\mathcal{R}$. A cycle W will be called μ-reduced iff all of the cyclic conjugates of the word W are μ-reduced. Cycles which are $\frac{1}{2}$-reduced are called Dehn-reduced cycles.

The other types of 'reduced' words and cycles will be defined using relator metrics. Let $\mathcal{R}$ be a set of measured relators, and let μ be a non-negative real number. A word W is called μ-free if it is reduced in the free group and does not contain strictly more than μ of a relator in $\mathcal{R}$ as measured by the relator metric. Specifically, there cannot exist a word U and a general relator R such that U is a subword of W which is readable in R with $d_R(U) > \mu$. The cycle of W is called μ-free iff all of the cyclic conjugates of the word W are μ-free.

Next, let $\mathcal{R}$ be a set of measured relators, and let μ be a real number strictly between 0 and 1. A word W is called μ-complement-free if it is reduced in the free group and there does not exist a subword of W with a complement in a particular relator measuring less than μ. Specifically, there must not exist words U and V and a general relator R such that U is a subword of W, UV is a representative of R, and $d_R(V) < \mu$. A cycle is called μ-complement-free iff all of its cyclic conjugates are μ-complement-free. Notice that by the properties of relator metrics, $d_R(V) < \mu$ means that $d_R(U) > 1 - \mu$. Thus a word or cycle W which is not μ-complement-free is not $(1-\mu)$-free. Or, said differently, if W is $(1-\mu)$-free then it is also μ-complement-free.

Both the normalized graph metric and relator metrics will be retained since graph metrics are necessary to complete the inductive step in Dehn's algorithm, but relator metrics are more convenient for various constructions.

5.7 Automorphism Groups

Under certain conditions, the automorphism groups of general relators are easy to describe. One such condition involves what are called crucial cones. Let R be a general relator, and consider an open cone in ∂R which is not contained in the boundary of any other open cone in ∂R. Let B be this open cone together with all of the open cones in its boundary which are not contained in any open cones in R except those which are themselves contained in this particular open cone. Topologically speaking, B is the complement of the closure of the complement of a maximal closed cone in R. In the case of a traditional relator, this is a description of the open edges in the boundary of the relator. Notice that this collection of open cones B contains an open cone of maximum rank and that all of B is contained in the closure of this open cone. Moreover, the closure of B is itself either an edge or a general relator S contained in ∂R. Let B_0 be a lift of this B to R^∞, and let S_0 be the lift of S which contains B_0. If the removal of B_0 disconnects the ends of R^∞ then define B^+ to be the intersection of S_0 with the positive end $R^{+\infty}$. Similarly, let B^- be the intersection of S_0 with the negative end $R^{-\infty}$. If the removal of B_0 not only disconntects the ends of R^∞ but it is also true that B^+ and B^- do not contain any loops which have nontrivial winding number in ∂S, then B_0 is called a crucial cone in R^∞ By extension, B will be called a crucial cone in ∂R. In a traditional relator, all of the open edges of the boundary of the 2-cell are crucial. The terminology derives from the fact that B is crucial for the completion of loops in ∂R, in the sense that every representative of R must contain at least one vertex or edge in B.

A slight generalization of this goes as follows: let C be a subcomplex of ∂R such that $g(C) = C$ for all $g \in \mathrm{Aut}(R)$, and let $C^\infty = f^{-1}(C)$ where $f : R^\infty \to \partial R$ is the usual covering map. So long as C is connected and C contains a representative of R, it can be proved that C^∞ has at most two ends, denoted $C^{+\infty}$ and $C^{-\infty}$, using the same proof as before. Such a subcomplex C will be called a core of the general relator R. As above, let B be the complement of the closure of the complement of a maximal open cone in C, let S be the closure of B, let B_0 be a lift of B to C^∞, let S_0 be the lift of S to C^∞ which contains B_0, and let B^+ and B^- be the intersections of S_0 with $C^{+\infty}$ and $C^{-\infty}$ respectively. If B_0 disconnects the ends of C^∞, and B^+ and B^- do not contain any loops with nontrivial winding number when viewed as paths in S_0, then B_0 will be called a crucial cone in C^∞ and B will be called a crucial cone in C. In the following, R will be said to have a crucial cone in its boundary if there exists some core C and some B for which B is a crucial cone of C.

Now suppose that R is a general relator which has a crucial cone and fix a core C which contains one. Since the crucial cones in C^∞ disconnect the ends of C^∞, and since by construction they are pairwise disjoint, by Lemma 5.12 they are totally ordered. The proofs of the lemmas stated earlier in this section are unchanged if C and C^∞ are used instead of ∂R and R^∞. One way to see this is to notice that since C is a subcomplex, the cone over C is a general relator in its own right. Next, since all of the lifts of a crucial cone in C are crucial cones in C^∞, there is no first or last crucial cone in C^∞ and the total ordering of the crucial cones in C^∞ is order isomorphic to $(\mathbf{Z}, <)$. If this ordering is imposed on the crucial cones in C, there is a well-defined notion of a next and a previous crucial cone along the boundary ∂R in a particular direction, although it is no longer, strictly speaking, an ordering. If there is only one crucial cone in ∂R, then this next crucial cone in ∂R may turn out to the original crucial cone itself.

