GROUP THEORY AND INTRODUCTION TO RINGS

NOTES FOR THE COURSE ALGEBRA 3, MATH 370

MCGILL UNIVERSITY, FALL 2004, VERSION: January 13, 2005

EYAL Z. GOREN

©All rights reserved to the author.

Contents

Part 1. Basic Concepts and Key Examples

Groups are among the most rudimentary forms of algebraic structures. Because of their simplicity, in terms of their definition, their complexity is large. For example, vector spaces, which have very complex definition, are easy to classify; once the field and dimension are known, the vector space is unique up to isomorphism. In contrast, it is difficult to list all groups of a given order, or even obtain an asymptotic formula for that number.

In the study of vector spaces the objects are well understood and so one focuses on the study of maps between them. One studies canonical forms (e.g., the Jordan canonical form), diagonalization, and other special properties of linear transformations (normal, unitary, nilpotent, etc.). In contrast, at least in the theory of finite groups on which this course focuses, there is no comparable theory of maps. A theory exist mostly for maps into matrix groups (such maps are called linear representation and will not be studied in this course).

While we shall define such maps (called homomorphisms) between groups in general, there will be a large set of so called simple groups[1] for which there are essentially no such maps: the image of a simple group under a homomorphism is for all practical purposes just the group itself. The set of atoms is large, infinite in fact. The classification of all simple groups was completed in the second half of the 20-th century and has required thousands of pages of difficult math.

Thus, our focus - apart from the three isomorphism theorems - will be on the structure of the objects themselves. We will occupy ourselves with understanding the structure of subgroups of a finite group, with groups acting as symmetries of a given set and with special classes of groups (cyclic, simple, abelian, solvable, etc.).

1. First definitions

1.1. **Group.** A *group* G is a non-empty set with a function

$$m : G \times G \longrightarrow G,$$

where we usually abbreviate $m(g,h)$ to $g \star h$ or simply gh, such that the following hold:

(1) *(Associativity)* $f(gh) = (fg)h$ for all $f, g, h \in G$. [2]
(2) *(Identity)* There is an element $e \in G$ such that for all $g \in G$ we have $eg = ge = g$.
(3) *(Inverse)* For every $g \in G$ there is an element $h \in G$ such that $gh = hg = e$.

It follows quite easily from associativity that given any n elements $g_1, \ldots, g_n$ of G we can put parentheses as we like in $g_1 \star \cdots \star g_n$ without changing the final outcome. For that reason we allow ourselves to write simply $g_1 \cdots g_n$ (though the actual computation of such product is done by successively multiplying two elements, e.g. $(((g_1g_2)(g_3g_4))g_5)$ is a way to compute $g_1g_2g_3g_4g_5$.)

The identity element is unique: if e' has the same property then $e' = ee' = e$. Sometimes we will denote the identity element by 1 (or by 0 is the group is commutative - see below). The element h provided in axiom (3) is unique as well: if h' has the same property then $hg = e = h'g$ and so $hgh = h'gh$, thus $h = he = hgh = h'gh = h'e = h'$. We may therefore denote this h unambiguously by g^{-1}. Note that if h is the inverse of g then g is the inverse of h and so $(g^{-1})^{-1} = g$. Another useful identity is $(fg)^{-1} = g^{-1}f^{-1}$. It is verified just by checking that $g^{-1}f^{-1}$ indeed functions as $(fg)^{-1}$ and it does: $(g^{-1}f^{-1})(fg) = g^{-1}(f^{-1}f)g = g^{-1}eg = g^{-1}g = e$.

We define by induction $g^n = g^{n-1}g$ for $n > 0$ and $g^n = (g^{-n})^{-1}$ for $n < 0$. Also $g^0 = e$, by definition. One proves that $g^{n+m} = g^n g^m$ for any $n, m \in \mathbb{Z}$.

---

[1] A more appropriate name might be "atomic groups", but the terminology is too deeply rooted to deviate from it.
[2] In fuller notation $m(f, m(g,h)) = m(m(f,g),h)$.

A group is called of *finite order* if it has finitely many elements. It is called *abelian* if it is *commutative*: $gh = hg$ for all $g, h \in G$.

1.2. **Subgroup and order.** A *subgroup* H of a group G is a non-empty subset of G such that (i) $e \in H$, (ii) if $g, h \in H$ then $gh \in H$, and (iii) if $g \in H$ then also $g^{-1} \in H$. One readily checks that in fact H is a group. One checks that $\{e\}$ and G are always subgroups, called the *trivial subgroups*. We will use the notation

$$H < G$$

to indicate that H is a subgroup of G.

One calls a subgroup H *cyclic* if there is an element $h \in H$ such that $H = \{h^n : n \in \mathbb{Z}\}$. Note that $\{h^n : n \in \mathbb{Z}\}$ is always a cyclic subgroup. We denote it by $\langle h \rangle$. The *order* of an element $h \in G$, $o(h)$, is defined to be the minimal positive integer n such that $h^n = e$. If no such n exists, we say h has infinite order.

Lemma 1.2.1. *For every $h \in G$ we have $o(h) = \sharp\langle h \rangle$.*

Proof. Assume first that $o(h)$ is finite. Since for every n we have $h^{n+o(h)} = h^n h^{o(h)} = h^n$ we see that $\langle h \rangle = \{e, h, h^2, \dots, h^{o(h)-1}\}$. Thus, also $\sharp\langle h \rangle$ is finite and at most $o(h)$.

Suppose conversely that $\sharp\langle h \rangle$ is finite, say of order n. Then the elements $\{e = h^0, h, \dots, h^n\}$ cannot be distinct and thus for some $0 \leq i < j \leq n$ we have $h^i = h^j$. Therefore, $h^{j-i} = e$ and we conclude that $o(h)$ is finite and $o(h)$ is at most $\sharp\langle h \rangle$. This concludes the proof. □

Corollary 1.2.2. *If h has a finite order n then $\langle h \rangle = \{e, h, \dots, h^{n-1}\}$ and consists of precisely n elements (that is, there are no repetitions in this list.)*

It is ease to check that if $\{H_\alpha : \alpha \in J\}$ is a non-empty set of subgroups of G then $\cap_{\alpha \in J} H_\alpha$ is a subgroup as well. Let $\{g_\alpha : \alpha \in I\}$ be a set consisting of elements of G (here I is some index set). We denote by $\langle \{g_\alpha : \alpha \in I\} \rangle$ the minimal subgroup of G containing $\{g_\alpha : \alpha \in I\}$. It is clearly the intersection of all subgroups of G containing $\{g_\alpha : \alpha \in I\}$.

Lemma 1.2.3. *The subgroup $\langle \{g_\alpha : \alpha \in I\} \rangle$ is the set of all finite expressions $h_1 \cdots h_t$ where each h_i is some g_α or g_α^{-1}.*

Proof. Clearly $\langle \{g_\alpha : \alpha \in I\} \rangle$ contains each g_α hence all the expressions $h_1 \cdots h_t$ where each h_i is some g_α or g_α^{-1}. Thus, it is enough to show that the set of all finite expressions $h_1 \cdots h_t$, where each h_i is some g_α or g_α^{-1}, is a subgroup. Clearly e (equal to the empty product, or to $g_\alpha g_\alpha^{-1}$ if you prefer) is in it. Also, from the definition it is clear that it is closed under multiplication. Finally, since $(h_1 \cdots h_t)^{-1} = h_t^{-1} \cdots h_1^{-1}$ it is also closed under taking inverses. □

We call $\langle \{g_\alpha : \alpha \in I\} \rangle$ *the subgroup of G generated by* $\{g_\alpha : \alpha \in I\}$; if it is equal to G, we say that $\{g_\alpha : \alpha \in I\}$ are *generators* for G.

2. Main examples

2.1. **$\mathbb{Z}$, $\mathbb{Z}/n\mathbb{Z}$ and $(\mathbb{Z}/n\mathbb{Z})^\times$.** The set of integers $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, 3, \dots\}$, with the addition operation, is an infinite abelian group. It is cyclic; both 1 and -1 are generators.

The group $\mathbb{Z}/n\mathbb{Z}$ of integers modulo n, $\{0, 1, 2, \dots, n-1\}$, with addition modulo n, is a finite abelian group. The group $\mathbb{Z}/n\mathbb{Z}$ is a cyclic group with generator 1. In fact (see the section on cyclic groups), an element x generates $\mathbb{Z}/n\mathbb{Z}$ if and only if $(x, n) = 1$.

Consider $(\mathbb{Z}/n\mathbb{Z})^\times = \{a \in \mathbb{Z}/n\mathbb{Z} : (a, n) = 1\}$ with multiplication. It is a group whose order is denoted by $\phi(n)$ (the function $n \mapsto \phi(n)$ is call *Euler's phi function*). To see it is a group, note that multiplication is associative and if $(a, n) = 1, (b, n) = 1$ then also $(ab, n) = 1$ (thus, we do indeed get an operation on $\mathbb{Z}/n\mathbb{Z}^\times$). The congruence class 1 is the identity and the existence of inverse follows from finiteness: given $a \in \mathbb{Z}/n\mathbb{Z}^\times$ consider the function $x \mapsto ax$. It is injective: if $ax = ay$ then $a(x - y) = 0 \pmod{n}$, that is (using the same letters to denote integers in these congruence classes) $n | a(x - y)$. Since $(a, n) = 1$ we

conclude that $n|(x-y)$ that is, $x = y$ in $\mathbb{Z}/n\mathbb{Z}$. It follows that $x \mapsto ax$ is also surjective and thus there is an element x such that $ax = 1$.

2.2. **The dihedral group D_{2n}.** Let $n \geq 3$. Consider the linear transformations of the plane that take a regular polygon with n sides, symmetric about zero, unto itself. One easily sees that every such symmetry is determine by its action of the vertices $1, 2$ (thought of as vectors, they form a basis) and that it takes these vertices to the vertices $i, i+1$ or $i+1, i$, where $1 \leq i \leq n$ (and the labels of the vertices are read modulo n). One concludes that every such symmetry is of the form $y^a x^b$ for suitable and unique $a \in \{0, 1\}, b \in \{1, \ldots, n\}$, where y is the reflection fixing 1 (so takes $n, 2$ to $2, n$) and x is the rotation taking $1, 2$ to $2, 3$. One finds that $y^2 = e = x^n$ and that $yxy = x^{-1}$. All other relations are consequences of these.

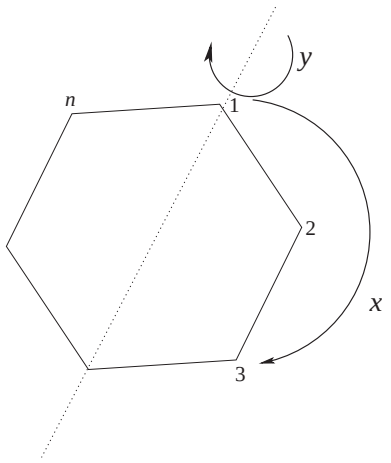


FIGURE 2.1. Symmetries of a regular Polygon with n vertices.

The Dihedral group is thus a group of order $2n$ generated by a reflection y and a rotation x satisfying $y^2 = x^n = xyxy = e$. This makes sense also for $n = 1, 2$.

2.3. **The symmetric group S_n.** Consider the set S_n consisting of all injective (hence bijective) functions, called *permutations*,

$$\sigma : \{1, 2, \ldots, n\} \longrightarrow \{1, 2, \ldots, n\}.$$

We define

$$m(\sigma, \tau) = \sigma \circ \tau.$$

This makes S_n into a group, whose identity e is the identity function $e(i) = i, \forall i$.

We may describe the elements of S_n in the form of a table:

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix}.$$

This defines a permutation σ by the rule $\sigma(a) = i_a$.

Another device is to use the notation $(i_1 i_2 \ldots i_s)$, where the i_j are distinct elements of $\{1, 2, \ldots, n\}$. This defines a permutation σ according to the following convention: $\sigma(i_a) = i_{a+1}$ for $1 \leq a < s$, $\sigma(i_s) = i_1$, and for any other element x of $\{1, 2, \ldots, n\}$ we let $\sigma(x) = x$. Such a permutation is called a *cycle*. One can easily prove the following facts:

(1) Disjoint cycles commute.
(2) Every permutation is a product of disjoint cycles (uniquely up to permuting the cycles and omitting cycles of length one).
(3) The order of $(i_1 i_2 \ldots i_s)$ is s.
(4) If $\sigma_1, \ldots, \sigma_t$ are disjoint cycles of orders $r_1, \ldots, r_t$ then the order of $\sigma_1 \circ \cdots \circ \sigma_t$ is the least common multiple of $r_1, \ldots, r_t$.
(5) The symmetric group has order $n!$.

Example 2.3.1. The order of the permutation (1 2 3 4) is 4. Indeed, it is not trivial and $(1\ 2\ 3\ 4)^2 = (1\ 3)(2\ 4)$, $(1\ 2\ 3\ 4)^3 = (4\ 3\ 2\ 1)$, $(1\ 2\ 3\ 4)^4 = 1$.

The permutation $\left(\begin{smallmatrix}1&2&3&4&5&6\\6&1&3&5&4&2\end{smallmatrix}\right)$ is equal to the product of cycles (1 6 2)(4 5). It is of order 6.

2.3.1. *The sign of a permutation, and realizing permutations as linear transformations.*

Lemma 2.3.2. *Let $n \geq 2$. Let S_n be the group of permutations of $\{1, 2, \ldots, n\}$. There exists a surjective homomorphism*[3] *of groups*

$$\mathrm{sgn} : S_n \longrightarrow \{\pm 1\}$$

(called the 'sign'). It has the property that for every $i \neq j$,

$$\mathrm{sgn}(\ (ij)\ ) = -1.$$

Proof. Consider the polynomial in n-variables[4]

$$p(x_1, \ldots, x_n) = \prod_{i<j}(x_i - x_j).$$

Given a permutation σ we may define a new polynomial

$$\prod_{i<j}(x_{\sigma(i)} - x_{\sigma(j)}).$$

Note that $\sigma(i) \neq \sigma(j)$ and for any pair $k < \ell$ we obtain in the new product either $(x_k - x_\ell)$ or $(x_\ell - x_k)$. Thus, for a suitable choice of sign $\mathrm{sgn}(\sigma) \in \{\pm 1\}$, we have[5]

$$\prod_{i<j}(x_{\sigma(i)} - x_{\sigma(j)}) = \mathrm{sgn}(\sigma)\prod_{i<j}(x_i - x_j).$$

We obtain a function

$$\mathrm{sgn} : S_n \longrightarrow \{\pm 1\}.$$

This function satisfies $\mathrm{sgn}(\ (k\ell)\ ) = -1$ (for $k < \ell$): Let $\sigma = (k\ell)$ and consider the product

$$\prod_{i<j}(x_{\sigma(i)} - x_{\sigma(j)}) = (x_\ell - x_k) \prod_{\substack{i<j\\ i\neq k, j\neq \ell}} (x_i - x_j) \prod_{\substack{k<j\\ j\neq \ell}} (x_\ell - x_j) \prod_{\substack{i<\ell\\ i\neq k}} (x_i - x_k).$$

(This corresponds to the cases (i) $i = k, j = \ell$; (ii) $i = k, j \neq \ell (\Rightarrow j > k)$; (iii) $i \neq k, j = \ell (\Rightarrow i < \ell)$; (iv) $i \neq k, j \neq \ell$.) Counting the number of signs that change we find that

$$\prod_{i<j}(x_{\sigma(i)} - x_{\sigma(j)}) = (-1)(-1)^{\sharp\{j:k<j<\ell\}}(-1)^{\sharp\{i:k<i<\ell\}} \prod_{i<j}(x_i - x_j) = -\prod_{i<j}(x_i - x_j).$$

It remains to show that sgn is a group homomorphism. We first make the innocuous observation that for *any* variables $y_1, \ldots, y_n$ and for *any* permutation σ we have

$$\prod_{i<j}(y_{\sigma(i)} - y_{\sigma(j)}) = \mathrm{sgn}(\sigma)\prod_{i<j}(y_i - y_j).$$

Let τ be a permutation. We apply this observation for the variables $y_i := x_{\tau(i)}$. We get

$$\begin{aligned}
\mathrm{sgn}(\tau\sigma)p(x_1, \ldots, x_n) &= p(x_{\tau\sigma(1)}, \ldots, x_{\tau\sigma(n)}) \\
&= p(y_{\sigma(1)}, \ldots, y_{\sigma(n)}) \\
&= \mathrm{sgn}(\sigma)p(y_1, \ldots, y_n) \\
&= \mathrm{sgn}(\sigma)p(x_{\tau(1)}, \ldots, x_{\tau(n)}) \\
&= \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau)p(x_1, \ldots, x_n).
\end{aligned}$$

This gives

$$\mathrm{sgn}(\tau\sigma) = \mathrm{sgn}(\tau)\mathrm{sgn}(\sigma).$$

---

[3] That means $\mathrm{sgn}(\sigma\tau) = \mathrm{sgn}(\sigma)\mathrm{sgn}(\tau)$

[4] For $n = 2$ we get $x_1 - x_2$. For $n = 3$ we get $(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$.

[5] For example, if $n = 3$ and σ is the cycle (123) we have

$$(x_{\sigma(1)} - x_{\sigma(2)})(x_{\sigma(1)} - x_{\sigma(3)})(x_{\sigma(2)} - x_{\sigma(3)}) = (x_2 - x_3)(x_2 - x_1)(x_3 - x_1) = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3).$$

Hence, $\mathrm{sgn}(\ (1\ 2\ 3)\ ) = 1$.

□

Calculating sgn **in practice.** Recall that every permutation σ can be written as a product of disjoint cycles

$$\sigma = (a_1 \dots a_\ell)(b_1 \dots b_m) \dots (f_1 \dots f_n).$$

Claim: $\operatorname{sgn}(a_1 \dots a_\ell) = (-1)^{\ell-1}$.
Corollary: $\operatorname{sgn}(\sigma) = (-1)^{\sharp \text{ even length cycles}}$.

Proof. We write

$$(a_1 \dots a_\ell) = \underbrace{(a_1 a_\ell) \dots (a_1 a_3)(a_1 a_2)}_{\ell-1 \text{ transpositions}}.$$

Since a transposition has sign -1 and sgn is a homomorphism, the claim follows. □

A Numerical example. Let $n = 11$ and

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 5 & 4 & 3 & 1 & 7 & 8 & 10 & 6 & 9 \end{pmatrix}.$$

Then

$$\sigma = (1\ 2\ 5)(3\ 4)(6\ 7\ 8\ 10\ 9).$$

Now,

$$\operatorname{sgn}(\ (1\ 2\ 5)\ ) = 1, \quad \operatorname{sgn}(\ (3\ 4)\ ) = -1, \quad \operatorname{sgn}(\ (6\ 7\ 8\ 10\ 9)\ ) = 1.$$

We conclude that $\operatorname{sgn}(\sigma) = -1$.

Realizing S_n as linear transformations. Let $\mathbb{F}$ be any field. Let $\sigma \in S_n$. There is a unique linear transformation

$$T_\sigma : \mathbb{F}^n \longrightarrow \mathbb{F}^n,$$

such that

$$T(e_i) = e_{\sigma(i)}, \quad i = 1, \dots n,$$

where, as usual, $e_1, \dots, e_n$ are the standard basis of $\mathbb{F}^n$. Note that

$$T_\sigma \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_{\sigma^{-1}(1)} \\ x_{\sigma^{-1}(2)} \\ \vdots \\ x_{\sigma^{-1}(n)} \end{pmatrix}.$$

(For example, because $T_\sigma x_1 e_1 = x_1 e_{\sigma(1)}$, the $\sigma(1)$ coordinate is x_1, namely, in the $\sigma(1)$ place we have the entry $x_{\sigma^{-1}(\sigma(1))}$.) Since for every i we have $T_\sigma T_\tau(e_i) = T_\sigma e_{\tau(i)} = e_{\sigma\tau(i)} = T_{\sigma\tau} e_i$, we have the relation

$$T_\sigma T_\tau = T_{\sigma\tau}.$$

The matrix representing T_σ is the matrix (a_{ij}) with $a_{ij} = 0$ unless $i = \sigma(j)$. For example, for $n = 4$ the matrices representing the permutations (12)(34) and (1 2 3 4) are, respectively

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

Otherwise said,[6]

$$T_\sigma = \left(e_{\sigma(1)} \mid e_{\sigma(2)} \mid \ldots \mid e_{\sigma(n)}\right) = \begin{pmatrix} e_{\sigma^{-1}(1)} \\ \hline e_{\sigma^{-1}(2)} \\ \hline \vdots \\ \hline e_{\sigma^{-1}(n)} \end{pmatrix}.$$

It follows that

$$\begin{aligned} \operatorname{sgn}(\sigma)\det(T_\sigma) &= \operatorname{sgn}(\sigma)\det\left(e_{\sigma(1)} \mid e_{\sigma(2)} \mid \ldots \mid e_{\sigma(n)}\right) \\ &= \det\left(e_1 \mid e_2 \mid \ldots \mid e_n\right) \\ &= \det(I_n) \\ &= 1. \end{aligned}$$

Recall that $\operatorname{sgn}(\sigma) \in \{\pm 1\}$. We get

$$\det(T_\sigma) = \operatorname{sgn}(\sigma).$$

2.3.2. *Transpositions and generators for S_n.* Let $1 \leq i < j \leq n$ and let $\sigma = (ij)$. Then σ is called a transposition. Let T be the set of all transpositions (T has $n(n-1)/2$ elements). Then T generates S_n. In fact, also the transpositions $(12), (23), \ldots, (n-1\ n)$ alone generate S_n.

2.3.3. *The alternating group A_n.* Consider the set A_n of all permutations in S_n whose sign is 1. They are called the *even* permutations (those with sign -1 are called *odd*). We see that $e \in A_n$ and that if $\sigma, \tau \in A_n$ also $\sigma\tau$ and σ^{-1}. This follows from $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau), \operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)^{-1}$.

Thus, A_n is a group. It is called the *alternating group*. It has $n!/2$ elements (use multiplication by (12) to create a bijection between the odd and even permutations). Here are some examples

n	A_n
2	$\{1\}$
3	$\{1, (123), (132)\}$
4	$\{1, (123), (132), (124), (142), (134), (143), (234), (243), (12)(34), (13)(24), (14)(23)\}$

2.3.4. *A useful formula for conjugation.* Let $\sigma, \tau \in S_n$. There is a nice formula for $\tau\sigma\tau^{-1}$ (this is called conjugating σ by τ). If σ is written as a product of cycles then the permutation $\tau\sigma\tau^{-1}$ is obtained by applying τ to the numbers appearing in the cycles of σ. That is, if σ takes i to j then $\tau\sigma\tau^{-1}$ takes $\tau(i)$ to $\tau(j)$. Indeed,

$$\tau\sigma\tau^{-1}(\tau(i)) = \tau(\sigma(i)) = \tau(j).$$

Here is an example: say $\sigma = (1\ 4)(2\ 5)(3\ 7\ 6)$ and $\tau = (1\ 2\ 3\ 4)(6\ 7)$ then $\tau\sigma\tau^{-1} = (\tau(1)\ \tau(4))\ (\tau(2)\ \tau(5))\ (\tau(3)\ \tau(7)\ \tau(6)) = (2\ 1)(3\ 5)(4\ 6\ 7)$.

2.4. **Matrix groups and the quaternions.** Let R be a commutative ring with 1. We let $\mathrm{GL}_n(R)$ denote the $n \times n$ matrices with entries with R, whose determinant is a unit in R.

Proposition 2.4.1. *$\mathrm{GL}_n(R)$ is a group under matrix multiplication.*

Proof. Multiplication of matrices is associative and the identity matrix is in $\mathrm{GL}_n(R)$. If $A, B \in \mathrm{GL}_n(R)$ then $\det(AB) = \det(A)\det(B)$ gives that $\det(AB)$ is a unit of R and so $AB \in \mathrm{GL}_n(R)$. The adjoint matrix satisfies $\mathrm{Adj(A)}A = \det(A)I_n$ and so every matrix A in $\mathrm{GL}_n(R)$ has an inverse equal to $\det(A)^{-1}\mathrm{Adj(A)}$. Note that $A^{-1}A = Id$ implies that $\det(A^{-1}) = \det(A)^{-1}$, hence an invertible element of R. Thus A^{-1} is in $\mathrm{GL}_n(R)$. □

[6]This gives the interesting relation $T_{\sigma^{-1}} = T_\sigma^t$. Because $\sigma \mapsto T_\sigma$ is a group homomorphism we may conclude that $T_\sigma^{-1} = T_\sigma^t$. Of course for a general matrix this doesn't hold.

Proposition 2.4.2. *If R is a finite field of q elements then $\mathrm{GL}_n(R)$ is a finite group of cardinality $(q^n-1)(q^n-q)\cdots(q^n-q^{n-1})$.*

Proof. To give a matrix in $\mathrm{GL}_n(R)$ is to give a basis of R^n (consisting of the columns of the matrix). The first vector v_1 in a basis can be chosen to be any non-zero vector and there are q^n-1 such vectors. The second vector v_2 can be chosen to be any vector not in $\mathrm{Span}(v_1)$; there are q^n-q such vectors. The third vector v_3 can be chosen to be any vector not in $\mathrm{Span}(v_1,v_2)$; there are q^n-q^2 such vectors. And so on. □

Exercise 2.4.3. Prove that the set of upper triangular matrices in $\mathrm{GL}_n(\mathbb{F})$, where $\mathbb{F}$ is any field, forms a subgroup of $\mathrm{GL}_n(F)$. It is also called a Borel subgroup.

Prove that the set of upper triangular matrices in $\mathrm{GL}_n(\mathbb{F})$ with 1 on the diagonal, where $\mathbb{F}$ is any field, forms a subgroup of $\mathrm{GL}_n(F)$. It is also called a unipotent subgroup.

Calculate the cardinality of these groups when $\mathbb{F}$ is a finite field of q elements.

Consider the case $R=\mathbb{C}$, the complex numbers, and the set of eight matrices

$$\left\{\pm\begin{pmatrix}1&0\\0&1\end{pmatrix},\pm\begin{pmatrix}i&0\\0&-i\end{pmatrix},\pm\begin{pmatrix}0&1\\-1&0\end{pmatrix},\pm\begin{pmatrix}0&i\\i&0\end{pmatrix}\right\}.$$

One verifies that this is a subgroup of $\mathrm{GL}_2(\mathbb{C})$, called the *Quaternion group*. One can use the notation

$$\pm 1, \pm i, \pm j, \pm k$$

for the matrices, respectively. Then we have

$$i^2=j^2=k^2=-1,\ ij=-ji=k,\ jk=i,\ ki=j.$$

2.5. **Groups of small order.** One can show that in a suitable sense (up to isomorphism, see § 8.1) the following is a complete list of groups for the given orders. (In the middle column we give the abelian groups and in the right column the non-abelian groups).

order	abelian groups	non-abelian groups
1	$\{1\}$	
2	$\mathbb{Z}/2\mathbb{Z}$	
3	$\mathbb{Z}/3\mathbb{Z}$	
4	$\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z}$	
5	$\mathbb{Z}/5\mathbb{Z}$	
6	$\mathbb{Z}/6\mathbb{Z}$	S_3
7	$\mathbb{Z}/7\mathbb{Z}$	
8	$\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/8\mathbb{Z}$	D_8, Q
9	$\mathbb{Z}/3\mathbb{Z}\times\mathbb{Z}/3\mathbb{Z}$, $\mathbb{Z}/9\mathbb{Z}$	
10	$\mathbb{Z}/10\mathbb{Z}$	D_{10}
11	$\mathbb{Z}/11\mathbb{Z}$	
12	$\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/6\mathbb{Z}$, $\mathbb{Z}/12\mathbb{Z}$	D_{12}, A_4, T

In the following table we list for every n the number $G(n)$ of subgroups of order n (this is taken from J. Rotman/*An introduction to the theory of groups*):

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
$G(n)$	1	1	1	2	1	2	1	5	2	2	1	5	1	2	1	14	1	5	1

n	20	21	22	23	24	25	26	27	28	29	30	31	32
$G(n)$	5	2	2	1	15	2	2	5	4	1	4	1	51

2.6. **Direct product.** Let G, H be two groups. Define on the cartesian product $G \times H$ multiplication by

$$m : (G \times H) \times (G \times H) \longrightarrow G \times H, \quad m((a,x),(b,y)) = (ab, xy).$$

This makes $G \times H$ into a group, called the *direct product* (also direct sum) of G and H.

One checks that $G \times H$ is abelian if and only if both G and H are abelian. The following relation among orders hold: $o(a,x) = \mathrm{lcm}(o(a), o(x))$. It follows that if G, H are cyclic groups whose orders are co-prime then $G \times H$ is also a cyclic group.

Example 2.6.1. If $H_1 < H, G_1 < G$ are subgroups then $H_1 \times G_1$ is a subgroup of $H \times G$. However, not every subgroup of $H \times G$ is of this form. For example, the subgroups of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ are $\{0\} \times \{0\}, \{0\} \times \mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/2\mathbb{Z} \times \{0\}, \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ and the subgroup $\{(0,0),(1,1)\}$ which is *not* a product of subgroups.

3. Cosets

Let G be a group and H a subgroup of G. A *left coset* of H in G is a subset S of G of the form

$$gH := \{gh : h \in H\}$$

for some $g \in G$. A *right coset* is a subset of G of the form

$$Hg := \{hg : h \in H\}$$

for some $g \in G$. For brevity we shall discuss only left cosets but the discussion with minor changes applies for right cosets too.

Example 3.0.2. Consider the group S_3 and the subgroup $H = \{1, (12)\}$. The following table lists the left cosets of H. For an element g, we list the coset gH in the middle column, and the coset Hg in the last column.

g	gH	Hg
1	$\{1,(12)\}$	$\{1,(12)\}$
(12)	$\{(12),1\}$	$\{(12),1\}$
(13)	$\{(13),(123)\}$	$\{(13),(132)\}$
(23)	$\{(23),(132))\}$	$\{(23),(123))\}$
(123)	$\{(123),(13)\}$	$\{(123),(23)\}$
(132)	$\{(132),(23)\}$	$\{(132),(13)\}$

The first observation is that the element g such that $S = gH$ is not unique. In fact, $gH = kH$ if and only if $g^{-1}k \in H$. The second observation is that two cosets are either equal or disjoint; this is a consequence of the following lemma.

