MATH411-22S2 LECTURE NOTES

Contents

1. Introduction

Galois theory is named after Evariste Galois (1811-1832). The roots of a polynomial $f(x) \in k[x]$ over a field k have symmetries which form a group. Galois theory studies such groups and how they can be used to describe the structure of the splitting field of $f(x)$ over k. The goal of this course is to develop this theory (and the necessary group theory) and use this to prove the following two theorems:

Theorem 1.1 (Rufini 1799, Abel 1824)**.** *A general polynomial of degree at least 5 has no solution in radicals*

Theorem 1.2 (Fundamental Theorem of Algebra)**.** *Every nonconstant polynomial over the complex numbers has a root.*

Along the way we will need to learn a fair amount of group theory.

Theorem 1.3 (Quadratic formula)**.** *The roots of* $f(x) = x^2 + ax + b \in \mathbb{Q}[x]$ *are given by* $x = \frac{a \pm \sqrt{a^2 - 4b}}{2}$

Corollary 1.4. *All quadratic extensions of* $\mathbb{Q}$ *are of the form* $\mathbb{Q}(\sqrt{d})$ *for some* $d \in \mathbb{Q}$*.*

Theorem 1.5 (Cubic Formula, 1500's)**.** *Let* $f(x) = x^3 + ax^2 + bx + c \in \mathbb{Q}[x]$ *and set*

$$\Delta_1 = \dots$$

Corollary 1.6. *Any cubic polynomial* $f(x) \in \mathbb{Q}[x]$ *splits in a field* K *of the form*

$$\mathbb{Q} \subset \mathbb{Q}(\sqrt{d}) = K_1 \subset K_1(\sqrt[3]{e}) = K$$

where $d \in \mathbb{Q}$ and $e \in K_1$

Definition 1.7. *We say $K/\mathbb{Q}$ is a radical extension if K is obtained by a sequence of extensions of the form*

$$\mathbb{Q} = K_1 \subset K_1(\sqrt[n_1]{a_1}) = K_2 \subset K_2(\sqrt[n_2]{a_2}) = K_3 \cdots K_m(\sqrt[n_m]{a_m}) = K$$

for some $a_i \in K_i$ and $n_i \in \mathbb{Z}_{\geqslant 2}$.

The quadratic and cubic formulas above show that every degree 2 or 3 polynomial splits over a radical extension. This means you can give a formula for the roots of the polynomial which only require n-th roots and the usual arithmetic operations. There is also a quartic formula (Ferrari 1540) which shows that the same is true for degree 4 polynomials. The Abel-Rufini theorem mentioned above states that a general quintic polynomial does not split over any radical extension. Galois theory gives a characterization of the polynomials that split in a radical extension by looking at the group of symmetries of the roots (a concept that we will cover in more depth later).

Here is a sketch of the idea: The roots of a polynomial have symmetries. These symmetries form a group, called the Galois group of the polynomial. For example the roots $x = \frac{a \pm \sqrt{a^2-4b}}{2}$ of $f(x) = x^2 + ax + b$ have a symmetry encoded in the $\pm$. The corresponding group is $\mu_2 = \{\pm 1\}$ under multiplication. More generally, the n-th roots of an element of a field have symmetries described by a group μ_n. Consequently polynomials that split in radical extensions have Galois groups that are of a rather particular type (called solvable groups). Galois proved that there are groups which are not solvable (for example the alternating group A_5) and moreover that there are quintic polynomials which have such nonsolvable groups as their Galois groups. Such polynomials cannot possibly split in any radical extension.

2. Groups

Definition 2.1. *A group is a set G together with a binary operation*

$$G \times G \to G\,, (g, h) \mapsto g \cdot h$$

satisfying the following axioms:

(1) (Identity) There is an element $1 \in G$ such that $1 \cdot g = g \cdot 1 = g$ for all $g \in G$,
(2) (Inverses) For all $g \in G$ there is some $h \in G$ such that $gh = hg = 1$, and
(3) (Associativity) For all $f, g, h \in G$, $f \cdot (g \cdot h) = (f \cdot g) \cdot h$.

Definition 2.2. *A homomorphism of groups is a function $\phi : G \to H$ such that $\phi(g_1 \cdot g_2) = \phi(g_1) \cdot \phi(g_2)$. If ϕ is a bijection it is called an isomorphism.*

Note that in this definition the $\cdot$ is being used to denote two different group operations.
Here are some examples of groups:

(1) The integers $\mathbb{Z}$ under addition forms a group.
(2) The integers modulo n, $\mathbb{Z}/n\mathbb{Z}$ under addition form a group
(3) The set $F^\times$ of nonzero elements in a field F under multiplication form a group.
(4) The set of $n \times n$ invertible matrices $\mathrm{GL}_n(F)$ forms a group under matrix multiplication. In the case $n = 1$ this recovers $F^\times$ from the previous example.
(5) Given a set X, the symmetric group $\mathrm{Sym}(X)$ is the set of all bijections of X with itself with the group operation being composition of maps. One typically uses S_n to denote the group $\mathrm{Sym}(\{1, \ldots, n\})$.

