MATH 6991 (1B) – Group Presentations and Group Representations

1 Basics of Presentations

1.1 Reading and exercises

1. Read Gallian, all of ch. 26.
2. Example 2 , page 437, is particularly important.
3. Do exercises page 447: 1,2,3,4,5,6,7,8,9,11,12,13,15,17,18,19,20,25.

1.2 Free Groups

1. On a first reading, it is sensible to think of the constructions in Chapter 26 in an intuitive way. Nevertheless, it is worth noting how certain objects can be more precisely described.
2. At first, the elements of $S = \{a, b, c, \ldots\}$ are mere symbols. S^{-1} is a set disjoint from S but in 1-1 correspondence with S, say

$$S^{-1} = \{a^{-1}, b^{-1}, c^{-1}, \ldots\}.$$

   For now, a^{-1} is merely a new typographic symbol for the element of S^{-1} which corresponds to a. We could just as well write $\hat{a}$, or A, or a_0, etc. These symbols as yet have no algebraic meaning, so that

$$aa^{-1}b^{-1}b$$

   is merely a four letter word. Of course, the algebraic structure comes soon.
3. A word from S is any *formal finite string*, say

$$w = x_1 x_2 \ldots x_k$$

where each $x_j \in S \cup S^{-1}$.

We include the empty word

$$e =$$

We multiply two words by juxtaposing them, i.e. by writing one next to the other. Example:

$$\begin{aligned} w &= abb^{-1} \\ u &= bca \\ e &= \\ wu &= abb^{-1}bca \\ uw &= bcaabb^{-1} \\ w^2 &= abb^{-1}abb^{-1} \\ eu &= bca \\ we &= abb^{-1}. \end{aligned}$$

Two words are *equal* when they have exactly the same symbols in exactly the same positions.

It is best to calculate with words in this intuitive way. However, if you are so inclined, you may want to compare the more rigorous constructions which follow.

4. **A More Precise Description of These Things**

Suppose 1 is a new symbol not in $S \cup S^{-1}$, and let

$$S^* = S \cup S^{-1} \cup \{1\}.$$

Definition. A **word** w in S is any sequence in S^* whose entries are 1 precisely for all positions past some slot k. That is, a word is a function

$$w : \mathbb{N} \to S^*,$$

where, for some integer $k \geq 0$,

$$\begin{cases} w(j) = 1, & \text{for} \quad j > k; \\ w(j) \neq 1, & \text{for} \quad j \leq k. \end{cases}$$

In short,

$$w = [x_1, x_2, \ldots, x_k, 1, 1, 1, \ldots]$$

where $x_j \in S \cup S^{-1}$, for $1 \leq j \leq k$.

We abbreviate this by writing the word in the much more convenient form

$$w = x_1 x_2 \ldots x_k.$$

The cut-off value k is called the **length** of the word w.

Notice that the empty word

$$e = [1, 1, 1, \ldots]$$

has length 0.

5. Also, one can more precisely define the **juxtaposition** of two words:

$$\begin{array}{rcll} w & = & x_1 \dots x_k & (\text{all } \; x_j \in S \cup S^{-1}) \\ u & = & y_1 \dots y_m & (\text{all } \; y_j \in S \cup S^{-1}) \end{array} .$$

   Then the juxtaposed word

$$wu = x_1 \dots x_k y_1 \dots y_m$$

   can be interpreted as a certain function $wu : \mathbb{N} \to S^*$.

   Note that

$$\text{length } (wu) = k + m = \text{ length } (w) + \text{ length } (u).$$

6. Anyway, we have

   - the set $W(S)$ of all words;
   - juxtaposition, an *associative binary operation* on $W(S)$;
   - an *identity* element e (for juxtaposition).

   Thus $W(S)$ is a **monoid**.

7. We don't quite have a group – no inverses! To manufacture inverses, we have to 'factor out' by an equivalence relation. For example, we really do want

$$aa^{-1} \sim e,$$

   or passing to equivalence classes,

$$\overline{a}\,\overline{a}^{-1} = \overline{e}.$$

   We normally abuse this cumbersome notation simply by writing

$$aa^{-1} = e.$$

8. One can likewise give a more precise description of the equivalence relation on page 436 of Gallian. For words $u, v \in W(S)$, we say $u \sim v$ if there is a finite list of words

$$u = u_1, u_2, \ldots, u_n = v \qquad (n \geq 1)$$

such that for $1 \leq i < n$ there are words $w_i, z_i \in W(S)$ and symbols $a_i \in S \cup S^{-1}$ so that either

$$u_i = w_i a_i^{\epsilon_i} a_i^{-\epsilon_i} z_i \text{ and } u_{i+1} = w_i z_i$$

or

$$u_i = w_i z_i \text{ and } u_{i+1} = w_i a_i^{\epsilon_i} a_i^{-\epsilon_i} z_i.$$

(Here $\epsilon_i = \pm 1$, and we agree that $(a_i^{-1})^{-1} = a_i$.)

9. **Remark**. In Gallian, the set of all equivalence classes $\overline{u}$ comprises the free group F. Another approach used by some authors is to employ **reduced words**.

   A word $w \in W(S)$ is **reduced** if it contains no substrings of the form $a_i^{\epsilon_i} a_i^{-\epsilon_i}$, where $a_i \in S \cup S^{-1}$, $\epsilon_i = \pm 1$. For example.

aba^{-1}	is reduced
$a^{-1}ab$	is not reduced.

   Then one may prove

   - each equivalence class $\overline{u}$ contains a unique reduced word u_0, say. In fact, u_0 is always the (unique) shortest word in $\overline{u} = \overline{u}_0$;
   - one defines F to be the set of all reduced words only. If u, v are reduced, then the product in F is defined this way:

$$u \cdot v = \text{unique reduced word in the class } \overline{uv}.$$

   One can check then that F is a group, free over S.

10. Theorem 26.2 (Universal Mapping Property) is not stated as broadly as it should be. A more general version follows.

    Let F be the free group constructed over S. Define a function

    $$j : S \hookrightarrow F$$

    by $j(a) := \overline{a}$, $\forall a \in S$. As a word in S, we note that a has length 1, and that the class $\overline{a}$ contains all sorts of words, such as

    $$aaa^{-1} \ , \ a^{-1}aaa^{-1}a \ , \ bb^{-1}acc^{-1} \ , \ \ldots \ .$$

    Remark: After the Theorem, we will prove that the function j must be 1–1. Since the a's and $\overline{a}$'s therefore pair off, we may safely identify each $a \in S$ with the corresponding $\overline{a} \in F$. Effectively, we may think of S as a subset of F.

    In fact, if you are very fussy, you can literally do this:

    - remove the subset $j(S) = \{\overline{a} : a \in S\}$ from $\{\overline{u} : u \in W(S)\}$, and replace it by S. Now, as a set

      $$F = [\{\overline{u} : u \in W(S)\} \ - \ j(S)] \ \cup S \ .$$

    - correctly redefine the binary operation on this new F, so as to be consistent with the old definition $\overline{u} \ \ \overline{v} := \overline{uv}$.

    This is possible precisely because j is 1–1. Now F is a group in which S really is a subset.

