Classification of Surfaces

Richard Koch

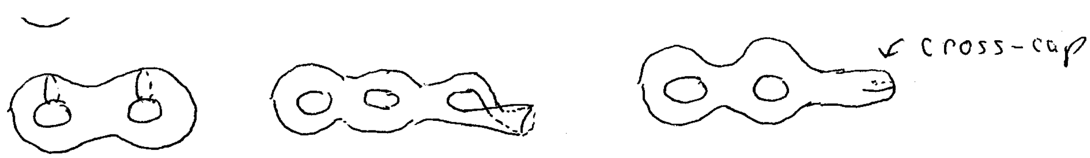
November 20, 2005

1 Introduction

We are going to prove the following theorem:

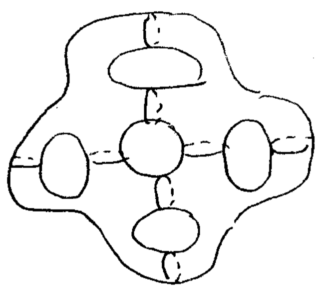
Theorem 1 *Let $\mathcal{S}$ be a compact connected 2-dimensional manifold, formed from a polygon in the plane by gluing corresponding sides of the boundary together. Then $\mathcal{S}$ is homeomorphic to exactly one of the following:*

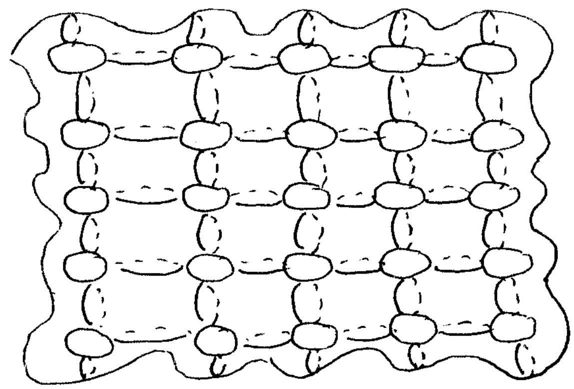
- $T^2\#\ldots\#T^2$*, that is, a sphere or g-holed torus*
- $T^2\#\ldots\#T^2\#RP^2$*, that is, a connected sum of a g-holed torus and a projective space*
- $T^2\#\ldots\#T^2\#K$*, that is, a a connected sum of a g-holed torus and a Klein bottle*



It can be proved that any compact connected 2-dimensional manifold can be obtained in this way, so we are actually classifying *all* compact connected 2-dimensional manifolds. But our comments about the general case involve some hand wringing, while our proof of the above theorem is completely rigorous.

The truth of the above theorem is far from obvious. For example, consider the following two-dimensional objects. It is not immediately obvious that these have the above form. However, it is easy to show that each of these objects can be obtained as a quotient object of a polygon by gluing corresponding sides together. Namely, cut the object into pieces, and mark the boundaries to explain how they are to be glued back together. Then arrange these pieces in the plane, deforming the pieces if necessary to make them fit, and marking the remaining unconnected edges along the boundary of this region.

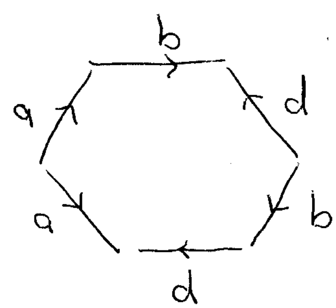




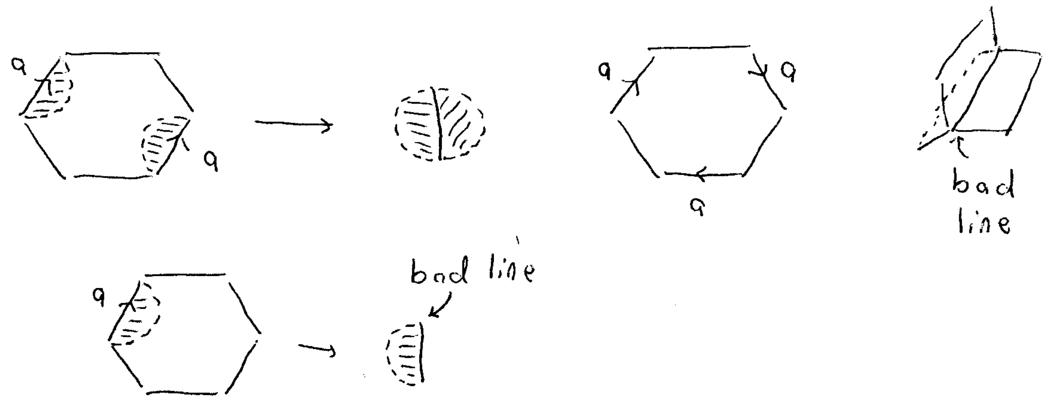
Our proof of the theorem comes from *Introduction to Topology* by Solomon Lefschetz, and from *Lehrbuch der Topologie* by H. Seifert and W. Threlfall. Lefschetz was one of the twentieth centurie's greatest topologists, although the Princeton graduate students liked to say "Lefschetz never stated an incorrect theorem, nor gave a correct proof." Lefschetz is one of the characters in the book and movie *A Beautiful Mind*. Seifert is also a very important topologist, and Seifert and Threlfall's book is often quoted. I highly recommend it.

2 Preliminaries

Suppose we have a polygon with sides labeled a, b, and so forth. We can read off the sides clockwise along the boundary, obtaining a string called the *symbol* of the diagram. This symbol can be replaced with a cyclic permutation without changing the polygon. For example, the diagram below has symbol $abd^{-1}bda^{-1}$ and also symbol $d^{-1}bda^{-1}ab$.