(6) For any $n \geqslant 3$, the dihedral group D_n is the group of symmetries of a regular n-gon. This can be described in terms of "generators and relations" as the group $D_n = \langle \rho, \sigma \ : \ \rho^n = \sigma^2 = 1\,,\, \rho\sigma = \sigma\rho^{-1} \rangle$. As a set $D_n = \{1, \rho, \ldots, \rho^{n-1}, \sigma, \rho\sigma, \ldots, \rho^{n-1}\sigma\}$ since every "word" in the "letters" ρ and σ can be reduced to one of these $2n$ words using the relations given. For example, in D_4 we have

$$\begin{aligned} \sigma\rho\sigma\rho^3\sigma^{-32}\rho &= \sigma\rho\sigma\rho^3\rho && (\text{since } \sigma^{-32} = (\sigma^2)^{-16} = 1^{-16} = 1) \\ &= \sigma\rho\sigma && (\text{since } \rho^3\rho = \rho^4 = 1) \\ &= \sigma^2\rho^{-1} && (\text{since } \rho\sigma = \sigma\rho^{-1}) \\ &= \rho^3 && (\text{since } \sigma^2 = 1 \text{ and } \rho^{-1} = \rho^3) \end{aligned}$$

Here are some examples of group homomorphisms:

(1) Multiplication by n defines a homomorphism $\mathbb{Z} \to \mathbb{Z}$.
(2) Addition by n does not define a homomorphism $\mathbb{Z} \to \mathbb{Z}$.
(3) If X and Y are sets of the same cardinality, then there is a bijection $X \to Y$. This induces an isomorphism $\mathrm{Sym}(X) \simeq \mathrm{Sym}(Y)$.
(4) The formula $\det(AB) = \det(A)\det(B)$ familiar from linear algebra shows that the determinant gives a homomorphism $\det : \mathrm{GL}_n(F) \to F^\times$.
(5) Labeling the vertices of an n-gon from 1 to n, any element of D_n gives a permutation of $\{1, \ldots, n\}$. This defines a homomorphism $D_n \to S_n$ which is injective.

Definition 2.3. *A subgroup of a group G is a subset $H \subset G$ such that the group operation of G restricts to give a group operation $H \times H \to H$. We write $H < G$ to denote that H is a subgroup of G. Given a subset $X < G$ we use $\langle X \rangle$ to denote the smallest subgroup of G containing all elements of X.*

(1) The integers $n\mathbb{Z} \subset \mathbb{Z}$ that are divisible by n form a subgroup of $\mathbb{Z}$.
(2) The upper-triangular invertible matrices form a subgroup of $\mathrm{GL}_n(F)$.
(3) Given a homomorphism $\phi : G \to H$, one obtains the follwing subgroups:
   (a) The image of ϕ is $\mathrm{im}(\phi) = \{h \in H \ : \ \exists g \in G \text{ such that } \phi(g) = h\}$, and
   (b) The kernel or ϕ is $\ker(\phi) = \{g \in G \ : \ \phi(g) = 1\}$.

Theorem 2.4 (Cayley's Theorem)**.** *For any group G there is an injective homomorphism $\phi : G \to \mathrm{Sym}(G)$. Consequently, every group G may be viewed as a subgroup of a symmetric group.*

Proof. Consider the map $\Phi : G \to \mathrm{Sym}(G)$ sending $g \in G$ to the permutation ϕ_g of G given by $\phi_g(h) = gh$. Now check that Φ is an injective group homomorphism. □

Remark 2.5. *Given a finite group G this shows $G < S_n$ for some n, but the n delivered by the proof is not minimal. For example, $D_n < S_n$ by the examples above, but the proof gives an embedding of D_n in S_{2n}. A more extreme example is S_n. It is a subgroup of itself, but the proof makes it a subgroup of $S_{n!}$ which has $n!!$ elements.*

3. Cosets, normal subgroups and quotient groups

Definition 3.1. *Given a subgroup $H < G$, a (left) coset of H in G is a subset of G of the form*

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

The set of cosets of H in G is denoted G/H. The index *of H in G is the cardinality of G/H and is denoted $[G : H]$.*

Example 3.2. *The subgroup $\langle\rho\rangle < D_n$ of rotations has 2 cosets. They are $\langle\rho\rangle = \{1, \rho, \ldots, \rho^{n-1}\}$ and $\sigma\langle\rho\rangle = \{\sigma, \rho\sigma, \ldots, \rho^{n-1}\sigma\}$. Note that the coset leaders are not unique. For example, $\sigma\langle\rho\rangle = \rho\sigma\langle\rho\rangle$. The subgroup $K = \langle\sigma\rangle < D_n$ has n cosets, $\langle\sigma\rangle$, $\rho\langle\sigma\rangle, \ldots, \rho^{n-1}\langle\sigma\rangle$, each with 2 elements.*