Lemma 3.0.3. *Define a relation $g \sim k$ if $\exists h \in H$ such that $gh = k$. This is an equivalence relation such that the equivalence class of g is precisely gH.*

Proof. Since $g = ge$ and $e \in H$ the relation is reflexive. If $gh = k$ for some $h \in H$ then $kh^{-1} = g$ and $h^{-1} \in H$. Thus, the relation is symmetric. Finally, if $g \sim k \sim \ell$ then $gh = k, kh' = \ell$ for some $h, h' \in H$ and so $g(hh') = \ell$. Since $hh' \in H$ we conclude that $g \sim \ell$ and the relation is transitive. □

Thus, pictorially the cosets look like that:

Aside. One should note that in general $gH \neq Hg$; The table above provides an example. Moreover, $(13)H$ is not a right coset of H at all. A difficult theorem of P. Hall asserts that given a finite group G and a subgroup H one can find a set $g_1, \ldots, g_d$ such that $g_1H, \ldots, g_dH$ are precisely the lest cosets of H and $Hg_1, \ldots, Hg_d$ are precisely the right cosets of H.

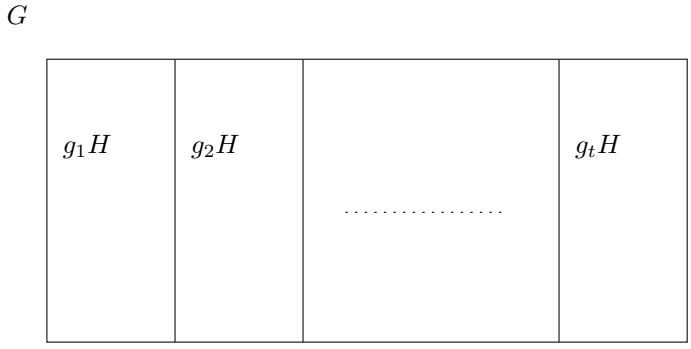


FIGURE 3.1. Cosets of a subgroup H of a group G.

4. Lagrange's theorem

Theorem 4.0.4. *Let $H < G$. The group G is a disjoint union of left cosets of H. If G is of finite order then the number of left cosets of H in G is $|G|/|H|$. We call the number of left cosets the index of H in G and denote it by $[G : H]$.*

Proof. We have seen that there is an equivalence relation whose equivalence classes are the cosets of H. Recall that different equivalence classes are disjoint. Thus,

$$G = \cup_{i=1}^{s} g_i H,$$

a disjoint union of s cosets, where the g_i are chosen appropriately. We next show that for every $x, y \in G$ the cosets xH, yH have the same number of elements.

Define a function

$$f : xH \longrightarrow yH, \qquad f(xh) = yh.$$

Note that f is well defined ($xh = xh' \Rightarrow h = h'$), injective ($f(xh) = yh = yh' = f(xh') \Rightarrow h = h' \Rightarrow xh = xh'$) and surjective as every element of yH has the form yh for some $h \in H$ hence is the image of xh. Thus, $|G| = s \cdot |H|$ and $s = [G : H]$. □

Corollary 4.0.5. *If G is a finite group then $|H|$ divides $|G|$.*

Remark 4.0.6. The converse does not hold. The group A_4, which is of order 12, does not have a subgroup of order 6.

Corollary 4.0.7. *If G is a finite group then $o(g) \,|\, |G|$ for all $g \in G$.*

Proof. We saw that $o(g) = |\langle g \rangle|$. □

Remark 4.0.8. The converse does not hold. If G is not a cyclic group then there is no element $g \in G$ such that $o(g) = |G|$.

Corollary 4.0.9. *If the order of G is a prime number then G is cyclic.*

Proof. From Corollary 4.0.7 we deduce that every element different from the identity has order equal to $|G|$. Thus, every such element generates the group. □

5. Cyclic groups

Let G be a finite cyclic group of order n, $G = \langle g \rangle$.

Lemma 5.0.10. *We have $o(g^a) = n/\gcd(a, n)$.*

Proof. Note that $g^t = g^{t-n}$ and so $g^t = e$ if and only if $n|t$ (cf. Corollary 1.2.2). Thus, the order of g^a is the minimal r such that ar is divisible by n. Clearly $a \cdot n/\gcd(a, n)$ is divisible by n so the order of g^a is less or equal to $n/\gcd(a, n)$. On the other hand if ar is divisible by n then, because $n = \gcd(a, n) \cdot n/\gcd(a, n)$, r is divisible by $n/\gcd(a, n)$. □

Corollary 5.0.11. *The element g^a generates G, $\langle g^a\rangle = G$, if and only if $(a,n)=1$. Thus, the number of generators of G is*

$$\varphi(n) := \sharp\{1 \le a \le n : (a,n)=1\}.$$

This function is called Euler's phi function.

Proposition 5.0.12. *For every $h|n$ the group G has a unique subgroup of order h. This subgroup is cyclic.*

Proof. We first show that every subgroup is cyclic. Let H be a non trivial subgroup. Then there is a minimal $0 < a < n$ such that $g^a \in H$ and hence $H \supseteq \langle g^a\rangle$. Let $g^r \in H$. We may assume that $r > 0$. Write $r = ka + k'$ for $0 \le k' < a$. Note that $g^{r-ka} \in H$. The choice of a then implies that $k' = 0$. Thus, $H = \langle g^a\rangle$.

Since $\gcd(a,n) = \alpha a + \beta n$ we have $g^{\gcd(a,n)} = (g^n)^\beta (g^a)^\alpha \in H$. Thus, $g^{a-\gcd(a,n)} \in H$. Therefore, by the choice of a, $a = \gcd(a,n)$; that is, $a|n$. Thus, every subgroup is cyclic and of the form $\langle g^a\rangle$ for $a|n$. Its order is n/a. We conclude that for every $b|n$ there is a unique subgroup of order b and it is cyclic, generated by $g^{n/b}$. □

Lemma 5.0.13. *We have*

$$n = \sum_{d|n} \varphi(d).$$

(The summation is over positive divisors of n, including 1 and n.)[7]

Proof. Let G be a cyclic group of order n. Then we have

$$\begin{aligned} n &= |G| \\ &= \sum_{1\le d\le n} \sharp\{g \in G : o(g) = d\} \\ &= \sum_{d|n} \sharp\{g \in G : o(g) = d\}, \end{aligned}$$

where we have used that the order of an element divides the order of the group.

Now, if $h \in G$ has order d it generates a subgroup of order d. Such subgroup being unique, it follows that all the elements of order d generate the same subgroup. That subgroup is a cyclic group of order d and thus has $\varphi(d)$ generators that are exactly the elements of order d. The formula follows. □

Proposition 5.0.14. *Let G be a finite group of order n such that for $h|n$ the group G has at most one subgroup of order h then G is cyclic.*

Proof. Consider an element $g \in G$ of order h. The subgroup $\langle g\rangle$ it generates is of order h and has $\varphi(h)$ generators. We conclude that every element of order h must belong to this subgroup (because there is a unique subgroup of order h in G) and that there are exactly $\varphi(h)$ elements of order h in G.

On the one hand $n = \sum_{d|n}^n \{\text{num. elts. of order } d\} = \sum_{d|n} \phi(d)\epsilon_d$, where ϵ_d is 1 if there is an element of order d and is zero otherwise. On the other hand $n = \sum_{d|n} \phi(d)$. We conclude that $\epsilon_d = 1$ for all $d|n$ and, in particular, $\epsilon_n = 1$ and so there is an element of order n. This element is a generator of G. □

Corollary 5.0.15. *Let $\mathbb{F}$ be a finite field then $\mathbb{F}^\times$ is a cyclic group.*

Proof. Let q be the number of element of $\mathbb{F}$. To show that for every h dividing $q-1$ there is at most one subgroup of order h we note that every element in that subgroup with have order dividing h and hence will solve the polynomial $x^h - 1$. That is, the h elements in that subgroup must be the h solutions of $x^h - 1$. In particular, this subgroup is unique. □

[7] This function has the following additional properties:

- If n and m are relatively prime then $\phi(nm) = \phi(n)\phi(m)$.
  (This can be proved as follows. Using the Chinese Remainder Theorem $\mathbb{Z}/nm\mathbb{Z} \cong \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ as rings. Now calculate the unit groups of both sides.)
- If p is a prime $\varphi(p^a) = p^a - p^{a-1}$.
  (This follows directly from the definition. Putting the two properties together, we find that:)
- $\varphi(n) = n\prod_{p|n}(1-1/p)$ if p^a is the highest power of p dividing n.

Remark 5.0.16. Though the groups $(\mathbb{Z}/p\mathbb{Z})^\times$ are cyclic for every prime p that doesn't mean we know an explicit generator. Artin's primitive root conjecture states that 2 is a generator for infinitely many primes p (the conjecture is the same for any prime number instead of 2). Work starting with R. Murty and R. Gupta and continued with K. Murty and Heath-Brown had shown that for infinitely many primes p either $2, 3$ or 5 are a primitive root.

6. Constructing subgroups

6.1. **Commutator subgroup.** Let G be a group. Define its *commutator subgroup* G', or $[G, G]$, to be the subgroup generated by $\{xyx^{-1}y^{-1}; x, y \in G\}$. An element of the form $xyx^{-1}y^{-1}$ is called a *commutator*. We use the notation $[x, y] = xyx^{-1}y^{-1}$. It is not true in general that every element in G' is a commutator, though every element is a product of commutators.

Example 6.1.1. We calculate the commutator subgroup of S_3. First, note that every commutator is an even permutation, hence contained in A_3. Next, $(12)(13)(12)(13) = (123)$ is in S_3'. It follows that $S_3' = A_3$.

6.2. **Centralizer subgroup.** Let H be a subgroup of G. We define its *centralizer* $C_G(H)$ to be the set $\{g \in G : gh = hg, \forall h \in H\}$. One checks that it is a subgroup of G called *the centralizer of H in G*.

Given an element $h \in G$ we may define $C_G(h) = \{g \in G : gh = hg\}$. It is a subgroup of G called the centralizer of h in G. One checks that $C_G(h) = C_G(\langle h \rangle)$ and that $C_G(H) = \cap_{h \in H} C_G(h)$.

Taking $H = G$, the subgroup $C_G(G)$ is the set of elements of G such that each of them commutes with every other element of G. It has a special name; it is called the *center* of G and denoted $Z(G)$.

Example 6.2.1. We calculate the centralizer of (12) in S_5. We first make the following useful observation: $\tau\sigma\tau^{-1}$ is the permutation obtained from σ by changing its entries according to τ. For example: $(1234)[(12)(35)](1234)^{-1} = (1234)[(12)(35)](1432) = (1234)(1453) = (23)(45)$ and $(23)(45)$ is obtained from $(12)(35)$ by changing the labels $1, 2, 3, 4, 5$ according to the rule (1234).

Using this, we see that the centralizer of (12) in S_5 is just $S_2 \times S_3$ (Here S_2 are the permutations of $1, 2$ and S_3 are the permutations of $3, 4, 5$. Viewed this way they are subgroups of S_5).

6.3. **Normalizer subgroup.** Let H be a subgroup of G. Define the *normalizer* of H in G, $N_G(H)$, to be the set $\{g \in G : gHg^{-1} = H\}$. It is a subgroup of G. Note that $H \subset N_G(H), C_G(H) \subset N_G(H)$ and $H \cap C_G(H) = Z(H)$.

7. Normal subgroups and quotient groups

Let $N < G$. We say that N is a *normal* subgroup if for all $g \in G$ we have $gN = Ng$; equivalently, $gNg^{-1} = N$ for all $g \in G$; equivalently, $gN \subset Ng$ for all $g \in G$; equivalently, $gNg^{-1} \subset N$ for all $g \in G$. We will use the notation $N \triangleleft G$ to signify that N is a normal subgroup of G. Note that an equivalent way to say that $N \triangleleft G$ is to say that $N < G$ and $N_G(N) = G$.

Example 7.0.1. The group A_3 is normal in S_3. If $\sigma \in A_3$ and $\tau \in S_3$ then $\tau\sigma\tau^{-1}$ is an even permutation because its sign is $\mathrm{sgn}(\tau)\mathrm{sgn}(\sigma)\mathrm{sgn}(\tau^{-1}) = \mathrm{sgn}(\tau)^2\mathrm{sgn}(\sigma) = 1$. Thus, $\tau A_3 \tau^{-1} \subset A_3$.

The subgroup $H = \{1, (12)\}$ is not a normal subgroup. Use the table above to see that $(13)H \neq H(13)$.

Let $N \triangleleft G$. Let G/N denote the set of left cosets of N in G. We show that G/N has a natural structure of a group; it is called the *quotient group* of G by N.

Given two cosets aN and bN we define

$$aN \star bN = abN.$$

We need to show this is well defined: if $aN = a'N$ and $bN = b'N$ then we should have $abN = a'b'N$. Now, we know that for a suitable $\alpha, \beta \in N$ we have $a'\alpha = a, b'\beta = b$. Thus, $a'b'N = a\alpha b\beta N = abb^{-1}\alpha b\beta N = ab(b^{-1}\alpha b)N$. Note that since $N \lhd G$ and $\alpha \in N$ also $b^{-1}\alpha b \in N$ and so $ab(b^{-1}\alpha b)N = abN$.

One checks easily that $N = eN$ is the identity of G/N and that $(gN)^{-1} = g^{-1}N$. (Note that $(gN)^{-1}$ - the inverse of the element gN in the group G/N is also the set $\{(gn)^{-1} : n \in N\} = Ng^{-1} = g^{-1}N$.)

Definition 7.0.2. A group is called *simple* if its only normal subgroups are the trivial ones $\{e\}$ and G.

Remark 7.0.3. We shall later prove that A_n is a simple group for $n \geq 5$. By inspection one find that also A_n is simple for $n \leq 3$. On the other hand A_4 is not simple. The "Klein 4 group" $V := \{1, (12)(34), (13)(24), (14)(23)\}$ is a normal subgroup of A_4.

Recall the definition of the commutator subgroup G' of G from §6.1. In particular, the notation $[x, y] = xyx^{-1}y^{-1}$. One easily checks that $g[x, y]g^{-1} = [gxg^{-1}, gyg^{-1}]$ and that $[x, y]^{-1} = [y, x]$. Hence, also $g[x, y]^{-1}g^{-1} = [gxg^{-1}, gyg^{-1}]^{-1}$.

Proposition 7.0.4. *The subgroup G' is normal in G. The group G/G' is abelian (it is called the* abelianization *of G). Furthermore, if G/N is abelian then $N \supseteq G'$.*

Proof. We know that $G' = \{[x_1, y_1]^{\epsilon_1} \cdots [x_r, y_r]^{\epsilon_r} : x_i, y_i \in G, \epsilon_i = \pm 1\}$. It follows that

$$gG'g^{-1} = \{[gx_1g^{-1}, gy_1g^{-1}]^{\epsilon_1} \cdots [gx_rg^{-1}, gy_rg^{-1}]^{\epsilon_r} : x_i, y_i \in G, \epsilon_i = \pm 1\} \subseteq G',$$

hence $G' \lhd G$.

For every $x, y \in G$ we have $xG' \cdot yG' = xyG' = xy(y^{-1}x^{-1}yx)G' = yxG' = yG' \cdot xG'$. Thus, G/G' is abelian. If G/N is abelian then for every $x, y \in G$ we have $xN \cdot yN = yN \cdot xN$. That is, $xyN = yxN$; equivalently, $x^{-1}y^{-1}xyN = N$. Thus, for every $x, y \in G$ we have $xyx^{-1}y^{-1} \in N$. So N contains all the generators of G' and so $N \supseteq G'$. □

Lemma 7.0.5. *Let B and N be subgroups of G, $N \lhd G$.*

1. *$B \cap N$ is a normal subgroup of B.*
2. *$BN := \{bn : b \in B, n \in N\}$ is a subgroup of G. Also, NB is a subgroup of G. In fact, $BN = NB$.*
3. *If $B \lhd G$ then $BN \lhd G$ and $B \cap N \lhd G$.*
4. *If B and N are finite then $|BN| = |B||N|/|B \cap N|$. The same holds for NB.*

Proof.

(1) $B \cap N$ is a normal subgroup of B: First $B \cap N$ is a subgroup of G, hence of B. Let $b \in B$ and $n \in B \cap N$. Then $bnb^{-1} \in B$ because $b, n \in B$ and $bnb^{-1} \in N$ because $N \lhd G$.

(2) $BN := \{bn : b \in B, n \in N\}$ is a subgroup of G: Note that $ee = e \in BN$. If $bn, b'n' \in BN$ then $bnb'n' = [bb'][\{(b')^{-1}nb'\}n'] \in BN$. Finally, if $bn \in BN$ then $(bn)^{-1} = n^{-1}b^{-1} = b^{-1}[bn^{-1}b^{-1}] \in BN$.

Note that $BN = \cup_{b \in B} bN = \cup_{b \in B} Nb = NB$.

(3) If $B \lhd G$ then $BN \lhd G$: We saw that BN is a subgroup. Let $g \in G$ and $bn \in BN$ then $gbng^{-1} = [gbg^{-1}][gng^{-1}] \in BN$, using the normality of both B and N. If $x \in B \cap N, g \in G$ then $gxg^{-1} \in B$ and $gxg^{-1} \in N$, because both are normal. Thus, $gxg^{-1} \in B \cap N$, which shows $B \cap N$ is a normal subgroup of G.

(4) If B and N are finite then $|BN| = |B||N|/|B \cap N|$: Define a map of *sets*,

$$f : B \times N \longrightarrow BN, \qquad (b, n) \overset{f}{\mapsto} bn.$$

to prove the assertion it is enough to prove that every fibre $f^{-1}x$, $x \in BN$, has cardinality $|B \cap N|$.

Suppose that $x = bn$, then for every $y \in B \cap N$ we have $(by)(y^{-1}n) = bn$. This shows that $f^{-1}(x) \supseteq \{(by, y^{-1}n) : y \in B \cap N\}$, a set of $|B \cap N|$ elements. On the other hand, if $bn = b_1n_1$ then $y_1 = b_1^{-1}b = n_1n^{-1}$ and hence $y_1 \in B \cap N$. Let $y = y_1^{-1}$ then $(by)(y^{-1}n) = b_1n_1$. Thus, $f^{-1}(x) = \{(by, y^{-1}n) : y \in B \cap N\}$. [8]

□

Remark 7.0.6. In general, if B, N are subgroups of G (that are not normal) then BN need not be a subgroup of G. Indeed, consider the case of $G = S_3$, $B = \{1, (12)\}, N = \{1, (13)\}$ then $BN = \{1, (12), (13), (132)\}$

[8] Note that we do need to assume BN is a subgroup. In particular, we do not need to assume that B or N are normal.

which is not a subgroup of S_3. Thus, in general $\langle B, N\rangle \supset BN$ and equality does not hold. We can deduce though that

$$|\langle B, N\rangle| \geq \frac{|B| \cdot |N|}{|B \cap N|}.$$

This is a very useful formula. Suppose, for example, that $(|B|, |N|) = 1$ then $|B \cap N| = 1$ because $B \cap N$ is a subgroup of both B and N and so by Lagrange's theorem: $|B \cap N|$ divides both $|B|$ and $|N|$. In this case then $|\langle B, N\rangle| \geq |B| \cdot |N|$. For example, any subgroup of order 3 of A_4 generates A_4 together with the Klein group.

Simple Groups.
A group G is called simple if it has no non-trivial normal subgroups. Every group of prime order is simple. A group of odd order, which is not prime, is not simple (Theorem of Feit and Thompson). The classification of all finite simple groups is known. We shall later prove that the alternating group A_n is a simple group for $n \geq 5$.
Another family of simple groups is the following: Let $\mathbb{F}$ be a finite field and let $\mathrm{SL}_n(\mathbb{F})$ be the $n \times n$ matrices with determinant 1. It's a group. Let T be the diagonal matrices with all elements on the diagonal being equal (hence the elements of T are in bijection with solutions of $x^n = 1$ in $\mathbb{F}$); it is the center of $\mathrm{SL}_2(\mathbb{F})$. Let $\mathrm{PSL}_n(\mathbb{F}) = \mathrm{SL}_n(\mathbb{F})/T$. This is a simple group for $n \geq 2$ and any $\mathbb{F}$, the only exceptions being $n = 2$ and $\mathbb{F} \cong \mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/3\mathbb{Z}$. (See Rotman, op. cit., §8).
One can gain some understanding about the structure of a group from its normal subgroups. If $N \lhd G$ then we have a *short exact sequence*

$$1 \longrightarrow N \longrightarrow G \longrightarrow G/N \longrightarrow 1.$$

(That means that all the arrows are group homomorphisms and the image of an arrow is exactly the kernel of the next one.) Thus, might hope that the knowledge of N and G/N allows to find the properties of G. This works best when the map $G \longrightarrow G/N$ has a section, i.e., there is a homomorphism $f : G/N \longrightarrow N$ such that $\pi_N \circ f = Id$. Then G is a *semi-direct product*. We'll come back to this later in the course.

Part 2. The Isomorphism Theorems

8. Homomorphisms

8.1. **Basic definitions.** Let G and H be two groups. A *homomorphism* $f : G \longrightarrow H$ is a function satisfying $f(ab) = f(a)f(b)$. It is a consequence of the definition that $f(e_G) = e_H$ and that $f(a^{-1}) = f(a)^{-1}$.

A homomorphism is called an *isomorphism* if it is $1:1$ and surjective. In that case, the set theoretic inverse function f^{-1} is also automatically is a homomorphism. Thus, f is an isomorphism if and only if there exists a homomorphism $g : H \longrightarrow G$ such that $h \circ g = id_G, g \circ h = id_H$.

Two groups, G and H, are called *isomorphic* if there exists an isomorphism $f : G \longrightarrow H$. We use the notation $G \cong H$. For all practical purposes two isomorphic groups should be considered as the same group.

Example 8.1.1. The sign map $\mathrm{sgn} : S_n \longrightarrow \{\pm 1\}$ is a surjective group homomorphism.

Example 8.1.2. Let G be a cyclic group of order n, say $G = \langle g \rangle$. The group G is isomorphic to $\mathbb{Z}/n\mathbb{Z}$: Indeed, define a function $f : G \longrightarrow \mathbb{Z}/n\mathbb{Z}$ by $f(g^a) = a$. Note that f is well defined because if $g^a = g^b$ then $n|(b-a)$. It is a homomorphism: $g^a g^b = g^{a+b}$. It is easy to check that f is surjective. It is injective, because $f(g^a) = 0$ implies that $n|a$ and so $g^a = g^0 = e$ in the group G.

The *kernel* $\mathrm{Ker}(f)$ of a homomorphism $f : G \longrightarrow H$ is by definition the set

$$\mathrm{Ker}(f) = \{g \in G : f(g) = e_H\}.$$

For example, the kernel of the sign homomorphism $S_n \longrightarrow \{\pm 1\}$ is the alternating group A_n.

Example 8.1.3. We have an isomorphism $S_3 \cong D_6$ coming from the fact that a symmetry of a triangle (an element of D_6) is completely determined by its action on the vertices.

Example 8.1.4. The Klein V-group $\{1, (12)(34), (13)(24), (14)(23)\}$ is isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ by $(12)(34) \mapsto (0,1),\ (13)(24) \mapsto (1,0),\ (14)(23) \mapsto (1,1)$.

Lemma 8.1.5. *The set* $\mathrm{Ker}(f)$ *is a normal subgroup of* G*;* f *is injective if and only if* $\mathrm{Ker}(f) = \{e\}$*. For every* $h \in H$ *the preimage* $f^{-1}(h) := \{g \in G : f(g) = h\}$ *is a coset of* $\mathrm{Ker}(f)$*.*

Proof. First, since $f(e) = e$ we have $e \in \mathrm{Ker}(f)$. If $x, y \in \mathrm{Ker}(f)$ then $f(xy) = f(x)f(y) = ee = e$ so $xy \in \mathrm{Ker}(f)$ and $f(x^{-1}) = f(x)^{-1} = e^{-1} = e$ so $x^{-1} \in \mathrm{Ker}(f)$. That shows that $\mathrm{Ker}(f)$ is a subgroup. If $g \in G, x \in \mathrm{Ker}(f)$ then $f(gxg^{-1}) = f(g)f(x)f(g^{-1}) = f(g)ef(g)^{-1} = e$. Thus, $\mathrm{Ker}(f) \lhd G$.

If f is injective then there is a unique element x such that $f(x) = e$. Thus, $\mathrm{Ker}(f) = \{e\}$. Suppose that $\mathrm{Ker}(f) = \{e\}$ and $f(x) = f(y)$. Then $e = f(x)f(y)^{-1} = f(xy^{-1})$ so $xy^{-1} = e$. That is $x = y$ and f is injective.

More generally, note that $f(x) = f(y)$ iff $f(x^{-1}y) = e$ iff $x^{-1}y \in \mathrm{Ker}(f)$ iff $y \in x\mathrm{Ker}(f)$. Thus, if $h \in H$ and $f(x) = h$ then the fibre $f^{-1}(h)$ is precisely $x\mathrm{Ker}(f)$. □

Lemma 8.1.6. *If* $N \lhd G$ *then the canonical map* $\pi_N : G \longrightarrow G/N$*, given by* $\pi_N(a) = aN$*, is a surjective homomorphism with kernel* N*.*

Proof. We first check that $\pi = \pi_N$ is a homomorphism: $\pi(ab) = abN = aNbN = \pi(a)\pi(b)$. Since every element of G/N is of the form aN for some $a \in G$, π is surjective. Finally, $a \in \mathrm{Ker}(\pi)$ iff $\pi(a) = aN = N$ (the identity element of G/N) iff $a \in N$. □

Corollary 8.1.7. *A subgroup* $N < G$ *is normal if and only if it is the kernel of a homomorphism.*

8.2. **Behavior of subgroups under homomorphisms.** Let $f : G \longrightarrow H$ be a group homomorphism.

Proposition 8.2.1. *If $A < G$ then $f(A) < H$, in particular $f(G) < H$. If $B < H$ then $f^{-1}(B) < G$. Furthermore, if $B \triangleleft H$ then $f^{-1}(B) \triangleleft G$. If, moreover, f is surjective then $A \triangleleft G$ implies $f(A) \triangleleft H$.*

Proof. Since $f(e) = e$, $e \in f(A)$. Furthermore, the identities $f(x)f(y) = f(xy), f(x)^{-1} = f(x^{-1})$ show that $f(A)$ is closed under multiplication and inverses. Thus, $f(A)$ is a subgroup.

Let $B < H$. Since $f(e) = e$ we see that $e \in f^{-1}(B)$. Let $x, y \in f^{-1}(B)$ then $f(xy) = f(x)f(y) \in B$ because both $f(x)$ and $f(y)$ are in B. Thus, $xy \in f^{-1}(B)$. Also, $f(x^{-1}) = f(x)^{-1} \in B$ and so $x^{-1} \in f^{-1}(B)$. This shows that $f^{-1}(B) < G$.

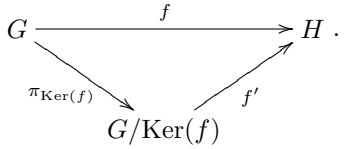
Suppose now that $B \triangleleft H$. Let $x \in f^{-1}(B), g \in G$. Then $f(gxg^{-1}) = f(g)f(x)f(g)^{-1}$. Since $f(x) \in B$ and $B \triangleleft H$ it follows that $f(g)f(x)f(g)^{-1} \in B$ and so $gxg^{-1} \in f^{-1}(B)$. Thus, $f^{-1}(B) \triangleleft G$.

The last claim follows with similar arguments. □

Remark 8.2.2. It is not necessarily true that if $A \triangleleft G$ then $f(A) \triangleleft H$. For example, consider $G = \{1, (12)\}$ with its embedding into S_3.

9. The first isomorphism theorem

Theorem 9.0.3. (The First Isomorphism Theorem) *Let $f : G \longrightarrow H$ be a homomorphism of groups. There is an injective homomorphism $f' : G/\mathrm{Ker}(f) \longrightarrow H$ such that the following diagram commutes:*



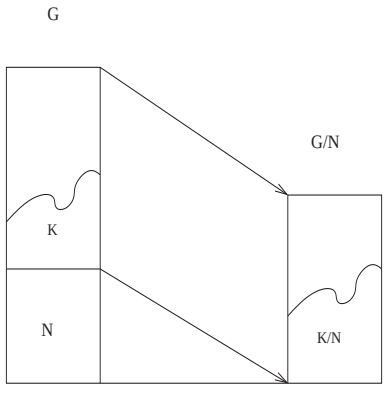
In particular, $G/\mathrm{Ker}(f) \cong f(G)$.

Proof. Let $N = \mathrm{Ker}(f)$. We define f' by

$$f'(aN) = f(a).$$

The map f' is well defined: if $aN = bN$ then $a = bn$ for some $n \in N$. Then $f'(aN) = f(a) = f(bn) = f(b)f(n) = f(b) = f'(bN)$. Therefore, f' is well defined. Now $f'(aNbN) = f'(abN) = f(ab) = f(a)f(b) = f'(aN)f'(bN)$, which shows f' is a homomorphism. If $f'(aN) = f(a) = e$ then $a \in N$ and so $aN = N$. That is, f' is injective and surjective onto its image. We conclude that $f' : G/N \longrightarrow f(G)$ is an isomorphism.

Finally, $f'(\pi_N(a)) = f'(aN) = f(a)$ so $f' \circ \pi_N = f$. Therefore, the diagram commutes. □



Example 9.0.4. Let us construct two homomorphisms

$$f_i : D_8 \longrightarrow S_2.$$

We get the first homomorphism f_1 be looking at the action of the symmetries on the axes $\{a, b\}$. Thus, $f_1(x) = (ab), f_1(y) = 1$ (x permutes the axes, while y fixes the axes – though not point-wise). Similarly,

if we let A, B be the lines whose equation is $a = b$ and $a = -b$, then D_8 acts as permutations on $\{A, B\}$ and we get a homomorphism $f_2 : D_8 \longrightarrow S_2$ such that $f_2(x) = (AB), f_2(y) = (AB)$.