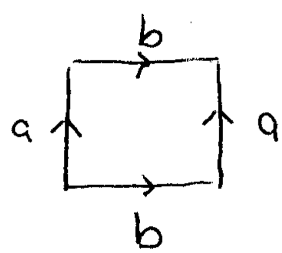
We wish to glue corresponding sides together. To get a manifold, each side must occur exactly twice along the boundary, as the following pictures show. There *might* also be a condition at vertices to insure that the quotient space is locally Euclidean at the vertices, but the proof will show that no such vertex condition is needed.



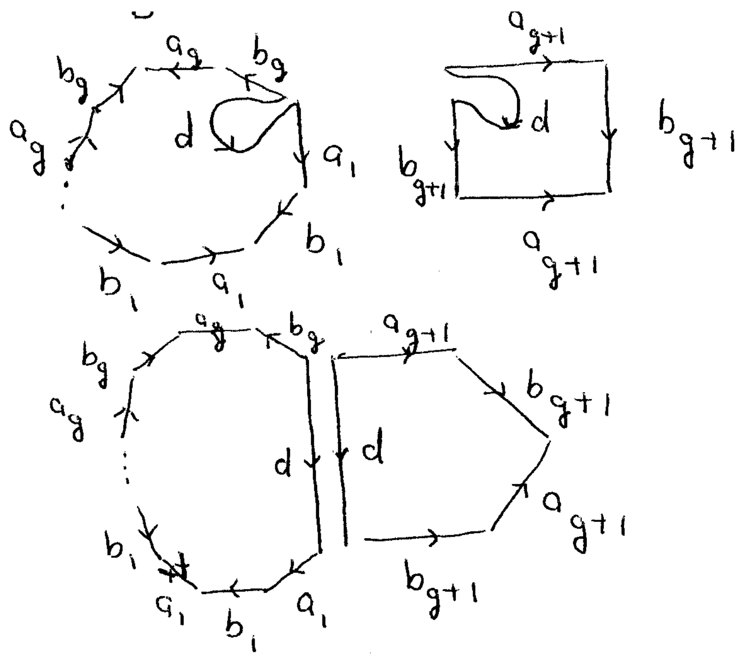
3 The Standard Models

Theorem 2 *The g-holed torus $T^2\#\ldots\#T^2$ can be represented by a polygon with $4g$ sides and symbol $a_1b_1a_1^{-1}b_1^{-1}\ldots a_gb_ga_g^{-1}b_g^{-1}$.*

Proof: The following picture of T^2 has symbol $aba^{-1}b^{-1}$.

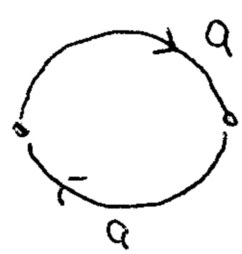


Inductively assume the above theorem is true for a g-holed torus; we will prove it for the connected sum of this torus with one additional T^2. The picture below shows each object with a disk removed. The boundary of this disk is labeled d. To form their connected sum, we must glue the two objects together along d. This is shown in the second picture below, which produces the same sort of symbol with one additional $a_{g+1}b_{g+1}a_{g+1}^{-1}b_{g+1}^{-1}$.

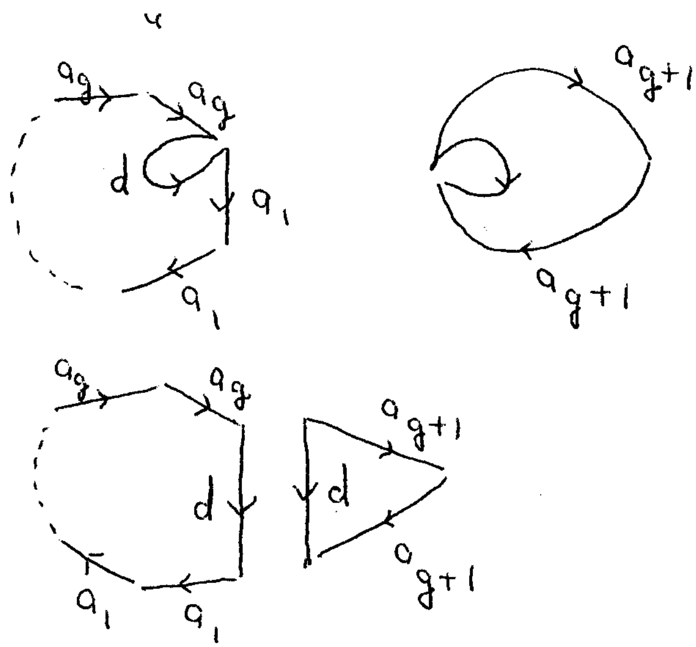


Theorem 3 *The connected sum of g copies of projective space $RP^2\#\ldots\#RP^2$ can be represented by a polygon with $2g$ sides and symbol $a_1a_1a_2a_2\ldots a_ga_g$.*

Proof: We prove this by induction; it is certainly true when $g=1$ because the object below describes our standard method of forming projective space.