Lemma 3.3. *Two cosets are either equal or disjoint.*

Theorem 3.4 (Lagrange's Theorem)**.** *Let H be a subgroup of G. Then $|G| = |H| \times |G/H|$ or, equivalently in index notation,*

$$[G : 1] = [G : H][H : 1] .$$

Proof. From the lemma it follows that the cosets of H give a partition of G. For any $g \in G$, the map $H \to gH$ sending h to gh is a bijection (if $gh = gh'$ canceling g gives $h = h'$). Hence we see that G is a disjoint union of its cosets each of which has cardinality $|H|$. □

Corollary 3.5. *The order of an element $g \in G$ divides the order of $|G|$.*

Proof. Apply the theorem with $H = \langle g\rangle$. □

We now address the question of whether the set G/H of coests of H itself forms a group. The naive way of making this set into a group is to define a binary operation by the rule

$$g_1H \cdot g_2H = (g_1g_2)H . \tag{3.1}$$

It is straightforward to check that this operation is associative, there is an identity (the coset H) and inverses $((gH)^{-1}) = g^{-1}H$). There is a problem though. This operation is not well defined! It depends on the choice of coset leader.

For example, consider $H = \langle\sigma\rangle < D_3$. The coset $C = \{\rho, \rho\sigma\}$ is represented by both ρ and $\rho\sigma$ (i.e., $\rho H = \rho\sigma H$). We would like to be able to compute $C \cdot C$ using whatever representatives we like. However, the computations

$$\rho H \cdot \rho H = \rho^2 H \quad \text{and} \quad \rho\sigma H \cdot \rho H = \rho\sigma\rho H = \sigma\rho^{-1}\rho H = \sigma H = H$$

show that we get different results depending on the choice. It turns out that the operation is only well defined when the subgroup in question is normal.

Definition 3.6. *We say that $H < G$ is a normal subgroup and write $H \lhd G$ if $gH = Hg$ for all $g \in G$. Equivalently, H is normal if $gHg^{-1} := \{ghg^{-1} \ : \ h \in H\} = H$ for all $g \in G$.*

Example 3.7. *Any subgroup of an abelian group (commutative group) is normal.*

Example 3.8. *The subgroup $\langle\sigma\rangle < D_n$ is not normal because $\rho\langle\sigma\rangle = \{\rho, \rho\sigma\} \neq \{\rho, \sigma\rho = \rho^{n-1}\sigma\} = \langle\sigma\rangle\rho$. The subgroup $\langle\rho\rangle$ is normal in $< D_n >$. This can be checked directly or by noting more generally that any subgroup of index 2 is normal. Indeed if H has index 2 there are exactly two disjoint cosets. So for any $g \in G$ we have the fact that $g \in gH$ and $g \in Hg$ implies that the two must coincide.*

Lemma 3.9. *If $H \lhd G$ is a normal subgroup then the operation in (3.1) is well defined and hence endows G/H with the structure of a group.*

Proof. Let $g_1, g_2 \in G$ be such that $g_1H = g_2H$ and let gH be any other coset. It suffices to show that $g_1gH = g_2gH$. Using that H is normal and associativity we have

$$g_1gH = g_1(gH) = g_1(Hg) = (g_1H)g = (g_2H)g = g_2(Hg) = g_2(gH) = g_2gH\,.$$

□

Lemma 3.10. *Suppose $\phi : G \to H$ is a homomorphism. Then $\ker(\phi)$ is a normal subgroup of G.*

Proof. Let $g \in G$. Check that $g\ker(\phi)$ and $\ker(\phi)g$ are both equal to the set of elements $\{x \in G \ : \ \phi(x) = \phi(g)$. □

4. The homomorphism and isomorphism theorems

Theorem 4.1 (Homomorphism Theorem)**.** *Let $\phi : G \to H$ be a homomorphism of groups. Then ϕ factors as*

$$\begin{array}{ccc} G & \xrightarrow{\quad\phi\quad} & H \\ \downarrow & & \uparrow \\ G/\ker(\phi) & \xrightarrow{\quad\sim\quad} & \mathrm{image}(\phi)\,, \end{array}$$

where the surjective map is $g \mapsto g\ker(\phi)$ and the isomorphism is $g\ker(\phi) \mapsto \phi(g)$.