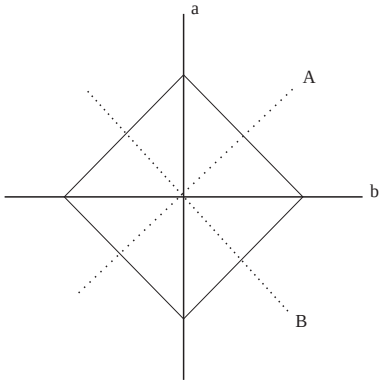
The homomorphism f_i is surjective and therefore the kernel $N_i = \mathrm{Ker}(f_i)$ has 4 elements. We find that $N_1 = \{1, x^2, y, x^2y\}$ and $N_2 = \{1, x^2, xy, x^3y\}$. By the first isomorphism theorem we have $D_8/N_i \cong S_2$.

Now, quite generally, if $g_i : G \longrightarrow H_i$ are group homomorphisms then $g : G \longrightarrow H_1 \times H_2$, defined by $g(r) = (g_1(r), g_2(r))$ is a group homomorphism with kernel $\mathrm{Ker}(g_1) \cap \mathrm{Ker}(g_2)$. One uses the notation $g = (g_1, g_2)$. Applying this to our situation, we get a homomorphism

$$f = (f_1, f_2) : D_8 \longrightarrow S_2 \times S_2,$$

whose kernel is $\{1, x^2\}$. It follows that the image of f has 4 elements and hence f is surjective. That is,

$$D_8/\langle x^2 \rangle \cong S_2 \times S_2.$$



10. The second isomorphism theorem

Theorem 10.0.5. *Let G be a group. Let $B < G, N \triangleleft G$. Then*

$$BN/N \cong B/(B \cap N).$$

Proof. Recall from Lemma 7.0.5 that $B \cap N \triangleleft B$. We define a function

$$f : BN \longrightarrow B/B \cap N, \qquad f(bn) = b \cdot B \cap N.$$

We need first to show it is well defined. Recall from the proof of Lemma 7.0.5 that if $bn = b'n'$ then $b' = by$ for some $y \in B \cap N$. Therefore, $b \cdot B \cap N = by \cdot B \cap N = b' \cdot B \cap N$ and f is well defined.

We show now that f is a homomorphism. Note that $(bn)(b_1n_1) = (bb_1)(b_1^{-1}nb_1)n_1$ and so $f(bn \cdot b_1n_1) = bb_1 \cdot B \cap N = b \cdot B \cap N \cdot b_1 \cdot B \cap N = f(b)f(b_1)$, which shows f is a homomorphism. Moreover, f is surjective: $b \cdot B \cap N = f(b)$.

The kernel of f is $\{bn : f(b) = e, b \in B, n \in N\} = \{bn : b \in B \cap N, b \in B, n \in N\} = (B \cap N)N = N$. By the First Isomorphism Theorem $BN/N \cong B/B \cap N$. □

Remark 10.0.6. This is often used as follows: Let $f : G \longrightarrow H$ be a group homomorphism with kernel N. Let $B < G$. What can we say about the image of B under f? Well $f(B) = f(BN)$ and $f : BN \longrightarrow H$ has kernel N. We conclude that $f(B) \cong BN/N \cong B/(B \cap N)$.

In fact, this idea gives another proof of the theorem. Consider the homomorphism $\pi : G \longrightarrow G/N$. Its restriction to BN is a homomorphism with kernel N and so, by the First Isomorphism Theorem, $f(BN) \cong BN/N$. The restriction of f to B is also a group homomorphism with kernel $B \cap N$. Thus, $f(B) \cong B/(B \cap N)$. But, $f(B) = f(BN)$ and we are done.

11. The third isomorphism theorem

Theorem 11.0.7. *Let $f : G \longrightarrow H$ be a surjective homomorphism of groups.*

(1) *f induces a bijection:*

$$\{\text{subgps of } G \text{ containing } \mathrm{Ker}(f)\} \quad \leftrightarrow \quad \{\text{subgps of } H\}.$$

Given by $G_1 \mapsto f(G_1)$, $G_1 < G$, and in the other direction by $H_1 \mapsto f^{-1}(H_1)$, $H_1 < H$.

(2) *Suppose that $\mathrm{Ker}(f) < G_1 < G_2$. Then $G_1 \lhd G_2$ if and only if $f(G_1) \lhd f(G_2)$. Moreover, in that case,*

$$G_2/G_1 \cong f(G_2)/f(G_1).$$

(3) *Let $N < K < G$ be groups, such that $N \lhd G, K \lhd G$. Then*

$$(G/N)/(K/N) \cong G/K.$$

Proof. We proved in general (Prop. 8.2.1) that if $G_1 < G$ then $f(G_1) < H$ and if $H_1 < H$ then $f^{-1}(H_1) < G$. Since f is a surjective map we have $f(f^{-1}(H_1)) = H_1$. We need to show that if $\mathrm{Ker}(f) < G_1$ then $f^{-1}(f(G_1)) = G_1$. Clearly $f^{-1}(f(G_1)) \supseteq G_1$. Let $x \in f^{-1}(f(G_1))$ then $f(x) \in f(G_1)$. Choose then $g \in G_1$ such that $f(g_1) = f(x)$ and write $x = g(g^{-1}x)$. Note that $f(g^{-1}x) = e_H$ and so $g^{-1}x \in \mathrm{Ker}(f) \subseteq G_1$. Thus, $x = g(g^{-1}x) \in G_1$.

Consider the restriction of f to G_2 as a surjective group homomorphism $f : G_2 \longrightarrow f(G_2)$. We proved under those conditions that if $G_1 \lhd G_2$ then $f(G_1) \lhd f(G_2)$. If $f(G_1) \lhd f(G_2)$ then we also proved that $f^{-1}(f(G_1)) \lhd G_2$. Since $G_1 \supset \mathrm{Ker}(f)$ we have $f^{-1}(f(G_1)) = G_1$.

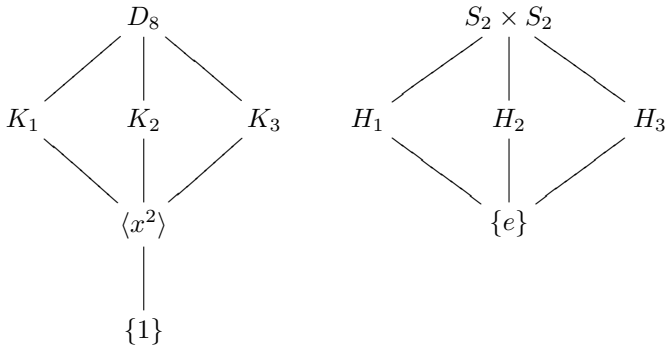
It remains to show that if $\mathrm{Ker}(f) < G_1 \lhd G_2$ then $G_2/G_1 \cong f(G_2)/f(G_1)$. The homomorphism obtained by composition

$$G_2 \longrightarrow f(G_2) \longrightarrow f(G_2)/f(G_1),$$

is surjective and has kernel $f^{-1}(f(G_1)) = G_1$. The claim now follows from the First Isomorphism Theorem.

We apply the previous results in the case where $H = G/N$ and $f : G \longrightarrow G/N$ is the canonical map. We consider the case $G_1 = K, G_2 = G$. Then $G/K \cong f(G)/f(K) = (G/N)/(K/N)$. □

Example 11.0.8. Consider again the group homomorphism $f : D_8 \longrightarrow S_2 \times S_2$ constructed in Example 9.0.4. Using the third isomorphism theorem we conclude that the graph of the subgroups of D_8 containing $\langle x^2 \rangle$ is exactly that of $S_2 \times S_2$ (analyzed in Example 2.6.1). Hence we have:



We'll see later that this does not exhaust the list of subgroups of D_8. Here we have
$K_1 = \langle x \rangle$,
$K_2 = \langle y, x^2 \rangle$,
$K_3 = \langle xy, x^2 \rangle$
and
$H_1 = f(K_1) = \{(1,1), ((ab), (AB))\}$,
$H_2 = f(K_2) = \{(1,1), (1, (AB))\}$,
$H_3 = f(K_3) = \{(1,1), ((ab), 1)\}$.

Example 11.0.9. Let $\mathbb{F}$ be a field and let $N = \{\mathrm{diag}[f, f, \ldots, f] : f \in \mathbb{F}^\times\}$ be the set of diagonal matrices with the same non-zero element in each diagonal entry. We proved in an assignment that $N = Z(\mathrm{GL}_n(\mathbb{F}))$ and is therefore a normal subgroup. The quotient group

$$\mathrm{PGL}_n(\mathbb{F}) := \mathrm{GL}_n(\mathbb{F})/N$$

is called the projective linear group.

Let $\mathbb{P}^{n-1}(\mathbb{F})$ be the set of equivalence classes of non-zero vectors in $\mathbb{F}^n$ under the equivalence $v \sim w$ if there is $f \in \mathbb{F}^*$ such that $fv = w$; that is, the set of lines through the origin. The importance of the group $\mathrm{PGL}_n(\mathbb{F})$ is that it acts as automorphisms on the projective $n-1$-space $\mathbb{P}^{n-1}(\mathbb{F})$.

Let

$$\pi : \mathrm{GL}_n(\mathbb{F}) \longrightarrow \mathrm{PGL}_n(\mathbb{F})$$

be the canonical homomorphism. The function

$$\det : \mathrm{GL}_n(\mathbb{F}) \longrightarrow \mathbb{F}^*$$

is a group homomorphism, whose kernel, the matrices with determinant one, is denoted $\mathrm{SL}_n(\mathbb{F})$. Consider the image of $\mathrm{SL}_n(\mathbb{F})$ in $\mathrm{PGL}_n(\mathbb{F})$; it is denoted $\mathrm{PSL}_n(\mathbb{F})$. We want to analyze it and the quotient $\mathrm{PGL}_n(\mathbb{F})/\mathrm{PSL}_n(\mathbb{F})$.

The group $\mathrm{PSL}_n(\mathbb{F})$ is equal to $\pi(\mathrm{SL}_n(\mathbb{F})) = \pi(\mathrm{SL}_n(\mathbb{F})N)$ and is therefore isomorphic to $\mathrm{SL}_n(\mathbb{F})N/N \cong \mathrm{SL}_n(\mathbb{F})/\mathrm{SL}_n(\mathbb{F}) \cap N = \mathrm{SL}_n(\mathbb{F})/\mu_n(\mathbb{F})$, where by $\mu_N(\mathbb{F})$ we mean the group $\{f \in \mathbb{F}^\times : f^n = 1\}$ (where we identify f with $\mathrm{diag}[f, f, \ldots, f]$). Therefore,

$$\mathrm{PSL}_n(\mathbb{F}) \cong \mathrm{SL}_n(\mathbb{F})/\mu_n(\mathbb{F}).$$

We have $\mathrm{PGL}_n(\mathbb{F})/\mathrm{PSL}_n(\mathbb{F}) \cong (\mathrm{GL}_n(\mathbb{F})/N)/(\mathrm{SL}_n(\mathbb{F})N/N) \cong \mathrm{GL}_n(\mathbb{F})/\mathrm{SL}_n(\mathbb{F})N$. Let $\mathbb{F}^{\times(n)}$ be the subgroup of $\mathbb{F}^\times$ consisting of the elements $\{f^n : f \in \mathbb{F}^\times\}$. Under the isomorphism $\mathrm{GL}_n(\mathbb{F})/\mathrm{SL}_n(\mathbb{F}) \cong \mathbb{F}^\times$ the subgroup $\mathrm{SL}_n(\mathbb{F})N$ corresponds to $\mathbb{F}^{\times(n)}$. We conclude that

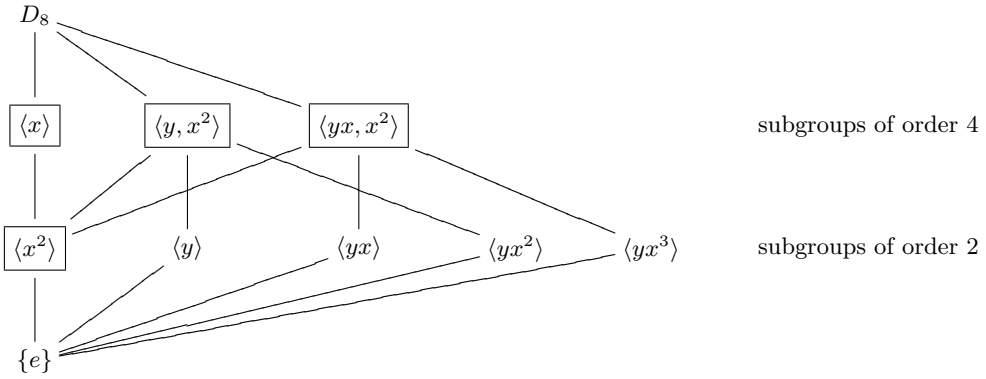
$$\mathrm{PGL}_n(\mathbb{F})/\mathrm{PSL}_n(\mathbb{F}) \cong \mathbb{F}^\times/\mathbb{F}^{\times(n)}.$$

12. The lattice of subgroups of a group

Let G be a group. Consider the set $\Lambda(G)$ of all subgroups of G. Define an order on this set by $A \leq B$ if A is a subgroup of B. This relation is transitive and $A \leq B \leq A$ implies $A = B$. That is, the relation is really an order.

The set $\Lambda(G)$ is a lattice. Every two elements A, B have a minimum $A \cap B$ (that is if $C \leq A, C \leq B$ then $C \leq A \cap B$) and a maximum $\langle A, B\rangle$ - the subgroup generated by A and B (that is $C \geq A, C \geq B$ then $C \geq \langle A, B\rangle$). If $A \in \Lambda(G)$ then let $\Lambda_A(G)$ to be the set of all elements in $\Lambda(G)$ that are greater or equal to A. It is a lattice in its own right. We have the property that $\boxed{\text{if } N \lhd G \text{ then } \Lambda_N(G) \cong \Lambda(G/N) \text{ as lattices}}$ – This is the Third Isomorphism Theorem.

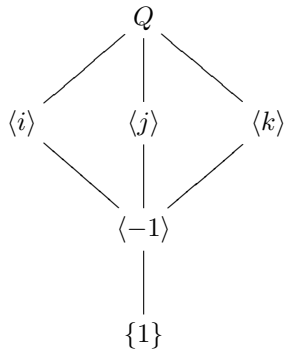
Here is the lattice of subgroups of D_8. Normal subgroup are boxed.



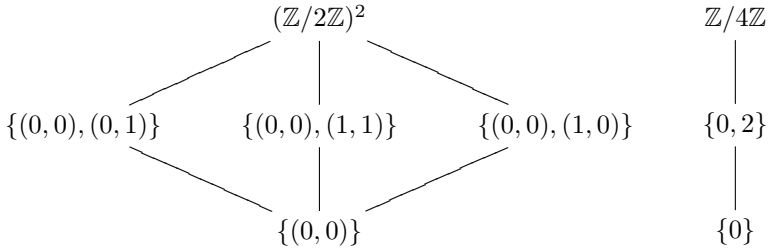
How to prove that these are all the subgroups? Note that every proper subgroup has order 2 or 4 by Lagrange's theorem. If it is cyclic then it must be one of the above, because the diagram certainly contains all cyclic subgroups. Else, it can only be of order 4 and every element different from e has order 2. It is east to verify that any two distinct elements of order 2 generate one of the subgroups we have listed.

There are at least two ways in which one uses this concept:

- To examine whether two groups can be isomorphic. Isomorphic groups have isomorphic lattices of subgroups. For example, the groups D_8 and Q both have 8 elements. The lattice of subgroups of Q is



  We conclude that Q and D_8 are not isomorphic.
- To recognize quotients. Consider for example $D_8/\langle x^2 \rangle$. This is a group of 4 elements. Let us give ourselves that there are only two groups of order 4 up to isomorphism and those are $(\mathbb{Z}/2\mathbb{Z})^2$ and $\mathbb{Z}/4\mathbb{Z}$. The lattice of subgroups for them are



  We conclude that $D_8/\langle x^2 \rangle \cong (\mathbb{Z}/2\mathbb{Z})^2$.

Part 3. Group Actions on Sets

13. Basic definitions

Let G be a group and let S be a non-empty set. We say that G *acts on* S if we are given a function

$$G \times S \longrightarrow S, \quad (g, s) \longmapsto g \star s,$$

such that;

(i) $e \star s = s$ for all $s \in S$;

(ii) $(g_1 g_2) \star s = g_1 \star (g_2 \star s)$ for all $g_1, g_2 \in G$ and $s \in S$.

Given an action of G on S we can define the following sets. Let $s \in S$. Define the *orbit* of s

$$\mathrm{Orb}(s) = \{g \star s : g \in G\}.$$

Note that $\mathrm{Orb}(s)$ is a subset of S, equal to all the images of the element s under the action of the elements of the group G. We also define the *stabilizer* of s to be

$$\mathrm{Stab}(s) = \{g \in G : g \star s = s\}.$$

Note that $\mathrm{Stab}(s)$ is a subset of G. In fact, it is a subgroup, as the next Lemma states.

One should think of every element of the group as becoming a symmetry of the set S. We'll make more precise later. For now, we just note that every element $g \in G$ defines a function $S \longrightarrow S$ by $s \mapsto gs$. This function, we'll see later, is bijective.

14. Basic properties

Lemma 14.0.10. (1) *Let $s_1, s_2 \in S$. We say that s_1 is related to s_2, i.e., $s_1 \sim s_2$, if there exists $g \in G$ such that*

$$g \star s_1 = s_2.$$

This is an equivalence relation. The equivalence class of s_1 is its orbit $Orb(s_1)$.

(2) *Let $s \in S$. The set $Stab(s)$ is a subgroup of G.*

(3) *Suppose that both G and S have finitely many elements. Then*

$$|Orb(s)| = \frac{|G|}{|Stab(s)|}.$$

Proof. (1) We need to show reflexive, symmetric and transitive. First, we have $e \star s = s$ and hence $s \sim s$, meaning the relation is reflexive. Second, if $s_1 \sim s_2$ then for a suitable $g \in G$ we have $g \star s_1 = s_2$. Therefore

$$\begin{aligned} & g^{-1} \star (g \star s_1) = g^{-1} \star s_2 \\ \Rightarrow \quad & (g^{-1} g) \star s_1 = g^{-1} \star s_2 \\ \Rightarrow \quad & e \star s_1 = g^{-1} \star s_2 \\ \Rightarrow \quad & s_1 = g^{-1} \star s_2 \\ \Rightarrow \quad & g^{-1} \star s_2 = s_1 \\ \Rightarrow \quad & s_2 \sim s_1. \end{aligned}$$

It remains to show transitive. If $s_1 \sim s_2$ and $s_2 \sim s_3$ then for suitable $g_1, g_2 \in G$ we have

$$g_1 \star s_1 = s_2, \quad g_2 \star s_2 = s_3.$$

Therefore,

$$(g_2 g_1) \star s_1 = g_2 \star (g_1 \star s_1) = g_2 \star s_2 = s_3,$$

and hence $s_1 \sim s_3$.

Moreover, by the very definition the equivalence class of an element s_1 of S is all the elements of the form $g \star s_1$ for some $g \in G$, namely, $\mathrm{Orb}(s_1)$.

(2) Let $H = \mathrm{Stab}(s)$. We have to show that: (i) $e \in H$; (2) If $g_1, g_2 \in H$ then $g_1 g_2 \in H$; (iii) If $g \in H$ then $g^{-1} \in H$.

First, by definition of group action we have

$$e \star s = s.$$

Therefore $e \in H$. Next suppose that $g_1, g_2 \in H$, i.e.,

$$g_1 \star s = s, \quad g_2 \star s = s.$$

Then

$$(g_1 g_2) \star s = g_1 \star (g_2 \star s) = g_1 \star s = s.$$

Thus, $g_1 g_2 \in H$. Finally, if $g \in H$ then $g \star s = s$ and so

$$\begin{aligned} & g^{-1} \star (g \star s) = g^{-1} \star s \\ \Rightarrow \quad & (g^{-1} g) \star s = g^{-1} \star s \\ \Rightarrow \quad & e \star s = g^{-1} \star s \\ \Rightarrow \quad & s = g^{-1} \star s, \end{aligned}$$

and therefore $g^{-1} \in H$.

(3) We claim that there exists a bijection between the left cosets of H and the orbit of s. If we show that, then by Lagrange's theorem,

$$|\mathrm{Orb}(s)| = \text{no. of left cosets of } H = \text{index of } H = |G|/|H|.$$

Define a function

$$\{\text{left cosets of } H\} \xrightarrow{\phi} \mathrm{Orb}(s),$$

by

$$\phi(gH) = g \star s.$$

We claim that ϕ is a well defined bijection. First

Well-defined: Suppose that $g_1 H = g_2 H$. We need to show that the rule ϕ would give the same result whether we take the representative g_1 or the representative g_2 to the coset, that is, we need to show

$$g_1 \star s = g_2 \star s.$$

Note that $g_1^{-1} g_2 \in H$, i.e., $(g_1^{-1} g_2) \star s = s$. We get

$$\begin{aligned} g_1 \star s &= g_1 \star ((g_1^{-1} g_2) \star s) \\ &= (g_1 (g_1^{-1} g_2)) \star s \\ &= g_2 \star s. \end{aligned}$$

ϕ is surjective: Let $t \in \mathrm{Orb}(s)$ then $t = g \star s$ for some $g \in G$. Thus,

$$\phi(gH) = g \star s = t,$$

and we get that ϕ is surjective.

ϕ is injective: Suppose that $\phi(g_1 H) = \phi(g_2 H)$. We need to show that $g_1 H = g_2 H$. Indeed,

$$\begin{aligned} & \phi(g_1 H) = \phi(g_2 H) \\ \Rightarrow \quad & g_1 \star s = g_2 \star s \\ \Rightarrow \quad & g_2^{-1} \star (g_1 \star s) = g_2^{-1} \star (g_2 \star s) \\ \Rightarrow \quad & (g_2^{-1} g_1) \star s = (g_2^{-1} g_2) \star s \\ \Rightarrow \quad & (g_2^{-1} g_1) \star s = e \star s \\ \Rightarrow \quad & (g_2^{-1} g_1) \star s = s \\ \Rightarrow \quad & g_2^{-1} g_1 \in \mathrm{Stab}(s) = H \\ \Rightarrow \quad & g_1 H = g_2 H. \end{aligned}$$

□

Corollary 14.0.11. *The set S is a disjoint union of orbits.*

Proof. The orbits are the equivalence classes of the equivalence relation $\sim$ defined in Lemma 14.0.10. Any equivalence relation partitions the set into disjoint equivalence classes. □

We have in fact seen that every orbit is in bijection with the cosets of some group. If H is any subgroup of G let us use the notation G/H for its cosets (note though that if H is not normal this is not a group, but just a set). We saw that if $s \in S$ then there is a natural bijection $G/Stab(s) \leftrightarrow Orb(s)$. Thus, the picture of S is as follows

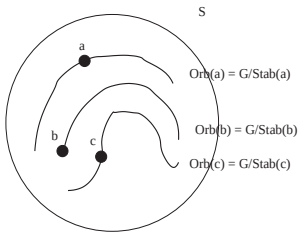


FIGURE 14.1. The set decomposes into orbits; each is the cosets of a subgroup.

15. Some examples

Example 15.0.12. The group S_n acts on the set $\{1, 2, \dots, n\}$. The action is transitive, i.e., there is only one orbit. The stabilizer of i is $S_{\{1,2,\dots,i-1,i+1,\dots,n\}} \cong S_{n-1}$.

Example 15.0.13. The group $\mathrm{GL}_n(\mathbb{F})$ acts on $\mathbb{F}^n$, and also $\mathbb{F}^n - \{0\}$. The action is transitive on $\mathbb{F}^n - \{0\}$ and has two orbits on $\mathbb{F}^n$. The stabilizer of 0 is of course $\mathrm{GL}_n(\mathbb{F})$; the stabilizer of a non-zero vector v_1 can be written in a basis $w_1, w_2, \dots, w_n$ with $w_1 = v_1$ as the matrices of the shape

$$\begin{pmatrix} 1 & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \dots & \vdots \\ 0 & * & \dots & * \end{pmatrix}.$$

Example 15.0.14. Let H be a subgroup of G then we have an action

$$H \times G \longrightarrow G, \qquad (h, g) \mapsto hg.$$

In this example, H is "the group" and G is "the set". Here the orbits are right cosets of H and the stabilizers are trivial. Since $G = \coprod Orb(g_i) = \coprod Hg_i$ we conclude that $|G| = \sum_i |Orb(g_i)| = \sum_i |H|/|Stab(g_i)| = \sum_i |H|$ and therefore that $|H| \mid |G|$ and that $[G : H]$, the number of cosets, is $|G|/|H|$. We have recovered Lagrange's theorem.

Example 15.0.15. Let H be a subgroup of G. Let $S = \{gH : g \in G\}$ be the set of left cosets of H in G. Then we have an action

$$G \times S \longrightarrow S, \qquad (a, bH) \mapsto abH.$$

Here there is a unique orbit (we say G acts *transitively*). The stabilizer of gH is the subgroup gHg^{-1}.

Example 15.0.16. Let $G = \mathbb{R}/2\pi\mathbb{Z}$. It acts on the sphere by rotations: an element $\theta \in G$ rotates the sphere by angle θ around the north-south axes. The orbits are latitude lines and the stabilizers of every point is trivial, except for the poles whose stabilizer is G. See Figure 15.1.

Example 15.0.17. Let G be the dihedral group D_{16}. Recall that G is the group of symmetries of a regular octagon in the plane.

$$G = \{e, x, x^2, \dots, x^7, y, yx, yx^2, \dots, yx^7\},$$

where x is the rotation clockwise by angle $2\pi/8$ and y is the reflection through the y-axis. We have the relations

$$x^8 = y^2 = e, \quad yxy = x^{-1}.$$

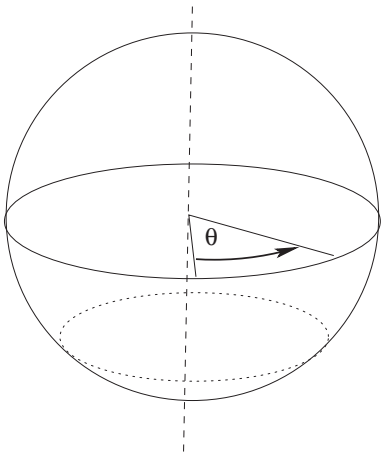


FIGURE 15.1. Action on the sphere by rotation.

We let S be the set of colorings of the octagon (= necklaces laid on the table) having 4 red vertices (rubies) and 4 green vertices (sapphires). The group G acts on S by its action on the octagon.

For example, the coloring s_0 in Figure 15.2 is certainly preserved under x^2 and under y. Therefore, the stabilizer of s_0 contains at least the set of eight elements

$$\{e, x^2, x^4, x^6, y, yx^2, yx^4, yx^6\}. \tag{15.1}$$

Remember that the stabilizer is a subgroup and, by Lagrange's theorem, of order dividing $16 = |G|$. On

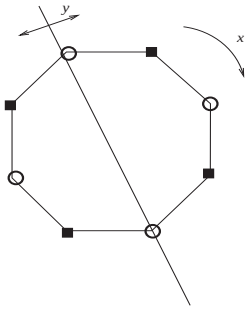


FIGURE 15.2. A necklace with 4 rubies and 4 sapphires.

the other hand, $\mathrm{Stab}(s_0) \neq G$ because $x \notin \mathrm{Stab}(s_0)$. It follows that the stabilizer has exactly 8 elements and is equal to the set in (15.1).

According to Lemma 14.0.10 the orbit of s_0 is in bijection with the left cosets of $\mathrm{Stab}(s_0) = \{e, x^2, x^4, x^6, y, yx^2, yx^4, yx^6\}$. By Lagrange's theorem there are two cosets. For example, $\mathrm{Stab}(s_0)$ and $x\mathrm{Stab}(s_0)$ are distinct cosets. The proof of Lemma 14.0.10 tells us how to find the orbit: it is the set

$$\{s_0, xs_0\},$$

portrayed in Figure 15.3.

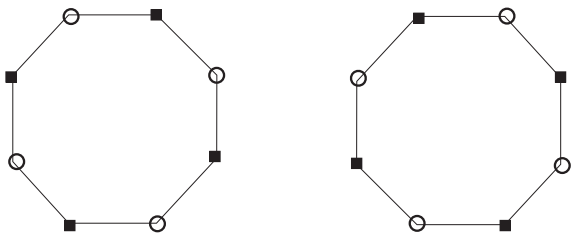


FIGURE 15.3. The orbit of the necklace.

16. Cayley's theorem

Theorem 16.0.18. *Every finite group of order n is isomorphic to a subgroup of S_n.*

We first prove a lemma that puts group actions in a different context. Let A be a finite set. Let Σ_A be the set of bijective functions $A \longrightarrow A$. Then, Σ_A is a group. In fact, if we let $s_1, \dots, s_n$ be the elements of A, we can identify bijective functions $A \longrightarrow A$ with permutations of $\{1, \dots, n\}$ and we see that $\Sigma_A \cong S_n$.

Lemma 16.0.19. *To give an action of a group G on a set A is equivalent to giving a homomorphism $G \longrightarrow \Sigma_A$. The kernel of this homomorphism is $\cap_{a \in A} Stab(a)$.*

Proof. An element g define a function $\phi_g : A \longrightarrow A$ by $\phi_g(a) = ga$. We have ϕ_e being the identity function. Note that $\phi_h\phi_g(a) = \phi_h(ga) = hga = \phi_{hg}(a)$ for every a and hence $\phi_h\phi_g = \phi_{hg}$. In particular, $\phi_g\phi_{g^{-1}} = \phi_{g^{-1}}\phi_g = Id$. This shows that every ϕ_g is a bijection and the map

$$\Psi : G \longrightarrow \Sigma_A, \qquad g \overset{\Psi}{\mapsto} \phi_g,$$

is a homomorphism. (Conversely, given such a homomorphism Ψ, define a group action by $g \star a := \Psi(g)(a)$.)