Theorem – Universal Mapping Property. Let F be the free group constructed over S. Then for any group G and any function

$$\varphi : S \to G$$

(not necessarily a homomorphism), there *exists a unique* homomorphism

$$\hat{\varphi} : F \to G$$

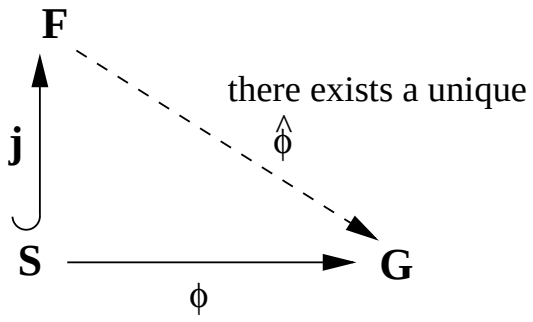
which extends φ:

$$\hat{\varphi}(\overline{a}) = \varphi(a) \ ,$$

for all $a \in S$, that is

$$\hat{\varphi}(j(a)) = \varphi(a) \ \ .$$

Thus there exists a unique homomorphism $\hat{\varphi}$ making the following diagram *commute*, meaning that $\hat{\varphi} \circ j = \varphi$:



Proof: Suppose $\overline{u} \in F$, where

$$u = a_1^{\epsilon_1} \ldots a_k^{\epsilon_k} \in W(S),$$

$a_j \in S$, $\epsilon_j = \pm 1$. It is easy to check from the definitions that

$$\overline{a}_1^{\epsilon_1} \cdot \overline{a}_2^{\epsilon_2} = \overline{a_1^{\epsilon_1} a_2^{\epsilon_2}}$$

for various choices of the signs ϵ_j. Similarly, and more generally we have

$$\overline{a}_1^{\epsilon_1} \cdot \overline{a}_2^{\epsilon_2} \cdot \ldots \cdot \overline{a}_k^{\epsilon_k} = \overline{a_1^{\epsilon_1} a_2^{\epsilon_2} \ldots a_k^{\epsilon_k}} = \overline{u}.$$

(Here '$\cdot$' indicates multiplication in F.) Thus, if the homomorphism $\hat{\varphi}$ does exist we are forced to conclude that

$$\begin{aligned} \hat{\varphi}(\overline{u}) &= \hat{\varphi}(\overline{a}_1^{\epsilon_1} \cdot \ldots \cdot \overline{a}_k^{\epsilon_k}) \\ &= \hat{\varphi}(\overline{a}_1)^{\epsilon_1} * \ldots * \hat{\varphi}(\overline{a}_k)^{\epsilon_k} \\ &= \varphi(a_1)^{\epsilon_1} * \ldots * \varphi(a_k)^{\epsilon_k}, \end{aligned}$$

where '$*$' indicates multiplication in G. In short, $\hat{\varphi}$ is uniquely determined if it exists at all, and we see how it must be defined. Indeed, if $u = a_1^{\epsilon_1} \ldots a_k^{\epsilon_k} \in W(S)$, then $\overline{u} \in F$ and we do attempt to <u>define</u>

$$\hat{\varphi}(\overline{u}) = \varphi(a_1)^{\epsilon_1} * \ldots * \varphi(a_k)^{\epsilon_k}.$$

We must show that $\hat{\varphi}$ is well-defined. For example, suppose $b \in S$ and

$$v = a_1^{\epsilon_1} \ldots a_j^{\epsilon_j} b^1 b^{-1} a_{j+1}^{\epsilon_{j+1}} \ldots a_k^{\epsilon_k}.$$

Thus $u \sim v$ and $\overline{u} = \overline{v}$. But

$$\begin{aligned} \hat{\varphi}(\overline{v}) &= \varphi(a_1)^{\epsilon_1} * \ldots * \varphi(a_j)^{\epsilon_j} * \varphi(b) * \varphi(b)^{-1} * \varphi(a_{j+1})^{\epsilon_{j+1}} * \ldots * \varphi(a_k)^{\epsilon_k} \\ &= \varphi(a_1)^{\epsilon_1} * \ldots * \varphi(a_j)^{\epsilon_j} * 1 * \varphi(a_{j+1})^{\epsilon_{j+1}} * \ldots * \varphi(a_k)^{\epsilon_k} \\ &= \varphi(\overline{u}). \end{aligned}$$

(Here we use '1' as the identity in G.) Iterating this argument we conclude that $\hat{\varphi}$ is well-defined (i.e. independent of the word u used to represent $\overline{u}$). It is easy now to check that $\hat{\varphi}$ is a homomorphism, and that

$$\hat{\varphi}(j(a)) = \hat{\varphi}(\overline{a}) = \varphi(a)$$

for all $a \in S$. □

11. **Corollary.** The function $j : S \to F$ is 1-1.

    Proof. Suppose $a \neq b$ in S. We must show $\overline{a} \neq \overline{b}$. Let $G = (Z_2, +)$ and define $\varphi : S \to G$ by

$$\begin{aligned} \varphi(a) &= 0 \\ \varphi(x) &= 1, \text{ if } x \neq a. \end{aligned}$$

    In particular, $\varphi(b) = 1$. From the Theorem we have

$$\begin{aligned} \hat{\varphi}(\overline{a}) &= \varphi(a) = 0 \\ \hat{\varphi}(\overline{b}) &= \varphi(b) = 1. \end{aligned}$$

    Therefore, $\overline{a} \neq \overline{b}$, since $\hat{\varphi}$ is a well-defined function. □

    We have noted that this allows us to consider S to be a subset of F.

12. This universal property is crucial and motivates a definition for what it rally means for a group to be free over a subset S:

    Definition. Let F be a group with a subset S. Then F is **free** over S if for any group G and function $\varphi : S \to G$ there exists a unique homomorphism

$$\hat{\varphi} : F \to G$$

    which extends φ.

13. **Remarks**. We say that F has been defined by a **universal property**. This does not mean that any such group F, with a specified subset, actually exists. But our construction using $W(S)$, and the theorem above, shows that free groups exist in abundance, namely over any set.

14. **Exercises** – typical of universal properties.

(a) The trivial group $\{e\}$ is free over the subset $\emptyset$.

(b) The group $(Z, +)$ is free over $\{1\}$.

(c) If F is free over S, then S generates F. (This is tricky: You need to use $G = \langle S \rangle$, the subgroup of F actually generated by S. Then show $G = F$ is forced.)

(d) If F_1 is free over S_1, F_2 is free over S_2, and there is a bijection $\alpha : S_1 \to S_2$, then

$$F_1 \simeq F_2$$

are isomorphic as groups.

(e) Definition. The **rank** of F is $|S|$, the cardinality of S. To show that this is well-defined, prove that if F is free over two subsets S and T, then

$$|S| = |T|.$$

Remark: this is very much analogous to the theorem that the dimension of a vector space is well-defined.

(f) If F is free over S, then $S \cap S^{-1} = \emptyset$.

(g) If $x = x_1 \dots x_k$ is a *reduced product* in F (i.e. all $x_j \in S \cup S^{-1}$, but no $x_j x_{j+1} = e$) then

$$x \neq e.$$

(h) Suppose a group F is generated by a subset S such that $S \cap S^{-1} = \phi$ and such that there are no non-trivial relations over S:

$$x_j \in S \cup S^{-1},\ x_j x_{j+1} \neq e \Rightarrow x_1 \dots x_k \neq e.$$

Prove that F is free over S.

1.3 Generators and Relations

1. One and the same group can be represented by various means, eg. permutations, matrices, geometrical symmetries. Each method has its advantages and natural applications. The same is true for 'generators and relations'.

2. **Example-Cyclic Groups**

   By definition, a cyclic group G can be generated by a single element 'a' (even though G might be described in terms of several generators). Thus

   $$G = \{a^n : n \in Z\} = \langle a \rangle.$$

   (Why?)

   Suppose $a^d = e$ for some positive integer d. Then each integer

   $$n = dq + r \ ,$$

   for an appropriate remainder r, satisfying $0 \leq r < d$. Thus

   $$\begin{aligned} a^n &= a^{dq+r} = (a^d)^q \cdot a^r = e^q \cdot a^r \\ &= a^r. \end{aligned}$$

   Hence actually

   $$G = \{e = a^0, a = a^1, a^2, \ldots, a^{d-1}\}.$$

   If these d elements are distinct then

   $$|G| = d.$$

   This happens precisely when d is the smallest positive integer such that

   $$a^d = e,$$

   namely if a has order d. (Proof ?)