Proof. The vertical map is obviously surjective and is a group homomorphism by the definition of the group law in $G/\ker(\phi)$. The inclusion is likewise obviously an injective homomorphism. We consider $\tilde{\phi} : G/\ker(\phi) \to \mathrm{im}(\phi)$ given by $\tilde{\phi}(g\ker(\phi)) = \phi(g)$. This is well defined because if $g\ker(\phi) = h\ker(\phi)$, then $g = hk$ for some $k \in \ker(\phi)$ in which case $\phi(g) = \phi(hk) = \phi(h)\phi(k) = \phi(h)$. It is injective since if $\phi(g\ker(\phi)) = 1$, then $\phi(g) = 1$ in which case $g \in \ker(\phi)$ so $g\ker(\phi) = \ker(\phi)$ which is the identity in $G/\ker(\phi)$. It is obviously surjective. □

Example 4.2. *$\phi : \mathbb{Z} \to S_n$ sending a to $(12\ldots n)^a$.*

Example 4.3. *$\mathrm{GL}_n(\mathbb{F}_q)/\mathrm{SL}_n(\mathbb{F}_q) \simeq \mathbb{F}_q^\times$ via the determinant map.*

Theorem 4.4 (Isomorphism Theorem)**.** *Let $N \lhd G$ and $H < G$ be subgroups of G. Then $HN < G$, $H \cap N \lhd H$ and the map*

$$H/(H \cap N) \ni h(H \cap N) \mapsto hN \in HN/N$$

is an isomorphism.

Proof. We begin by showing that HN is a subgroup. Say $h_1, h_2 \in H$ and $n_1, n_2 \in N$ we need to check that $h_1n_1h_2n_2 \in HN$. Using that N is normal $Nh_2 = h_2N$, so there is some $n_1' \in N$ such that $n_1h_2 = h_2n_1'$. Then $h_1n_1h_2n_2 = h_1h_2n_1'n_2 \in HN$ as required. The same idea is used to show HN is closed under inverses and hence is a subgroup.

Now we show that $H \cap N$ is normal in H. Let $h \in H$. Then $hN = Nh$, so $hN \cap H = Nh \cap H$. Now if $hn \in (hN \cap H)$, then $hn \in H$ so $n \in h^{-1}H = H$ showing that $hN \cap H = h(N \cap H)$. Similarly $Nh \cap H = (N \cap H)h$.

Now consider the map ϕ defined in the statement. It is well-defined since if $h_1(H \cap N) = h_2(H \cap N)$, then there is $n \in H \cap N$ such that $h_1 = h_2n$ in which case $h_1N = h_2nN = h_2N$.

It is injective since if $hN = N$, then $h \in N$ in which case $h(H \cap N) = H \cap N$. It is surjective since hN is the image of $h(H \cap N)$. Finally we check that it is a homomorphism:

$$\phi(h_1(H\cap N)h_2(H\cap N) = \phi(h_1h_2(H\cap N)) = h_1h_2N = h_1Nh_2N = \phi(h_1(H\cap N))\phi(h_2(H\cap N))\,.$$

□

Example 4.5. *$D_4 = \langle \rho, \sigma | \rho^4 = \sigma^2 = 1, \rho\sigma = \sigma\rho^3 \rangle$ contains the subgroups $H = \langle \sigma \rangle$ and $N = \langle \rho \rangle$. We have $HN = \{\sigma^a\rho^b\} = D_4$ while $H \cap N = \{1\}$. We have an isomorphism $H/H \cap N = H = \{1, \sigma\} \simeq \{\langle\rho\rangle, \sigma\langle\rho\rangle\} = D_4/\langle\rho\rangle = HN/N$.*

5. The correspondence theorem

Theorem 5.1 (Correpsondence Theorem)**.** *Let $N \lhd G$ be a normal subgroup. The map*

$$\Phi : \{H \ : \ N < H < G\} \to \{\tilde{H} \ : \ \tilde{H} < G/N\} \quad H \mapsto H/N$$

is a bijection which

(1) preserves inclusions (i.e. $H < H' \Leftrightarrow \Phi(H) < \Phi(H')$,
(2) preserves index (i.e., $[G : H] = [G/N : \Phi(H)]$), and
(3) preserves normality (i.e., H is normal in G if and only if $\Phi(H)$ is normal in G/N.

All of this can be phrased by saying that the subgroup lattice of G/N is identical to the part of the subgroup lattice of G above N.