The kernel of this homomorphism is the elements g such that ϕ_g is the identity, i.e., $\phi_g(a) = a$ for all $a \in A$. That is, $g \in Stab(a)$ for every $a \in A$. The set of such elements g is just $\cap_{a \in A} \mathrm{Stab}(a)$. □

Proof. (*of Theorem*) Consider the action of G on itself by multiplication (Example 15.0.14), $(g, g') \mapsto gg'$. Recall that all stabilizers are trivial. Thus this action gives an injective homomorphism

$$G \longrightarrow \Sigma_G \cong S_n,$$

where $n = |G|$. □

16.1. **Applications to construction of normal subgroups.** Let G be a group and H a subgroup of finite index n. Consider the action of G on the set of cosets G/H of H and the resulting homomorphism

$$\Psi : G \longrightarrow \Sigma_{G/H} \cong \Sigma_n,$$

where $n = [G : H]$. The kernel K of Ψ is

$$\cap_{a \in G/H} \mathrm{Stab}(a) = \cap_{g \in G} \mathrm{Stab}(gH) = \cap_{g \in G} gHg^{-1}.$$

Being a kernel of a homomorphism, K is normal in G and is contained in H. Furthermore, since the resulting homomorphism $G/K \longrightarrow S_n$ is injective we get that $|G/K| = [G : K]$ divides $[G : H]! = |S_n|$. In particular, we conclude that every subgroup H of G contains a subgroup K which is normal in G and of index at most $[G : H]!$. Thus, for example, a simple infinite group has no subgroups of finite index.

In fact, the formula $K = \cap_{g \in G} gHg^{-1}$ shows that K is the maximal subgroup of H which is normal in G. Indeed, if $K' \lhd G, K' < H$ then $K = gKg^{-1} \subset gHg^{-1}$ and we see that $K' \subseteq K$.

17. The Cauchy-Frobenius formula

17.1. **A formula for the number of orbits.**

Theorem 17.1.1. (CFF) *Let G be a finite group acting on a finite set S. Let N be the number of orbits of G in S. Define*

$$I(g) = |\{s \in S : g \star s = s\}|$$

(the number of elements of S fixed by the action of g). Then

$$N = \frac{1}{|G|} \sum_{g \in G} I(g). \tag{17.1}$$

Remark 17.1.2. If $N = 1$ we say that G acts *transitively* on S. It means exactly that: For every $s_1, s_2 \in S$ there exists $g \in G$ such that $g \star s_1 = s_2$.

Proof. We define a function

$$T : G \times S \longrightarrow \{0, 1\}, \quad T(g, s) = \begin{cases} 1 & g \star s = s \\ 0 & g \star s \neq s \end{cases}.$$

Note that for a fixed $g \in G$ we have

$$I(g) = \sum_{s \in S} T(g, s),$$

and that for a fixed $s \in S$ we have

$$|\mathrm{Stab}(s)| = \sum_{g \in G} T(g, s).$$

Let us fix representatives $s_1, \ldots, s_N$ for the N disjoint orbits of G in S. Now,

$$\begin{aligned} \sum_{g \in G} I(g) &= \sum_{g \in G} \left(\sum_{s \in S} T(g, s) \right) = \sum_{s \in S} \left(\sum_{g \in G} T(g, s) \right) \\ &= \sum_{s \in S} |\mathrm{Stab}(s)| = \sum_{s \in S} \frac{|G|}{|\mathrm{Orb}(s)|} \\ &= \sum_{i=1}^{N} \sum_{s \in \mathrm{Orb}(s_i)} \frac{|G|}{|\mathrm{Orb}(s)|} = \sum_{i=1}^{N} \sum_{s \in \mathrm{Orb}(s_i)} \frac{|G|}{|\mathrm{Orb}(s_i)|} \\ &= \sum_{i=1}^{N} \frac{|G|}{|\mathrm{Orb}(s_i)|} \cdot |\mathrm{Orb}(s_i)| = \sum_{i=1}^{N} |G| \\ &= N \cdot |G|. \end{aligned}$$

□

Corollary 17.1.3. *Let G be a finite group acting transitively on a finite S. Suppose that $|S| > 1$. Then there exists $g \in G$ without fixed points.*

Proof. By contradiction. Suppose that every $g \in G$ has a fixed point in S. That is, suppose that for every $g \in G$ we have

$$I(g) \geq 1.$$

Since $I(e) = |S| > 1$ we have that

$$\sum_{g \in G} I(g) > |G|.$$

By Cauchy-Frobenius formula, the number of orbits N is greater than 1. Contradiction. □

17.2. **Applications to combinatorics.**

Example 17.2.1. How many roulettes with 11 wedges painted 2 blue, 2 green and 7 red are there when we allow rotations?

Let S be the set of painted roulettes. Let us enumerate the sectors of a roulette by the numbers $1, \ldots, 11$. The set S is a set of $\binom{11}{2}\binom{9}{2} = 1980$ elements (choose which 2 are blue, and then choose out of the nine left which 2 are green).

Let G be the group $\mathbb{Z}/11\mathbb{Z}$. It acts on S by rotations. The element 1 rotates a painted roulette by angle $2\pi/11$ anti-clockwise. The element n rotates a painted roulette by angle $2n\pi/11$ anti-clockwise. We are interested in N – the number of orbits for this action. We use **CFF**.

The identity element always fixes the whole set. Thus $I(0) = 1980$. We claim that if $1 \leq i \leq 10$ then i doesn't fix any element of S. Indeed, suppose that $1 \leq i \leq 10$ and i fixes s. Then so does $\langle i \rangle = \mathbb{Z}/11\mathbb{Z}$ (the stabilizer is a subgroup). But any coloring fixed under rotation by 1 must be single colored! Contradiction.

Applying **CFF** we get

$$N = \frac{1}{11}\sum_{n=0}^{10} I(n) = \frac{1}{11}\cdot 1980 = 180.$$

Example 17.2.2. How many roulettes with 12 wedges painted 2 blue, 2 green and 8 red are there when we allow rotations?

Let S be the set of painted roulettes. Let us enumerate the sectors of a roulette by the numbers $1,\dots,12$. The set S is a set of $\binom{12}{2}\binom{10}{2} = 2970$ elements (choose which 2 are blue, and then choose out of the ten left which 2 are green).

Let G be the group $\mathbb{Z}/12\mathbb{Z}$. It acts on S by rotations. The element 1 rotates a painted roulette by angle $2\pi/12$ anti-clockwise. The element n rotates a painted roulette by angle $2n\pi/12$ anti-clockwise. We are interested in N – the number of orbits for this action. We use **CFF**.

The identity element always fixes the whole set. Thus $I(0) = 2970$. We claim that if $1 \le i \le 11$ and $i \ne 6$ then i doesn't fix any element of S. Indeed, suppose that i fixes a painted roulette. Say in that roulette the r-th sector is blue. Then so must be the $i+r$ sector (because the r-th sector goes under the action of i to the $r+i$-th sector). Therefore so must be the $r+2i$ sector. But there are only 2 blue sectors! The only possibility is that the $r+2i$ sector is the same as the r sector, namely, $i = 6$.

If i is equal to 6 and we enumerate the sectors of a roulette by the numbers $1,\dots,12$ we may write i as the permutation

$$(1\ 7)(2\ 8)(3\ 9)(4\ 10)(5\ 11)(6\ 12).$$

In any coloring fixed by $i = 6$ the colors of the pairs $(1\ 7), (2\ 8), (3\ 9), (4\ 10), (5\ 11)$ and $(6\ 12)$ must be the same. We may choose one pair for blue, one pair for green. The rest would be red. Thus there are $30 = 6\cdot 5$ possible choices. We summarize:

element g	$I(g)$
0	2970
$i \ne 6$	0
$i = 6$	30

Applying **CFF** we get that there are

$$N = \frac{1}{12}(2970 + 30) = 250$$

different roulettes.

Example 17.2.3. In this example S is the set of necklaces made of four rubies and four sapphires laid on the table. We ask how many necklaces there are when we allow rotations and flipping-over.

We may talk of S as the colorings of a regular octagon, four vertices are green and four are red. The group $G = D_{16}$ acts on S and we are interested in the number of orbits for the group G.

The results are the following

element g	$I(g)$
e	70
x, x^3, x^5, x^7	0
x^2, x^6	2
x^4	6
yx^i for $i = 0,\dots,7$	6

We explain how the entries in the table are obtained:

The identity always fixes the whole set S. The number of elements in S is $\binom{8}{4} = 70$ (chossing which 4 would be green).

The element x cannot fix any coloring, because any coloring fixed by x must have all sections of the same color (because $x = (1\ 2\ 3\ 4\ 5\ 6\ 7\ 8)$). If x^r fixes a coloring s_0 so does $(x^r)^r = x^{(r^2)}$ because the

stabilizer is a subgroup. Apply that for $r = 3, 5, 7$ to see that if x^r fixes a coloring so does x , which is impossible. [9]

Now, x^2 written as a permutation is $(1\ 3\ 5\ 7)(2\ 4\ 6\ 8)$. We see that if, say 1 is green so are $3, 5, 7$ and the rest must be red. That is, all the freedom we have is to choose whether the cycle $(1\ 3\ 5\ 7)$ is green or red. This gives us two colorings fixed by x^2. The same rational applies to $x^6 = (8\ 6\ 4\ \ 2)(7\ 5\ 3\ 1)$.

Consider now x^4. It may written in permutation notation as $(1\ 5)(2\ 6)(3\ 7)(4\ 8)$. In any coloring fixed by x^4 each of the cycles $(1\ 5)(2\ 6)(3\ 7)$ and $(4\ 8)$ must be single colored. There are thus $\binom{4}{2} = 6$ possibilities (Choosing which 2 out of the four cycles would be green).

It remains to deal with the elements yx^i. We recall that these are all reflections. There are two kinds of reflections. One may be written using permutation notation as

$$(i_1\ i_2)(i_3\ i_4)(i_5\ i_6)$$

(with the other two vertices being fixed. For example $y = (2\ 8)(3\ 7)(4\ 6)$ is of this form). The other kind is of the form

$$(i_1\ i_2)(i_3\ i_4)(i_5\ i_6)(i_7\ i_8).$$

(For example $yx = (1\ 8)(2\ 7)(3\ 6)(4\ 5)$ is of this sort). Whatever is the case, one uses similar reasoning to deduce that there are 6 colorings preserved by a reflection.

One needs only apply **CFF** to get that there are

$$N = \frac{1}{16}(70 + 2 \cdot 2 + 6 + 8 \cdot 6) = 8$$

distinct necklaces.

17.3. **The game of 16 squares.** [10] Sam Loyd (1841-1911) was America's greatest puzzle expert and invented thousands of ingenious and tremendously popular puzzles.

In this game, we are given a 4×4 box with 15 squares numbered $1, 2, \ldots, 15$ and one free spot. At every step one is allowed to move an adjacent square into the vacant spot. For example

$$\begin{array}{|cccc|} \hline 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & \\ \hline \end{array} \mapsto \begin{array}{|cccc|} \hline 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & & 15 \\ \hline \end{array} \mapsto \begin{array}{|cccc|} \hline 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & & 12 \\ 13 & 14 & 11 & 15 \\ \hline \end{array} \mapsto \begin{array}{|cccc|} \hline 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & & 10 & 12 \\ 13 & 14 & 11 & 15 \\ \hline \end{array} \mapsto \begin{array}{|cccc|} \hline 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 14 & 10 & 12 \\ 13 & & 11 & 15 \\ \hline \end{array}$$

Can one pass from the original position to the position below?

$$\begin{array}{|cccc|} \hline 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 15 & 14 & \\ \hline \end{array}$$

It turns out that the answer is no. Can you prove it? Apparently, the puzzle was originally marketed with the tiles in the impossible position with the challenge to rearrange them into the initial position!

17.4. **Rubik's cube.** [11] In the case of the Rubik cube there is a group G acting on the cube. The group G is generated by 6 basic moves a, b, c, d, e, f (each is a rotation of a certain "third of the cube") and could be thought of as a subgroup of the symmetric group on $54 = 9 \times 6$ letters. It is called the cube group. The order of the Cube Group is $2^{27} \cdot 3^{14} \cdot 5^3 \cdot 7^2 \cdot 11 = 43,252,003,274,489,856,000$, while the order of S_{54} is 230843697339241380472092742683027581083278564571807941132288000000000000.

[9] $x^{(3^2)} = x^9 = x$ because $x^8 = e$, etc.

[10] This doesn't have much to do with group theory. At least an elementary solution is available with no notions from groups. It is given here for sheer fun and as illustration of "acting on a set".

[11] Also known as the Hungarian cube.

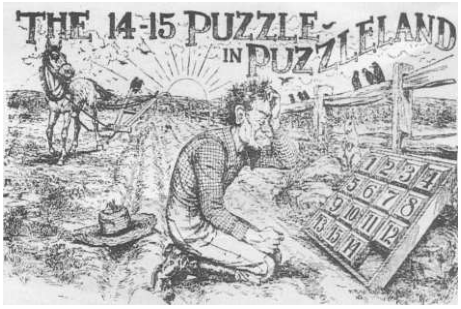


FIGURE 17.1. Loyd's $14-15$ puzzle.

One is usually interested in solving the cube. Namely, reverting it to its original position. Since the current position was gotten by applying an element τ of G, in group theoretic terms we attempt to find an algorithm of writing every G in terms of the generators a, b, c, d, e, f since then also τ^{-1} will have such an expression, which is nothing else than a series of moves that return the cube to its original position. It is natural to deal with the set of generators $a^{\pm 1}, b^{\pm 1}, \ldots, f^{\pm 1}$ (why do 3 times a when you can do a^{-1}??). A common question is what is the maximal number of basic operations that may be required to return a cube to its original position. Otherwise said, what is the diameter of the Cayley graph? But more than that, is there a simple algorithm of finding for every element of G an expression in terms of the generators?

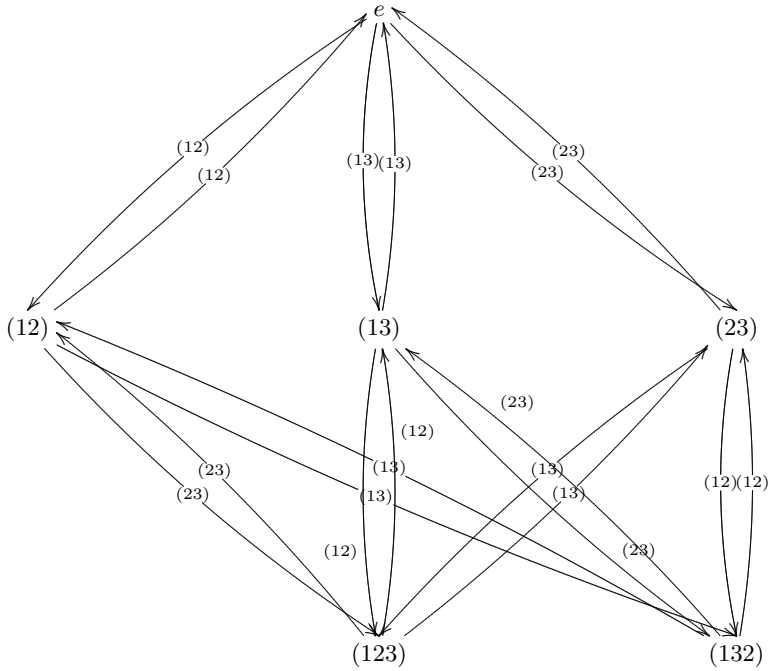
The Cayley graph.

Suppose that $\{g_\alpha : \alpha \in I\}$ are generators for G. We define an oriented graph taking as vertices the elements of G and taking for every $g \in G$ an oriented edge from g to gg_α. If we forget the orientation, the property of $\{g_\alpha : \alpha \in I\}$ being a set of generators is equivalent to the graph being connected.

Suppose that the set of generators consists of n elements. Then, by definition, from every vertex we have n vertices emanating and also n arriving. We see therefore that all Cayley graphs are regular graphs. This, in turn, gives a systematic way of constructing regular graphs.

Suppose we take as a group the symmetric group (see below) S_n and the transpositions as generators. One can think of a permutation as being performed in practice by successively swapping the places of two elements. Thus, in the Cayley graph, the distance between a permutation and the identity (the distance is defined as the minimal length of a path between the two vertices) is the minimal way to write a permutation as a product of transpositions, and could be thought of as a certain measure of the complexity of a transposition.

The figure below gives the Cayley graph of S_3 with respect to the generating set of transpositions. It is a 3-regular oriented graph and a 6 regular graph.



Now, since the Cayley graph of G has 12 edges emanating from each vertex (and is connected by definition of the cube group) it follows that to reach all positions one is forced to allow at least $\log_{12}|G| \sim 18.2$, thus at least 19, moves.[12]

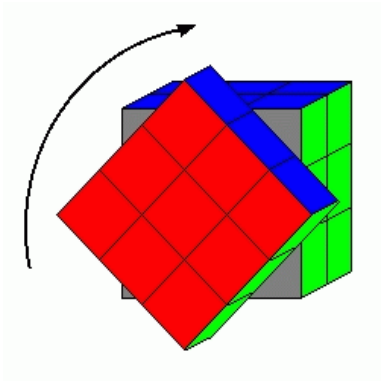


FIGURE 17.2. The Rubik Cube.

[12]There is a subtle point we are glossing over here. It is that perhaps there are operations that move the cube but leave the overall coloring fixed ("we move the pieces but in the end it looks the same"). That is, is the stabilizer of every position of the cube trivial? It seems that the answer is yes; note that it is enough to prove that for the original position (as stabilizers of elements in the same orbit are conjugate subgroups). Here, it seems that the key point is to consider the corner pieces and then the edge pieces. **STOP PRESS**: Keith Conrad tells me that's wrong. It is possible to bring the cube back to the same initial setting only that the center pieces are rotated. He says "... in fact, the stabilizers are not trivial. It is possible to return the cube to the original position with the center faces rotated. This can be seen if you make a mark across the boundary between the center face and an adjacent face, and then mess up and try to solve the cube." (something I would never try doing!) So the argument above should be modified. This will have to wait till I teach the course again....

Part 4. The Symmetric Group

18. Conjugacy classes

Let $\sigma \in S_n$. We write σ as a product of disjoint cycles:

$$\sigma = \sigma_1\sigma_2\cdots\sigma_r.$$

Since disjoint cycles commute, the order does not matter and we may assume that the length of the cycles is non-decreasing. Namely, if we let $|(i_1 i_2 \dots i_t)| = t$ (we shall call it the length of the cycle; it is equal to its order as an element of S_n), then

$$|\sigma_1| \le |\sigma_2| \le \cdots \le |\sigma_r|.$$

We may also allow cycles of length 1 (they simple stand for the identity permutation) and then we find that

$$n = |\sigma_1| + |\sigma_2| + \cdots + |\sigma_r|.$$

We therefore get a partition $p(\sigma)$ of the number n, that is, a set of non-decreasing positive integers $1 \le a_1 \le a_2 \le \cdots \le a_r$ such that $n = a_1 + a_2 + \cdots + a_r$. Note that every partition is obtained from a suitable σ.

Lemma 18.0.1. *Two permutations, σ and ρ, are conjugate (namely there is a τ such that $\tau\sigma\tau^{-1} = \rho$) if and only if $p(\sigma) = p(\rho)$.*

Proof. Recall the formula we used before, if $\sigma(i) = j$ then $(\tau\sigma\tau^{-1})(\tau(i)) = \tau(j)$. This implies that for every cycle $(i_1\ i_2 \dots i_t)$ we have

$$\tau(i_1\ i_2 \dots i_t)\tau^{-1} = (\tau(i_1)\ \tau(i_2)\dots\tau(i_t)).$$

In particular, since $\tau\sigma\tau^{-1} = (\tau\sigma_1\tau^{-1})(\tau\sigma_2\tau^{-1})\cdots(\tau\sigma_r\tau^{-1})$, a product of disjoint cycles, we get that $p(\sigma) = p(\tau\sigma\tau^{-1})$.

Conversely, suppose that $p(\sigma) = p(\rho)$. Say

$$\begin{aligned}\sigma &= \sigma_1\sigma_2\dots\sigma_r \\ &= (i_1^1\dots i_{t(1)}^1)(i_1^2\dots i_{t(2)}^2)\dots(i_1^r\dots i_{t(r)}^r),\end{aligned}$$

and

$$\begin{aligned}\rho &= \rho_1\rho_2\dots\rho_r \\ &= (j_1^1\dots j_{t(1)}^1)(j_1^2\dots j_{t(2)}^2)\dots(j_1^r\dots j_{t(r)}^r).\end{aligned}$$

Define τ by

$$\tau(i_b^a) = j_b^a,$$

then $\tau\sigma\tau^{-1} = \rho$. □

Corollary 18.0.2. *Let $p(n)$ be the number of partitions of n.*[13] *There are $p(n)$ conjugacy classes in S_n.*

Next, we discuss conjugacy classes in A_n. Note that if $\sigma \in A_n$ then since $A_n \lhd S_n$ also $\tau\sigma\tau^{-1} \in A_n$. That is, all the S_n-conjugacy classes of elements of A_n are in A_n. However, we would like to consider the A_n-conjugacy classes of elements of A_n.

Lemma 18.0.3. *The S_n-conjugacy class of an element $\sigma \in A_n$ is a disjoint union of $[S_n : A_nC_{S_n}(\sigma)]$ A_n-conjugacy classes. In particular, it is one A_n-conjugacy class if there is an odd permutation commuting with σ and is two A_n-conjugacy class if there is no odd permutation commuting with σ. In the latter case, the S_n-conjugacy class of σ is the disjoint union of the A_n-conjugacy class of σ and the A_n-conjugacy class of $\tau\sigma\tau^{-1}$, where τ can be chosen to be any odd permutation.*

[13] Since $2 = 2 = 1+1,\ 3 = 3 = 1+2 = 1+1+1,\ 4 = 4 = 2+2 = 1+3 = 1+1+2 = 1+1+1+1,\ 5 = 5 = 2+3 = 1+4 = 1+1+3 = 1+2+2 = 1+1+1+2 = 1+1+1+1+1\dots$ we get $p(1) = 1, p(2) = 2, p(3) = 3, p(4) = 5, p(5) = 7, p(6) = 11, \dots$. The function $p(n)$ is asymptotic to $\frac{e^{\pi\sqrt{2n/3}}}{4n\sqrt{3}}$.

Proof. Let A be the S_n-conjugacy class of σ. Write $A = \coprod_{\alpha\in J} A_\alpha$, a disjoint union of A_n-conjugacy classes. We first note that S_n acts on the set $B = \{A_\alpha : \alpha \in J\}$. Indeed, if A_α is the A_n-conjugacy class of σ_α, and $\rho \in S_n$ then define $\rho A_\alpha \rho^{-1}$ to be the A_n-conjugacy class of $\rho\sigma_\alpha\rho^{-1}$. This is well defined: if σ'_α is another representative for the A_n-conjugacy class of σ_α then $\sigma'_\alpha = \tau\sigma_\alpha\tau^{-1}$ for some $\tau \in A_n$. It follows that $\rho\sigma'_\alpha\rho^{-1} = \rho\tau\sigma_\alpha\tau^{-1}\rho^{-1} = (\rho\tau\rho^{-1})(\rho\sigma_\alpha\rho^{-1})(\rho\tau\rho^{-1})^{-1}$ is in the A_n-conjugacy class of $\rho\sigma_\alpha\rho^{-1}$ (because $\rho\tau\rho^{-1} \in A_n$).

The action of S_n is transitive on B. Consider the A_n-conjugacy class of σ and denote it by A_0. The stabilizer of A_0 is just $A_nC_{S_n}(\sigma)$. Indeed, $\rho A_0\rho^{-1} = A_0$ if and only if $\rho\sigma\rho^{-1}$ is in the same A_n-conjugacy class as σ. Namely, if and only if $\rho\sigma\rho^{-1} = \tau\sigma\tau^{-1}$ for some $\tau \in A_n$, equivalently, $(\tau^{-1}\rho)\sigma = \sigma(\tau^{-1}\rho)$, that is $(\tau^{-1}\rho) \in C_{S_n}(\sigma)$ which is to say that $\rho \in A_nC_{S_n}(\sigma)$.

We conclude that the size of B is the length of the orbit of A_0 and hence is of size $[S_n : A_nC_{S_n}(\sigma)]$. Since $[S_n : A_n] = 2$, we get that $[S_n : A_nC_{S_n}(\sigma)] = 1$ or 2, with the latter happening if and only if $A_n \supseteq C_{S_n}(\sigma)$. That is, if and only if σ does not commute with any odd permutation. Moreover, the orbit consists of the A_n-conjugacy classes of the elements $g\sigma$, g running over a complete set of representatives for the cosets of $A_nC_{S_n}(\sigma)$ in S_n. □

19. The simplicity of A_n

In this section we prove that A_n is a simple group for $n \neq 4$. The cases where $n < 4$ are trivial; for $n = 4$ we have seen it fails (the Klein 4-group is normal). We shall focus on the case $n \geq 5$ and prove the theorem inductively. We therefore first consider the case $n = 5$.

We make the following general observation:

Lemma 19.0.4. *Let $N \triangleleft G$ then N is a disjoint union of G-conjugacy classes.*

Proof. Distinct conjugacy classes, being orbits for a group action, are always disjoint. If N is normal and $n \in N$ then its conjugacy class $\{gng^{-1} : g \in G\}$ is contained in N. □

Let us list the conjugacy classes of S_5 and their sizes.

Conjugacy classes in S_5

cycle type	representative	size of conjugacy class	order	even?
5	(12345)	24	5	✓
1+4	(1234)	30	4	×
1+1+3	(123)	20	3	✓
1+ 2+ 2	(12)(34)	15	2	✓
1 + 1 + 1 + 2	(12)	10	2	×
1 + 1+ 1+ 1+ 1	1	1	1	✓
2+ 3	(12)(345)	20	6	×

Let τ be a permutation commuting with (12345). Then

$$(12345) = \tau(12345)\tau^{-1} = (\tau(1)\ \tau(2)\ \tau(3)\ \tau(4)\ \tau(5))$$

and so τ is the permutation $i \mapsto i+n$ for $n = \tau(1) - 1$. In particular, $\tau = (12345)^{n-1}$ and so is an even permutation. We conclude that the S_5-conjugacy class of (12345) breaks into two A_5-conjugacy classes, with representatives $(12345), (21345)$.

One checks that (123) commutes with the odd permutation (45). Therefore, the S_5-conjugacy class of (123) is also an A_5-conjugacy class. Similarly, the permutation (12)(34) commutes with the odd permutation (12). Therefore, the S_5-conjugacy class of (12)(34) is also an A_5-conjugacy class. We get the following table for conjugacy classes in A_5.

Conjugacy classes in A_5

cycle type	representative	size of conjugacy class	order
5	(12345)	12	5
5	(21345)	12	5
1+1+3	(123)	20	3
1+ 2+ 2	(12)(34)	15	2
1 + 1+ 1+ 1+ 1	1	1	1

If $N \lhd A_5$ then $|N|$ divides 60 and is the sum of 1 and some of the numbers in $(12, 12, 20, 15)$. One checks that this is impossible unless $N = A_5$. We deduce

Lemma 19.0.5. *The group A_5 is simple.*

Theorem 19.0.6. *The group A_n is simple for $n \geq 5$.*

Proof. The proof is by induction on n. We may assume that $n \geq 6$. Let N be a normal subgroup of A_n and assume $N \neq \{1\}$.

First step: There is a permutation $\rho \in N, \rho \neq 1$ and $1 \leq i \leq n$ such that $\rho(i) = i$.

Indeed, let $\sigma \in N$ be a non-trivial permutation and write it as a product of disjoint non-trivial cycles, $\sigma = \sigma_1\sigma_2 \dots \sigma_s$, say in decreasing length. Suppose that σ_1 is $(i_1 i_2 \dots i_r)$, where $r \geq 3$. Then conjugating by the transposition $\tau = (i_1 i_2)(i_5 i_6)$, we get that $\tau\sigma\tau^{-1}\sigma \in N$, $\tau\sigma\tau^{-1}\sigma(i_1) = i_1$ and if $r > 3$ $\tau\sigma\tau^{-1}\sigma(i_2) = i_4 \neq i_2$. If $r = 3$ then $\sigma = (i_1 i_2 i_3)(i_4 \dots) \dots$. Take $\tau = (i_1 i_2)(i_3 i_4)$ then $\tau\sigma\tau^{-1}\sigma(i_1) = i_1$ and $\tau\sigma\tau^{-1}\sigma(i_2) = \tau\sigma(i_4) \in \{i_3, i_5\}$. Thus, $\tau\sigma\tau^{-1}\sigma$ is a permutation of the kind we were seeking.

It still remains to consider the case where each σ_i is a transposition. Then, if $\sigma = (i_1 i_2)(i_3 i_4)$ then σ moves only 4 elements and thus fixes some element and we are done, else $\sigma = (i_1 i_2)(i_3 i_4)(i_5 i_6) \dots$. Let $\tau = (i_1 i_2)(i_3 i_5)$ then $\tau\sigma\tau^{-1}\sigma = (i_2 i_1)(i_5 i_4)(i_3 i_6) \dots (i_1 i_2)(i_3 i_4)(i_5 i_6) \cdots = (i_3 i_5)(i_4 i_6) \dots$ and so is a permutation of the sort we were seeking.

Second step: $N = A_n$.

Consider the subgroups $G_i = \{\sigma \in A_n : \sigma(i) = i\}$. We note that each G_i is isomorphic to A_{n-1} and hence is simple. By the preceding step, for some i we have that $N \cap G_i$ is a non-trivial normal subgroup of G_i, hence equal to G_i.