   This cyclic group of order d might be denoted C_d. Recall that

   $$C_d \simeq (Z_d, +).$$

3. **Example**. Consider complex numbers of norm 1 (unit circle). These form a multiplicative group with identity $e = 1$.

   The case $d = 6$ is typical, so suppose $G = \langle a \rangle$, where $a^6 = 1$. There are several possibilities for 'a' (how many?). For example, perhaps

   (a) $a = 1$; $1^6 = 1$; $G = \langle 1 \rangle$ has order 1.

   (b) $a = -1$; $(-1)^6 = 1$; $G = \langle -1 \rangle$ has order 2.

   (c) $a = e^{2\pi i/3} = -\dfrac{1}{2} + i\,\dfrac{\sqrt{3}}{2}$ satisfies $a^6 = 1$, but G has order 3.

   (d) $a = e^{2\pi i/6} = \dfrac{1}{2} + i\,\dfrac{\sqrt{3}}{2}$ satisfies $a^6 = 1$; now G does have order 6.

   Thus the fact that $G = \langle a \rangle$, where $a^6 = 1$, does not imply $|G| = 6$. However, we did exhibit a for which G does have the largest possible order 6. (Each other case involves a subgroup, so has order dividing 6.)

   Conclusion. The largest group with

   (i) a single generator 'a'

   (ii) satisfying a single relation $a^6 = e$

   is indeed the cyclic group C_6. These data determine the group since from this we can recreate all elements of the group, as well as their multiplication table. (Why?)

   We therefore say that C_6 has the **presentation**

$$C_6 = \langle a \mid \ a^6 = e \ \rangle.$$

   More generally, the cyclic group of order n is

$$C_n = \langle a \mid \ a^n = e \, \rangle.$$

4. We shall do various examples which prompt the question

What does it mean to derive one relation in a group G from others?

Suppose G is generated by a subset S. Thus a typical element in G is a word

$$w = x_1 x_2 \dots x_k, \quad x_j \in S \cup S^{-1}.$$

Any relation (i.e. equation) in G can be written as equality between two words in these generators, say

$$x_1 \dots x_k = y_1 \dots y_\ell$$

or

$$v = u$$

or

$$vu^{-1} = e$$

or

$$x_1 \dots x_k y_\ell^{-1} \dots y_1^{-1} = e.$$

In brief, any relation in G amounts to writing

$$w = e$$

for some word w in the elements of $S \cup S^{-1}$.

Intuitively, we say that

$$w_1 = e, \dots, w_m = e$$

are **defining relations** for G if every relation in G can be 'derived' from these defining relations. What does this mean?

We start with

(i) defining relations, say

$$w_1 = e, \ldots, w_m = e$$

which are particular to G; and

(ii) trivial relations, true in any group G,

$$z = z$$

or

$$zz^{-1} = e, \quad \text{for all words } z.$$

In a group, we are limited algebraically to multiplication and taking inverses. Hence about all we can do is take inverses and multiply equal things by equal things, as in

$$\begin{aligned} a &= b \qquad (\text{i.e. } ab^{-1} = e) \\ c &= d \qquad (\text{i.e. } cd^{-1} = e). \end{aligned}$$

Thus

$$ac = bd$$

or

$$[a(cd^{-1})a^{-1}] \cdot [ab^{-1}] = e.$$

Note that

$$\begin{aligned} a(cd^{-1})a^{-1} &= \text{conjugate of a known relation} \\ ab^{-1} = e(ab^{-1})e^{-1} &= \text{another known relation.} \end{aligned}$$

We thus arrive at a new relation, perhaps a little lengthier, but still in the form

$$\tilde{w} = e.$$

In this way we further enhance our supply of known relations. This process may be iterated.

Conclusion. Any relation derived from

$$w_1 = e, \ldots, w_m = e$$

has the form

$$[z_1 \, {w_{j_1}}^{\epsilon_1} \, {z_1}^{-1}] \, [z_2 \, {w_{j_2}}^{\epsilon_2} \, {z_2}^{-1}] \, \ldots \, [z_r \, {w_{j_r}}^{\epsilon_r} \, {z_r}^{-1}] = e$$

where the $\epsilon_t = \pm 1, \;\; z_t \in G, \;\; j_t \in \{1, \ldots, m\}$.

5. If we let F be the free group over the set S, then the collection of all words such as

$$[z_1 \, {w_{j_1}}^{\epsilon_1} \, {z_1}^{-1}] \, \ldots \, [z_r \, {w_{j_r}}^{\epsilon_r} \, {z_r}^{-1}]$$

form a normal subgroup N of F. Indeed, the collection of all such words is clearly closed under multiplication and the taking of inverses. Furthermore, since the z_k's are allowed to vary over the whole group F, we have

$$N \lhd F \;\; .$$

From another point of view, N is the intersection of all normal subgroups containing the *relator set*

$$R = \{w_1, \ldots, w_m\} \; .$$

(Hence, N is the smallest normal subgroup containing R. Consequently, N is called the **normal closure** of R in F.)

Thus if $x_1, \ldots, x_k \in S \cup S^{-1}$, we conclude that

$$x_1 \ldots x_k = e \qquad [\text{as a } \underline{\text{product}} \text{ in } G]$$

if-f

$$x_1 \ldots x_k \in N \qquad [\text{as a } \underline{\text{word}} \text{ in } F].$$

In other words, it makes sense to define

$$G \simeq F/N.$$

This should motivate the definition of generators and relations on pages 438-439 of Gallian.

1.4 A Brief Look Back

The upshot of our discussion above is that we can legitimately treat the elements of any set S as generators of a free group F. The elements of F are words in the generators $a, b. \ldots \in S$, and their inverses. We multiply these words subject to <u>no</u> special relations, other than merely letting one generator cancel its inverse:

$$bb^{-1} = e = a^{-1}a, \text{ if } a, b \in S.$$

A free group thus has no defining relations. More generally, if S is any set of objects, and if R is any set of words over $S \cup S^{-1}$, then there does exist a group

$$G = \langle\, S \mid R \,\rangle,$$

consisting of all words in S, subject only to the relations $w = e$, for all $w \in R$.

1.5 Finitely Presented groups

Typically, both $S = \{a_1, \ldots, a_n\}$ and $R = \{w_1, \ldots, w_t\}$ are finite, and we write

$$G = \langle\, a_1, \ldots, a_n \mid w_1 = e, \ldots, w_t = e \,\rangle.$$

Once again, the sensible point of view here is that G is the group generated by $a_1, \ldots, a_n$, subject only to the defining relations $w_1 = e, \ldots, w_t = e$. Thus any relation in G can be derived from these.

More formally, if F is the free group over S, then also the $w_j \in F$, so that R is a subset of F. Letting N be the normal closure of R, we have seen that we may explicitly define

$$G := F/N.$$

Actual Computing in G.

Strictly speaking in this definition, the generators of G are cosets $a_1N, \ldots, a_nN$, and the identity is N itself. A typical relation in G resembles

$$(a_1N)(a_2N)(a_1N)^{-1}(a_2N) = N$$

which holds if-f $a_1a_2a_1^{-1}a_2 \in N$ in F. We have seen that this means $a_1a_2a_1^{-1}a_2$ can be 'derived' as a product of conjugates of relators.

When working in G, as opposed to F, it is convenient to abbreviate a_jN by a_j, and $N = eN$ by e, so that in G itself we have

$$a_1a_2a_1^{-1}a_2 = e,$$

a relation which is derivable from $w_1 = e, \ldots, w_t = e$. (Just how this derivation proceeds is a deep algorithmic question.) Anyway, we are quite justified in working with generators and relations in such intuitive ways.