Example 5.2. *Consider the subgroups of $\mathbb{Z}$ which contain $n\mathbb{Z}$ and compare with the subgroups of $\mathbb{Z}/n\mathbb{Z}$ for $n = 2, 4$ and 6.*

Example 5.3. *Work out the details for $N = \langle\rho^2\rangle \lhd D_4$.*

6. Examples

Example 6.1. *$L \subset \mathbb{R}^3$ a line through the origin is a subgroup. What are its cosets?*

Example 6.2. *Consider a linear map $L : \mathbb{R}^3 \to \mathbb{R}^3$. What does the homomorphism theorem say? How is this like the rank-nullity theorem?*

Example 6.3. *Consider a pair of planes in $\mathbb{R}^3$. What does the isomorphism theorem say?*

Example 6.4. *Consider the subgroup $N = \{1, (12)(34), (13)(24), (14), (23)\} < S_4$. Show that N is normal and identify the quotient group S_4/N.*

7. Direct products of groups

Given groups G, H one can form their direct product

$$G \times H = \{(g, h) \ : \ g \in G, h \in H\}$$

This forms a group under the operation $(g, h) \cdot (g', h') = (gg', hh')$.

Note that $G \times H$ contains the two subgroups

$$\tilde{G} = \{(g, 1_H) \ : \ g \in G\}$$
$$\tilde{H} = \{(1_H, h) \ : \ h \in H\}\,,$$

which are isomorphic to G and H respectively. Moreover, one easily checks that

(1) $\tilde{G} \cap \tilde{H} = \{1\}$

(2) $\tilde{G}\tilde{H} = G \times H$
(3) $\tilde{G}$ and $\tilde{H}$ are normal subgroups of $G \times H$.

The following theorem gives a converse to these observations.

Theorem 7.1. *Let $H, K < G$. Then $G \simeq H \times K$ if and only if*
(1) $H, K \lhd G$,
(2) $HK = G$, and
(3) $H \cap K = \{1\}$.

Proof. Suppose H, K satisfy the conditons give and define $\phi : H \times K \to G$ by $(h, k) \mapsto hk$. This map is surjective since $HK = G$. It is injective since $H \cap K = \{1\}$. Indeed, if $\phi(h, k) = 1$, then $hk = 1$ which implies $h = k^{-1}$ showing that h lies in both H and K. To see that ϕ is a homomorphism consider that

$$\phi((h_1, k_1)(h_2, k_2)) = \phi(h_1h_2, k_1k_2) = h_1h_2k_1k_2\,, \text{ while}$$
$$\phi(h_1, k_1)\phi(h_2, k_2) = h_1k_1h_2k_2\,.$$

It therefore suffices to prove the following lemma: □

Lemma 7.2. *If $H, K \lhd G$ are normal subgroups with trivial intersection then their elements commute.*

Proof. Let $g = hkh^{-1}k^{-1}$ To show $hk = kh$ is the same as showing $g = 1$. Consider that $g = h(kh^{-1}k^{-1}) \in h(kHk^{-1}) = hH = H$ and $g = (hkh^{-1})k^{-1} \in (hKh^{-1})k^{-1} = Kk^{-1} = K$. Hence $g \in H \cap K = \{1\}$. □

Example 7.3. $\mathbb{Z}/6 \simeq 2\mathbb{Z}/6\mathbb{Z} \times 3\mathbb{Z}/6\mathbb{Z}$

Example 7.4. *$S_3 = D_3$ is not the direct product of its subgroups $\langle\rho\rangle \simeq \mathbb{Z}/3$ and $\langle\sigma\rangle \simeq \mathbb{Z}/2\mathbb{Z}$.*

Example 7.5. *If G is abelian and $\phi : G \to \mathbb{Z}$ is a surjective homomorphism, then $G \simeq \mathbb{Z} \times \ker(\phi)$.*

8. Finitely generated abelian groups

Definition 8.1. *Elements $x_1, \ldots, x_n \in M$ form a basis for the abelian group M if*
(1) $M = \langle x_1, \ldots, x_n\rangle$ and
(2) Whenever $a_1x_1 + \cdots + a_nx_n = 0$ with $a_i \in \mathbb{Z}$ we have $a_ix_i = 0$ for all i.

Lemma 8.2. *If $x_1, \ldots, x_n$ is a basis for M, then $M \simeq \langle x_1\rangle \times \cdots \times \langle x_n\rangle$.*

Proof. Note that $x_2, \ldots, x_n$ must form a basis for $\langle x_2, \ldots, x_n\rangle$. So by induction it suffices to prove that $M \simeq \langle x_1\rangle \times \langle x_2, \ldots, x_n\rangle$. For this it suffices to note that $H = \langle x_1\rangle$ and $K = \langle x_2, \ldots, x_n\rangle$ are normal subgroups which generate M and have trivial intersection and then use Theorem 7.1. □

Theorem 8.3. *Every finitely generated abelian group M has a basis.*

The proof of the theorem requires the following lemma.