Next, note that $(12)(34)G_1(12)(34) = G_2$ and, similarly, all the groups G_i are conjugate in A_n to each other. It follows that $N \supseteq \langle G_1, G_2, \dots, G_n \rangle$. Now, every element in S_n is a product of (usually not disjoint) transpositions and so every element σ in A_n is a product of an even number of transpositions,

$\sigma = \lambda_1\mu_1 \dots \lambda_r\mu_r$ (λ_i, μ_i transpositions). Since $n > 4$ every product $\lambda_i\mu_i$ belongs to some G_j and we conclude that $\langle G_1, G_2, \dots, G_n \rangle = A_n$.

□

Part 5. p-groups, Cauchy's and Sylow's Theorems

20. The class equation

Let G be a finite group. G acts on itself by conjugation: $g \star h = ghg^{-1}$. The class equation is the partition of G to orbits obtained this way. The orbits are called in this case *conjugacy classes*. Note that the stabilizer of $h \in G$ is $C_G(h)$ and so its orbit has length $[G : C_G(h)]$. Note thus the elements with orbit of length 1 are precisely the elements in the center $Z(G)$ of G. We get

$$|G| = |Z(G)| + \sum_{\text{reps.} x \notin Z(G)} \frac{|G|}{|C_G(x)|}. \tag{20.1}$$

Remark 20.0.7. One can prove that for every $n > 0$ there are only finitely many finite groups with exactly n conjugacy classes. (One uses the following fact: Given $n > 0$ and a rational number q there are only finitely many n-tuples $(c_1, \ldots, c_n)$ of natural numbers such that $q = \frac{1}{c_1} + \cdots + \frac{1}{c_n}$.)

For example, the only group with one conjugacy class is the trivial group $\{1\}$; the only group with two conjugacy classes is $\mathbb{Z}/2\mathbb{Z}$; the only groups with 3 conjugacy classes are $\mathbb{Z}/3\mathbb{Z}$ and S_3.

21. p-groups

Let p be a prime. A finite group G is called a p-group if its order is a positive power of p.

Lemma 21.0.8. *Let G be a finite p group. Then the center of G is not trivial.*

Proof. We use the class equation 20.1. Note that if $x \notin Z(G)$ then $C_G(x) \neq G$ and so the integer $\frac{|G|}{|C_G(x)|}$ is divisible by p. Thus, the left hand side of

$$|G| - \sum_{\text{reps.} x \notin Z(G)} \frac{|G|}{|C_G(x)|} = |Z(G)|$$

is divisible by p, hence so is the right hand side. In particular $|Z(G)| \geq p$. □

Theorem 21.0.9. *Let G be a finite p group, $|G| = p^n$.*

1. *For every normal subgroup $H \triangleleft G$, $H \neq G$, there is a subgroup $K \triangleleft G$ such that $H < K < G$ and $[K : H] = p$.*
2. *There is a chain of subgroups $H_0 = \{1\} < H_1 < \cdots < H_n = G$, such that each $H_i \triangleleft G$ and $|H_i| = p^i$.*

Proof. (1) The group G/H is a p group and hence its center is a non-trivial group. Take an element $e \neq x \in Z(G/H)$; its order is p^r for some r. Then $y = x^{p^{r-1}}$ has exact order p. Let $K' = \langle y \rangle$. It is a normal subgroup of G/H of order p (y commutes with any other element). Let $K = \pi_H^{-1}(K')$. By the Third Isomorphism Theorem K is a normal subgroup of G, $K/H \cong K'$ so $[K : H] = p$.

(2) The proof just given shows that every p group has a normal subgroup of p elements. Now apply repeatedly the first part.

□

21.1. Examples of p groups.

21.1.1. *Groups of order p.* We proved in the assignments that every such group is cyclic, thus isomorphic to $\mathbb{Z}/p\mathbb{Z}$.

21.1.2. *Groups of order p^2.* We shall prove in the assignments that every such group is commutative. It then follows from the structure theorem for finite abelian groups that such a group is either isomorphic to $\mathbb{Z}/p^2\mathbb{Z}$ or to $(\mathbb{Z}/p\mathbb{Z})^2$.

21.1.3. *Groups of order p^3.* First, there are the abelian groups $\mathbb{Z}/p^3\mathbb{Z}$, $\mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$ and $(\mathbb{Z}/p\mathbb{Z})^3$.

We shall prove in the assignments that if G is not abelian then $G/Z(G)$ cannot be cyclic. It follows that $Z(G) \cong \mathbb{Z}/p\mathbb{Z}$ and $G/Z(G) \cong (\mathbb{Z}/p\mathbb{Z})^2$. One example of such a group is provided by the matrices

$$\begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix},$$

where $a, b, c \in \mathbb{F}_p$. Note that if $p \geq 3$ then every element in this group is of order p (use $(I+N)^p = I+N^p$), yet the group is non-abelian. (This group, using a terminology to be introduced later, is a semi-direct product $(\mathbb{Z}/p\mathbb{Z})^2 \rtimes \mathbb{Z}/p\mathbb{Z}$.) More generally the upper unipotent matrices in $\mathrm{GL}_n(\mathbb{F}_p)$ are a group of order $p^{n(n-1)/2}$ in which every element has order p if $p \geq n$. Notice that these groups are non-abelian.

Getting back to the issue of non-abelian groups of order p^3, one can prove that there is precisely one additional non-abelian group of order p^3. It is generated by two elements x, y satisfying: $x^p = y^{p^2} = 1, xyx^{-1} = y^{1+p}$. (This group is a semi-direct product $(\mathbb{Z}/p^2\mathbb{Z}) \rtimes \mathbb{Z}/p\mathbb{Z}$.)

21.2. **A few words on free groups.** Let $x_1, \ldots, x_d$ be formal symbols. The *free group on* $x_1, \ldots, x_d$ is the set of expressions (called "words") $y_1 \ldots y_t$, where each y_i is a symbol x_j or x_j^{-1}, taken under the equivalence relation generated by the following basic equivalence: if v, w are words then

$$vx_jx_j^{-1}w \sim vw, \quad vx_j^{-1}x_jw \sim vw.$$

We remark that the empty word is allowed. We define multiplication of two words v, w by putting them together into one word

$$v \star w = vw.$$

One checks that this is well defined on equivalence classes, that it is an associative operation, that the (equivalence class of the) empty word is the identity, and that every element has an inverse: $(y_1 \ldots y_t)^{-1} = y_t^{-1} \ldots y_1^{-1}$. We thus get a group, called the free group of rank d, denoted $\mathscr{F}(d)$. It has the following properties:

(1) given a group G, and d elements $s_1, \ldots s_d$ in G, there is a unique group homomorphism $f : \mathscr{F}(d) \longrightarrow G$ such that $f(x_i) = s_i$. Indeed, one first defines for a word $y_1 \ldots y_t$, $y_i = x_{n(i)}^{e_i}, e_i \in \{\pm 1\}$, $f(y_1 \cdots y_t) = s_{n(1)}^{e_1} \cdots s_{n(t)}^{e_t}$. One checks that equivalent words have the same image and so gets a well defined function $\mathscr{F}(d) \to G$. It is easy to verify it is a homomorphism.
(2) if G is a group generated by d elements there is a surjective group homomorphism $\mathscr{F}(d) \longrightarrow G$. This follows immediately from the previous point. If $s_1, \ldots, s_d$ are generators take the homomorphism taking x_i to s_i.
(3) if $w_1, \ldots w_r$ are words in $\mathscr{F}(d)$, let N be the minimal normal subgroup containing all the w_i (such exists!). The group $\mathscr{F}(d)/N$ is also denoted by $\langle x_1, \ldots, x_d | w_1, \ldots, w_r \rangle$ and is said to be given by the generators $x_1, \ldots x_d$ and relations $w_1, \ldots, w_r$. For example, one can prove that $\mathbb{Z} \cong \mathscr{F}(1)$, $\mathbb{Z}/n\mathbb{Z} \cong \langle x_1 | x_1^n \rangle$, $\mathbb{Z}^2 \cong \langle x_1, x_2 | x_1x_2x_1^{-1}x_2^{-1} \rangle$, $S_3 \cong \langle x_1, x_2 | x_1^2, x_2^3, (x_1x_2)^2 \rangle$, and more generally $D_{2n} = \langle x, y | x^n, y^2, yxyxy \rangle$.
(4) if $d = 1$ then $\mathscr{F}(d) \cong \mathbb{Z}$ but if $d > 1$ then $\mathscr{F}(d)$ is a non-commutative infinite group. In fact, for every k, S_k is a homomorphic image of $\mathscr{F}(d)$ if $d \geq 2$.

21.2.1. *Some famous problems in group theory.* Fix positive integers d, n. *The Burnside problem* asks if a group generated by d elements in which every element x satisfies $x^n = 1$ is finite. Every such group is a quotient of the following group $B(d, n)$: it is the free group $\mathscr{F}(d)$ generated by $x_1, \ldots, x_d$ moded out by the minimal normal subgroup containing the expressions f^n where f is an element of $\mathscr{F}(d)$. It turns out that in general the answer is negative; $B(d, n)$ is infinite for $d \geq 2, n \geq 4381$, n odd. There are some instances where it is finite: $d \geq 2, n = 2, 3, 4, 6$.

One can then ask, is there a finite group $B_0(d, n)$ such that every finite group G, generated by d elements and in which $f^n = 1$ for every element $f \in G$, is a quotient of $B_0(d, n)$? E. Zelmanov, building on the work of many others, proved that the answer is yes. He received the 1994 Fields medal for this.

The *word problem* asks whether there is an algorithm (guaranteed to stop in finite time) that determines whether a finitely presented group, that is a group gives by generators and relations as $\langle x_1, \ldots, x_d | w_1, \ldots, w_r \rangle$ for some integers d, r, is the trivial group or not. It is known that the answer to this question (and almost any variation on it!) is NO. This has applications to topology. It is known that every finitely presented

group is the fundamental group of a manifold[14] of dimension 4. It then follows that there is no good classification of 4-manifolds. If one can decide if a manifold X is isomorphic to the 4-dimensional sphere or not, one can decide the question of whether the fundamental group of X is isomorphic to that of the sphere, *which is the trivial group*, and so solve the word problem.

22. Cauchy's Theorem

One application of group actions is to provide a simple proof of an important theorem in the theory of finite groups.

Theorem 22.0.1. (Cauchy) *Let G be a finite group of order n and let p be a prime dividing n. Then G has an element of order p.*

Proof. Let S be the set consisting of p-tuples $(g_1, \dots, g_p)$ of elements of G, considered up to cyclic permutations. Thus if T is the set of p-tuples $(g_1, \dots, g_p)$ of elements of G, S is the set of orbits for the action of $\mathbb{Z}/p\mathbb{Z}$ on T by cyclic shifts . One may therefore apply **CFF** and get

$$|S| = \frac{n^p - n}{p} + n.$$

Note that $n \not| |S|$.

Now define an action of G on S. Given $g \in G$ and $(g_1, \dots, g_p) \in S$ we define

$$g(g_1, \dots, g_p) = (gg_1, \dots, gg_p).$$

This is a *well defined* action .

Since the order of G is n, since $n \not| |S|$, and since S is a disjoint union of orbits of G, there must be an orbit $\mathrm{Orb}(s)$ whose size is not n. However, the size of an orbit is $|G|/|\mathrm{Stab}(s)|$, and we conclude that there must an element $(g_1, \dots, g_p)$ in S with a non-trivial stabilizer. This means that for some $g \in G$, such that $g \neq e$, we have

$$(gg_1, \dots, gg_p) \text{ is equal to } (g_1, \dots, g_p) \text{ up to a cyclic shift.}$$

This means that for some i we have

$$(gg_1, \dots, gg_p) = (g_{i+1}, g_{i+2}, g_{i+3}, \dots, g_p, g_1, g_2, \dots, g_i).$$

Therefore, $gg_1 = g_{i+1}$, $g^2 g_1 = gg_{i+1} = g_{2i+1}, \dots, g^p g_1 = \cdots = g_{pi+1} = g_1$ (we always read the indices mod p). That is, there exists $g \neq e$ with

$$g^p = e.$$

□

23. Sylow's Theorems

Let G be a finite group and let p be a prime dividing its order. Write $|G| = p^r m$, where $(p, m) = 1$. By a p-subgroup of G we mean a subgroup whose order is a power of p. By a maximal p subgroup of G we mean a p-subgroup of G not contained in a strictly larger p-subgroup.

Theorem 23.0.2. *Every maximal p-subgroup of G has order p^r (such a subgroup is called a* Sylow p-subgroup*) and such a subgroup exists. All Sylow p-subgroups are conjugate to each other. The number n_p of Sylow p-subgroups satisfies: (i) $n_p | m$; (ii) $n_p \equiv 1 \pmod{p}$.*

[14] A manifold of dimension 4 is a space that locally looks like $\mathbb{R}^4$. The fundamental group is a topological construction that associate a group to any topological space. The group has as its elements equivalent classes of closed loops in the space, starting and ending at some arbitrarily chosen point, where if we can deform, within the space, one loop to another we consider them as the same element of the fundamental group.

Remark 23.0.3. To say that P is conjugate to Q means that there is a $g \in G$ such that $gPg^{-1} = Q$. Recall that the map $x \mapsto gxg^{-1}$ is an automorphism of G. This implies that P and Q are isomorphic as groups.

Another consequence is that to say there is a unique p-Sylow subgroup is the same as saying that a p-Sylow is normal. This is often used this way: given a finite group G the first check in ascertaining whether it is simple or not is to check whether the p-Sylow subgroup is unique for some p dividing the order of G. Often one engages in combinatorics of counting how many p-Sylow subgroups can be, trying to conclude there can be only one for a given p and hence getting a normal subgroup.

We first prove a lemma that is an easy case of Cauchy's Theorem 22.0.1:

Lemma 23.0.4. *Let A be a finite abelian group, let p be a prime dividing the order of A. Then A has an element of order p.*

Proof. We prove the result by induction on $|A|$. Let N be a maximal subgroup of A, distinct from A. If p divides the order of N we are done by induction. Otherwise, let $x \notin N$ and let $B = \langle x \rangle$. By maximality the subgroup BN is equal to A. On the other hand $|BN| = |B| \cdot |N|/|B \cap N|$. Thus, p divides the order of B. That is the order of x is pa for some a and so the order of x^a is precisely p. □

Proposition 23.0.5. *There is a p-subgroup of G of order p^r.*

Proof. We prove the result by induction on the order of G. Assume first that p divides the order of $Z(G)$. Let x be an element of $Z(G)$ of order p and let $N = \langle x \rangle$, a normal subgroup. The order of G/N is $p^{r-1}m$ and by induction it has a p-subgroup H' of order p^{r-1}. Let H be the preimage of H'. It is a subgroup of G such that $H/N \cong H'$ and thus H has order $|H'| \cdot |N| = p^r$.

Consider now the case where p does not divide the order of $Z(G)$. Consider the class equation

$$|G| = |Z(G)| + \sum_{\text{reps.} x \notin Z(G)} \frac{|G|}{|C_G(x)|}.$$

We see that for some $x \notin Z(G)$ we have that p does not divide $\frac{|G|}{|C_G(x)|}$. Thus, p^r divides $|C_G(x)|$. The subgroup $C_G(x)$ is a *proper* subgroup of G because $x \notin Z(G)$. Thus, by induction $C_G(x)$, and hence G, has a p-subgroup of order p^r. □

Lemma 23.0.6. *Let P be a maximal p-subgroup and Q any p-subgroup then*

$$Q \cap P = Q \cap N_G(P).$$

Proof. Since $P \subset N_G(P)$ also $Q \cap P \subset Q \cap N_G(P)$. Let $H = Q \cap N_G(P)$. Then, since $P \lhd N_G(P)$ we have that HP is a subgroup of $N_G(P)$. Its order is $|H| \cdot |P|/|H \cap P|$ and so a power of p. Since P is a maximal p-subgroup we must have $HP = P$ and thus $H \subset P$. □

Proof. (Of Theorem) Let P be a Sylow subgroup of G. Such exists by Proposition 23.0.5. Let

$$S = \{P_1, \ldots, P_a\}$$

be the set of conjugates of $P = P_1$. That is, the subgroups gPg^{-1} one gets by letting g vary over G. Note that for a fixed g the map $P \longrightarrow gPg^{-1}$, $x \mapsto gxg^{-1}$ is a group isomorphism. Thus, every P_i is a Sylow p-subgroup. Our task is to show that every maximal p-subgroup is an element of S and find out properties of a.

Let Q be any p-subgroup of G. The subgroup Q acts by conjugation on S. The size of $Orb(P_i)$ is $|Q|/|\text{Stab}_Q(P_i)|$. Now $\text{Stab}_Q(P_i) = Q \cap N_G(P_i) = Q \cap P_i$ by Lemma 23.0.6. Thus, the orbit consists of one element if $Q \subset P_i$ and is a proper power of p otherwise.

Take first Q to be P_1. Then, the orbit of P_1 has size 1. Since P_1 is a maximal p-subgroup it is not contained in any other p-subgroup, thus the size of every other orbit is a power of p. It follows, using that S is a disjoint union of orbits, that $a = 1 + tp$ for some t. Note also that $a = |G|/|N_G(P)|$ and thus divides $|G|$.

We now show that all maximal p-subgroups are conjugate. Suppose, to the contrary, that Q is a maximal p-subgroup which is not conjugate to P. Thus, for all i, $Q \neq P_i$ and so $Q \cap P_i$ is a proper subgroup of Q. It follows then that S is a union of disjoint orbit all having size a proper power of p. Thus, $p|a$. This is a contradiction. □

23.1. **Examples and applications.**

23.1.1. *p-groups.* Every finite p-group is of course the only p-Sylow subgroup (trivial case).

23.1.2. $\mathbb{Z}/6\mathbb{Z}$. In every abelian group the p-Sylow subgroups are normal and unique. The 2-Sylow subgroup is $\langle 3\rangle$ and the 3-Sylow subgroup is $\langle 2\rangle$.

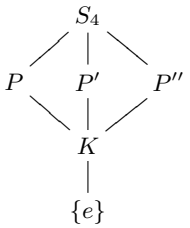
23.1.3. S_3. Consider the symmetric group S_3. Its 2-Sylow subgroups are given by $\{1,(12)\},\{1,(13)\},\{1,(23)\}$. Note that indeed $3|3=3!/2$ and $3\equiv 1 \pmod 2$. It has a unique 3-Sylow subgroup $\{1,(123),(132)\}$. This is expected since $n_3|2=3!/3$ and $n_3\equiv 1 \pmod 3$ implies $n_3=1$.

23.1.4. S_4. We want to find the 2-Sylow subgroups. Their number $n_2|3=24/8$ and is congruent to 1 modulo 2. It is thus either 1 or 3. Note that every element of S_4 has order $1,2,3,4$. The number of elements of order 3 is 8 (the 3-cycles). Thus, we cannot have a unique subgroup of order 8 (it will contain any element of order 2 or 4). We conclude that $n_2=3$. One such subgroup is $D_8\subset S_4$; the rest are conjugates of it.

Further, $n_3|24/3$ and $n_3\equiv 1 \pmod 3$. If $n_3=1$ then that unique 3-Sylow would need to contain all 8 element of order 3 but is itself of order 3. Thus, $n_3=4$.

Remark 23.1.1. A group of order 24 is never simple, though it does not mean that one of the Sylow subgroups is normal, as the example of S_4 shows. However, consider the representation of a group G of order 24 on the cosets of P, where P is one of its 2-Sylow subgroup. It gives us, as we have seen in the past, a normal subgroup of G, contained in P, whose index divides $6=[G:P]!$ and hence is non-trivial.

Suppose that $G=S_4$ now and call this subgroup K. Then, we see that $|K|=4$; it is preserved under conjugation hence is a subgroup of all three 2-Sylow subgroups, say P,P',P''. We have the following picture



23.1.5. *Groups of order pq.* Let $p<q$ be primes. Let G be a group of order pq. Then $n_q|p$, $n_q\equiv 1 \pmod q$. Since $p<q$ we have $n_q=1$ and the q-Sylow subgroup is normal (in particular, G is never simple). Also, $n_p|q$, $n_p\equiv 1 \pmod p$. Thus, either $n_p=1$, or $n_p=q$ and the last possibility can happen only for $q\equiv 1 \pmod p$.

We conclude that if $p\nmid(q-1)$ then both the p-Sylow subgroup P and the q-Sylow subgroup Q are normal. Note that the order of $P\cap Q$ divides both p and q and so is equal to 1. Let $x\in P, y\in Q$ then $[x,y]=(xyx^{-1})y^{-1}=x(yx^{-1}y^{-1})\in P\cap Q=\{1\}$. Thus, PQ, which is equal to G, is abelian.

We shall later see that whenever $p|(q-1)$ there is a non-abelian group of order pq (in fact, unique up to isomorphism). The case of S_3 falls under this.

23.1.6. *Groups of order p^2q.* Let G be a group of order p^2q, where p and q are distinct primes. We prove that G is not simple:

If $q<p$ then $n_p\equiv 1 \pmod p$ and $n_p|q<p$, which implies that $n_p=1$ and the p-Sylow subgroup is normal.

Suppose that $p<q$, then $n_q\equiv 1 \pmod q$ and $n_q|p^2$, which implies that $n_q=1$ or p^2. If $n_q=1$ then the q-Sylow subgroup is normal. Assume that $n_q=p^2$. Each pair of the p^2 q-Sylow subgroups intersect only at the identity (since q is prime). Hence they account for $1+p^2(q-1)$ elements. Suppose that there were 2 p-Sylow subgroups. They intersect at most at a subgroup of order p. Thus, they contribute at least $2p^2-p$ new elements. All together we got at least $1+p^2(q-1)+2p^2-p=p^2q+p^2-p+1>p^2q$ elements. That's a contradiction and so $n_p=1$; the p-Sylow subgroup is normal.

Remark 23.1.2. A theorem of Burnside states that a group of order p^aq^b with $a+b>1$ is not simple. You will prove in the assignments that groups of order pqr ($p<q<r$ primes) are not simple. Note that $|A_5|=60=2^2\cdot 3\cdot 5$ and A_5 is simple. A theorem of Feit and Tompson says that a finite simple group is either of prime order, or of even order.

23.1.7. $\mathrm{GL}_n(\mathbb{F})$. Let $\mathbb{F}$ be a finite field with q elements, q a power of a prime p. The order of $\mathrm{GL}_n(\mathbb{F})$ is $(q^n-1)(q^n-q)\cdots(q^n-q^{n-1}) = q^{(n-1)n/2}(q^n-1)(q^{n-1}-1)\cdots(q-1)$. Thus, a p-Sylow has order $q^{(n-1)n/2}$. One such subgroup consists of the upper triangular matrices with 1 on the diagonal (the unipotent group):

$$\begin{pmatrix} 1 & * & \dots & * \\ 0 & 1 & \cdots & * \\ & & \ddots & \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Part 6. Finitely Generated Abelian Groups, Semi-direct Products and Groups of Low Order

24. The structure theorem for finitely generated abelian groups

The structure theorem will be proved in the next semester as a corollary of the structure theorem for modules over a principal ideal domain. That same theorem will also yield the Jordan canonical form of a matrix.

Theorem 24.0.3. *Let G be a finitely generated abelian group. Then there exists a unique non-negative integer r and integers $1 < n_1|n_2|\dots|n_t$ ($t \geq 0$) such that*

$$G \cong \mathbb{Z}^r \times \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_t\mathbb{Z}.$$

Remark 24.0.4. The integer r is called the *rank* of G. The subgroup in G that corresponds to $\mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_t\mathbb{Z}$ under such an isomorphism is canonical (independent of the isomorphism). It is the subgroup of G of elements of finite order, also called the *torsion subgroup* of G and sometime denoted G_{tor}.

On the other hand, the subgroup corresponding to $\mathbb{Z}^r$ is not canonical and depends very much on the isomorphism.

A group is called *free abelian group* if it is isomorphic to $\mathbb{Z}^r$ for some r (the case $t = 0$ in the theorem above). In this case, elements $x_1, \dots, x_r$ of G that correspond to a basis of $\mathbb{Z}^r$ are called a basis of G; every element of G has the form $a_1x_1 + \cdots + a_rx_r$ for unique integers $a_1, \dots, a_r$.

Remark 24.0.5. The Chinese remainder theorem gives that if $n = p_1^{a_1} \cdots p_s^{a_s}$, p_i distinct primes, then

$$\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/p_1^{a_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_s^{a_s}\mathbb{Z}.$$

Thus, one could also write an isomorphism $G \cong \mathbb{Z}^r \times \prod_i \mathbb{Z}/p_i^{b_i}\mathbb{Z}$.

We shall also prove the following corollary in greater generality next semester.

Corollary 24.0.6. *Let G, H be two free abelian groups of rank r. Let $f : H \longrightarrow G$ be a homomorphism such that $G/f(H)$ is a finite group. There are bases $x_1, \dots, x_r$ of G and $y_1, \dots, y_r$ of H and integers $1 \leq n_1|\dots|n_r$ such that $f(y_i) = n_ix_i$.*

Example 24.0.7. Let G be a finite abelian p group, $|G| = p^n$. Then $G \cong \mathbb{Z}/p_1^{a_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_s^{a_s}\mathbb{Z}$ for unique a_i satisfying $1 \leq a_1 \leq \cdots \leq a_s$ and $a_1 + \cdots + a_s = n$. It follows that the number of isomorphism groups of finite abelian groups of order p^n is $p(n)$ (the partition function of n).

25. Semi-direct products

Given two groups B, N we have formed their direct product $G = N \times B$. Identifying B, N with their images $\{1\} \times B, N \times \{1\}$ in G, we find that: (i) $G = NB$, (ii) $N \lhd G, B \lhd G$, (iii) $N \cap B = \{1\}$. Conversely, one can easily prove that if G is a group with subgroups B, N such that: (i) $G = NB$, (ii) $N \lhd G, B \lhd G$, (iii) $N \cap B = \{1\}$, then $G \cong N \times B$. The definition of a semi-direct product relaxes the conditions a little.

Definition 25.0.8. Let G be a group and let B, N be subgroups of G such that: (i) $G = NB$; (ii) $N \cap B = \{1\}$; (iii) $N \lhd G$. Then we say that G is a *semi-direct product* of N and B.

Let N be any group. Let $\mathrm{Aut}(N)$ be the set of automorphisms of the group N. It is a group in its own right under composition of functions.

Let B be another group and $\phi : B \longrightarrow \mathrm{Aut}(N), b \mapsto \phi_b$ be a homomorphism (so $\phi_{b_1b_2} = \phi_{b_1} \circ \phi_{b_2}$). Define a group

$$G = N \rtimes_\phi B$$

as follows: as a set $G = N \times B$, but the group law is defined as

$$(n_1, b_1)(n_2, b_2) = (n_1 \cdot \phi_{b_1}(n_2), b_1b_2).$$

We check associativity:

$$\begin{aligned}[(n_1,b_1)(n_2,b_2)](n_3,b_3) &= (n_1\cdot\phi_{b_1}(n_2),b_1b_2)(n_3,b_3)\\ &= (n_1\cdot\phi_{b_1}(n_2)\cdot\phi_{b_1b_2}(n_3),b_1b_2b_3)\\ &= (n_1\cdot\phi_{b_1}(n_2\cdot\phi_{b_2}(n_3)),b_1b_2b_3)\\ &= (n_1,b_1)(n_2\cdot\phi_{b_2}(n_3),b_2b_3)\\ &= (n_1,b_1)[(n_2,b_2)(n_3,b_3)].\end{aligned}$$

The identity is clearly $(1_N,1_B)$. The inverse of (n_2,b_2) is $(\phi_{b_2^{-1}}(n_2^{-1}),b_2^{-1})$. Thus G is a group. The two injections

$$N\longrightarrow G,\quad n\mapsto(n,1);\qquad B\longrightarrow G,\quad b\mapsto(1,b),$$

are group homomorphisms. We identify N and B with their images in G. We claim that G is a semi-direct product of N and B.

Indeed, clearly the first two properties of the definition hold. It remains to check that N is normal and it's enough to verify that $B\subset N_G(N)$. According to the calculation above:

$$(1,b)(n,1)(1,b^{-1})=(\phi_b(n),1).$$

We now claim that every semi-direct product is obtained this way: Let G be a semi-direct product of N and B. Let $\phi_b:N\longrightarrow N$ be the map $n\mapsto bnb^{-1}$. This is an automorphism of N and the map

$$\phi:B\longrightarrow\mathrm{Aut}(N)$$

is a group homomorphism. We claim that $N\rtimes_\phi B\cong G$. Indeed, define a map

$$(n,b)\mapsto nb.$$

It follows from the definition that the map is surjective. It is also bijective since $nb=1$ implies that $n=b^{-1}\in N\cap B$ hence $n=1$. It remains to check that this is a group homomorphism, but $(n_1\cdot\phi_{b_1}(n_2),b_1b_2)\mapsto n_1\phi_{b_1}(n_2)b_1b_2=n_1b_1n_2b_1^{-1}b_1b_2=(n_1b_1)(n_2b_2)$.

Proposition 25.0.9. *A semi-direct product $N\rtimes_\phi B$ is the direct product $N\times B$ if and only if $\phi:B\longrightarrow\mathrm{Aut}(N)$ is the trivial homomorphism.*

Proof. Indeed, that happens iff for all $(n_1,b_1),(n_2,b_2)$ we have $(n_1\phi_{b_1}(n_2),b_1b_2)=(n_1n_2,b_1b_2)$. That is, iff for all b_1,n_2 we have $\phi_{b_1}(n_2)=n_2$, which implies $\phi_{b_1}=id$ for all b_1. That is, ϕ is the trivial homomorphism. □

Example 25.0.10. The Dihedral group D_{2n} is a semi-direct product. Take $N=\langle x\rangle\cong\mathbb{Z}/n\mathbb{Z}$ and $B=\langle y\rangle\cong\mathbb{Z}/2\mathbb{Z}$. Then $D_{2n}\cong\mathbb{Z}/n\mathbb{Z}\rtimes_\phi\mathbb{Z}/2\mathbb{Z}$ with $\phi_1=-1$.