1.6 The Substitution Theorem

This is the key theorem for working with presentations like

$$G = \langle\, a_1, \ldots, a_m \mid w_1 = e, \ldots, w_t = e \,\rangle.$$

The theorem asserts that in a *natural* way, G is the largest group of any sort having m generators satisfying the given defining relations. In fact, any such group H is isomorphic to a quotient of G (perhaps even to G itself). Von Dyck's theorem is actually a special case of the Substitution Theorem.

Our notation is unchanged:

$$\begin{aligned} S &= \{a_1, \ldots, a_n\} \text{ generates the free group } F \\ R &= \{w_1, \ldots, w_t\} \text{ is a set of relators (in } F) \\ N &= \text{normal closure of } R \text{ in } F \\ G &= F/N. \end{aligned}$$

The Substitution Theorem.

Suppose $G = \langle a_1, \ldots, a_n \mid w_1 = e, \ldots, w_t = e \rangle$. Let H be any group (abstract, matrices, symmetries, permutation, etc.) having generators $\tilde{a}_1, \ldots, \tilde{a}_n$ 'satisfying' in H the relations for G.

(More precisely, we assume that if $\tilde{w}_j \in H$ is that product of various $\tilde{a}_k$'s and their inverses which corresponds to $w_j \in G$, then in fact $\tilde{w}_j = 1$, the identity in H, for $1 \leq j \leq t$.)

Then there exists a unique onto homomorphism

$$\eta : G \longrightarrow H$$

such that

$$a_k \longmapsto \tilde{a}_k, \quad 1 \leq k \leq n.$$

Proof. Let u be any word in N, say

$$u = [z_1\ w_{j_1}^{\epsilon_1}\ z_1^{-1}]\ \ [z_2\ w_{j_2}^{\epsilon_2}\ z_2^{-1}] \ldots,$$

where the $w_{j_i} \in R,\ \epsilon_i = \pm 1$, and the z_i are *arbitrary* elements of F. Thus the corresponding element of H is

$$\begin{aligned} \tilde{u} &= [\tilde{z}_1\ \tilde{w}_{j_1}^{\epsilon_1}\ \tilde{z}_1^{-1}]\ \ [\tilde{z}_2\ \tilde{w}_{j_2}^{\epsilon_2}\ \tilde{z}_2^{-1}]\ \ldots \\ &= [\tilde{z}_1\ 1^{\epsilon_1}\ \tilde{z}_1^{-1}]\ \ [\tilde{z}_2\ 1^{\epsilon_2}\ \tilde{z}_2]\ \ldots \\ &= 1, \end{aligned}$$

by the 'substitution hypothesis'. In short,

$$u \in N \Rightarrow \tilde{u} = 1 \in H.$$

Now define

$$\varphi : S \longrightarrow H$$

by mapping each $a_j \longmapsto \tilde{a}_j$. By the universal property of the free group F, there exists a unique homomorphism

$$\tilde{\varphi} : F \longrightarrow H$$

also mapping each $a_j \longmapsto \tilde{a}_j$. Let

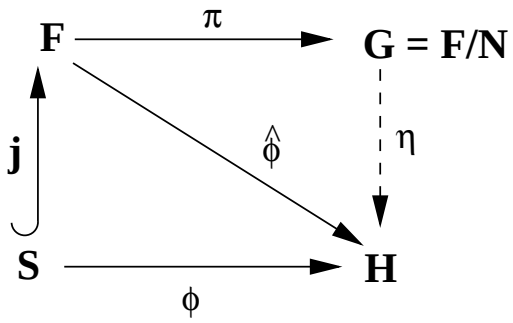
$$\pi : F \longrightarrow F/N = G$$

be the natural homomorphism. Thus

$$\pi(a_j) = a_j N,$$

(which, as we have noted before, we generally write briefly as just a_j, when working in G).

The diagram



suggests that we define

$$\eta : G \longrightarrow H$$

by

$$\eta(\pi(w)) = \hat{\varphi}(w), \quad \forall w \in F.$$

Notice that if $w = a_1^{\epsilon_1} \dots a_k^{\epsilon_k}$, then

$$\begin{aligned} \eta(\pi(w)) &= \hat{\varphi}(a_1^{\epsilon_1} \dots a_k^{\epsilon_k}) \\ &= \hat{\varphi}(a_1)^{\epsilon_1} \dots \hat{\varphi}(a_k)^{\epsilon_k} \\ &= \tilde{a}_1^{\epsilon_1} \dots \tilde{a}_k^{\epsilon_k} = \tilde{w}. \end{aligned}$$

The key thing now is to check that η is well-defined. But

$$\begin{aligned} \pi(w) = \pi(z) &\Rightarrow \pi(wz^{-1}) \in N \\ &\Rightarrow \overset{\sim}{wz^{-1}} = 1 \ \ (\text{in } \ H) \\ &\Rightarrow \tilde{w}\tilde{z}^{-1} = 1 \\ &\Rightarrow \tilde{w} = \tilde{z}. \end{aligned}$$

In short, η is well-defined. It is now easy to check that η is an onto homomorphism. Remembering that when working in G we abbreviate $a_k N$ by just a_k, we have

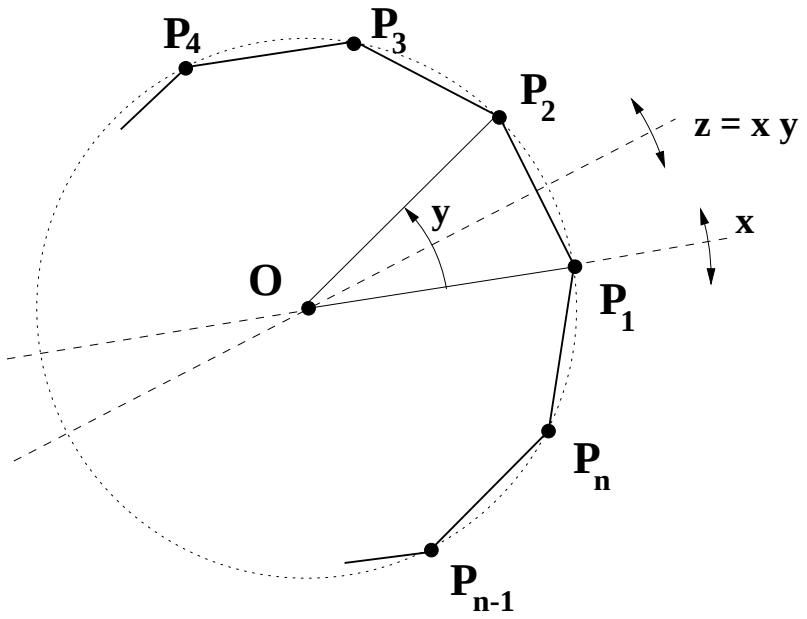
$$\begin{aligned} \eta(a_k) &= \eta(a_k N) \text{ (to be precise)} \\ &= \eta(\pi(a_k)) \\ &= \hat{\varphi}(a_k) \\ &= \tilde{a}_k, \end{aligned}$$

as required. It's easy to check that these conditions uniquely specify η. □

1.7 Dihedral Groups

Definition. The dihedral group D_n is the symmetry group of a regular n-sided polygon $\{n\}$, for $n \geq 3$. (The presentation derived below will provide a sensible definition in the cases $n = 1, 2$, as well.)

The vertices of the regular n-gon $\{n\}$ are n equally spaced points $P_1, \ldots, P_n$ on a circle, joined consecutively by edges $P_1P_2, P_2P_3, \ldots, P_nP_1$. It is useful to take subscripts mod n.



It is easy to see, and not hard to prove, that our $\{n\}$ is symmetric by

(a) rotation y through $2\pi/n$ about the centre O of the circle,

(b) reflection x in the line OP_1.