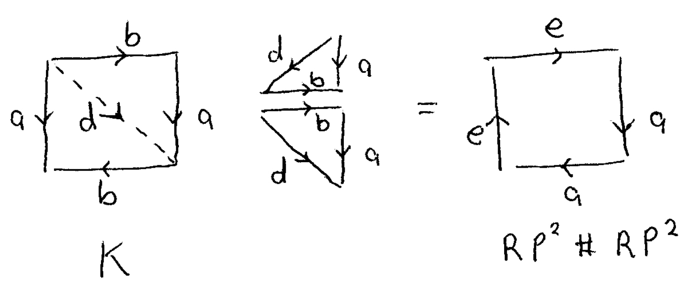


Assume the theorem is true for a connected sum of g copies of projective space. We will show it true when we add one additional copy of projective space. To form this last connected sum, we must remove a disk from each object and glue them together along the boundary of these disks. This process is shown below.

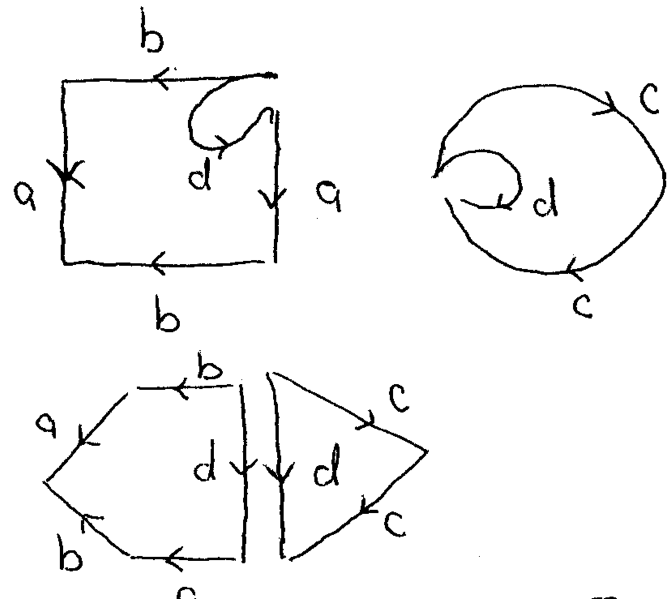


Theorem 4 *The following objects are homeomorphic:* $RP^2 \# RP^2 \# RP^2 \approx K \# RP^2 \approx T^2 \# RP^2$.

Proof: The following picture shows that $K \approx RP^2 \# RP^2$, and thus proves the first half of the assertion.

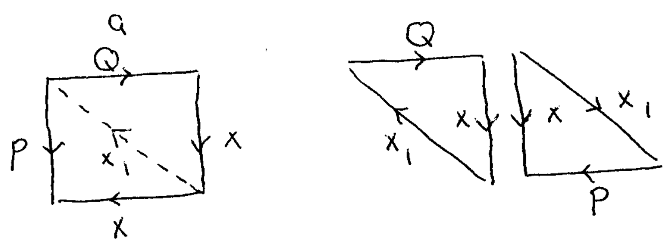


The symbol for $RP^2\#RP^2\#RP^2$ is $aabbcc$ by theorem three. The following picture shows that the symbol for $T^2\#RP^2$ is $aba^{-1}b^{-1}cc$. We must prove that these symbols give isomorphic objects.



Lemma 1 *If x represents a side and P and Q represent sequences of sides, then $xxP^{-1}Q \approx x_1Px_1Q$ for an appropriate side x_1.*

Proof: See the picture below.



Remark: Using this result, we can prove $aba^{-1}b^{-1}cc \approx aabbcc$ algebraically. Each line of the argument below is an application of the above lemma. To make the argument clearer, we place parentheses around P and Q. Sometimes we will cyclically permute the symbols; in that case the original placement is at the end of one line and the permutation is at the beginning of the next line. We start with $aba^{-1}b^{-1}cc = cc(ab)(a^{-1}b^{-1})$.

$$cc(ab)(a^{-1}b^{-1}) \to c_1(ab)^{-1}c_1(a^{-1}b^{-1}) = c_1b^{-1}a^{-1}c_1a^{-1}b^{-1}$$
$$b^{-1}(a^{-1}c_1a^{-1})b^{-1}(c_1) \to b_1b_1(a^{-1}c_1a^{-1})^{-1}(c_1) = b_1b_1ac_1^{-1}ac_1$$
$$a(c_1^{-1})a(c_1b_1b_1) \to a_1a_1(c_1)(c_1b_1b_1) = a_1a_1c_1c_1b_1b_1$$

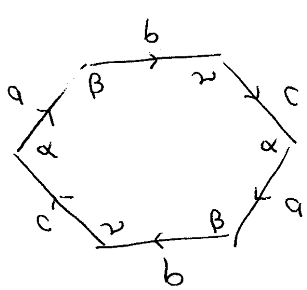
Theorem 5 *The space $RP^2\#\ldots\#RP^2$ is homeomorphic to $T^2\#\ldots\#T^2\#RP^2$ if the number of copies of RP^2 is odd, and to $T^2\#\ldots\#T^2\#K$ if the number of copies of RP^2 is even.*

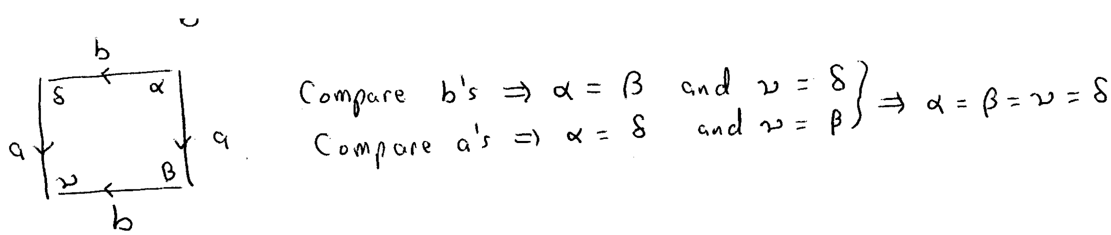
Proof: Write $RP^2\#\ldots\#RP^2 = (RP^2\#RP^2)\#(RP^2\#RP^2)\#\ldots\#(RP^2\#RP^2)\#M$ where M is either one copy of RP^2 or a connected sum of two copies of RP^2. Using theorem 4, we can replace of pair $RP^2\#RP^2$ by T^2 since there is at least one additional RP^2 at the end. The remaining M is either RP^2 or $RP^2\#RP^2 \approx K$.