25.1. **Application to groups of order pq.** We have seen in § 23.1.5 that if $p<q$ and $p\not|(q-1)$ then every group of order pq is abelian. Assume therefore that $p|(q-1)$.

Proposition 25.1.1. *If $p|(q-1)$ there is a unique non-abelian group, up to isomorphism, of order pq.*

Proof. Let G be a non-abelian group of order pq. We have seen that in every such group G the q-Sylow subgroup Q is normal. Let P be any p-Sylow subgroup. Then $P\cap Q=\{1\}$ and $G=QP$. Thus, G is a semi-direct product of Q and P.

It is thus enough to show that there is a non-abelian semi-direct product and that any two such products are isomorphic. We may consider the case $Q=\mathbb{Z}/q\mathbb{Z},P=\mathbb{Z}/p\mathbb{Z}$.

Lemma 25.1.2. $\mathrm{Aut}(Q)=(\mathbb{Z}/q\mathbb{Z})^\times$.

Proof. Since Q is cyclic any group homomorphism $f:Q\longrightarrow H$ is determined by its value on a generator, say 1. Conversely, if $h\in H$ is of order dividing q then there is such a group homomorphism with $f(1)=h$. Take $H=Q$. The image of f is the cyclic subgroup $\langle h\rangle$ and thus f is surjective (equivalently, bijective) iff h is a generator. Thus, any element $h\in(\mathbb{Z}/q\mathbb{Z})^\times$ determines an automorphism f_h of Q by $a\mapsto ah$. Note that $f_h(f_g)(a)=f_h(ag)=agh=f_{hg}(a)$ and so the association $h\leftrightarrow f_h$ is a group isomorphism $(\mathbb{Z}/q\mathbb{Z})^\times\cong\mathrm{Aut}(Q)$. □

Since $(\mathbb{Z}/q\mathbb{Z})^\times$ is a cyclic group of order $q-1$ (Corollary 5.0.15), and since $p|(q-1)$, there is an element h of exact order p in $(\mathbb{Z}/q\mathbb{Z})^\times$. Let ϕ be the homomorphism determined by $\phi_1 = f_h$ and let $G = Q \rtimes_\phi P$. We claim that G is not abelian.

$$(n,0)(0,b) = (n,b), \quad (0,b)(n,0) = (\phi_b(n), b).$$

The two are always equal only if $\phi_b(n) = n$ for all b and n, i.e., $\phi_b = 1$ for all b, but choosing $b = 1$ we get $\phi_1 = h$ and thus a contradiction.

We now show that G is unique up to isomorphism. If H is another such semi-direct product then $H = \mathbb{Z}/q\mathbb{Z} \rtimes_\psi \mathbb{Z}/p\mathbb{Z}$, where ψ_1 is an element of order p (if it is the identity H is abelian) and thus $\psi_1 = \phi_1^r = \phi_r$ for some r prime to p.

Define a map

$$\mathbb{Z}/q\mathbb{Z} \rtimes_\psi \mathbb{Z}/p\mathbb{Z} \longrightarrow \mathbb{Z}/q\mathbb{Z} \rtimes_\phi \mathbb{Z}/p\mathbb{Z}, \quad (n,b) \mapsto (n, rb).$$

This function is easily checked to be injective, hence bijective. We check it is a group homomorphism:

In G we have $(n_1, rb_1)(n_2, rb_2) = (n_1 + \phi_{rb_1}(n_2), r(b_1+b_2)) = (n_1 + \psi_{b_1}(n_2), r(b_1+b_2))$ which is the image of $(n_1 + \psi_{b_1}(n_2), b_1 + b_2)$, the product $(n_1, b_1)(n_2, b_2)$ in H. □

Example 25.1.3. Is there a non-abelian group of order 165 containing $\mathbb{Z}/55\mathbb{Z}$?

In such a group G, the subgroup $\mathbb{Z}/55\mathbb{Z}$ must be normal (its index is the minimal one dividing the order of G). Since there is always a 3-Sylow, we conclude that G is a semi-direct product $\mathbb{Z}/55\mathbb{Z} \rtimes \mathbb{Z}/3\mathbb{Z}$. This is determined by a homomorphism $\mathbb{Z}/3\mathbb{Z} \to \mathrm{Aut}(\mathbb{Z}/55\mathbb{Z}) \cong (\mathbb{Z}/55\mathbb{Z})^\times$. The right hand side has order $\varphi(55) = 4 \cdot 10$. Thus, the homomorphism is trivial and G is a direct product. It follows that G must be commutative.

Cases where two semi-direct products are isomorphic.
It is useful to generalize the last argument. Consider a homomorphism $\phi : B \longrightarrow \mathrm{Aut}(N)$ and consider the group

$$G = N \rtimes_\phi B.$$

Consider two automorphisms $f : N \to N, g : B \to B$. Let S be G considered as a set and consider the self map

$$S \longrightarrow S, \quad (n,b) \mapsto (f(n), g(b)).$$

We may define a new group law on S by

$$\begin{aligned}(n_1,b_1) \star (n_2,b_2) &= f \circ g \ \ (f^{-1}(n_1), g^{-1}(b_1))(f^{-1}(n_2), g^{-1}(b_2)) \\ &= f \circ g \ \ (f^{-1}(n_1) \cdot [\phi(g^{-1}(b_1))](f^{-1}(n_2)), g^{-1}(b_1)g^{-1}(b_2)) \\ &= (n_1 \cdot f([\phi(g^{-1}(b_1))](f^{-1}(n_2))), b_1 b_2)\end{aligned}$$

Clearly, S with the new group law is isomorphic as groups to G.

This suggests the following, define an action of $\mathrm{Aut}(B) \times \mathrm{Aut}(N)$ on $\mathrm{Hom}(B, \mathrm{Aut}(N))$ via the embedding $\mathrm{Aut}(B) \times \mathrm{Aut}(N) \longrightarrow \mathrm{Aut}(B) \times \mathrm{Aut}(\mathrm{Aut}(N))$. That is, $g \in \mathrm{Aut}(B)$ acts by $\phi \mapsto \phi \circ g$ and $f \in \mathrm{Aut}(N)$ acts by $\phi \mapsto c_f \circ \phi$, where c_f is conjugation by f. That is, $(c_f \circ \phi)(b) = f\phi(b)f^{-1}$. Then, we see that every orbit for this action gives isomorphic groups $N \rtimes_\phi B$. Note that the action of $\mathrm{Aut}(B) \times \mathrm{Aut}(N)$ on $\mathrm{Hom}(B, \mathrm{Aut}(N))$ factors through $\mathrm{Aut}(B) \times \mathrm{Aut}(N)/Z(\mathrm{Aut}(N))$.

26. Groups of low, or simple, order

26.1. **Groups of prime order.** We have seen in Corollary 4.0.9 that all such groups are cyclic. By Example 8.1.2 the unique cyclic group up to isomorphism of order p is $\mathbb{Z}/p\mathbb{Z}$.

26.2. **Groups of order p^2.** You proved in Assignment 7 that every such group is abelian. By the structure theorem it is either isomorphic to $\mathbb{Z}/p^2\mathbb{Z}$ or to $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$.

26.3. **Groups of order pq, $p < q$.** This case was discussed in § 25.1 above. We summarize the results: there is a unique abelian group of order pq. If $p \nmid (q-1)$ then every group of order pq is abelian. If $p|(q-1)$ there is a unique non-abelian group up to isomorphism; it can be taken as any non trivial semi-direct product $\mathbb{Z}/p\mathbb{Z} \ltimes \mathbb{Z}/q\mathbb{Z}$.

27. Groups of order $1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 13, 14, 15$

The results about groups of prime order and of order $pq, p \leq q$ allow us to determine the following possibilities:

order	abelian groups	non-abelian groups
1	$\{1\}$	
2	$\mathbb{Z}/2\mathbb{Z}$	
3	$\mathbb{Z}/3\mathbb{Z}$	
4	$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z}$	
5	$\mathbb{Z}/5\mathbb{Z}$	
6	$\mathbb{Z}/6\mathbb{Z}$	S_3
7	$\mathbb{Z}/7\mathbb{Z}$	
9	$\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$, $\mathbb{Z}/9\mathbb{Z}$	
10	$\mathbb{Z}/10\mathbb{Z}$	D_{10}
11	$\mathbb{Z}/11\mathbb{Z}$	
13	$\mathbb{Z}/13\mathbb{Z}$	
14	$\mathbb{Z}/14\mathbb{Z}$	D_{14}
15	$\mathbb{Z}/15\mathbb{Z}$	

28. Groups of order 8

We know already the structure of abelian groups of order 8: $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$, $\mathbb{Z}/8\mathbb{Z}$. We also know two non-isomorphic non-abelian groups of order 8: D_8, Q (in Q there are six elements of order 4, while in D_8 there are two).

We prove that every non-abelian group G of order 8 is isomorphic to either D_8 or Q. Suppose that G has a non-normal subgroup of order 2, then the kernel of the coset representation $G \longrightarrow S_4$ is trivial. Thus, G is a 2-Sylow subgroup of S_4, but so is D_8. Since all 2-Sylow subgroups are conjugate, hence isomorphic, we conclude that $G \cong D_8$.

Thus, assume that G doesn't have a non-normal subgroup of order 2. Consider the center $Z(G)$ of G. We claim that the center has order 2. Indeed, otherwise $G/Z(G)$ is of order 2 hence cyclic. But $G/Z(G)$ is never cyclic (seen in assignments).

We now claim that $Z(G) = \{1, z\}$ is the unique subgroup of G of order 2. Indeed, if $\{1, h\} = H < G$ of order 2 it must be normal by hypothesis. Then, for every $g \in G$, $ghg^{-1} = h$, i.e. $h \in Z(G)$. It follows that every element x in G apart from 1 or z has order 4, and so every such x satisfies $x^2 = z$. Rename z to -1 and the rest of the elements (which are of order 4 so come in pairs) are then $i, i^{-1}, j, j^{-1}, k, k^{-1}$. Since $i^2 = j^2 = k^2 = -1$ we can write $i^{-1} = -i$, etc.

Note that the subgroup $\langle i, j\rangle$ must be equal to G and so i and j do not commute. Thus, $ij \neq 1, -1, i, -i, j, -j$ (for example, $ij = -i$ implies that $j = (-i)^2 = -1$ and so commutes with i). Without loss of generality $ij = k$ and then $ji = -k$ (because the only other possibility is $ji = k$ which gives $ij = ji$). We therefore get the relations (the new ones are easy consequences):

$$G = \{\pm 1, \pm i, \pm j, \pm k\}, \quad i^2 = j^2 = k^2 = -1,\ ij = -ji = k.$$

This determines completely the multiplication table of G which is identical to that of Q. Thus, $G \cong Q$.

29. Groups of order 12

We know that the abelian groups are $\mathbb{Z}/12\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/6\mathbb{Z}$. We are also familiar with the groups A_4 and D_{12}. One checks that in A_4 there are no elements of order 6 so these two groups are not isomorphic.

Note that in A_4 the 4-Sylow subgroup is normal (it is $\{1,(12)(34),(13)(24),(14)(23)\}$), and the 3-Sylow is not. Note that in D_{12} the 3-Sylow is normal (it is $\{1,x^2,x^4\}$, the rest are 6 reflections and the rotations x,x^3,x^5).

In a non-abelian group of order $12=2^23$, either the 3-Sylow is normal or the 2-Sylow is normal, but not both (if both are, prove the group is abelian).

We conclude that each non-abelian group is the semi direct product of a group of order 4 and a group of order 3. Indeed, one checks that $A_4=(\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z})\rtimes\mathbb{Z}/3\mathbb{Z}$, $D_{12}=(\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z})\ltimes\mathbb{Z}/3\mathbb{Z}$. Let us then consider a semi-direct product $\mathbb{Z}/4\mathbb{Z}\ltimes\mathbb{Z}/3\mathbb{Z}$ (show that every semi-direct product $\mathbb{Z}/4\mathbb{Z}\rtimes\mathbb{Z}/3\mathbb{Z}$ is actually a direct product and so is commutative). Here $1\in\mathbb{Z}/4\mathbb{Z}$ acts on $\mathbb{Z}/3\mathbb{Z}$ as multiplication by -1. This gives a non-abelian group with a cyclic group of order 4 that is therefore not isomorphic to the previous groups. Call it T.

The proof that these are all the non-abelian groups of order 12 is easy given the results of § 25.1. We already know that every such group is a non-trivial semi-direct product $(\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z})\rtimes\mathbb{Z}/3\mathbb{Z},(\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z})\ltimes\mathbb{Z}/3\mathbb{Z}$ or $\mathbb{Z}/4\mathbb{Z}\ltimes\mathbb{Z}/3\mathbb{Z}$.

A non-trivial homomorphism $\mathbb{Z}/3\mathbb{Z}\longrightarrow\mathrm{Aut}(\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z})=\mathrm{GL}_2(\mathbb{F}_2)\cong S_3$ corresponds to an element of order 3 in S_3. All those elements are conjugate and by § 25.1 all these semi-direct products are isomorphic.

A non-trivial homomorphism $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}\longrightarrow\mathrm{Aut}(\mathbb{Z}/3\mathbb{Z})\cong\mathbb{Z}/2\mathbb{Z}$ is determined by its kernel which is a subgroup of order 2 = line in the 2-dimensional vector space $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ over $\mathbb{Z}/2\mathbb{Z}$. The automorphism group of $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}$ acts transitively on lines and by § 25.1 all these semi-direct products are isomorphic.

A non-trivial homomorphism $\mathbb{Z}/4\mathbb{Z}\longrightarrow\mathrm{Aut}(\mathbb{Z}/3\mathbb{Z})\cong\mathbb{Z}/2\mathbb{Z}$ is uniquely determined.

Part 7. Composition series, the Jordan-Hölder theorem and solvable groups

30. Composition series

30.1. **Two philosophies.** In the study of finite groups one can sketch two broad philosophies:

The first one, that we may call the *"Sylow philosophy"* (though such was not made by Sylow, I believe), is given a finite group to study its p-subgroups and then study how they fit together. Sylow's theorems guarantee that the size of p-subgroup is as big as one can hope for, guaranteeing the first step can be taken. The theory of p-groups, the second step, is a beautiful and powerful theory, which is quite successful. I know little about a theory that tells us how p-groups fit together.[15]

The second philosophy, that one may call the *"Jordan-Hölder philosophy"*, suggests given a group G to find a non-trivial normal subgroup N in G and study the possibilities for G given N and G/N. The first step then is to hope for the classification of all finite simple groups. Quite astonishingly, this is possible and was completed towards the end of the last (20^{th}) century.

The second step is figuring out how to create groups G from two given subgroups N and H such that N will be a normal subgroup of G and H isomorphic to G/H. There is a lot known here. We have seen one machinery, the semi-direct product $N \rtimes H$.

30.2. **Composition series.** Let G be a finite group. A *composition series* for G is a sequence of distinct subgroups

$$\underline{G}_\bullet := \{e\} = G_0 \lhd G_1 \lhd \ldots \lhd G_t = G$$

(note that $G_i \lhd G_{i+1}$ but we do not require that $G_i \lhd G$), such that G_{i+1}/G_i is a simple group for every i. If the series just satisfies the normality+ distinct condition, but without requiring the quotient to be simple, then we call it a *normal series*.

Given a normal series $\underline{G}_\bullet = \{e\} = G_0 \lhd G_1 \lhd \ldots \lhd G_t = G$, we say that a normal series $\underline{H}_\bullet = \{e\} = H_0 \lhd H_1 \lhd \ldots \lhd H_s = G$ is a refinement of $\underline{G}_\bullet$ if all the groups G_i appear among the groups H_j. Then, a composition series is a normal series that cannot be refined. (The statement that we can form $G_i \lhd H \lhd G_{i+1}$ with distinct quotients is equivalent via the third isomorphism theorem with the statement that G_{i+1}/G_i is not simple).

One call the quotient groups G_i/G_{i-1} of a composition series, the *composition factors*.

31. The Jordan-Hölder theorem

Theorem 31.0.1. *Let G be a finite group, $G \neq \{e\}$. Then G has a composition series. Moreover, if $\{e\} = G_0 \lhd G_1 \lhd \ldots \lhd G_t = G$ and $\{e\} = H_0 \lhd H_1 \lhd \ldots \lhd H_s = G$ are composition series. Let $\mathfrak{g}_i = G_i/G_{i-1}$ and $\mathfrak{h}_i = H_i/H_{i-1}$ be the composition factors. Then $s = t$ and there is a permutation π of $\{1, 2, \ldots, s\}$ such that*

$$\mathfrak{g}_i \cong \mathfrak{h}_{\pi(i)}, \quad \forall i.$$

That is, the composition factors (with their multiplicities) are uniquely determined.

The proof is not particularly difficult, but will not be covered in this course. It can be found, for example, in the book J. Rotman/*Introduction to the theory of groups.*

[15] The class of nilpotent groups turns out to be the same as the class of groups that are a direct product of their p-Sylow subgroups.

32. Solvable groups

A finite group G is called *Solvable* (or *Soluble*) if it has a normal series such that the composition factors are abelian. It is not hard to prove that G is solvable if and only if there is a composition series $\underline{G}_\bullet$ such that the quotient groups G_i/G_{i-1} are cyclic of prime order.

Example 32.0.2. (1) Every abelian group is solvable.
(2) Every p-group G is solvable. Indeed we proved that there is a normal series $G = G_0 \supset G_1 \supset \cdots \supset G_r = \{e\}$ such that $G_i \lhd G_{i-1}$ (even $G_i \lhd G$ but that is not needed right now) and G_{i-1}/G_i is of order p, hence cyclic abelian.
(3) The group S_3 is solvable. It has the series $S_3 \supset A_3 \supset \{e\}$ with quotients isomorphic to $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$.

Proposition 32.0.3. *Let G be a finite group and $N \lhd G$. Then G is solvable if and only if N and G/N are solvable.*

Proof. Assume that N and G/N are solvable,

$$G/N = H_0' \supset H_1' \supset \cdots \supset H_r' = \{e\}, \quad N = N_0 \supset N_1 \supset \cdots \supset N_s = \{e\},$$

with abelian quotients. Let $H_i = \pi_N^{-1}(H_i')$. Then we have a sequence of groups

$$G = H_0 \supset H_1 \supset \cdots \supset H_r = N_0 \supset N_1 \supset \cdots \supset N_s = \{e\}.$$

It follows from the third isomorphism theorem that $H_i \lhd H_{i-1}$ and $H_{i-1}/H_i \cong H_{i-1}'/H_i'$ and in particular is abelian. Thus, G is solvable.

Let G be solvable,

$$G = G_0 \supset G_1 \supset \cdots \supset G_n = \{e\},$$

with abelian quotients. Let N be a subgroup of G. Consider the series

$$N = N \cap G_0 \supset N \cap G_1 \ldots N \cap G_n = \{e\}.$$

We claim that $N \cap G_i \lhd N \cap G_{i-1}$ and that the quotient is abelian (it follows then that N is solvable; no need to assume N is normal). Consider the homomorphism $f : G_{i-1} \longrightarrow G_{i-1}/G_i$ and its restriction

$$g := f|_{N \cap G_i} : N \cap G_{i-1} \longrightarrow G_{i-1}/G_i.$$

By the first isomorphism theorem $\mathrm{Ker}(g) = \mathrm{Ker}(f) \cap (N \cap G_{i-1}) = N \cap G_i$ is normal in $N \cap G_{i-1}$ and $N \cap G_{i-1}/N \cap G_i = N \cap G_{i-1}/\mathrm{Ker}(g) \cong \mathrm{Im}(g)$. Since the image of g is a subgroup of the abelian group G_{i-1}/G_i, it is abelian.

Assume now that N is normal and let $\pi := \pi_N : G \longrightarrow G/N$ be the canonical map. We have a sequence of subgroups

$$G/N = \pi(G_0) \supset \pi(G_1) \supset \cdots \supset \pi(G_n) = \{e\}.$$

We claim that $\pi(G_i) \lhd \pi(G_{i-1})$ and that $\pi(G_{i-1})/\pi(G_i)$ is abelian. Indeed, let $x \in \pi(G_{i-1}), y \in \pi(G_i)$. We need to prove that $xyx^{-1} \in \pi(G_i)$. Choose $X \in G_{i-1}, Y \in G_i$ such that $\pi(X) = x, \pi(Y) = y$. Then $XYX^{-1} \in G_i$, because $G_i \lhd G_{i-1}$, and $\pi(XYX^{-1}) = xyx^{-1}$. It follows that $xyx^{-1} \in \pi(G_i)$.

Consider now the induced homomorphism $f : G_{i-1} \xrightarrow{\pi} \pi(G_{i-1}) \longrightarrow \pi(G_{i-1})/\pi(G_i)$. It is surjective. The kernel of f contains G_i. We can therefore argue as follows: $\pi(G_{i-1})/\pi(G_i) \cong G_{i-1}/\mathrm{Ker}(f) \cong (G_{i-1}/G_i)/(\mathrm{Ker}(f)/G_i)$ and so $\pi(G_{i-1})/\pi(G_i)$ is a quotient of the abelian group G_{i-1}/G_i and hence abelian.

□

Example 32.0.4. *Every group of order less than* 60 *is solvable.* To show that we argue by induction on the order of the group. Using Proposition 32.0.3, it is enough to prove that a non-abelian group of order less than 60 is not simple.[16] We know already (by results proven in class and in assignments) that groups of order p are abelian and of order p^a $(a > 1)$, pq, pqr and p^2q are not simple. The numbers less than 60 not of this form are $24, 36, 40, 48, 54, 56$. We saw that groups of order 24 (in class) $36, 40, 48, 54$ (in an assignment) are not simple. It remains to show that a group G of order $56 = 2^3 \cdot 7$ is not simple.

Suppose that the 7-Sylow of G is not normal. Then there are 8 7-Sylow subgroups. These already account for a set S consisting of $1 + (7-1) \times 8 = 49$ distinct elements of G. If P is a 2-Sylow subgroup then $P \cap S = \{e\}$ and it follows that $P = G \setminus S \cup \{e\}$. Since this holds for any 2-Sylow subgroup, we conclude that P is the unique 2-Sylow subgroup and hence normal.

[16] Note that A_5 is a simple non-abelian group of order 60 and hence non-solvable.

The motivation for the study of solvable groups comes from Galois theory. Let $f(x) = x^n + a_{n_1}x^{n-1} + \cdots + a_0$ be an irreducible polynomial with rational coefficients. In Galois theory one associates to f a finite group $G_f \subseteq S_n$, called the Galois group of f. One of Galois's main achievements is to prove that one can solve f in radicals (meaning, express the solutions of f using operations as taking roots, adding and multiplying) if and on if G_f is a solvable group.

It follows that there are formulas in radicals to solve equations of degree ≤ 4 (every group that can possibly arise as G_f has order less than 60, hence is solvable). On the other hand, one can produce easily an equation f of degree 5 such that $G_f = S_5$, hence is not solvable.

Part 8. Rings

33. Basic definitions

Definition 33.0.5. A *ring* R is an abelian group together with a multiplication map,

$$R \times R \longrightarrow R, \quad (x, y) \mapsto xy,$$

and an element $1 \in R$, such that the following holds:

(1) (Associativity) $(xy)z = x(yz)$ for all $x, y, z \in R$.
(2) (Distributivity) $x(y + z) = xy + xz$ and $(x + y)z = xz + yz$ for all $x, y, z \in R$.
(3) (Identity) $1x = x1 = x$ for all $x \in R$.

Note that we insist on R having a (specified) identity element 1. In that our conventions differ from Dummit and Foote's.

Two formal and easy consequences of the definition are:

(1) $0x = x0 = 0$;
(2) $(-1)x = -x = x(-1)$.

A ring is called *commutative* if $xy = yx$ for all $x, y \in R$. A non-zero element $x \in R$ is called a *zero divisor* if for some non-zero element y we have $xy = 0$ or $yx = 0$. A non zero commutative ring with no zero divisors if called an *integral domain*.

An element $x \in R$ is called a *unit* if $\exists y \in R$ such that $xy = yx = 1$. The units form a group under multiplication denoted $R^\times$.

Example 33.0.6. Let k be a field and V a vector space over k. One easily verifies that the collection of linear maps from V to itself, $\mathrm{End}(V)$, is a ring, where multiplication is composition of linear maps. If V has finite dimension then if $x, y \in \mathrm{End}(V)$ and $xy = 1$ then also $yx = 1$ and so x is a unit. However, suppose that $V = \{(a_1, a_2, \dots) : a_i \in k\}$ and x is the linear map $(a_1, a_2, a_3, \dots) \mapsto (a_2, a_3, \dots)$. Then x is not injective and so there is no $y$such that $yx = 1$. On the other hand, if y is the linear map $(a_1, a_2, a_3, \dots) \mapsto (0, a_1, a_2, \dots)$ then $xy = 1$. This example explains why we insist on $xy = yx = 1$ in the definition of a unit.

A non zero ring R is called a *division ring* (or a *skew field*) if $R^\times = R - \{0\}$, i.e., every non-zero element is a unit. If R is also commutative then R is called a *field*.

A subset $I \subseteq R$ is called a *two-sided ideal* of R (or simply, an *ideal* of R) , denoted $I \lhd R$, if I is a subgroup and for all $r \in R$ we have both inclusions

$$Ir \subseteq I, \quad rI \subseteq I.$$

A left (resp. right) ideal is defined the same only that one requires just $rI \subseteq I$ (resp. $Ir \subseteq I$).

Proposition 33.0.7. *Let R be a ring and $I \lhd R$ a two sided ideal. The quotient group R/I has a canonical ring structure given by*

$$(r + I)(s + I) = rs + I,$$

with identity element $1 + I$.

Proof. We first check that multiplication is well defined. Any two other representatives for the cosets are of the form $r + i_1, s + i_2$ for $i_1, i_2 \in I$. Then $(r + i_1 + I)(s + i_2 + I)$ is equal by definition to $(r + i_1)(s + i_2) + I = rs + i_1 s + r i_2 + i_1 i_2 + I = rs + I$, using that I is an ideal.

The rest of the axioms follow mechanically from the fact that they hold in R. For example, letting $\overline{r} = r + I$, we have

$$\begin{aligned} \overline{r}(\overline{x} + \overline{y}) &= \overline{r} \cdot \overline{x + y} \\ &= \overline{r(x + y)} \\ &= \overline{rx + ry} \\ &= \overline{rx} + \overline{ry} \\ &= \overline{r} \cdot \overline{x} + \overline{r} \cdot \overline{y}. \end{aligned}$$

□

34. Key Examples of Rings

34.1. **The zero ring.** This is the ring $R = \{0\}$. Note that in this ring $1 = 0$. This is the case excluded when defining integral domains, fields or division rings. Note that to say that R is a non-zero ring (i.e., R is not the zero ring) is equivalent to saying that $1 \neq 0$ in R.

34.2. **The integers and the integers modulo n.** The primal example is the integers

$$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, 3, \dots\}$$

with the usual addition and multiplication. This is an integral domain and $\mathbb{Z}^\times = \{\pm 1\}$.

If R is any commutative ring and $r \in R$ we can define $(r) = Rr = rR = \{ra : a \in R\}$.

Lemma 34.2.1. *The set (r) is an ideal, called a* principal *ideal.*

Proof. We first check it is a subgroup. Indeed, $0 = 0r \in Rr$, if $ar, br \in Rr$ then $ar + br = (a+b)r$ is in Rr and $-(ar) = -1(ar) = (-1 \cdot a)r = (-a)r \in Rr$.

Next, let $ar \in Rr$ and $b \in R$ then $b(ar) = (ba)r \in Rr$ and $(ar)b = abr \in Rr$ (here we use the commutativity of R in an essential way). Thus, Rr is an ideal. □

As an application, we find the ideals $(n) = \mathbb{Z}n = \{\dots, -2n, -n, 0, n, 2n, 3n, \dots\}$ of $\mathbb{Z}$. One can prove that every ideal of $\mathbb{Z}$ has such a form. This is an example of PID, as we shall see later.

Using Proposition 33.0.7 we find the rings

$$\mathbb{Z}/(n) = \mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \overline{1}, \dots, \overline{n-1}\}$$

(already familiar to us as abelian groups), where we let $\overline{i} = i + n\mathbb{Z}$. Note that this is a commutative ring with n elements. If n is not prime, say $n = ab$, then $\overline{a}\overline{b} = \overline{n} = \overline{0}$ and we find that $\mathbb{Z}/n\mathbb{Z}$ has zero divisors. If, on the other hand, $n = p$ is prime $\mathbb{Z}/p\mathbb{Z}$ doesn't have zero divisors because $\overline{a}\overline{b} = \overline{0}$ implies that $p|ab$ and so, w.l.o.g., $p|a$, giving $\overline{a} = \overline{0}$. It follows from the next proposition that $\mathbb{Z}/p\mathbb{Z}$ is a field.

Proposition 34.2.2. *Let R be an integral domain with finitely many elements then R is a field.*

Proof. Let $a \in R$ be a non zero element. The map $R \longrightarrow R, x \mapsto ax$ is injective: $ax = ay \Rightarrow a(x-y) = 0 \Rightarrow x = y$. Since R is a finite set, the map is also surjective and so there is an x such that $ax = 1$. □

The units of $\mathbb{Z}/n\mathbb{Z}$ are $\mathbb{Z}/n\mathbb{Z}^\times = \{\overline{a} : 1 \leq a < n, (a,n) = 1\}$. This is a set familiar to us; recall that its cardinality is denoted $\varphi(n)$.

34.3. **Matrices over R.** Let R be a commutative ring. Then $M_n(R)$ denote the $n \times n$ matrices with entries in R under matrix addition and multiplication. This is a ring whose units are denoted $\mathrm{GL}_n(R)$; a matrix is invertible in R if and only if its determinant belongs to $R^\times$. Indeed, the usual determinant properties show that for any commutative ring if $AB = I$ then $\det(A) \cdot \det(B) = 1$ and hence $\det(B) \in R^\times$. Conversely, we have $A \cdot \mathrm{adj(A)} = \det(\mathrm{A}) \cdot \mathrm{I}$ and so, if $\det(A) \in R^\times$ we have an inverse: $A^{-1} = \det(A)^{-1}\mathrm{adj(A)}$.