Taking P_1 as a base vertex, we may enumerate all symmetries of $\{n\}$ as follows. Think of $\{n\}$ as a piece of cardboard. We may flip it over, or not; that is we may apply x, or not. Having done so, we may rotate P_1 to any of the n vertices $P_1, \ldots, P_n$. Thus every symmetry is of the form

$$xy^j \text{ or } y^j, \quad 0 \leq j \leq n-1,$$

and so

$$|D_n| = 2n.$$

Now we have a geometrical description of the group D_n. But by looking at the effect of x, y on the n vertices, we also obtain an isomorphic permutation group, generated by

$$\begin{array}{l}(2\ n)(3\ n-1)(4\ n-2)....\\ (1\ 2\ 3....n)\end{array}$$

corresponding to x, y respectively.

Notice that $z = xy$ is the reflection which interchanges P_1, P_2. (We apply symmetries left to right, here first x, then y.) Thus D_n satisfies the relations

$$x^2 = y^n = (xy)^2 = 1,$$

but conceivably also further independent relations. In fact, we shall immediately see that this is not the case. Indeed, consider the presentation

$$G = \langle\ a, b\ \mid\ a^2 = b^n = (ab)^2 = e\ \rangle.$$

Note that $a^2 = e$. Thus in any word in G we may assume that a occurs to the first power, or not at all:

$$\begin{aligned} aaa &= a^2a = ea = a \\ aaaa &= e. \end{aligned}$$

We also have $abab = e$, so that

$$\begin{aligned} ba &= a^{-1}b^{-1} \\ &= ab^{-1}. \end{aligned}$$

Therefore, if ever b precedes a in some word, we may interchange a, b at the expense of the exponent on b:

$$\begin{aligned} w &= \ldots b^2 a \\ &= \ldots b(ba) \ldots \\ &= \ldots b(ab^{-1}) \ldots \\ &= \ldots ab^{-2} \ldots \quad . \end{aligned}$$

Thus, in any word we may move all a's to the front, b's to the end. But $a^2 = e = b^n$, so any word can be written

$$a^j b^k, \;\; 0 \leq j \leq 1, \;\; 0 \leq k \leq n-1.$$

Hence $|G| \leq 2n$. (Conceivably there are duplicates in this census.)

On the other hand, by the Substitution Theorem, there is an onto homomorphism

$$\varphi : G \longrightarrow D_n,$$

so that

$$G/\ker\varphi \simeq D_n,$$

and hence

$$\frac{|G|}{|\ker\varphi|} = 2n.$$

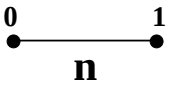
Since $|G| \leq 2n$, we conclude that $|G| = 2n$, so that $|\ker\varphi| = 1$, meaning that φ is an isomorphism. In short, we have proved that the geometrically defined group D_n has the presentation

$$\langle\, a, b \;\mid\; a^2 = b^n = (ab)^2 = e \,\rangle.$$

Equivalently, letting $r_0 = a$ (corresponding to x), and $r_1 = ab$ (corresponding to $z = xy$), so that $a = r_0,\ b = r_0 r_1$, we get

$$D_n \simeq \langle\, r_0, r_1 \mid r_0^2 = r_1^2 = (r_0 r_1)^n = e \,\rangle.$$

This says that D_n is (isomorphic to) the **Coxeter group** whose diagram is



1.8 Finitely Generated Abelian Groups

(a) Two elements a_1, a_2 in a group G commute if and only if

$$\begin{aligned} a_1 a_2 &= a_2 a_1 \quad , \text{ i.e.} \\ a_1 a_2 a_1^{-1} a_2^{-1} &= e \quad , \text{ i.e.} \\ [a_1, a_2] &= e \quad , \end{aligned}$$

using commutator notation.

(b) Suppose

$$G = \langle\, a_1, \ldots, a_n \mid \text{ abel; extra relations} \,\rangle$$

where 'abel' is shorthand for the list of all relations

$$[a_i, a_j] = e, \quad 1 \le i < j \le n.$$

Thus G is abelian, perhaps with various extra relations on the finitely many generators.

Because of commutativity, in any word w we may put all a_1 terms first, then all a_2's, etc. Hence each extra relation resembles

$$a_1^{k_1} \ldots a_n^{k_n} = e$$

for integers $k_j \in Z$. If there are t extra relations, we may therefore write them as

$$a_1^{k_{i1}} \ldots a_n^{k_{in}} = e, \quad 1 \le i \le t.$$

Thus the $t \times n$ integer matrix

$$K = [k_{ij}]$$

specifies the abelian group G.

(c) Example. Classify the abelian group

$$G = \langle\, a, b, c \mid \text{abel}, \;\; a^3 c^4 = b^3, \;\; a^6 b^2 c^4 = e \,\rangle.$$

The extra relations are

$$\begin{aligned} a^3b^{-3}c^4 &= e \\ a^6b^2c^4 &= e \end{aligned}$$

so that

$$K = \begin{bmatrix} 3 & -3 & 4 \\ 6 & 2 & 4 \end{bmatrix}.$$

Now we apply various reversible operations which leave G alone but which amount to altering K in a nice way.

(a) Clearly G is unchanged by

- permuting generators; this corresponds to permuting columns of K:

$$a^3b^{-3}c^4 = e \Rightarrow b^{-3}a^3c^4 = e.$$

(Note use of commutativity!)

- replacing a generator by its inverse; this multiplies a column of K by -1:

$$a^3b^{-3}c^4 = e \Rightarrow a^3b^{-3}(c^{-1})^{-4} = e.$$

- replacing generator x by xy, where y is a different generator; here we subtract the x column from the y column:

$$a^6b^2c^4 = e \Rightarrow a^{6-2}(ab)^2c^4 = e.$$

Iterating these operations, we may thus add or subtract any integer multiple of one column to another column.

(b) G is also unchanged by the analogous operations on the extra relations. Clearly, we can permute the relations or take their inverses:

$$\begin{aligned} a^3b^{-3}c^4 = e &\Rightarrow (a^3b^{-3}c^4)^{-1} = e \\ &\Rightarrow c^{-4}b^3a^{-3} = e \\ &\Rightarrow a^{-3}b^3c^{-4} = e. \end{aligned}$$

Again note the importance of commutativity.
Likewise we can replace one extra relation by its product with another:

$$\begin{aligned} a^3b^{-3}c^4 = e = a^6b^2c^4 \quad &\Rightarrow \quad (a^3b^{-3}c^4)(a^6b^2c^4) = e \\ &\Rightarrow \quad a^9b^{-1}c^8 = e. \end{aligned}$$

Hence we may permute rows of K, negate rows, or add any integer multiple of one row to another.

(c) These 'integral' row and column operations are precisely what are needed to transform K to Smith normal form:

$$\begin{aligned} K &= \begin{bmatrix} 3 & -3 & 4 \\ 6 & 2 & 4 \end{bmatrix} \\ &\sim \begin{bmatrix} 3 & -3 & 1 \\ 6 & 2 & -2 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 3 & -3 \\ -2 & 6 & 2 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 12 & -4 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 12 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \end{bmatrix} = \tilde{K}. \end{aligned}$$

In short, G has the presentation

$$G = \langle\, x, y, z |\ \text{abel},\ \ x^1 = e,\ \ y^4 = e \,\rangle.$$

(If you keep careful track of the row, column operations, you can see how to write x, y, z in terms of a, b, c and conversely.) In fact, $x = e$ is a superfluous generator, but z isn't! Thus

$$G = \langle x, y, z |\ \text{abel},\ \ y^4 = e \rangle.$$

In other words,

$$G \simeq Z_4 \times Z,$$

is a direct product of cyclic groups.