Remark: The proof of our main theorem thus reduces to showing that any symbol of a polygon can be converted to one of the two forms $a_1b_1a_1^{-1}b_1^{-1}\ldots a_gb_ga_g^{-1}b_g^{-1}$ or $a_1a_1a_2a_2\ldots a_ga_g$ and that the manifolds corresponding to different such symbols are not homeomorphic.

4 First Reduction

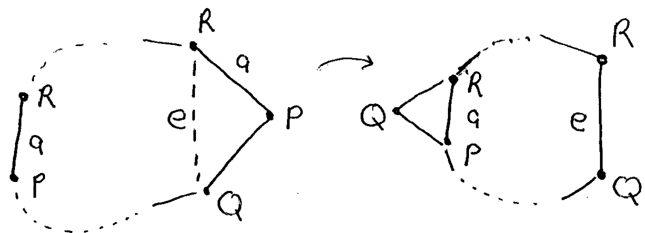
When the sides of a polygon represented by a symbol are identified, it may or may not happen that all vertices glue to the same point. For example, the vertices of the polygon $abcabc$ glue to three distinct points, but the vertices of $aba^{-1}b^{-1}$ glue to the same point, as the pictures below show.



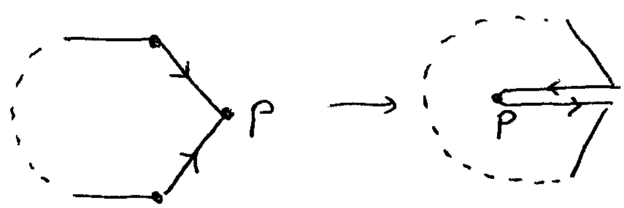


Theorem 6 *A polygon can always be replaced with an equivalent polygon in which all vertices glue to the same point.*

Proof: Suppose that one of the vertices after gluing is P but it is not the only vertex. We will produce a new diagram with one fewer P-vertex and one additional Q-vertex for $Q \neq P$. Start by tracing the boundary of the polygon until an edge joins a P-vertex to a Q-vertex for $Q \neq P$. Let a be the *previous* edge; in the picture R may or may not be a P-vertex. Cut as indicated and glue to the matching a edge as in the picture below. Notice that in the new polygon the number of P-vertices has been reduced by one.



Repeat this process until there is only one P-vertex. Then the two edges joining this vertex must have arrows pointing in to P, or else arrows pointing out from P because otherwise there would be another P-vertex at the end of one of these edges. We can join the edges as indicated below, completely eliminating the final P-vertex.

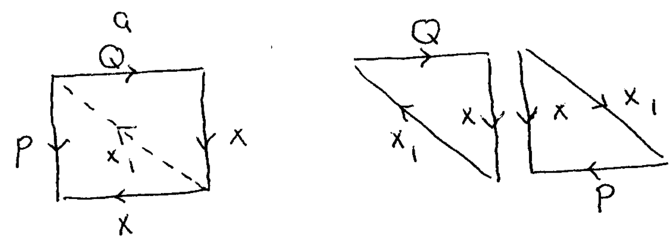


5 Reduction Techniques

The remaining reductions depend on two simple observations. The first has already been presented, but we list it here again:

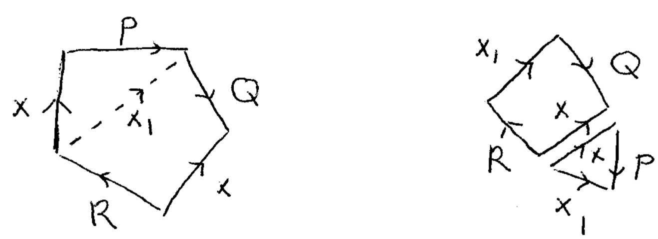
Lemma 1 *If x represents a side and P and Q represent sequences of sides, then for an appropriate side x_1:*

$$xxP^{-1}Q \approx x_1Px_1Q$$



Lemma 2 *If x represents a side and P, Q, and R represent sequences of sides, then for an appropriate side x_1:*

$$xPQx^{-1}R \approx x_1QPx_1^{-1}R$$



Remark: Notice that we can tack on an extra S at the front of these symbol sequences, since the sequences can be symmetrically permuted. For instance, $SxxP^{-1}Q = xxP^{-1}QS \approx x_1Px_1QS = Sx_1Px_1Q$.

Each side a in a symbol sequence appears twice, either in the form $aPaQ$ or in the form $aPa^{-1}Q$. The lemmas say that any sequence P of sides inside an a-a pair can be thrown outside if inverted, and any pair PQ of sequences of sides inside an a-a^{-1} pair can be commuted.