If $n > 1$ and R is not the zero ring, it is in a non-commutative ring and has zero divisors. Indeed

$$\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)^2 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right),$$

and

$$\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right), \qquad \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right).$$

34.4. **Polynomial and power series rings.** Let R be a commutative ring and x a formal symbol. The ring of polynomials over R, $R[x]$, is the expressions of the form $a_0 + a_1 x + \cdots + a_n x^n$, $a_i \in R$ (where n may be different for each expression). We allow zero coefficients; we may therefore define addition by

$$\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{n} b_i x^i = \sum_{i=0}^{n} (a_i + b_i) x^i.$$

Multiplication is defined by

$$(\sum_{i=0}^{n} a_i x^i)(\sum_{i=0}^{m} b_i x^i) = \sum_{i=0}^{m+n} (\sum_{j=0}^{i} a_j b_{i-j}) x^i.$$

In general, due to zero divisors, there is no elegant description of the units of this ring. However,

Proposition 34.4.1. *Let R be an integral domain. The units of $R[x]$ are $R^\times$.*

Proof. Suppose that $\sum_{i=0}^{n} a_i x^i$ is a unit, $a_n \neq 0$, and $\sum_{i=0}^{m} b_i x^i$ is the inverse and $a_m \neq 0$. The coefficient of x^{n+m} is $a_n b_m$, which is not zero because R is an integral domain. Thus, we must have $n + m = 0$ and so $n = m = 0$. That is, $\sum_{i=0}^{n} a_i x^i = a_0$, $\sum_{i=0}^{m} b_i x^i = b_0$, and $a_0 b_0 = 1$; that is $a_0 \in R^\times$. □

We may define two related rings: the ring $R[\![x]\!]$ of Taylor series, whose general element is $\sum_{i=0}^{\infty} a_i x^i$, $a_i \in R$, and the ring $R(\!(x)\!)$ of Laurent series, whose general element is $\sum_{i=N}^{\infty} a_i x^i$, $a_i \in R$, where N is an integer that depends on the element and may be negative. We have

$$R[x] \subset R[\![x]\!] \subset R(\!(x)\!).$$

Addition and multiplication are defined by the same formulas. We have

Proposition 34.4.2. *The units of $R[\![x]\!]$ are $\{\sum_{i=0}^{\infty} a_i x^i : a_0 \in R^\times\}$. The ring $R(\!(x)\!)$ is a field; every non-zero element is a unit.*

34.5. **Hamilton's quaternions.** Recall the quaternion group of 8 elements:

$$\left\{ \pm \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \pm \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right), \pm \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right), \pm \left(\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right) \right\} \subseteq M_2(\mathbb{C}).$$

We denoted these elements, respectively, $\pm 1, \pm i, \pm j, \pm k$. Let $\mathbb{F} \subseteq \mathbb{R}$ be a field, e.g., $\mathbb{F} = \mathbb{Q}, \mathbb{R}$. The *quaternion algebra over* $\mathbb{F}$ is the set

$$\{a + bi + cj + dk : a, b, c, d \in \mathbb{F}\},$$

with matrix multiplication and addition. Namely, the matrices

$$\begin{aligned} \left\{ \left(\begin{smallmatrix} a+bi & c+di \\ -c+di & a-bi \end{smallmatrix}\right) : a, b, c, d \in \mathbb{F} \right\} &= \left\{ \left(\begin{smallmatrix} a+bi & c+di \\ -\overline{(c+di)} & \overline{a+bi} \end{smallmatrix}\right) : a, b, c, d \in \mathbb{F} \right\} \\ &= \left\{ \left(\begin{smallmatrix} A & B \\ -\overline{B} & \overline{A} \end{smallmatrix}\right) : A = a + bi, B = c + di, a, b, c, d \in \mathbb{F} \right\} \\ &\left(\underset{\text{if } \mathbb{F}=\mathbb{R}}{=} \left\{ \left(\begin{smallmatrix} A & B \\ -\overline{B} & \overline{A} \end{smallmatrix}\right) : A, B \in \mathbb{C} \right\} \right) \end{aligned}$$

Definition 34.5.1. Let R be a ring. A subset $R_1 \subseteq R$ is called a *subring* if it is a subgroup of R, closed under multiplication and $1 \in R_1$.

It follows immediately that a subring is a ring in its own right.

Proposition 34.5.2. *The quaternions over $\mathbb{F}$, denoted $\mathbb{H}_\mathbb{F}$, are a subring of $M_2(\mathbb{C})$. Moreover, they form a non-commutative division ring.*

Proof. We note that if $z_1 = a_1 + b_1 i, z_2 = a_2 + b_2 i$, where $a_1, a_2, b_1, b_2 \in \mathbb{F}$ – we say that $z_i \in \mathbb{F}[i]$ – then $z_1 + z_2, z_1 z_2, \overline{z_1}$ are also in $\mathbb{F}[i]$. Using the usual properties of conjugation of complex numbers we find that

$$\left(\begin{smallmatrix} A & B \\ -\overline{B} & \overline{A} \end{smallmatrix}\right) + \left(\begin{smallmatrix} A' & B' \\ -\overline{B'} & \overline{A'} \end{smallmatrix}\right) = \left(\begin{smallmatrix} A+A' & B+B' \\ -\overline{(B+B')} & \overline{A+A'} \end{smallmatrix}\right),$$

which shows closure under addition. Also $0 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 0 \\ -\overline{0} & \overline{0} \end{smallmatrix}\right)$ is in $\mathbb{H}_\mathbb{F}$ and $-\left(\begin{smallmatrix} A & B \\ -\overline{B} & \overline{A} \end{smallmatrix}\right) = \left(\begin{smallmatrix} -A & -B \\ \overline{B} & -\overline{A} \end{smallmatrix}\right) = \left(\begin{smallmatrix} -A & -B \\ -\overline{-B} & \overline{-A} \end{smallmatrix}\right)$, which shows closure under additive inverse. Thus, $\mathbb{H}_\mathbb{F}$ is a subgroup of $M_2(\mathbb{C})$.

Note that $1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ -\overline{0} & \overline{1} \end{smallmatrix}\right)$ is in $\mathbb{H}_{\mathbb{F}}$ and

$$\left(\begin{smallmatrix} A & B \\ -\overline{B} & \overline{A} \end{smallmatrix}\right)\left(\begin{smallmatrix} C & D \\ -\overline{D} & \overline{C} \end{smallmatrix}\right) = \left(\begin{smallmatrix} AC-B\overline{D} & AD+B\overline{C} \\ -(AD+B\overline{C}) & \overline{AC-B\overline{D}} \end{smallmatrix}\right).$$

Hence, $\mathbb{H}_{\mathbb{F}}$ is closed under multiplication too.

Non-commutativity is familiar to us: $ij = -ji$ et cetera. To show $\mathbb{H}_{\mathbb{F}}$ is a division ring, note that if $M = \left(\begin{smallmatrix} A & B \\ -\overline{B} & \overline{A} \end{smallmatrix}\right)$ then $\det(M) = |A|^2 + |B|^2$ and so if $M \neq 0$ then $\det(M) \neq 0$. Now, $M^{-1} = \frac{1}{|A|^2+|B|^2}\left(\begin{smallmatrix} \overline{A} & -B \\ \overline{B} & A \end{smallmatrix}\right)$, which is again an element of $\mathbb{H}_{\mathbb{F}}$. □

34.6. **The ring of quotients.** The ring of quotients is a general construction that allows embedding a commutative integral domain in a field; moreover, that field is the smallest possible. A case to keep in mind is the ring $\mathbb{Z}$ and the field $\mathbb{Q}$. If $\mathbb{Z} \subset \mathbb{F}$ and $\mathbb{F}$ is a field, then for every non-zero $n \in \mathbb{Z}$ and $m \in \mathbb{Z}$ we have the element $m \times n^{-1}$ in $\mathbb{F}$. In this sense $\mathbb{Q} \subseteq \mathbb{F}$. This discussion also provides a clue as to how to construct the field of quotients.

Let R be a commutative integral domain. Define a relation on $R \times (R - \{0\})$ by

$$(a, b) \sim (c, d) \quad \text{if} \quad ad - bc = 0. \tag{34.1}$$

Theorem 34.6.1. *The relation (34.1) is an equivalence relation. One denotes the equivalence classes by $Q(R)$. The operations*

$$(a, b) + (c, d) = (ad + bc, bd), \qquad (a, b)(c, d) = (ac, bd),$$

are well defined. Under these operations $Q(R)$ is a field. The map $R \to Q(R)$, $r \mapsto (r, 1)$ is injective and R may be viewed as a subring of $Q(R)$.

Proof. Straight from the definition we get that $(a, b) \sim (a, b)$ and that if $(a, b) \sim (c, d)$ then $(c, d) \sim (a, b)$. Suppose that $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$. Then $d(af - be) = (ad - bc)f + (cf - de)b = 0$. Since $d \neq 0$, and R is an integral domain, we have that $af - be = 0$ and so $(a, b) \sim (e, f)$.

We denote from now on a pair (a, b) by $\frac{a}{b}$. Then $(a, b) \sim (c, d)$, that is $\frac{a}{b} \sim \frac{c}{d}$, if $ad - bc = 0$. The addition and multiplication rules are familiar:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

We verify that they are well defined. We need to show that if $\frac{a}{b} \sim \frac{a_1}{b_1}$, $\frac{c}{d} \sim \frac{c_1}{d_1}$, then $\frac{ad+bc}{bd} \sim \frac{a_1d_1+b_1c_1}{b_1d_1}$ and $\frac{ac}{bd} \sim \frac{a_1c_1}{b_1d_1}$. This amounts to the identities $(ad + bc)(b_1d_1) = (ab_1)dd_1 + bb_1(cd_1) = a_1bdd_1 + bb_1c_1d = (a_1d_1 + b_1c_1)(bd)$ and $(ac)(b_1d_1) = (ab_1)(cd_1) = a_1bc_1d = (a_1c_1)(bd)$.

One now checks that the operations are commutative, associative and distributive. The verification is formal and straightforward. For example: $\left(\frac{a}{b} + \frac{c}{d}\right)\frac{e}{f} = \frac{ad+bc}{bd} \cdot \frac{e}{f} = \frac{ade+bce}{bdf}$ and $\frac{a}{b} \cdot \frac{e}{f} + \frac{c}{d} \cdot \frac{e}{f} = \frac{ae}{bf} + \frac{ec}{df} = \frac{aedf+cebf}{bdff} \sim \frac{ade+bce}{bdf}$.

The zero element is the equivalence class of $\frac{0}{1}$ (it consists of the elements $\frac{0}{a}, a \in R$), and the identity element is the equivalence class of $\frac{1}{1}$ (it consists of the elements $\frac{a}{a}, a \in R, a \neq 0$) . The additive inverse of $\frac{a}{b}$ is $\frac{-a}{b}$. Indeed $\frac{a}{b} + \frac{-a}{b} = \frac{ab-ab}{b^2} = \frac{0}{b^2} \sim \frac{0}{1}$. It follows that $Q(R)$ is a commutative ring.

Finally, if $\frac{a}{b} \neq 0$ then $a \neq 0$ and so $\frac{b}{a}$ is defined. We have $\frac{a}{b} \cdot \frac{b}{a} = \frac{ab}{ab} \sim \frac{1}{1} = 1$. Thus, $Q(R)$ is a field. □

Example 34.6.2. We have $Q(\mathbb{Z}) = \mathbb{Q}$ and for any field $\mathbb{F}$ we have $Q(F[x]) = F(x)$. Also, for any commutative integral domain R we have $Q(R[x]) = Q(R)(x)$. In section 35 we shall see that, in a precise sense, if R is a field then $R = Q(R)$.

35. Ring homomorphisms and the isomorphism theorems

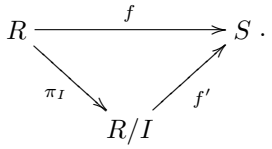
Definition 35.0.3. A *ring homomorphism* $f : R \to S$ is a function satisfying: (i) $f(r_1+r_2) = f(r_1)+f(r_2)$; (ii) $f(r_1r_2) = f(r_1)f(r_2)$ and (iii) $f(1_R) = 1_S$.

Example 35.0.4. Let $I \lhd R$ be a two sided ideal. The canonical map

$$\pi_R : R \longrightarrow R/I, \qquad \pi_I(a) = a + I,$$

is a ring homomorphism. Indeed, this is just $(a+I)+(b+I) = a+b+I$, $(a+I)(b+I) = ab+I$, and $1+I$ being the identity of R/I. We see that if we want π_I to be a ring homomorphism this forces the definition of addition and multiplication on the cosets R/I.

Theorem 35.0.5. *Let $f : R \longrightarrow S$ be a ring homomorphism. Then $J := \mathrm{Ker}(f)$ is a two sided ideal of R called the* kernel *of f. If I is a two sided ideal of R such that $I \subseteq J$ there is a unique ring homomorphism $f' : R/I \longrightarrow S$ such that the following diagram is commutative:*



The kernel of f' is J/I.

Context: Two sided ideals are in analogy to normal subgroups. We can take quotients by such ideals. If $f : R \to S$ is a ring homomorphism, $K \lhd S$ then $f^{-1}(K) \lhd R$. If f is *surjective* and $K \lhd R$ then $f(K) \lhd S$. In particular, it follows that $J/I = \pi_I(J)$ is an ideal of R/I (though this also follows from the first part of the Theorem applied to f').

Proof. We already know that $\mathrm{Ker}(f)$ is a subgroup of R. If $r \in R$ and $a \in \mathrm{Ker}(f)$ then $f(ra) = f(r)f(a) = f(r) \cdot 0 = 0$ and likewise $f(ar) = 0$. Thus, $\mathrm{Ker}(f)$ is an ideal of R.

Define $f' : R \to S$ by $f'(r+I) = f(r)$. This is well defined: if $i \in I$ then, because $I \subseteq \mathrm{Ker}(f)$, $f'(r+i+I) = f(r+i) = f(r)+f(i) = f(r)$. It follows immediately that f' is a ring homomorphism. For example, $f'((a+I)(b+I)) = f'(ab+I) = f(ab) = f(a)f(b) = f'(a+I)\, f'(b+I)$.

Note that $f'(\pi_I(a)) = f'(a+I) = f(a)$ so $f' \circ \pi_I = f$. Moreover, $f'(a+I) = 0$ iff $f(a) = 0$. Thus, $f'(a+I) = 0$ iff $a \in J$, and it follows that $\mathrm{Ker}(f') = J/I$.

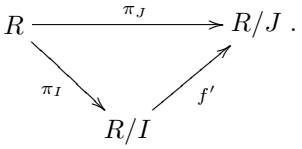
Finally, f' is unique. Suppose that $f'' : R/I \to S$ also satisfies $f'' \circ \pi_I = f$ then $f''(a+I) = f''(\pi_I(a)) = f(a) = f'(a+I)$ and so $f' = f''$. □

Corollary 35.0.6. *If f is surjective and $I = J$ we conclude that $f' : R/\mathrm{Ker}(f) \to S$ is an isomorphism, $R/\mathrm{Ker}(f) \cong S$.*

Corollary 35.0.7. *If $I \subset J$ are two sided ideals of R then*

$$(R/I)/(J/I) \cong R/J.$$

Proof. Apply the Theorem to the homomorphism $\pi_J : R \to R/J$. We get



We have $\mathrm{Ker}(f') = J/I$. By the previous Corollary, $(R/I)/\mathrm{Ker}(f') = (R/I)/(J/I) \cong R/J$. □

Remark 35.0.8. The only ideals of a division ring R (e.g., a field) are 0 and R. Thus, if R is a division ring, S is not the zero ring, and $f : R \to S$ is a ring homomorphism then f is injective. In particular, any ring homomorphism between fields is injective.

Proposition 35.0.9. *Let $f : R \to S$ be a surjective ring homomorphism. There is a bijection between ideals of R containing the kernel of f and ideals of S, given by $I \mapsto f(I)$ (with inverse $J \mapsto f^{-1}(J)$).*

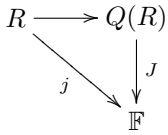
I leave a detailed proof to you. Note that we already know such a bijection exists on the level of subgroups. Thus, the only point to check is that it takes ideals to ideals, which is quite straight forward.

35.1. **The universal property of the ring of quotients.**

Theorem 35.1.1. *Let R be a commutative integral domain. There is a natural injective ring homomorphism*

$$R \longrightarrow Q(R), \qquad r \mapsto (r,1) = \frac{r}{1}.$$

Every element of R is invertible in $Q(R)$. If $\mathbb{F}$ is a field and $j : R \to \mathbb{F}$ is an injective ring homomorphism then there is a unique ring homomorphism $J : Q(R) \to \mathbb{F}$ rendering the following diagram commutative:



Proof. It follows straight from the definitions that $r \mapsto \frac{r}{1}$ is a ring homomorphism. It is injective since $\frac{r}{1} = 0$ iff $r = 0$. We may thus view R as a *subring* of $Q(R)$ as we shall usually do. If $r \in R$ is not zero then $r \cdot \frac{1}{r}$ (more precisely, $\frac{r}{1} \cdot \frac{1}{r}$) is just $1 = \frac{r}{r}$. Thus, every non-zero element of R is invertible in $Q(R)$.

Given j, define J by $J(\frac{r}{s}) = j(r)j(s)^{-1}$. This is well defined: First, if $j(s) \neq 0$ then $j(s)^{-1}$ exists and, second, if $\frac{r}{s} = \frac{r'}{s'}$ (thus $rs' = r's$) then $J(\frac{r}{s}) = j(r)j(s)^{-1} = j(r)j(s')j(s)^{-1}j(s')^{-1} = j(rs')j(s)^{-1}j(s')^{-1} = j(r's)j(s)^{-1}j(s')^{-1} = j(r')j(s)j(s)^{-1}j(s')^{-1} = j(r')j(s')^{-1} = J(\frac{r'}{s'})$.

It is easy to verify that J is a homomorphism and of course $j(r) = J(\frac{r}{1})$. □

35.2. **A useful lemma.**

Lemma 35.2.1. *Let R, S be commutative rings. Let $f : R \to S$ be a ring homomorphism. Let $s \in S$ be any element. There exists a unique ring homomorphism,*

$$F : R[x] \longrightarrow S,$$

such that $F(r) = f(r)$ for $r \in R$ and $F(x) = s$.

Proof. Define

$$F(\sum a_i x^i) = \sum f(a_i)s^i.$$

By definition, $F(r) = f(r)$ for $r \in R$ and $F(x) = s$. It is easy to check that F is a ring homomorphism. □

From now on, all rings are assumed to be commutative

36. More on ideals

Here are some easy properties of ideals:

- If $\{I_\alpha : \alpha \in A\}$ are ideals then so is $\cap_{\alpha\in A} I_\alpha$.
- If I, J are ideals then $I + J = \{i + j : i \in I, j \in J\}$ is an ideal.
- If I, J are ideals then IJ, defined as $\cap_{K\supseteq\{ij:i\in I,j\in J\}}^{K\lhd R} K$, is an ideal. It is the minimal ideal of R containing the set $\{ij : i \in I, j \in J\}$. Note that $IJ \subseteq I \cap J$; an equality does not hold in general. For example, take $I = J = 2\mathbb{Z}$ in the ring $\mathbb{Z}$.
- Let A be any subset of R. *The ideal generated by* A is defined to be $\cap_{K\supseteq A}^{K\lhd R} K$ and is denoted (A) or $\langle A\rangle$. For example, if $A = \{ij : i \in I, j \in J\}$ then $\langle A\rangle$ is the ideal IJ. A very important case is when A contains one element, $A = \{a\}$, then (a) is $Ra = aR$. A *principal ideal* is such an ideal, namely, of the form (a) for some $a \in R$.

  Lemma 36.0.2. *We have* $\langle A\rangle = \{\sum_{i=1}^N r_i a_i : r_i \in R, a_i \in A, N \geq 0\}$ *(by definition, the empty sum is equal to the zero element of* R*).*

  Proof. Certainly any ideal containing A contains the right hand side. Hence, it is enough to prove that the r.h.s. is an ideal. Indeed, given two finite sums we may assume that they involve the same elements $a_i \in A$ by adding zero coefficients $r_i = 0$. Then $\sum r_i a_i + \sum s_i a_i = \sum(r_i + s_i)a_i$ and $r(\sum r_i a_i) = \sum(rr_i)a_i$ and we are done. □

Example 36.0.3. In $\mathbb{Z}$ every ideal is principal, equal to (n) for some $n \in \mathbb{Z}$. The same holds in the ring $\mathbb{Z}[i]$ of Gaussian integers and in the ring of polynomials $\mathbb{F}[x]$ over a field $\mathbb{F}$. This will follow from the fact that the rings $\mathbb{Z}, \mathbb{Z}[i], \mathbb{F}[x]$ are all Euclidean.

In the ring $\mathbb{Z}[\sqrt{-6}]$ the ideal $(2, \sqrt{-6})$ is not principal. In the ring $\mathbb{Q}[x, y]$ (polynomials in two variables with rational coefficients) the ideal (x, y) is not principal.

Definition 36.0.4. An ideal $I \lhd R$ is called *prime* if $I \neq R$ and

$$ab \in I \Rightarrow a \in I \text{ or } b \in I.$$

An ideal $I \lhd R$ is called *maximal* if $I \neq R$ and if J is an ideal containing I then $J = I$ or $J = R$.

Proposition 36.0.5. *The following holds:*

1. *I is prime $\Leftrightarrow$ R/I is an integral domain.*
2. *I is maximal $\Leftrightarrow$ R/I is a field.*
3. *I is maximal $\Rightarrow$ I is prime.*
4. *Every ideal of R is contained in a maximal ideal.*

Proof. (1) I is prime iff $I \neq R$ and $\{ab \in I \Rightarrow a \in I \text{ or } b \in I\}$, i.e., iff R/I is not the zero ring and $\overline{a}\overline{b} = \overline{0} \Rightarrow \overline{a} = 0$ or $\overline{b} = 0$ (where $\overline{a} = a + I$, etc.). That is, I is prime iff R/I is an integral domain.

(2) Suppose that I is maximal. Let $a \notin I$ then $(I, a) = I + (a) = R$ so $1 = ri + sa$ for some $r, s \in R, i \in I$, which gives $\overline{1} = \overline{s} \cdot \overline{a}$. Since any non zero element of R/I is of the form $\overline{a}$ for some $a \notin I$ we conclude that every non-zero element of R/I is invertible and thus R/I is a field.

Suppose that R/I is a field. Let $J \supseteq I$ be an ideal. Then J/I is an ideal of R/I and so is either the zero ideal or equal to R/I. It follows that $J = I$ or $J = R$. Thus, I is a maximal ideal.

(3) If I is maximal R/I is a field, hence an integral domain and therefore I is prime.

(4) Let S be a poset – a partially ordered set. Namely, there is a relation $\preccurlyeq$ defined on S, which is transitive, reflexive and if $x \preccurlyeq y, y \preccurlyeq x$ then $x = y$. A chain in S is a subset S_0 such that if $x, y \in S_0$ then either $x \preccurlyeq y$ or $y \preccurlyeq x$. A subset S_0 has a supremum if there is an element $s \in S$ such that for all $s_0 \in S_0$ we have $s_0 \preccurlyeq s$ and if $t \in S$ and for all $s_0 \in S_0$ we have $s_0 \preccurlyeq t$ then $s \preccurlyeq t$.
Zorn's Lemma. *Let S be a poset in which any chain has a supremum. Then S has a maximal element, namely, an element $z \in S$ such that if $s \in S$ and $z \preccurlyeq s$ then $z = s$.*
The proof of this lemma is beyond the scope of this course. It is known to be equivalent to the Axiom of Choice of set theory. We apply the lemma as follows. Let S be the set of all ideals of R except the ideal R itself. This is a poset: $I \preccurlyeq J$ if $I \subseteq J$. Any chain of ideals $\{I_\alpha : \alpha \in A\}$ has a

supremum $\cup_{\alpha\in A} I_\alpha$ (this is indeed an ideal!). Hence ,by Zorn's lemma S, has a maximal element M . The construction gives that M is a maximal ideal of R.

□

Example 36.0.6. When is a principal ideal (r) prime? The first condition is that $(r) \neq R$. That is, r is not a unit. Secondly, if $ab \in (r)$, that is $ab = rc_1$ for some $c_1 \in R$ then $a \in (r)$ or $b \in (r)$, meaning $a = rc_2$ or $b = rc_3$ for some $c_i \in R$.

Let us say, for a general commutative ring R, that $f|g$ in R if $g = fc$ for some $c \in R$. We see that in this terminology, (r) is a prime ideal if r is not a unit and $r|ab \Rightarrow r|a$ or $r|b$. This is a property of prime numbers and motivates the terminology "prime" (but we also require $r \neq 0$).

In particular, the prime ideals of $\mathbb{Z}$ are precisely the ideals of the form (p), where p is a prime number. The ideal $(1+i)$ of $\mathbb{Z}[i]$ is maximal: $\mathbb{Z}[i]/(1+i) \cong (\mathbb{Z}[x]/(x^2+1))/((1+x, x^2+1)/(x^2+1)) \cong \mathbb{Z}[x]/(x^2+1, 1+x) = \mathbb{Z}[x]/(1+x, 2) \cong \mathbb{Z}/2\mathbb{Z}[x]/(1+x) \cong \mathbb{Z}/2\mathbb{Z}$.

The ideal $(x^2 - y^2)$ of $\mathbb{Q}[x, y]$ is not prime. We have $(x+y)(x-y) = x^2 - y^2$ and $x + y \notin (x^2 - y^2)$.

37. The Chinese Remainder Theorem

Let R be a commutative ring. Two ideals I, J of R are called co-prime if $I + J = R$; equivalently, we have $1 = i + j$ for some $i \in I, j \in J$.

Theorem 37.0.7. (The Chinese Remainder Theorem) *Let R be a commutative ring and $A_1, \ldots, A_k$ ideals of R, co-prime in pairs ($A_i + A_j = R$ for $i \neq j$). Then,*

$$R/(A_1 \cdots A_k) \cong R/A_1 \times \cdots \times R/A_k.$$

Proof. We define a map

$$f : R \longrightarrow R/A_1 \times R/A_2 \times \cdots \times R/A_k, \qquad r \mapsto (r + A_1, \ldots, r + A_k).$$

This is a ring homomorphism whose kernel is $A_1 \cap A_2 \cap \cdots \cap A_k \supseteq A_1 A_2 \cdots A_k$. We need to prove that this is actually an equality and that f is surjective. The key is the following Lemma:

Lemma 37.0.8. *For every i there is an element $e_i \in R$ such that*

$$e_i \equiv 1 \pmod{A_i}, \qquad e_i \equiv 0 \pmod{A_j}, \forall j \neq i.$$

Proof. (Lemma) Without loss of generality, $i = 1$. For each $j = 2, 3, \ldots, k$ write

$$1 = x_j + y_j, \qquad x_j \in A_1, y_j \in A_j.$$

Then

$$\begin{aligned} 1 &= (x_1 + y_1)(x_2 + y_2) \cdots (x_k + y_k) \\ &= \alpha + y_2 y_3 \ldots y_k. \end{aligned} \tag{37.1}$$

Here α is a sum of products, each involving at least one x_j, so $\alpha \in A_1$. Let

$$e_1 = 1 - \alpha.$$

Then $e_1 \equiv 1 \pmod{A_1}$ and $e_1 = y_2 y_3 \ldots y_k \equiv 0 \pmod{A_j}$ for $2 \leq j \leq k$. □

We now show that f is surjective. Given $(r_1, r_2, \cdots, r_k) \in R/A_1 \times R/A_2 \times \cdots \times R/A_k$ choose $s_i \in R$ such that $\overline{s_i} = s_i + A_i = r_i$. Then $f(s_1 e_1 + s_2 e_2 + \cdots + s_k e_k) = \sum_i s_i f(e_i) = \sum_i (0, \ldots, 0, \underset{i}{\overline{s_i}}, 0, \ldots, 0)) = (\overline{s_1}, \overline{s_2}, \ldots, \overline{s_k})$.

It remains to prove that $A_1 A_2 \cdots A_k \supseteq A_1 \cap A_2 \cap \cdots \cap A_k$. We prove that by induction on k. For $k = 1$ this is clear. Consider the case $k = 2$. We have $1 = x_2 + y_2$ as in Equation (37.1). Let $c \in A_1 \cap A_2$. Then $c = cx_2 + cy_2$. Note that $c \in A_2, x_2 \in A_1 \Rightarrow cx_2 \in A_1 A_2$ and $c \in A_1, y_2 \in A_2 \Rightarrow cy_2 \in A_1 A_2$. Thus, $c \in A_1 A_2$.

Assume now that $k > 2$. Let $B = A_2 \cap \cdots \cap A_k$. We know by induction that $B = A_2 \cdots A_k$. Note that A_1 and B are relatively prime, because by Equation (37.1)

$$1 = \alpha + y_2 \cdots y_k, \quad \alpha \in A_1, y_2 \cdots y_k \in B.$$

Using the case $k = 2$ we have that $A_1 B \supseteq A_1 \cap B$, i.e., $A_1 A_2 \cdots A_k \supseteq A_1 \cap A_2 \cap \cdots \cap A_k$. □

Remark 37.0.9. One may ask why is it important to prove that the kernel is $A_1A_2\cdots A_k$ and not just $A_1\cap A_2\cap\cdots\cap A_k$. The reason is that in general it is easier to calculate the product of ideals than their intersection. For example, if each A_i is principal, $A_i=(a_i)$, then $A_1A_2\cdots A_k=(a_1a_2\cdots a_k)$. This formula can be generalized. For example, if $A_1=(\{a_i\}_i), A_2=(\{b_j\}_j)$ then $A_1A_2=(\{a_ib_j\}_{i,j})$.