(d) Clearly, we can similarly prove that every finitely generated abelian group is a direct product of cyclic groups. A closer look at Smith normal form reveals quite a bit more:

- the crucial orders 1 for x, 4 for y, none for z are uniquely determined by K itself. In fact,

$$\begin{aligned} 1 &= \gcd\{\text{all } 1\times 1 \text{ subdeterminants of } K\} \\ &= \gcd\{\text{all entries of } K\} \\ 4 &= \gcd\{\text{all } 2\times 2 \text{ subdeterminants of } K\} \\ \text{none} &= \text{number of } 3\times 3 \text{ subdeterminants.} \end{aligned}$$

- the direct product structure is uniquely determined if we demand that each order divide the next.

Reference: Algebra, Vol. 1, P. M. Cohn.

1.9 Coset Enumeration

(a) The method is due to Todd and Coxeter (1936); see Coxeter and Moser for more details and examples. The summary below is based on J. Rotman, *An Introduction to the Theory of Groups*, 4th ed.

(b) **Enumerating The Cosets of the Subgroup H in**

$$G = \langle a_1, \ldots, a_n \mid w_1 = e, \ldots, w_t = e \rangle.$$

i. For each defining relation $w_j = e$ we require a **relation table**. For example, if

$$a^2 b^{-1} a c = e,$$

in which the word on the left has length 5, then we require a table with 6 columns separated by 5 lines headed by the letters:

a	a	b^{-1}	a	c	

Such tables have (for now) an unknown number of rows.

ii. For each generator a and its inverse we maintain an **auxiliary table** with two columns:

a	
	D
	D
	B
	D

a^{-1}	
	D
	D
	B
	D

It is helpful to mark the rows as arriving by definition (D) or bonus (B). In fact, sometimes we might want to indicate collapse (C), as well. Notice that if $a^2 = e$, then $a^{-1} = a$, so we need only one auxiliary table.

iii. To compute the order $|G|$, we must enumerate all elements. This amounts to counting cosets of the trivial subgroup $H = \{e\}$. In this case, no **subgroup generator tables** are needed. However, if we enumerate cosets of a larger subgroup H, such tables are necessary. For example, if the subgroup

$$H = \langle\ aba\ ,\ b^2\ \rangle$$

is generated by the two indicated elements, then we maintain a subgroup generator table for each:

a	b	a
1		1

b	b
1	1

Only one row will be needed, as shown.

iv. The entries $1, 2, 3 \ldots$ in any table are abbreviations for various cosets, starting with

$$1 := H.$$

Then $2 := 1a$ is an abbreviation for the coset Ha; and $3 := 2b^{-1}$ indicates the coset $H(ab^{-1})$.
In any table, the entries

a^ϵ	
$\vdots$	$\vdots$
j	k

indicate that

$$ja^\epsilon = k.$$

That is if $j = Hw$, then

$$k = H(wa^\epsilon).$$

v. If, for example,

$$a^2b^{-1}ac = e$$

is one of the defining relations, and if $j = Hw$ is any coset, then in the corresponding table there will be a row beginning and ending with j:

a	a	b^{-1}	a	c	
1	*	*	*	*	1
⋮					
j	*	*	*	*	j

This is because

$$j(a^2b^{-1}ac) = (Hw)e = H(we) = j.$$

(c) The actual running of the algorithm is quite delicate. Various points should be kept in mind:

- At each stage one must extract the maximum information from the various definitions and bonuses. This amounts to repeatedly scanning all tables for fillable slots.
- The general strategy, which can be made a bit more precise, is to fill in slots in the earliest defined rows.
- Coset collapse may occur. Here, a later coset is unexpectedly found to equal an earlier one. This forces updating of all tables, and may induce several bonuses.
- If $[G : H] = n < \infty$, the algorithm is guaranteed to stop. This happens when all rows in all tables are 'full', and then

$$[G : H] = \#\text{ of distinct cosets.}$$

(Basically, this follows from the Lemma below.)

- However, there is no predictable bound on the number of steps needed in the algorithm. For example, you might define cosets for 10^{10} centuries before finally reaching the crucial collapse which shows in fact that $|G| = |H| = 1$!
- Of course, if $[G : H] = \infty$, then the algorithm will never stop. When a computer implementation (as in GAP) runs out of time or space, one may reasonably suspect an infinite index. But you cannot be sure without further investigation.

(d) **Lemma**. Suppose $G = \langle\, S \,\rangle$.

i. If Y is a non-empty subset of G such that

$$Ya \subseteq Y \text{ for all } a \in S \cup S^{-1},$$

then $Y = G$.

ii. Suppose H is a subgroup of G, and suppose $Hw_1, \ldots, Hw_n$ is a finite list of cosets, closed under right multiplication by all $a \in S \cup S^{-1}$. Then

$$G = \bigcup_{j=1}^{n} Hw_j.$$

Conceivably there are duplicates in the list. But if there are m distinct cosets, then

$$[G : H] = m.$$

1.10 Exercises.

i. Compute the order of $\langle\, a, b \mid a^2 = b^2 = (ab)^2 \,\rangle$.
(Careful: we have $a^2b^{-2} = e$, not necessarily $a^2 = e$!) What group is this?

ii. Compute the orders of
(i) $\langle\, a, b \mid a^4 = b^2 = (ab)^3 = e \,\rangle$.
(ii) $\langle\, a, b \mid a^5 = b^2 = (ab)^3 = e \,\rangle$.
by enumerating cosets of H. Find a permutation representation for each group and identify each group.

iii. Compute the order of

$$G = \langle x, y, z, u, v \mid xy = z, yz = u, zu = v, uv = x, vx = y \rangle.$$

Hint: first enumerate cosets for $H = \langle x \rangle$, then calculate a bit.

iv. Find the order of

$$G = \langle\, a, b \mid\ ab^2 = e,\ a^2b^3 = e \,\rangle.$$

v. Identify the abelian group

$$G = \langle\, a, b, c \mid \text{ abel}, a^{10}b^8c^{12} = a^{16}b^8c^{20} = a^4b^4c^4 = e \,\rangle.$$

vi. Identify the abelian group

$$G = \langle\, a, b, c \mid \text{abel},\ a^{18}b^{24} = c^2, a^2 = b^{16}c^{22},\ a^{24}b^{48}c^{16} = e \,\rangle.$$

vii. Suppose

$$G = \langle a_1, \ldots, a_n | w_1 = e, \ldots, w_t = e \rangle$$

is a presentation in which every relator w_j has even length in the generators $a_1, \ldots, a_n$.
Show that the map

$$\begin{array}{rcl} \varphi : G & \longrightarrow & \{\pm 1\} \\ w & \mapsto & (-1)^{\ell(w)} \end{array}$$

is a well-defined homomorphism. (Here, $\ell(w)$ is the length of w, which may <u>not</u> be well-defined.) Show that G has an 'even' subgroup G^+ of index 2 (and hence normal) in G.

1.11 Symmetry – Groups Acting on Spaces

1. The symmetry group of a regular n-gon $\{n\}$ is the dihedral group D_n. We know that D_n has order $2n$ and presentation

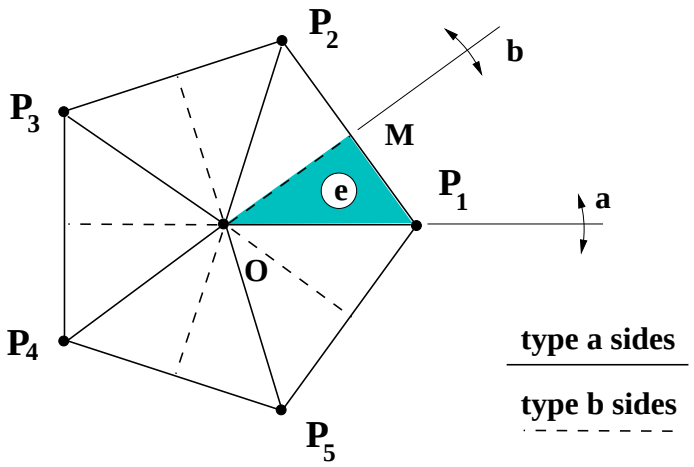
$$\langle\, a, b \mid a^2 = b^2 = (ab)^n = e \,\rangle.$$