6 Second Reduction

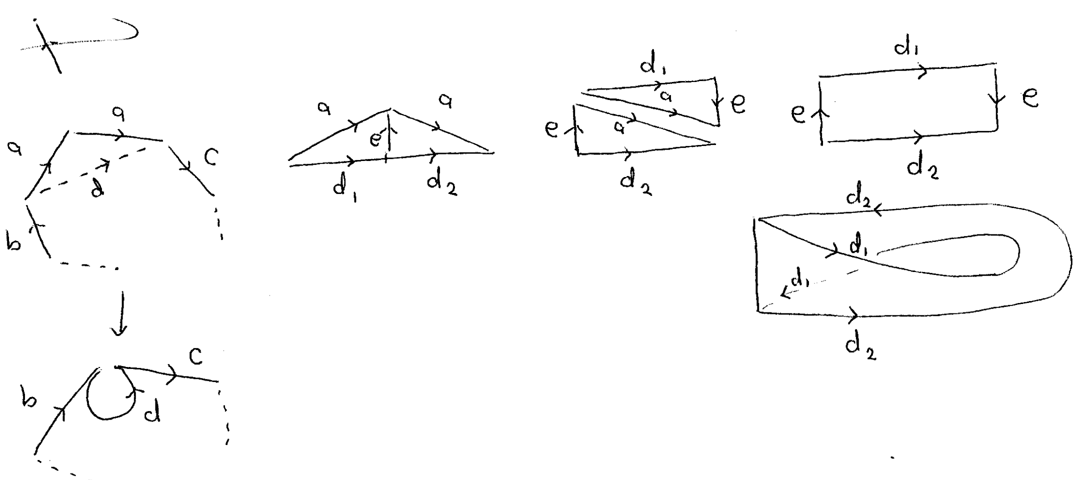
Theorem 7 *We can always reduce a symbol sequence to a form in which all a-a pairs appear side by side.*

Proof: If the sequence has the form $SxPxQ$, replace it with the sequence $Sx_1x_1P^{-1}Q$. This process will never destroy an already existing adjacent pair because both members of the pair can be assumed to belong to S or P or Q and the pair will still be adjacent in S or P^{-1} or Q. Continue until all a-a pairs are adjacent.

Remark: When the author of our book discusses connected sums, he leaves unproved a central fact: the connected sum $M\#N$ only depends on M and N, and not on the choices of the disks D_1 and D_2 to be removed from M and N or the gluing homeomorphism φ from the boundary of D_1 to the boundary of D_2. The omission of this fact leaves proofs which depend on connected sums somewhat shaky.

Consequently, in the reduction of surface symbols to the forms $a_1b_1a_1^{-1}b_1^{-1}\dots a_gb_ga_g^{-1}b_g^{-2}$ or $a_1a_1\dots a_ga_g$, we are not going to use connected sums. We only use connected sums to identify the ultimate canonical forms with familiar objects. This step could be replaced by a direct construction of these two symbols from standard models of our canonical objects.

If we were willing to use connected sums in the reduction argument, we could simplify that argument a little at this moment. Consider the picture below, which shows an adjacent pair a-a. Suppose we cut off this adjacent pair along the dotted edge d. The



picture shows that the piece which has been cut off is homeomorphic to a Mobius band. Suppose we connect the sides b and c which came just before and just after the a-a pair.

The new polygon would represent a simpler compact surface, and the picture shows that the original surface could be obtained from this simpler surface by cutting out a disk and gluing in a Mobius band. Said another way, if the simpler surface is M then the original surface is $M \# RP^2$.

We could cut out *all* adjacent a-a pairs, so the resulting simpler polygon would only contain a-a^{-1} pairs. An argument from the following section of these notes would then show that such a polygon represents a g-holed torus. Consequently the original polygon represents a connected sum of such a torus with several copies of RP^2.

But we will not do that, so at the moment our polygon contains a-a pairs and also a-a^{-1} pairs, and all of the a-a pairs are adjacent.

7 Third Reduction

Definition 1 *Two sides a and b in a symbol form a* crossed pair *if up to cyclic permutation the symbol has the form $PaQbRa^{-1}Sb^{-1}T$.*

Theorem 8 *A symbol can always be converted to a reduced form in each crossed pair consists of adjacent sides $aba^{-1}b^{-1}$ and each a-a pair consists of adjacent edges aa.*

Proof: Consider the symbol $PxQyRx^{-1}Sy^{-1}T$. We can assume that each adjacent pair aa and each previously simplified adjacent pair $aba^{-1}b^{-1}$ occurs in one of P, Q, R, S, or T. We will simplify the given symbol so x and y are replaced by adjacent sides $x_1y_1x_1^{-1}y_1^{-1}$. In the process we will not change the order of sides in P, Q, R, S, or T Hence this process can be continued until all crossed pairs are adjacent.

The reduction will use the $xPQx^{-1} \approx x_1QPx_1^{-1}$ simplification of lemma 2. To make the process easier to read, we surround P and Q with parentheses before applying the lemma. Occasionally we will cyclically permute symbols; in these cases, the original order is at the end of one line, and the permuted symbol is at the beginning of the next line.

$$Px(Q)(yR)x^{-1}Sy^{-1}T \rightarrow Px_1(yR)(Q)x_1^{-1}Sy^{-1}T$$

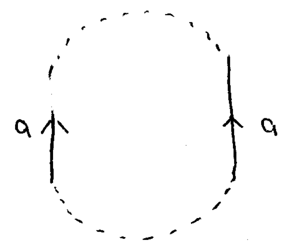
$$Px_1y(RQ)(x_1^{-1}S)y^{-1}T \rightarrow Px_1y_1(x_1^{-1}S)(RQ)y_1^{-1}T$$

$$y_1x_1^{-1}(SRQ)(y_1^{-1}TP)x_1 \rightarrow y_1x_2^{-1}(y_1^{-1}TP)(SRQ)x_2$$

$$x_1y_1x_2^{-1}y_1^{-1}TPSRQ$$

8 Eliminating Uncrossed Pairs

At this point, all a-a pairs are adjacent and all crossed pairs $a_1b_1a_1^{-1}b_1^{-1}$ are adjacent, but these groups of adjacent symbols may be separated by sequences of other unclassified sides. These sides must belong to a-a^{-1} pairs because all a-a pairs are adjacent. So suppose we have a symbol representing a polygon of the form below.