Lemma 8.4. *Let $x_1, \ldots, x_n$ be a generating set for M and $c_1, \ldots, c_n \geqslant 0$ integers with $\gcd(c_i) = 1$. Then there is a generating set $y_1, \ldots, y_n$ for M with $y_1 = c_1x_1 + \cdots + c_nx_n$.*

Proof. By induction on $s = \sum c_i$. For $s = 1$ the statement is clear. If $s < 1$, then there must be at least two c_i which are nonzero. We can assume $c_1 \geqslant c2 > 0$. Now consider the generating set $x_1, x_1 + x_2, x_3, \ldots, x_n$ and the relatively prime coefficients $d_1 = c_1 - c_2, d_2 = c_2, \ldots, d_k = c_k$. The sum of these is less than s, so the induction hypothesis gives a generating set $y_1, \ldots, y_k$ with $y_1 = d_1x_1 + d_2x_2 + \cdots d_kx_k = c_1x_1 + c_2x_2 + \cdots + c_kx_k$ as desired. □

Proof of the Theorem. We do this by induction on the number k of generators required. Among all generating sets of size k choose one $x_1, \ldots, x_k$ with the order of x_1 minimal. If they do not form a basis, then there is a relation $a_1x_1 + \cdots a_kx_k$ with a_ix_i not all 0. Changing the signs of the x_i if necessary we may assume $a_i \geqslant 0$ and $a_1 < \operatorname{ord}(x_1)$. Let $d = \gcd(a_i)$ and set $c_i = a_i/d$. Applying the lemma we find a generating set $y_1, \ldots, y_k$ with $y_1 = c_1x_1 + \cdots c_kx_k$. Then $dy_1 = 0$ gives a contradiction to the fact that x_1 had minimal order. □

Lemma 8.5 (Chinese Remainder Theorem)**.** *If $m = p_1^{n_1} \cdots p_k^{n_k}$ is the prime factorization of m, then $\mathbb{Z}/m\mathbb{Z} \simeq \mathbb{Z}/p_1^{n_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_k^{n_k}\mathbb{Z}$.*

Proof. By induction it suffices to prove that $\mathbb{Z}/a \times \mathbb{Z}/b \simeq \mathbb{Z}/ab$ when a, b are relatively prime. For this consider the subgroups $a\mathbb{Z}/ab$ and $b\mathbb{Z}/ab$ of $\mathbb{Z}/ab$. Since a and b are relatively prime they generate (e.g., use the Euclidean algorithm to find $r, s \in \mathbb{Z}$ such that $ra + sb = 1$). These groups also have trivial intersection since the orders of the elements in each are relatively prime. Hence Theorem 7.1 gives the result. □

Remark 8.6. *This actually gives an isomorphism of rings.*

Corollary 8.7. *Let M be a finite abelian group. Then there are integers $n_1, \ldots, n_k$ such that $M \simeq \mathbb{Z}/n_1\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}$. Moreover, the n_i are uniquely specified by the condition that $2 \leqslant n_1 \mid n_2 \mid \cdots \mid n_k$.*

Corollary 8.8. *A finite abelian group is cyclic if and only if for every $n \geqslant 0$ it has at most n elements of order dividing n.*

Proof. Write M as $H \times Z/n\mathbb{Z}$ where H is a product of cyclic groups of order dividing n. If M has at most n elements of order dividing n, then H must be trivial. Conversely it is clear that a cyclic group satisfies the condition. □

Example 8.9. *If F is a field, then any finite subgroup $G < F^\times$ is cyclic. This is because the elements of order dividing n in G are roots of the polynomial $x^n - 1$, which has at most n roots by unique factorization in $F[x]$.*

9. Jordan-Hölder theorem

Definition 9.1. *Let G be a group. A subnormal series for G is a sequence of normal subgroups*

$$G = G_0 \rhd G_1 \rhd G_2 \rhd \cdots \rhd G_n$$

with $G_i \neq G_{i+1}$. It is a compostion series if it admits no proper refinement.

Example 9.2. *There are two compostion series for $\mathbb{Z}/6\mathbb{Z}$. Namely,*

$$\mathbb{Z}/6\mathbb{Z} \rhd 2\mathbb{Z}/6\mathbb{Z} \rhd 1\,, \text{ and}$$
$$\mathbb{Z}/6\mathbb{Z} \rhd 3\mathbb{Z}/6\mathbb{Z} \rhd 1\,.$$

Example 9.3. *There is just one composition series for D_3:*

$$D_3 \rhd \langle \rho \rangle \rhd 1 \,.$$

Example 9.4.

$$D_4 \rhd \langle \rho \rangle \rhd 1$$

is a subnormal series, but not a composition series because it can be refined to

$$D_4 \lhd \langle \rho \rangle \rhd \langle \rho^2 \rangle \rhd 1$$

Note that it is not required that $\langle \rho^2 \rangle$ be a normal subgroup of D_4, only that it is a normal subgroup of the term preceding it in the series.