Corollary 37.0.10. *Let $a_1,\cdots,a_k$ be relatively prime integers – that is, $(a_i,a_j)=1$ for $i\neq j$. Then*

$$\mathbb{Z}/(a_1a_2\cdots a_k)\cong\mathbb{Z}/(a_1)\times\cdots\times\mathbb{Z}/(a_k).$$

In particular, given residues classes b_i mod a_i, there is an integer b, unique up to adding multiples of $a_1a_2\cdots a_k$ such that $b\equiv a_i \pmod{a_i}$ for all i.

Example 37.0.11. Find an integer congruent to 5 mod 7 and congruent to 10 mod 13. In the notation of the proof, we are looking for $5e_1+10e_2$. Write $1=2\cdot 7-13$ (this can be done using the Euclidean algorithm in general). Then $e_1=1-2\cdot 7=-13, e_2=1+13=14$. Then $b=5\cdot(-13)+10\cdot 14=75$ is congruent to 5 mod 7 and to 10 mod 13. Note that by modifying by a multiple of 7×13 we can get a small solution, namely -16. This is typical too.

Part 9. Euclidean, Principal Ideal and Unique Factorization Domains

38. Euclidean domain

Definition 38.0.12. Let R be an integral domain. We say that R is Euclidean if there is a function (called norm)

$$N : R \longrightarrow \mathbb{N} = \{0, 1, 2, \dots\}$$

such that for any $a, b \in R$, $b \neq 0$, there are $q, r \in R$ such that

$$a = qb + r,$$

with $r = 0$ or $N(r) < N(b)$.

Example 38.0.13. $R = \mathbb{Z}, N(a) = |a|$.

Example 38.0.14. Let $\mathbb{F}$ be a field. Define on $R = \mathbb{F}[x]$, $N(f(x)) = \deg(f(x))$ then $\mathbb{F}[x]$ is Euclidean. Indeed, write

$$a = a_N x^N + a_{N-1}x^{N-1} + \cdots + a_0, \quad a_N \neq 0$$

and

$$b = b_M x^M + b_{M-1}x^{M-1} + \cdots + b_0, \quad b_M \neq 0.$$

If $N < M$ take $q = 0$ and $r = a$. If $N \geq M$, let $q = q_{N-M}x^{N-M} + \cdots + q_0$, where the coefficients q_i are determined recursively by attempting to solve $a = qb$, i.e.,

$$a_N x^N + a_{N-1}x^{N-1} + \cdots + a_0 = (q_{N-M}x^{N-M} + \cdots + q_0)(b_M x^M + b_{M-1}x^{M-1} + \cdots + b_0).$$

That is, we solve recursively for the q_i:

$$\begin{aligned} q_{N-M}b_M &= a_N \\ q_{N-M-1}b_M + q_{N-m}b_{M-1} &= a_{N-1} \\ &\vdots \end{aligned}$$

Example 38.0.15. Let $R = \mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$. This is a subring of the complex numbers. Let

$$N(a + bi) = a^2 + b^2 = |a + bi|^2.$$

Given two elements $a + bi, c + di$ of R, let us write

$$a + bi = \frac{a + bi}{c + di}(c + di) = \left(\frac{ac + bd}{c^2 + d^2} + \frac{-ad + bc}{c^2 + d^2}i\right)(c + di).$$

Let $\alpha = \frac{ac+bd}{c^2+d^2}$ and $\beta = \frac{-ad+bc}{c^2+d^2}$. Find integers A, B such that

$$|\alpha - A| \leq 1/2, \quad |\beta - B| \leq 1/2.$$

Then

$$a + bi = (A + Bi)(c + di) + ((\alpha - A) + (\beta - B)i)(c + di).$$

Then $q = A + Bi$ and $r = ((\alpha - A) + (\beta - B)i)(c + di)$. Note that $q, r \in R$. Finally,

$$N(r) = [(\alpha - A)^2 + (\beta - B)^2]N(c + di) \leq \frac{1}{2}N(c + di) < N(c + di).$$

39. Principal ideal domain (PID)

Definition 39.0.16. An integral domain R in which every ideal is principal, i.e. of the form $(r) = Rr = rR$ for some $r \in R$, is called a principal ideal domain (PID).

Proposition 39.0.17. *Every Euclidean domain is a PID.*

Proof. Let $I \lhd R$ be an ideal. If $I = \{0\} = (0)$ there is nothing to prove. Else, choose $b \in I, b \neq 0$ such that $N(b)$ is minimal among the norms of the non-zero elements of I. Let $a \in I$ then we may write $a = qb + r$ with $r = 0$ or $N(r) < N(b)$. However, $r = a - qb \in I$ so $r = 0$ else we get a contradiction to the definition of b. That is, $a \in (b)$ and it follows that $I = (b)$. □

39.1. **Division and gcd's.** Let R be an integral domain and $a, b \in R$. We say that b divides a, $b|a$, if there exists $x \in R$ such that $a = bx$. We say that a and b are associates, $a \sim b$, if $a = bx$ and $x \in R^\times$.

Here are some easy consequences of the definitions:

- $c|b$ and $b|a \Rightarrow c|a$.
- $1|a$. $a|1 \Leftrightarrow a \in R^\times$.
- $b|a$ and $a|b \Leftrightarrow b \sim a$.
- $b|a_1, b|a_2 \Rightarrow b|(a_1 + a_2)$.
- $b|a \Rightarrow b|ac, \forall c \in R$.
- $a \sim b$ if and only if $a|b$ and $b|a$. Being associates is an equivalence relation.

Lemma 39.1.1. *$b|a \Leftrightarrow (a) \subseteq (b)$ ("to divide is to contain"). In particular, $a \sim b \Leftrightarrow (a) = (b)$.*

Proof. We have $b|a \Leftrightarrow a = bx \Leftrightarrow a \in (b) \Leftrightarrow (a) \subseteq (b)$. □

A greatest common divisor (g.c.d.) of two elements $a, b \in R$ is an element $d \in R$ having the following properties:

(1) $d|a$ and $d|b$;
(2) If $d'|a$ and $d'|b$ then $d'|d$.

Lemma 39.1.2. *A g.c.d., if it exists, is unique up to a unit. In that case, it will be denoted* $\gcd(a, b)$ *or, simply,* (a, b).

Proof. Assume that d is a g.c.d. of a and b. Say $a = da', b = db'$. Let $x \in R^\times$. Then $a = (dx)(x^{-1}a'), b = (dx)(x^{-1}b')$. Suppose that $d'|a, d'|b$ then $d = d'd''$ and so $dx = d'(d''x)$. It follows that dx is a g.c.d. too.

Conversely, say d and d' are both g.c.d.'s. Then $d|d', d'|d$. It follows that $d \sim d'$ and so differ by a unit. □

In general a g.c.d. need not exist. The following lemma provides a criterion for its existence. Note that this criterion is not necessary but only sufficient. For example, in the ring $\mathbb{Q}[x, y]$ we have $g.c.d.(x, y) = 1$ but $\langle x, y\rangle$ is not principal.

Lemma 39.1.3. *If the ideal $\langle a, b\rangle$ is principal, $\langle a, b\rangle = (d)$, then d is a g.c.d. of a, b.*

Proof. If $\langle a, b\rangle = (d)$ then $a \in (d), b \in (d)$ so $d|a, d|b$. If $d'|a, d'|b$ then $a, b \in (d')$ and so $\langle a, b\rangle \subseteq (d')$. Hence, $(d) \subseteq (d')$ and so $d'|d$. □

Corollary 39.1.4. *If R is a PID then any two elements of R have a g.c.d..*

39.2. **Calculation of g.c.d.'s – the Euclidean algorithm.** Let R be a Euclidean ring. Then R is a PID and hence any two elements $a, b \in R$ have a g.c.d.. The Euclidean algorithm provides means to calculate that g.c.d..

Theorem 39.2.1. *Let a, b be elements of the Euclidean ring R. Write*

$$\begin{aligned} a &= q_0 b + r_0 \\ b &= q_1 r_0 + r_1 \\ r_0 &= q_2 r_1 + r_2 \\ &\vdots \\ r_{n-1} &= q_{n+1} r_n \end{aligned}$$

Indeed, the process always stops. Moreover r_n is $\gcd(a,b)$.

Example 39.2.2. Let us calculate the g.c.d. of 1079 and 1131. We have

$$\begin{aligned} 1131 &= 1*1079+52\\ 1079 &= 20*52+39\\ 52 &= 1*39+13\\ 39 &= 3*13 \end{aligned}$$

Therefore, $13 = (1079, 1131)$.

Example 39.2.3. Let us calculate the g.c.d. of x^3-x and x^3+3x^2+x in $\mathbb{Q}[x]$. We have

$$\begin{aligned} x^3+3x^2+x &= 1*(x^3-x)+3x^2+2x\\ x^3-x &= (x/3-2/9)(3x^2+2x)-5x/9\\ 3x^2+2x &= -9/5(3x+2)(-5x/9) \end{aligned}$$

It follows that $\gcd(x^3-x, x^3+3x^2+x) = x$.

Remark 39.2.4. Let R be a PID. Then for every a, b we have $\langle a,b\rangle = \langle d\rangle$ for some $d \in R$. In the case R is Euclidean we have a method to find d. In the general case, we do not have a method.

Note that in the case of PID we have $\langle a,b\rangle = \langle d\rangle$ and so there are $x, y \in R$ such that $\gcd(a,b) = xa+yb$. In the Euclidean case the Euclidean algorithm also gives x, y by "solving back". An example will suffice to clarify how to do that. Refer back to Example 39.2.2. We have $13 = (1079, 1131)$. Moreover, $52 = 1*39+13$ and so $13 = 52-39$. Now, $1079 = 20*52+39$ and so $13 = 52-(1079-20*52) = 21*52-1079$. Use now that $1131 = 1*1079+52$ to get that $13 = 21*(1131-1079)-1079 = 21*1131-22*1079$.

39.3. **Irreducible and prime elements.**

Definition 39.3.1. Let R be an integral domain. Let r be an element of R, $r \neq 0$ and r not a unit.

(1) The element r is called irreducible if

$$r = ab \Longrightarrow r \sim a \text{ or } r \sim b.$$

(2) The element r is called prime if

$$r|ab \Longrightarrow r|a \text{ or } r|b.$$

Remark 39.3.2. Note that if r, s are associates then r is irreducible (prime) if and only if s is.

Note also that r is prime if and only if (r) is a non-zero prime ideal.

Lemma 39.3.3. *If r is prime then r is irreducible.*

Proof. Suppose that $r = ab$. Then $r|ab$ and so, without loss of generality, $r|a$. But $a|r$ and so $r \sim a$. □

Example 39.3.4. In general an irreducible element need not be prime. Consider the ring $\mathbb{Z}[\sqrt{-5}]$. We have the factorization

$$(1+\sqrt{-5})\cdot(1-\sqrt{-5}) = 2\cdot 3.$$

I claim that all these elements are irreducible. First, the units of this ring are just ± 1. Now, for example, if $2 = (a+b\sqrt{-5})(c+d\sqrt{-5})$ then $|2|^2 = (a^2+5b^2)(c^2+5d^2)$. From that we see that $a = \pm 2, b = 0$ and so $2 \sim a$. Similar arguments work for the rest.

On the other hand, none of these elements can be prime. For example, $2|(1+\sqrt{-5})\cdot(1-\sqrt{-5})$ but clearly $2 \nmid +\sqrt{-5}$. Or, if you prefer, $\mathbb{Z}[\sqrt{-5}]/(2) \cong \mathbb{Z}/2\mathbb{Z}[x]/(x^2+5)$. We have $(x+1)^2 = x^2+1 = x^2+5 = 0$ in this ring, which shows that we have zero divisors. Hence, (2) is not a prime ideal.

In contrast, in certain rings, such as $\mathbb{Z}$, the concepts of prime and irreducible are one. The following Proposition generalizes this.

Proposition 39.3.5. *If R is a PID (e.g., if R is Euclidean) then r is prime if and only if r is irreducible.*

Proof. A prime element is always irreducible by Lemma 39.3.3. We show the converse. Let r be an irreducible element. Suppose that $(r) \subseteq B \lhd R$. Since R is a PID, we have $B = (b)$ for some $b \in R$. Thus, $r = ab$ for some $a \in R$. But r is irreducible so $r \sim a$ (and so $b \in R^\times$)) of $r \sim b$. We see that, correspondingly, either $B = R$ or $B = (r)$. We conclude that (r) is a maximal ideal.

Since a maximal ideal is a prime ideal, it follows that (r) is a prime ideal and so r is a prime element. □

Corollary 39.3.6. *In a PID, every prime ideal is maximal.*

Proof. Every prime ideal if of the form (r), for r prime/irreducible. We saw that this implies (r) is maximal. □

Corollary 39.3.7. *Let $\mathbb{F}$ be a field. In the polynomial ring $\mathbb{F}[x]$ a polynomial is prime if and only if it is irreducible. The quotient ring $\mathbb{F}[x]/(f(x))$ is a field if and only if $f(x)$ is irreducible.*

Example 39.3.8. Let R be an integral domain which is not a field (e.g., $R = \mathbb{Z}$ or $R = \mathbb{F}[x]$, $\mathbb{F}$ a field). Then $R[y]$ is an integral domain that is not a PID.

Indeed, the ideal (y) is prime since $R[y]/(y) \cong R$ which is an integral domain. It is not a maximal ideal since R is not a field.

40. Unique factorization domain (UFD)

Definition 40.0.9. Let R be an integral domain. R is called a unique factorization domain (UFD) if for every $r \in R$, not zero and not a unit, the following holds:

(1) r can be written as a product of irreducible elements p_i,
$$r = p_1 p_2 \dots p_n.$$
(2) If $r = q_1 q_2 \dots q_m$ is another expression of r as a product of irreducible elements then $m = n$ and after re-indexing we have $p_i \sim q_i$ for all i.

Proposition 40.0.10. *Let R be a UFD and r an element of R. Then r is prime if and only if r is irreducible.*

Remark 40.0.11. Recall that a PID also has this property (Prop. 39.3.5. We shall prove below that a PID is UFD, so it all adds up!

Proof. A prime element is always irreducible (Lemma 39.3.3). Let $r \in R$ be irreducible. Suppose that $r|ab$. Then $ab = rx$. Write the irreducible decomposition of each element: $a = p_1 p_2 \cdots p_m, b = q_1 q_2 \cdots q_m, w = t_1 t_2 \cdots t_\ell$. Then $p_1 p_2 \cdots p_m q_1 q_2 \cdots q_m = r t_1 t_2 \cdots t_\ell$ is two expressions as product of irreducible elements. It follows that either $r \sim p_i$ for some i, or $r \sim q_j$ for some j. Thus, either $r|a$ or $r|b$. □

40.1. **A PID is a UFD.**

Theorem 40.1.1. *Let R be a PID then R is a UFD.*

We have thus the following situation

$$\boxed{R \text{ Euclidean}} \underset{\not\Leftarrow}{\Longrightarrow} \boxed{R \text{ PID}} \underset{\not\Leftarrow}{\Longrightarrow} \boxed{R \text{ UFD}}$$

In particular, we conclude:

Corollary 40.1.2. *Let $\mathbb{F}$ be a field then $\mathbb{F}[x]$ is UFD; every polynomial can be written as a product of irreducible polynomials uniquely (up to multiplication by units $= \mathbb{F}^\times$, and permuting the polynomials).*

Example 40.1.3. A UFD need not be a PID. We shall show below that R is a UFD implies that $R[x]$ is a UFD. Hence, $\mathbb{Q}[x, y]$ is a UFD but is not a PID (the ideal (x, y) is not principal).

Example 40.1.4. A PID needs not be Euclidean. I don't know an easy example. One can prove that $\mathbb{Z}[\frac{1+\sqrt{-19}}{2}]$ is a PID, but not Euclidean.

Proof. The first step is to prove that if $r \in R$ is not zero, or a unit, then r can be written as a product of irreducible elements.

Suppose not, then r is not irreducible and so $r = r_1 s_1$, where either r_1 or s_1 are not a product of irreducible elements, without loss of generality, r_1. Then $r_1 = r_2 s_2$, , where either r_2 or s_2 are not a product of irreducible elements (and are not associates of r_1), without loss of generality, r_2. Then $r_2 = r_3 s_3$, , where either r_3 or s_3 are not a product of irreducible elements (and are not associates of r_2), without loss of generality, r_3. And so on.

We get a chain of division: $\ldots r_3 | r_2 | r_1 | r$, where this is "true division"; any two elements are not associates. We thus get a strictly increasing chain of ideals:

$$(r) \subsetneq (r_1) \subsetneq (r_2) \subsetneq (r_3) \subsetneq \ldots.$$

Consider then $\cup_{i=1}^{\infty}(r_i)$. It is easy to check this is an ideal, and so, since R is a PID, of the form (g) for some $g \in R$. But then $g \in (r_i)$ for some i and we get $(r_i) = (g)$. It then follows that $(r_i) = (r_{i+1}) = (r_{i+2}) = \ldots$. This is a contradiction.

The second step is to prove this decomposition is unique. Say

$$r = p_1 \cdots p_n = q_1 \cdots q_m,$$

a product of irreducible elements and without loss of generality $m \geq n$. We prove the uniqueness by induction on n:

If $n = 1$ then we get a factorization of the irreducible element p_1. Then either q_1 or $q_2 \cdots q_m$ is a unit. It must thus be the case that $m = 1$ and $p_1 = q_1$.

Assume the result of $n - 1$. Since $p_n | q_1 \cdots q_m$ there is some i such that $p_n | q_i$ (in a PID an irreducible element is prime). Thus, since q_i is irreducible, $q_i = p_n x$ for some unit x. We get

$$p_1 \cdots p_{n-1} = (x q_1) q_2 \cdots \hat{q_i} \cdots q_m.$$

By induction $n - 1 = m - 1$ and $p_1, p_2, \ldots, p_{n-1}$ are the same as $x q_1, q_2, \ldots, \hat{q_i}, \ldots, q_m$ up to multiplication by units (or, what is the same, $q_1, q_2, \ldots, \hat{q_i}, \ldots, q_m$ up to multiplication by units). □

40.1.1. *Arithmetic in UFD's.* The unique factorization property allows us to do arithmetic in a UFD much like in $\mathbb{Z}$. For instance, we can define g.c.d.'s, l.c.m.'s and such.

Proposition 40.1.5. *Let R be a UFD. Let $x = p_1^{\alpha_1} \cdots p_n^{\alpha_n} \cdot u$, $y = p_1^{\beta_1} \cdots p_n^{\beta_n} \cdot v$, where the p_i are non-associated irreducible elements, α_i, β_i are non-negative integers and u, v are units. Then*

$$\gcd(x, y) = p_1^{s_1} \cdots p_n^{s_n}, \quad s_i = \min\{\alpha_i, \beta_i\},$$

and

$$\operatorname{lcm}(x, y) = p_1^{t_1} \cdots p_n^{t_n}, \quad t_i = \max\{\alpha_i, \beta_i\}.$$

The proposition follows immediately from the following result.

Lemma 40.1.6. *In the notation above, $z | x$ if and only if $z = p_1^{a_1} \cdots p_n^{a_n} w$ with $a_i \leq \alpha_i$ for all i and w a unit.*

Proof. Clearly every such z divides x: $x = p_1^{\alpha_1 - a_1} \cdots p_n^{\alpha_n - a_n} u w^{-1} z$. Conversely, if $z | x$, say $x = zt$ then write z and t as a product of irreducible elements. Say $z = p_1^{a_1} \cdots p_n^{a_n} q_1^{b_1} \cdots q_m^{b_m} w$ and $t = p_1^{a_1'} \cdots p_n^{a_n'} q_1^{b_1'} \cdots q_m^{b_m'} w'$, where we allow non-negative (including zero) exponents. Thus,

$$p_1^{\alpha_1} \cdots p_n^{\alpha_n} \cdot u = p_1^{a_1 + a_1'} \cdots p_n^{a_n + a_n'} q_1^{b_1 + b_1'} \cdots q_m^{b_m + b_m'} w w'.$$

Unique factorization gives that each $b_i = b_i' = 0$ (or, if you prefer, $m = 0$) and $a_i + a_i' = \alpha_i$. □

40.2. **Application: construction of fields.** Let $\mathbb{F}$ be a field.

Proposition 40.2.1. *Let $f(x)$ be a monic irreducible polynomial of $\mathbb{F}[x]$ of degree n. The ring $\mathbb{F}[x]/(f(x))$ is a field. Every element in this field can be written uniquely as the coset of a polynomial $a_0 + a_1 x + \cdots + a_{n-1} x^{n-1}$ with $a_i \in \mathbb{F}$. In particular, if $\mathbb{F}$ is a finite field of q elements then $\mathbb{F}[x]/(f(x))$ has q^n elements.*

Proof. Since $f(x)$ is irreducible $(f(x))$ is a maximal ideal. Thus, $\mathbb{F}[x]/(f(x))$ is a field. Given a polynomial $a(x)$ divide it with residue by $f(x)$. We have

$$a(x) = q(x)f(x) + r(x),$$

where $r(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$ with $a_i \in \mathbb{F}$ and the coefficients may well be all or some equal to zero. Note that $\overline{a(x)} = \overline{r(x)}$. This shows that every element has such a representation.

Now, if $a_0 + a_1x + \cdots + a_{n-1}x^{n-1} = b_0 + b_1x + \cdots + b_{n-1}x^{n-1}$ in $\mathbb{F}[x]/(f(x))$ then $a_0 + a_1x + \cdots + a_{n-1}x^{n-1} - (b_0 + b_1x + \cdots + b_{n-1}x^{n-1})$ belongs to the ideal $f(x)$. But every non-zero polynomial in this ideal has degree at least $n = \deg(f(x))$. Thus, $a_0 + a_1x + \cdots + a_{n-1}x^{n-1} = b_0 + b_1x + \cdots + b_{n-1}x^{n-1}$ in $\mathbb{F}[x]$. □

It is known that every finite field has p^n elements for some prime p and some integer $n \geq 1$. To show that for every p and n there is such field one can show that for every p and n there is an irreducible polynomial of degree n over $\mathbb{Z}/p\mathbb{Z}$. This usually requires some clever tricks. Here we content ourselves in proving that there is a field of p^2 elements for every p.

Assume first that $p \neq 2$. Consider the map

$$(\mathbb{Z}/p\mathbb{Z})^\times \longrightarrow (\mathbb{Z}/p\mathbb{Z})^\times, \qquad t \mapsto t^2.$$

The kernel of this map is $\{\pm 1\}$ and, since p is odd, consists of two elements. It follows that the image is of order $(p-1)/2$. Thus, there are $(p-1)/2+1$ elements in $\mathbb{Z}/p\mathbb{Z}$ that are squares (we count now also 0) and $(p-1)/2$ that aren't. Let t be a non-square. Then $x^2 - t$ is an irreducible polynomial.

For $p = 2$ we may take the polynomial $x^2 + x + 1$.

40.3. **Gauss' Lemma.**

Lemma 40.3.1. *Let I be an ideal of R, a commutative ring, let $IR[x]$ be the ideal generated by I in the polynomial ring $R[x]$. Then*

$$IR[x] = \{\sum_{n=0}^{N} a_n x^n : a_n \in I\}$$

and

$$R[x]/IR[x] \cong (R/I)[x].$$

Proof. By definition, $IR[x] = \{\sum_{n=0}^{N} i_n f_n(x) : i_n \in I, f_n(x) \in R[x]\}$. Clearly it contains $\{\sum_{n=0}^{N} a_n x^n : a_n \in I\}$. On the other hand, by expanding a sum $\sum_{n=0}^{N} i_n f_n(x)$, $i_n \in I$, $f_n(x) \in R[x]$, according to powers of x we get the other inclusion.

Now, define a homomorphism

$$R[x] \longrightarrow (R/I)[x], \quad f(x) \mapsto \overline{f(x)},$$

where if $f(x) = \sum a_i x^i$ then $\overline{f(x)} = \sum \overline{a_i} x^i$ (we use $\overline{a_i}$ to denote the coset $a_i + I$). The kernel is $\{\sum_{n=0}^{N} a_n x^n : \overline{a_n} = 0\} = \{\sum_{n=0}^{N} a_n x^n : a_n \in I\} = IR[x]$ and the map is clearly surjective. We conclude by the first isomorphism theorem. □

Lemma 40.3.2. (Gauss' lemma) *Let R be a UFD with field of fractions F, Let $f(x) \in R[x]$. If $f(x)$ is reducible in $F[x]$ then $f(x)$ is reducible in $R[x]$. More precisely, if $f(x) = A(x)B(x)$ in $F[x]$, a product of non-constant polynomials, then $f(x) = a(x)b(x)$ in $R[x]$ where $a(x)$ (resp., $b(x)$), is a constant multiple of $A(X)$ (resp., $B(X)$).*

Remark 40.3.3. Note that the contrapositive has to be taken with care. It is not, $f(x)$ irreducible in $R[x]$ implies that $f(x)$ is irreducible in $F[x]$. The issue is that the units of the rings are different. For example, $2 \in \mathbb{Z}$ is irreducible in $\mathbb{Z} \subset \mathbb{Z}[x]$ but is not irreducible in $\mathbb{Q} \subset \mathbb{Q}[x]$ simply because it is a unit in $\mathbb{Q}$ and a unit is not an irreducible element. See Corollary 40.3.5 below for the correct converse.

Example 40.3.4. A typical application of Gauss' lemma is the following. Let $f(x)$ be a monic polynomial with integer coefficients. Every rational root of f is an integer. Indeed, a root r gives a factorization $f(x) = (x-r)B(x)$ and hence a factorization into polynomials with integer coefficients $f(x) = [s(x-r)][tB(x)]$, where $s, t \in \mathbb{Q}$. Since f is monic we must have $s = \pm 1$ and so, sr and hence r are integers.

Proof. Suppose that $f(x) = A(x)B(x)$ in $F[x]$. Since the coefficients of A, B are fractions s/t, where $s, t \in R$, we can find a common denominator and so an equation

$$df(x) = A_1(X)B_1(X),$$

with $0 \neq d \in R, A_1(X), B_1(X) \in R[x]$. Note that A_1, B_1 are constant multiples of A, B.

If d is a unit, take $a(x) = d^{-1}A_1(x), b(x) = B(x)$. Else,

$$d = p_1 \cdots p_n,$$

a product of irreducible elements. Now, since p_1 is irreducible it is prime and so (p_1) is a prime ideal. In the ring $R/(p_1)[x]$, which is an integral domain, we have $0 = \overline{A_1(x)} \cdot \overline{B_1(x)}$ and thus, without loss of generality, $\overline{A_1(x)} = 0$. Lemma 40.3.1 gives that each coefficient of $A_1(x)$ is divisible by p_1. Hence, there is a polynomial $A_2(x) \in R[x]$ such that

$$p_2 \cdots p_n f(x) = A_2(x)B_1(x).$$

Continuing in such fashion, we find polynomials $a(x), b(x) \in R[x]$ such that $f(x) = a(x)b(x)$ and a, b are constant multiples of A, B. □

Corollary 40.3.5. *Let $f(x) \in R[x]$ be a polynomial such that the g.c.d. of its coefficients is 1, e.g., $f(x)$ is monic. Then $f(x)$ is irreducible in $R[x]$ if and only if $f(x)$ is irreducible in $F[x]$.*

Proof. One direction is Gauss' Lemma. Suppose then that $f(x)$ is reducible in $R[x]$, say $f(x) = a(x)b(x)$, where neither is a unit in $R[x]$. Note that $a(x)$ cannot be a constant, because this would imply that $a(x)$ divides the g.c.d. of the coefficients of $f(x)$ and hence that g.c.d. is not 1. Thus, $a(x)$ is also not a unit of $F[x]$. The same holds for $b(x)$ and thus f is reducible in $F[x]$. □

Example 40.3.6. It is good to keep the following example in mind. The polynomial $2x$ is reducible in $\mathbb{Z}[x]$ but is irreducible in $\mathbb{Q}[x]$.

40.4. R UFD $\Rightarrow$ $R[x]$ UFD.

Theorem 40.4.1. *Let R be a UFD then $R[x]$ is a UFD.*

Proof. Let $f(x) \in R[x]$ and write

$$f = df_1,$$

where the g.c.d. of the coefficients of f_1 is 1. Note that this decomposition is unique up to a unit, namely, up to $d \mapsto du, f_1 \mapsto f_1 u^{-1}$. Since d can be written as product of irreducible elements, unique up to being associate, and since irreducible elements of R are irreducible elements of $R[x]$, we may assume that the g.c.d. of the coefficients of f is 1.

let F be the quotient field of R. We use the fact that $F[x]$ is Euclidean, hence PID, hence UFD, to write

$$f(x) = P_1(x) \cdots P_n(x), \qquad P_i(x) \in F[x] \text{ irreducible.}$$

By Gauss' Lemma

$$f(x) = p_1(x) \cdots p_n(x), \qquad p_i(x) \in R[x],$$

where each p_i is a multiple of P_i, in particular irreducible in $F[x]$. Note that the g.c.d. of the coefficients of p_i must be 1 (because of our assumption of f). Corollary 40.3.5 gives that each p_i is irreducible in $R[x]$.

The decomposition of f is unique. If

$$f = q_1(x) \cdots q_m(x)$$

is another factorization into irreducible polynomials in $R(x)$ then each q_i has g.c.d. of its coefficients equal to 1, hence by Corollary 40.3.5 is irreducible in $F[x]$. Since $F[x]$ is a UFD, we must have, after re-indexing, that $m = n$ and $q_i \sim p_i$ for all i in $F[x]$, say $p_i = \frac{r_i}{s_i} q_i$. We get an equality in $R[x]$: $s_i p_i = r_i q_i$. The g.c.d. of the r.h.s. is r_i and is equal to that of the l.h.s. which is s_i. It follows that $r_i \sim s_i$ and so $p_i \sim q_i$ in $R[x]$. □

Corollary 40.4.2. *Let $\mathbb{F}$ be a field and $x_1, \ldots, x_n$ be variables. The ring of polynomials $\mathbb{F}[x_1, \ldots, x_n]$ is a UFD.*