Notice that whenever b is an edge in the top dotted sequence of this picture, the matching b edge must also belong to this top sequence. This certainly holds for z-z edges or crossed pairs because these sequences are adjacent, but it also holds for unclassified edges because if b is in the top sequence and b^{-1} is in the bottom sequence, then a-b represents a crossed edge.

Similarly, whenever b is an edge in the bottom dotted sequence of this picture, the matching b edge must also belong to this bottom sequence.

But then the starting point P of the side a and the ending point Q of the side a represent difference vertices in the quotient manifold, contradicting the first reduction of section four. Indeed, the Q's from both copies of side a are in the top dotted section, and all sides in this section have matching sides also in this section, so none of these sides can require us to glue Q to P. A similar remark holds for the bottom.

The unavoidable conclusion is that there are *no unclassified sides* and thus the complete symbol consists of adjacent a-a and adjacent $aba^{-1}b^{-1}$ sections, one after another, around the boundary of the polygon.

9 The Final Step

If there are no a-a pairs, then the symbol has the form $a_1b_1a_1^{-1}b_1^{-1}\dots a_gb_ga_g^{-1}b_g^{-1}$ and so is one of our canonical forms. If the symbol has no crossed pairs, then it is $a_1a_1\dots a_ga_g$ and so is the other of our canonical forms.

Suppose the symbol has both crossed pairs and a-a pairs. It then has the form $aaPbcb^{-1}c^{-1}Q$ up to cyclic permutation. Apply lemma 1 to convert this as follows:

$$aa(Pbc)(b^{-1}c^{-1}Q) \rightarrow a_1(Pbc)^{-1}a_1(b^{-1}c^{-1}Q) = a_1c^{-1}b^{-1}P^{-1}a_1b^{-1}c^{-1}Q$$

This new symbol has two new pairs $b^{-1}\ldots b^{-1}\ldots$ and $c^{-1}\ldots c^{-1}\ldots$ and one less crossed pair. Reduce using the method of section 6. If there are other crossed pairs, repeat the argument. Ultimately all pairs of sides will have the form aa.

QED.

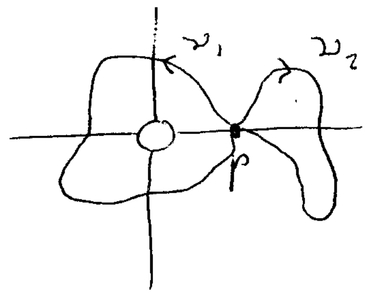
10 Uniqueness

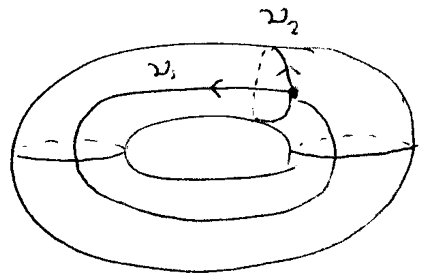
We must still prove that if two surfaces correspond to different canonical symbols, the surfaces are not homeomorphic. This task will be done in the second term of the course. I'd like to give a very rough sketch of the method now. I run the risk that if you understand too much of this paragraph, you won't take the second term because you already know it all.

Our textbook is called *Algebraic Topology*, but where's the algebra? It turns out that it is possible to assign to each topological space X a group $G(X)$. The assignment depends only on the topology of X, so if X and Y are homeomorphic, then the groups $G(X)$ and $G(Y)$ are isomorphic. It is much easier to decide if groups are isomorphic than to decide if topological spaces are homeomorphic; if we discover that $G(X)$ and $G(Y)$ are not isomorphic, then X and Y are not homeomorphic. Incidentally, if you have never heard of a group, you can relax, because they will be defined completely from scratch in 432/532.

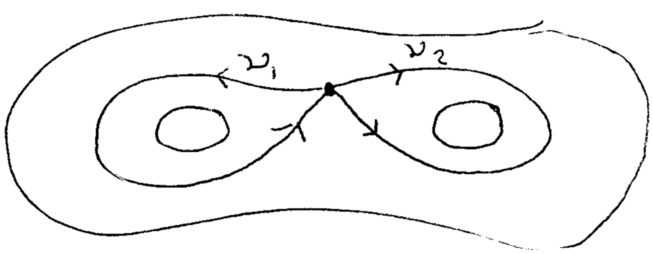
Many different sorts of group have been assigned to topological spaces. We will discuss a particular assignment, called the *fundamental group* and written $\pi_1(X)$. To obtain $\pi_1(X)$, we fix a point $p \in X$ and let $\pi_1(X)$ be the set of all paths which begin and end at p. If one of these paths can be deformed to another path through paths which all start and end at p, we call the paths homotopic and declare that they define the same element of $\pi_1(X)$. Thus $\pi_1(X)$ consists of homotopy classes of paths in X which begin and end at p.