If $\{n\}$ has vertices $P_1, P_2, \ldots$ and center O, then we may take

$$\begin{aligned} a &= \text{reflection in line } OP_1 \\ b &= \text{reflection in line } OM, \end{aligned}$$

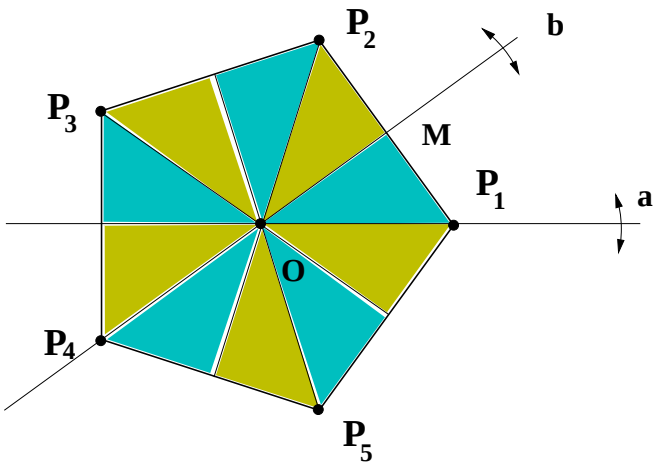
where M is the midpoint of P_1P_2.

2. The regular pentagon $\{5\}$ is typical (for n odd):



Let $\mathcal{F} = \triangle OMP_1$, considered here as a 'nice' subset of the convex regular pentagon $\mathcal{P} = \{5\}$. (By 'nice' we usually mean topologically equivalent, i.e. homeomorphic, to a disc.)

We say that $\mathcal{F}$ is a **fundamental region** for the group D_5 acting on $\mathcal{P}$, since by repeatedly applying a, b to $\mathcal{F}$ we completely fill out the pentagon without gaps and overlaps. In other words, $\mathcal{P}$ is the union of the images $\mathcal{F}^g := g(\mathcal{F})$, for $g \in D_5$; and two distinct images can intersect only at their boundaries, i.e. along an edge equivalent to either OP_1 (type a), or to OM (type b).



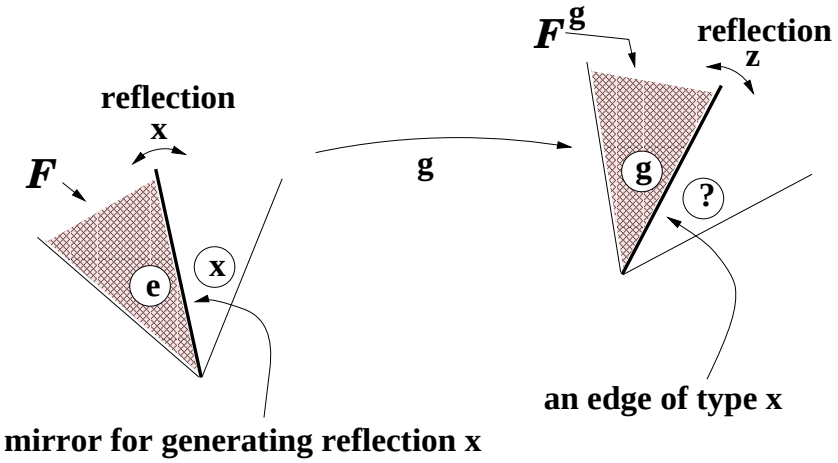
Visually, we can observe this by building an actual kaleidoscope with two mirrors inclined at $\pi/5$. Then if a cardboard triangle similar to $\triangle OMP_1$ is placed between the mirrors we see re-created the complete pentagon. It is a nice exercise in 3-dimensional optics to prove that what you actually do see is an image of our ideal 2-dimensional kaleidoscope.

Notice that each image of $\mathcal{F}$ has a type a edge, joining O to a vertex; and a type b edge joining O to a midpoint. Colour these red and green, respectively, say; and label $\mathcal{F}$ by 'e', indicating that $\mathcal{F}^e = \mathcal{F}$. Indeed, since $\mathcal{F}$ is a fundamental region,

$$\mathcal{F}^g = \mathcal{F} \Rightarrow g = e.$$

Likewise every copy $\mathcal{F}^g$ of $\mathcal{F}$ is labelled by a unique element $g \in D_5$. For example, $\triangle OMP_2 = \mathcal{F}^b$, so we label this region 'b'. And reflecting $\mathcal{F}$ in a, we get the region $\mathcal{F}^a$, which we label 'a'. Continuing this way, we label each copy of the fundamental region by a word in a, b. (We never require inverses since $a = a^{-1},\ b = b^{-1}$.)

In fact, there is a simple rule for inductively manufacturing this labelling. Suppose that we have already determined a region labelled 'g' (namely $\mathcal{F}^g$); and we wish next to label the region adjacent to 'g' across an edge of type x:



One way to do this is to first transform $\mathcal{F}^g$ back to $\mathcal{F}$ by applying g^{-1}; then apply reflection x; finally re-apply g. We see therefore that $z = g^{-1}xg$ is the reflection across the edge of type x which belongs to the region $\mathcal{F}^g$. Notice that z is conjugate to x, so that it is quite reasonable to say that z is a reflection of type x.

Anyway, the adjacent region is

$$(\mathcal{F}^g)^z = \mathcal{F}^{(gz)} = \mathcal{F}^{xg}.$$

Keeping in mind that we apply isometries left to right, we conclude that

"the region across an edge of type x from the region labelled 'g' should be labelled 'xg'."

Since $\mathcal{F}$ is a fundamental region, the number of copies of $\mathcal{F}$ is just the order of D_5, namely 10. However, a particular copy of $\mathcal{F}$ might be labelled in different ways. For instance, here

$$\text{`}babab\text{'} = \text{`}ababa\text{'}.$$

But the existence of two equivalent labellings amounts to a relation in the group, since two images can coincide only if the corresponding group elements are equal!

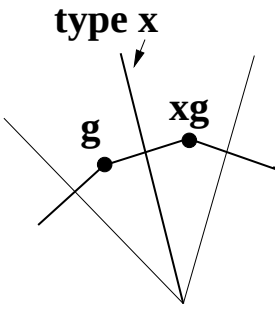
In other words, we may drop the quotes to get

$$babab = ababa \qquad \text{in } D_5, \text{ or}$$

$$\begin{aligned} ababababab &= e \\ (ab)^5 &= e. \end{aligned}$$

In short, defining relations in the abstract group are geometrically equivalent to tours proceeding face-to-face through the various images of $\mathcal{F}$, beginning and ending at the base region labelled 'e'. In this way, we may use topological and other geometric techniques to manufacture a presentation for a geometrically defined group, such as D_5.

3. Now select a typical point interior to $\mathcal{F}$ and label it 'e'. Its image in $\mathcal{F}^g$ is a similarly situated point, which we may as well label 'g', too. These 10 points are the vertices of the **Cayley graph** for D_5. Two such vertices are joined by an edge labelled (or coloured) x $(= a$ or $b)$, if the corresponding regions share an edge of type x:



In a sense, the Cayley graph is dual to the partition of $\mathcal{P}$ into fundamental regions.

In short, the vertices of the graph correspond to the elements of D_5; and two vertices g, h are joined by an edge, which is labelled x, precisely when $h = xg$, for $x = a$ or b. (Warning: this convention is 'reversed' below.) The structure of this particular Cayley graph is a bit simplified since a, b are involutions:

$$a^2 = b^2 = e.$$

4. More generally, any (finitely generated) group G has a **Cayley graph** Γ, as soon as we specify a set of generators $S = \{a, b, c \ldots\}$.