It should also be noted that the property of being a crucial cone is preserved under automorphisms of the relator R, and it is also preserved under automorphisms of R^∞. By one of the defining properties of C, an automorphism of R is also an automorphism of C. In addition, the total ordering of the crucial cones in C^∞ is either preserved or reversed under an automorphism depending on whether the automorphism preserves or reverses the ends of C^∞. This is immediate from the definition of the ordering. Thus, if R contains a crucial cone in its boundary there is a group homomorphism from $\mathrm{Aut}(R^\infty)$ to $\mathrm{Aut}(\mathbf{Z}, <)$ which is the infinite dihedral group $\mathbf{D}_\infty$. The homomorphism is given by observing the action of the automorphism on the linearly ordered crucial cones in C^∞. Similarly, if C contains exactly r crucial cones, then there is a group homomorphism from $\mathrm{Aut}(R)$ to the dihedral group of order $2r$, which is denoted $\mathbf{D}_{2r}$.

The kernels of these homomorphisms contain those automorphisms of R^∞ and ∂R which fix all of the crucial cones in C^∞ and C respectively. Notice that the morphism from $\mathrm{Aut}(R^\infty)$ to $\mathbf{D}_\infty$ also shows that the only elements in $\mathrm{Aut}(R^\infty)$ of finite order are either orientation-reversing or else they are automorphisms in the kernel of this map. The following lemmas will show that the kernels of these homomorphisms are 2-groups. Lemma 5.26 and Lemma 5.27 will be proved by simultaneous induction on the rank n of the general relator R under consideration. As a final note, a set of general relators is called closed under subcones if whenever S is a general relator contained in $R \in \mathcal{R}$, S is also in $\mathcal{R}$.

Lemma 5.26 *Let $\mathcal{R}$ be a set of general relators closed under subcones and suppose that all general relators in $\mathcal{R}$ have at least one crucial cone in their* *boundary. If H is a subgroup of the automorphism group of a general relator $R \in \mathcal{R}$, then for some r there is a group homomorphism from H to $\mathbf{D}_{2r}$, whose kernel is a 2-group. In particular, H is isomorphic to a cyclic or a dihedral group extended by a 2-group.*

Proof: Let C be a core of R and let r be the number of crucial cones in C. Since the presence of crucial cones in C already implies the existence of a group homomorphism $f : \mathrm{Aut}(R) \to \mathbf{D}_{2r}$, it only remains to show that the kernel of this map is a 2-group. Let B be a crucial cone in C, and let B_0 be a lift of B to C^∞. If g is an automorphism of R which fixes all crucial cones in C, then there is an automorphism g' of R^∞ which fixes B_0 and is equal to g when composed with the covering map from C^∞ to C.

Let S be the general relator which results from the closure of B_0 Define B^+ and B^- be the intersection of S with $C^{+\infty}$ and $C^{-\infty}$ respectively. Since B_0 is a crucial cone in C^∞, B_0 disconnects the two infinite ends of C^∞, and B^+ and B^- are thus disjoint. Moreover, since the action of g' either preserves the ends or switches them, it follows that either $g'(B^+) = B^+$ and $g'(B^-) = B^-$, or else $g'(B^+) = B^-$ and $g'(B^-) = B^+$. By definition, B^+ and B^- do not contain the lift loop in S with a nontrivial winding number. Thus the general relator S with subcomplexes B^+ and B^- satisfies all of the hypotheses of Lemma 5.27. Moreover, since the rank of S is strictly less than of R, Lemma 5.27 can be applied to show that g' is an automorphism whose order is a power of 2. This in turn implies that the order of g itself is a power of 2, and that the kernel of the homomorphism f is a 2-group. $\Box$

Lemma 5.27 *Let $\mathcal{R}$ be a set of general relators closed under subcones and suppose that all general relators in $\mathcal{R}$ have at least one crucial cone in their boundary. If R is a general relator in $\mathcal{R}$, B_1 and B_2 are disjoint subcomplexes of ∂R which do not contains any loops with a nontrivial winding number, and H is the subgroup of Aut(R) of automorphisms which either sends B_1 to B_2 and vice versa or else which fixes B_1 and B_2 separately, then H is a 2-group.*

Proof: Let C be a core of R and let r be the number of crucial cones in C. The presence of crucial cones in C already implies the existence of a group homomorphism $f : \mathrm{Aut}(R) \to \mathbf{D}_{2r}$. Let g be an automorphism in H, and consider g^2. By the definition of H, g^2 must fix both B_1 and B_2, and in addition, it is orientation-preserving. Since B_1 does not contain any loop with a nontrivial winding number, by Lemma 1.13 it is possible to lift B_1 to R^∞. Call this lift B_0. Corresponding to the automorphism g^2 of C, there is an automorphism of R^∞ (and C^∞) which fixes B_0. Since this automorphism of C^∞ is also orientation-preserving, but of finite order, it must fix all of the crucial cones in C^∞. As a consequence, g^2 must fix all of the crucial cones in C, and thus g^2 is a member of the kernel of the map from $\mathrm{Aut}(R)$ to $\mathbf{D}_{2r}$. By Lemma 5.26, the order of g^2 must be a power of 2. This shows that the order of g itself is a power of 2, and since g was chosen at random, the group H is a 2-group. $\Box$