For example, the picture on the left below shows $X = R^2 - \{0\}$. The paths γ_1 and γ_2 define elements of $\pi_1(X)$. The path γ_1 cannot be deformed to the path which is constantly equal to p because it is trapped by the origin. But γ_2 can be deformed to a constant path. Thus γ_1 and γ_2 define different elements of π. The picture on the right show two paths which represent different elements of the fundamental group of a torus.





We can "add" two paths by first traversing one of them and after that the second one. This gives a group structure to $\pi_1(X)$. Usually this group is not abelian. For example, in the picture below we can traverse γ_1 and then γ_2, but it is not possible to deform this to a new path which traverses γ_2 before it traverses γ_1.



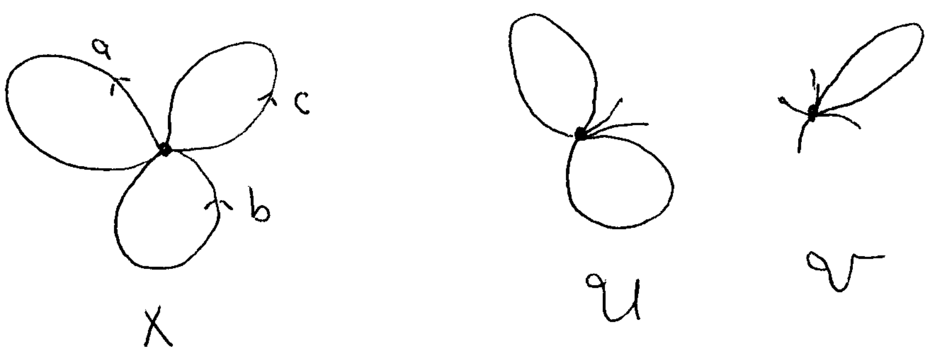
What is $\pi_1(S^1)$? The determination of this group is a high point of the second term. Perhaps we can guess the answer. One possible path is the path which traverses the circle exactly one counterclockwise. Call this path [1]. Then the sum of this path with itself traverses the circle twice. Call this element [2]. Etc. We can also traverse the circle once counterclockwise. Call this path $[-1]$. We can consider a constant path which just stays at p without moving; call this path [0]. These examples and related ones suggest that $\pi_1(S^1) = Z$.

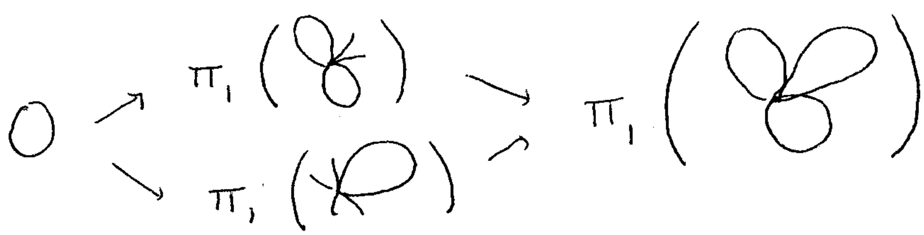
A central result in the subject is called the Seifert-Van Kampen theorem. Suppose $X = \mathcal{U} \cup \mathcal{V}$ where $\mathcal{U}$ and $\mathcal{V}$ are open, and suppose $\mathcal{U}, \mathcal{V}$, and $\mathcal{U} \cap \mathcal{V}$ are pathwise connected. We obtain a diagram

$$
\begin{array}{ccccc}
 & & \pi_1(\mathcal{U}) & & \\
 & \nearrow & & \searrow & \\
\pi_1(\mathcal{U} \cap \mathcal{V}) & & & & \pi_1(X) \\
 & \searrow & & \nearrow & \\
 & & \pi_1(V) & &
\end{array}
$$

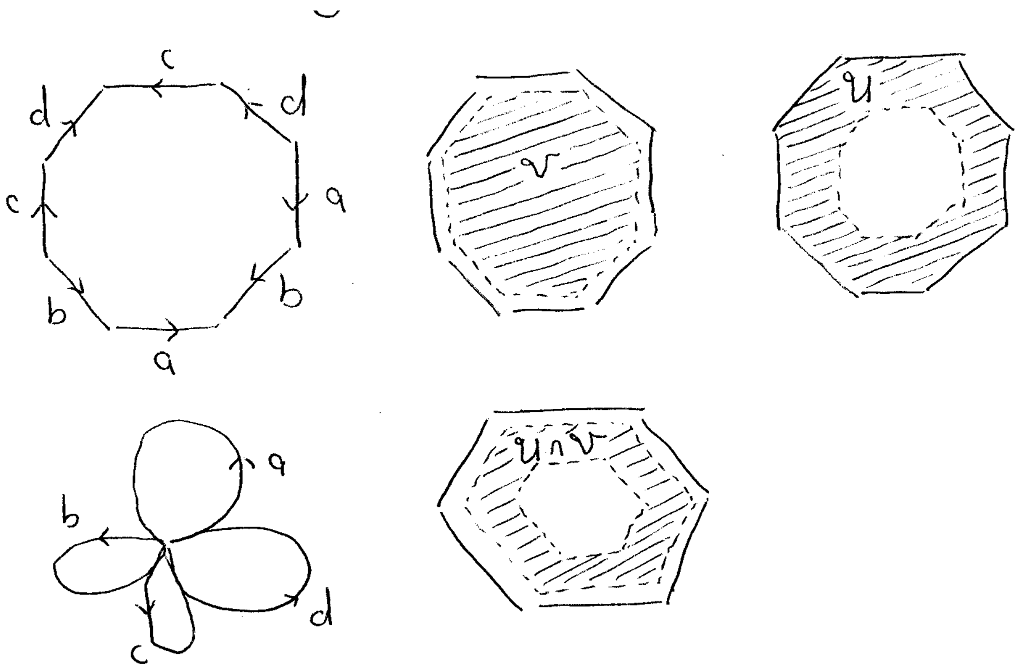
According to the theorem, the images of $\pi_1(\mathcal{U})$ and $\pi_1(\mathcal{V})$ in $\pi_1(X)$ generate $\pi_1(X)$. Moreover, according to the theorem all relations among these generators are consequences of identifying an element in $\pi_1(\mathcal{U})$ with an element in $\pi_1(\mathcal{V})$ if both elements come from a common element in $\pi_1(\mathcal{U} \cap \mathcal{V})$.