The vertices of Γ are just the elements of G. Two vertices $g, h \in G$ are joined by an edge labelled x, and directed from g to h, precisely when

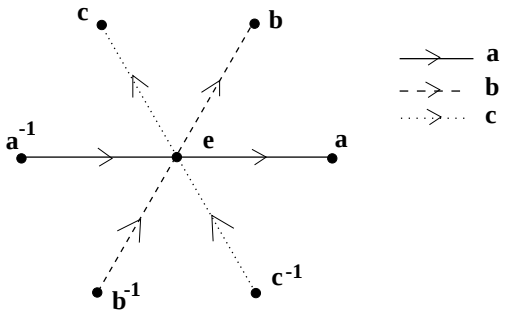
$$h = gx, \quad \text{for } \ x \in S.$$

(I've reversed the convention used just above for D_5; this won't matter.)

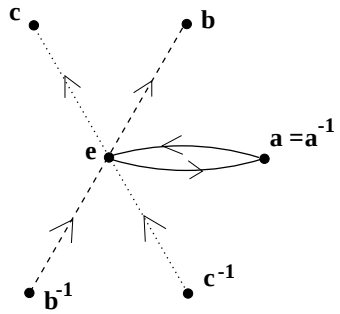
Thus Γ is a **digraph**. And for each generator $x \in S$, and group element g, there are edges labelled x

- entering g , from $k = gx^{-1}$
- leaving g , toward $h = gx$.

For example, if $S = \{a, b, c\}$, then every vertex has degree 6. The neighborhood of the identity is typical. It looks like



If a is an involution, then $a = a^{-1}$, and we really have:



In this case it is useful to collapse the multiple edge into 1, just as we did for D_5. Thus if G is generated by say 3 involutions, each vertex will have degree 3, rather than 6.

In practice, when there are a small number of generators, it is useful to suppress the edge labels and instead employ different colours or line styles (thick, thin, dotted, etc.).

5. The idea of a group acting on a space generalizes considerably. Suppose $\mathcal{P}$ is a metric space. An **isometry** of $\mathcal{P}$ is any bijection $g : \mathcal{P} \to \mathcal{P}$ such that g (hence also g^{-1}) is distance preserving. (It follows that g is a self-homeomorphism for $\mathcal{P}$.)

   The collection $\mathrm{Isom}(\mathcal{P})$ of all isometries on $\mathcal{P}$ with ordinary composition of functions, is clearly a group.

   Suppose G is a subgroup of $\mathrm{Isom}(\mathcal{P})$. We then say that G, or any group isomorphic to it, **acts on** $\mathcal{P}$. If $x \in \mathcal{P}$, $g \in G$, it is convenient to write $x^g := g(x)$, since we shall compose functions left to right. Likewise, if $\mathcal{F} \subseteq \mathcal{P}$ we write $\mathcal{F}^g$, instead of $g(\mathcal{F})$, for the image set.

   We say that the subset $\mathcal{F}$ is a **fundamental region** for G (acting on $\mathcal{P}$ if

   (i) $\mathcal{P} = \bigcup_{g \in G} \mathcal{F}^g$;

   (ii) for $g \neq h$, $u \in \mathcal{F}^g \cap \mathcal{F}^g \Rightarrow u \in \partial\mathcal{F}^g \cap \partial\mathcal{F}^h$.

   Normally, one makes further topological demands, such as that $\mathcal{F}$ be the closure of an open connected subset of $\mathcal{P}$, or even that $\mathcal{F}$ be homeomorphic to a closed ball.

   In our examples, these technicalities will never be of much concern.

1.12 Spaces of Constant Curvature

1. A fundamental result in differential geometry is that there exist exactly three kinds of spaces which fill satisfy a 'wish list' of desirable geometric properties: *homogeneous* (all points alike), *isotropic* (all directions alike), *simply connected* (no 'holes'), *constant curvature*, etc.

   In particular, there are three two–dimensional possibilites, sometimes called *surfaces of constant curvature*. However, the word surface is a little misleading: it is possible, sometimes preferable, to think of these geometries *intrinsically*, as if you were a bug living in the surface, with no higher–dimensional awareness.

2. The three nice geometries are **spherical** $\mathbb{S}^d$, **Euclidean** $\mathbb{R}^d$, and **hyperbolic** (or **Lobachevskian**) $\mathbb{H}^d$.

   - The spherical case $\mathbb{S}^2$ is somewhat peculiar in that we have a ready made model for the geometry in Euclidean space, namely an ordinary unit sphere, with great circles serving as the lines (or *geodesics*) in the geometry.
   - And of course, even more obviously, we can model the Euclidean plane $\mathbb{R}^2$ in ordinary space. The Euclidean plane *is* just a plane in ordinary space $\mathbb{R}^3$.
   - Actually, there is a subtle point to make here. All three geometries can be described in an intrinsic fashion, making no mention of Euclidean space. For example, we could set up axioms for each geometry and develop the theory without ever making mention of ordinary space. In particular, we can talk about distance, congruence, angles, etc. in a sensible way in all three geometries.

     The peculiar feature of the spherical and Euclidean cases is that we have ready made models in Euclidean space, and further

     distance in the model, as 'inherited' from $\mathbb{R}^3$,
     coincides with intrinsic distance in the geometry.

   - It is the hyperbolic plane $\mathbb{H}^2$ that is quite unfamiliar, perhaps because there is no distance–faithful model in ordinary space $\mathbb{R}^3$. It is possible to model the hyperbolic plane in a distance-faithful way as a surface embedded in some $\mathbb{R}^d$. (I think $d = 5$ works.)

But techniques such as obstruction theory in algebraic topology do prove that the hyperbolic plane cannot be isometrically embedded in $\mathbb{R}^3$.

However, there are several very nice distance–unfaithful models of the hyperbolic plane, right inside the Euclidean plane $\mathbb{R}^2$. In each case, the whole hyperbolic plane is represented by the interior of the unit circle. (The circle iteself is interesting; but its points do not belong to the model for $\mathbb{H}^2$.)

(i) in the **Klein model**, lines are 'open' chords of the circle (i.e. discard endpoints); angles are distorted, except at the centre. And, yes, distance is distorted. Each chord has infinite hyperbolic length.

(ii) in the **Poincaré model**, lines are open circular arcs or diameters, always perpendicular to the unit circle. (Again we don't include end points.) Here angles are exactly represented, and we say that the model is *conformal*. But again distance is distorted: a circular arc or diameter connecting points on the boundary has infinite hyperbolic length.

It is this model that underlies Escher's Circle Limit prints.

3. Spherical space $\mathbb{S}^d$ does have a somewhat irksome feature: two distinct geodesics (think great circles) intersect in two points (antipodes). The remedy is to identify all such antipodal pairs of points so as to get **elliptic space** (or **projective space**) $\mathbb{P}^d$. Now two distinct lines do intersect in one point. But the space is no longer simply connected. We won't pursue it here.

4. Here is a table comparing some fundamental features of the three two-dimensional geometries.

	$\mathbb{S}^2$	$\mathbb{R}^2$	$\mathbb{H}^2$
name	spherical	Euclidean	hyperbolic
total area	4π	∞	∞
	(think unit sphere)		
curvature $\kappa =$	$+1$	0	-1
angle sum in $\triangle ABC$	$\alpha+\beta+\gamma > \pi$	$\alpha+\beta+\gamma = \pi$	$\alpha+\beta+\gamma < \pi$
(with angles α, β, γ)			
area of $\triangle ABC$	$= (\ \alpha+\beta+\gamma) - \pi$	is indeterminate	$= \pi - (\alpha+\beta+\gamma)$
	(the angular excess)		(the angular defect)
Pythagoras (for $\triangle ABC$	$\cos(c) = \cos(a)\cos(b)$	$a^2+b^2=c^2$	$\cosh(c) = \cosh(a)\cosh(b)$
with $\angle C = \pi/2$)			