Let us apply this theorem to a *bouquet of circles* shown below. I claim that the fundamental group of such a bouquet is just the free group with one generator for each circle. This follows from the Seifert-Van Kampen theorem and induction. For example, apply the theorem to the indicated open sets $\mathcal{U}$ and $\mathcal{V}$ shown below. Then $\pi_1(\mathcal{U})$ is the free group on a and b by induction, $\pi_1(\mathcal{V}) = Z$ is the free group on c by the determination of the fundamental group of a circle, and $\pi_1(\mathcal{U} \cap \mathcal{V}) = 0$ because $\mathcal{U} \cap \mathcal{V}$ can be deformed to a point. The diagram then implies that $\pi_1(X)$ is the free group generated by a, b, c.





Finally, apply the Seifert-Van Kampen theorem to the identification polygon of a surface with $\mathcal{U}$ and $\mathcal{V}$ as indicated below.



Notice that after the boundary of this disk is glued together, this boundary becomes a bouquet of circles, with one circle for each pair of boundary edges. Notice also that $\mathcal{U}$ can be deformed to this bouquet by just pushing everything out to the boundary, notice that $\mathcal{V}$ can be deformed to a single point, and notice that $\mathcal{U} \cap \mathcal{V}$ can be deformed to a circle by pushing everything to the circle in the middle of this piece. Thus the diagram

$$\begin{array}{ccccc} & & \pi_1(\mathcal{U}) & & \\ & \nearrow & & \searrow & \\ \pi_1(\mathcal{U} \cap \mathcal{V}) & & & & \pi_1(X) \\ & \searrow & & \nearrow & \\ & & \pi_1(V) & & \end{array}$$

becomes

$$\begin{array}{ccccc} & & F(\text{edges}) & & \\ & \nearrow & & \searrow & \\ \mathbb{Z} & & & & \pi_1(\text{surface}) \\ & \searrow & & \nearrow & \\ & & 0 & & \end{array}$$

According to the Seifert-Van Kampen theorem, the fundamental group of the surface is generated by $\pi_1(\mathcal{U})$ and by $\pi_1(\mathcal{V})$. The first of these is generated by the edges of the polygon and the second group is trivial.

The relations are generated by $\pi_1(\mathcal{U} \cap \mathcal{V})$, which equals the fundamental group of a circle. There is just one generator of this group, and thus just one relation. This relation is obtained by reading off these edges as we go around the boundary of the polygon. In the orientable case, π_1 is generated by $a_1, b_1, \ldots, a_g, b_g$ with relation

$$a_1 b_1 a_1^{-1} b_1^{-1} \ldots a_g b_g a_g^{-1} | b_g^{-1} = 1$$

In the nonorientable case, π_1 is generated by $a_1, a_2, \ldots, a_g$ with relation

$$a_1 a_1 a_2 a_2 \ldots a_g a_g = 1$$

Unfortunately, π_1 is not abelian. It is *very* difficult to determine whether non-abelian groups defined by generators and relations are isomorphic. But there is a way to abelianize an arbitrary group. Roughly speaking, whenever a and b are elements of the group, we declare that $aba^{-1}b^{-1} = 1$, since this implies that $ab = ba$. Let us abelianize π_1. This new group is called the *first homology group* and often denoted $H_1(X)$.

It follows that H_1 of a torus is generated by *commuting* symbols $a_1, b_1, \ldots, a_g, b_g$ which satisfy the relation $a_1 b_1 a_1^{-1} b_1^{-1} \ldots a_g b_g a_g^{-1} | b_g^{-1} = 1$. If the a's and b's commute, this relation is automatically true and H_1 is an abelian group with $2g$ generators. This group is thus

$$Z \oplus Z \oplus \ldots \oplus Z = Z^{2g}$$

In the nonorientable case, the group H_1 is generated by commuting symbols $a_1, a_2, \ldots, a_g$ satisfying the relation $a_1 a_1 a_2 a_2 \ldots a_g a_g = 1$. Since these elements commute, this relation can be written $(a_1 a_2 \ldots a_g)^2 = 1$. It is easy to see that the elements

$$A_1 = a_1, A_2 = a_2, \ldots, A_{g-1} = a_{g-1}, A_g = a_1 a_2 \ldots a_g$$

also generate the group. These new elements commute and satisfy the single relation $A_g^2 = 1$. The corresponding group is thus

$$Z \oplus Z \oplus \ldots \oplus Z \oplus Z_2$$

By easy group theory, if $H_1(X)$ and $H_1(Y)$ are groups of the form Z^g or $Z^{g-1} \oplus Z_2$, the only way $H_1(X)$ and $H_1(Y)$ could be isomorphic is if both have the first form with the same g, or both have the second form with the same g. The uniqueness of canonical forms for fundamental polygons of compact surfaces follows immediately from this result.