Lemma 9.5. *Every finite group has at least one composition series.*

Proof. The finite group G has finitely many normal subgroups. Choose one that is maximal with respect to inclusion and set this as G_1. Now repeat with G_1 in place of G. This process must terminate since each group G_i has order smaller than the previous. Moreover this subnormal series admits no proper refinement since, by the correspondence theorem, any group $G_i \rhd H \rhd G_{i+1}$ corresponds to a normal subgroup of G_i containing G_{i+1}. By maximality of G_{i+1} we have that $H = G_i$ or $H = G_{i+1}$. □

As seen above, a group may have more than one composition series. The Jordan-Hölder theorem says that the composition factors G_i/G_{i+1} do not depend on the choice of composition series, hence are an invariant of the group.

Theorem 9.6 (Jordan-Hölder Theorem)**.** *Let G be a finite group and suppose*

$$G = G_0 \rhd G_1 \rhd G_2 \rhd \cdots \rhd G_n \;\; \textit{and}$$
$$G = H_0 \rhd H_1 \rhd H_2 \rhd \cdots \rhd H_m$$

are composition series for G. Then $m = n$ and there is a permutation $\sigma \in S_n$ such that $G_{i-1}/G_i \simeq H_{\sigma(i)-1}/H_{\sigma(i)}$ for all $i = 1, \ldots, n$.

Proof. By induction on the order of G, the case $|G| = 1$ being trivial. Given the two composition series set $K_2 = H_1 \cap G_1$ and take a composition series $K_2 \rhd K_1 \rhd K_2 \rhd \cdots \rhd K_\ell$ for K_2. Since K_2 is a maximal normal subgroup of H_1 and of G_1, we now have four composition series for G:

$$G \rhd H_1 \rhd H_2 \rhd \cdots \rhd H_m$$
$$G \rhd H_1 \rhd K_2 \rhd \cdots \rhd K_\ell$$
$$G \rhd G_1 \rhd K_2 \rhd \cdots \rhd K_\ell$$
$$G \rhd G_1 \rhd G_2 \rhd \cdots \rhd G_n \,.$$

We can apply the induction hypothesis to H_1 to see that $m = \ell$ and that the first two series have the same quotients. We can do the same with G_1 to see that $\ell = n$ and the last two series have the same quotients. Now the middle two series are the same after the K_2 term, and the first two quotients of each can be related using the isomorphism theorem. Namely we have

$$G/H_1 = H_1G_1/H_1 \simeq G_1/K_2 \quad \text{and} \quad G/G_1 = G_1H_1/G_1 \simeq H_1/K_2 \,.$$

Taken together this shows that the quotients of all four series are the same (up to permutation). □

Example 9.7. *For a finite abelian group the Jordan-Hölder factors are all of the form $\mathbb{Z}/p\mathbb{Z}$ (with multiplicity equal to the largest n such that p^n divides the order of the group).*

Example 9.8. *The quotients in the two composition series for $\mathbb{Z}/6\mathbb{Z}$ are cyclic groups of order 2 and 3. There are several composition series for D_4, but all have length 3 and all quotients are cyclic groups of order 2. We will see later in the course that in a group of order p^n with p prime, all of the quotients in a composition series are isomorphic to $\mathbb{Z}/p\mathbb{Z}$.*

We will see later that there are groups whose Jordan Hölder factors are not all abelian (and of prime order).

10. Group actions and the orbit-stabilizer theorem

Definition 10.1. *Let G be a group and X a set. A (left) group action of G on X is a map of sets*

$$G \times X \to X \ , \quad (g, x) \mapsto g \star x$$

satisfying:

(1) For all $x \in X$, $1 \star x = x$, and
(2) For all $x \in X$ and $g, h \in G$, $g \star (h \star x) = gh \star x$.

We say that G acts on X and write $G \circlearrowright X$.

Definition 10.2. *Suppose $G \circlearrowright X$.*
The orbit of $x \in X$ is the set $\mathrm{Orb}(x) = \{y \in X \ : \ \exists g \in G \text{ such that } g \star x = y\}$.
The stabilizer of $x \in X$ is the subgroup $\mathrm{Stab}(x) = \{g \in G \ : \ g \star x = x\}$.
When $\mathrm{Orb}(x) = X$ we say that the action is transitive.

Here are several examples:

(1) The group law $G \times G \to G$ is a group action of G on itself. This action is transitive and all stablizers are trivial (because of inverses).
(2) S_n acts on $\{1, 2, \ldots, n\}$ in the obvious way. For any i, $\mathrm{Stab}(i) \simeq S_{n-1}$ is the subgroup which permutes only the other $n-1$ symbols and $\mathrm{Orb}(i) = \{1, \ldots, n\}$, so the action is transitive.
(3) By definition D_n acts on the regular n-gon. This gives actions $D_n \circlearrowright \{1, \ldots, n\}$ (act on vertices) and $D_n \circlearrowright \mathbb{R}^2$ (center the n-gon at the origin and act on the plane by the corresponding rotations and reflections). The action $D_n \circlearrowright \{1, \ldots, n\}$ is the same as the restriction of the action $S_n \circlearrowright \{1, \ldots, n\}$ to the subgroup $D_n < S_n$. We leave it as an exercise to describe the orbits and stabilizers for these actions.
(4) $\mathrm{GL}_n(F)$ acts on F^n by $A \star \mathbf{v} = A\mathbf{v}$. This action is transitive and the stabilizer of $\mathbf{v}$ is the subgroup of matrices which have $\mathbf{v}$ as an eigenvector with eigenvalue 1. There is also an action of $\mathrm{GL}_n(F)$ on the set $\mathbb{P}^{n-1}(F)$ of lines through the origin in F^n.
(5) $\mu_2 = \{\pm 1\}$ acts on $\mathbb{Q}(\sqrt{2})$ by the rule $(-1) \star (a + b\sqrt{2}) = a - b\sqrt{2}$. Then $\mathrm{Orb}(z) = \{z, \overline{z}\}$ and $\mathrm{Stab}(z) = \begin{cases} \mu_2 & \text{if } z \in \mathbb{Q} \\ \{1\} & \text{otherwise} \end{cases}$

Now suppose $H < G$ is a subgroup.

(6) $G \circlearrowright G/H$ by the rule $g \star (g_0 H) = g g_0 H$. This action is transitive and the stabilzer of $g_0 H$ is the conjugate subgroup $g_0 H g_0^{-1}$.

(7) G acts on itself by conjugation: $g \star h = ghg^{-1}$. The orbit of $g \in G$ is called a conjugacy class. The stabilizer $\mathrm{Stab}(g) = \{h \in G \ : \ hg = gh\}$ is the subgroup of elements that commute with G; it is called the centralizer of g in G.
(8) Let $\mathrm{Sub}(G)$ be the set of subgroups of G. Then $G \circlearrowright \mathrm{Sub}(G)$ by the rule $g \star H = gHg^{-1}$.

Fact: The orbits of a group action $G \circlearrowright X$ give a partition of X.

Proposition 10.3. *Suppose $G \circlearrowright X$ transitively. Then for any $x \in X$ the map*

$$\phi : G/\mathrm{Stab}(x) \to X \ , \quad (g\,\mathrm{Stab}(x)) \mapsto g \star x$$

is an isomorphism of G-sets, i.e., is a bijection such that $\phi(g * y) = g \star \phi(y)$ where $g * y$ denotes the action on G on $G/\mathrm{Stab}(x)$ as in (6).

Proof. One easily checks that this map is well defined, injective, surjective and respects the G-actions. □

Corollary 10.4 (Orbit-Stabilizer Theorem)**.** *Suppose $G \circlearrowright X$ and $x \in X$. Then $|\,\mathrm{Orb}(x)| = [G : \mathrm{Stab}(x)]$.*

Proof. The action of G on X restricts to a transitive action of G on $\mathrm{Orb}(x)$. The proposition gives an isomorphism of $\mathrm{Orb}(x)$ with the coset space $G/\mathrm{Stab}(x)$. □

Example: Look at conjugacy classes in D_3.

11. The class equation and consequences for p-groups

Let G be a finite group and consider the action of G on itself by conjugation. The orbits are called conjugacy classes. The stabilizer of $g \in G$ under this action is the subgroup $\mathrm{Stab}(g) = C_G(g) = \{h \in G \ : \ gh = hg\}$ is called the centralizer of g in G. It is the set of elements of G that commute with G. The intersection of all of theses centralizers (as one ranges over all $g \in G$) is called the center of G and is denoted $Z(G)$. It is the set of elements of G that commute with all other elements.

Theorem 11.1 (Class equation)**.** *Suppose G is a finite group. Then*

(1) $|G| = \sum [G : C_G(g)]$, the sum ranging over a set of representatives g for the conjugacy classes, and
(2) $|G| = |Z(G)| + \sum [G : C_G(g)]$, the sum ranging over a set of representatives for the conjugacy classes of size at least 2.

Proof. The orbits of $G \circlearrowright G$ give a partition $G = \coprod \mathrm{Orb}(g)$, where the union ranges over a set of representatives for the orbits (which are conjugacy classes). By the orbit-stabilizer theorem, $|\,\mathrm{Orb}(g)| = [G : \mathrm{Stab}(g)] = [G : C_G(g)]$. This proves the first statement. For the second, use that $Z(G)$ is the set of elements with a conjugacy class of size 1 to decompose the sum. □

This theorem has several interesting corollaries for the structure of finite groups, and in particular for so-called p-groups. Here p is a prime number and G is a p-group if $|G|$ is a power of p.

Corollary 11.2 (Cauchy's Theorem)**.** *If $p \mid |G|$, then there is an element of order p in G.*

Proof. By induction on $|G|$. Case 1: If $p \nmid [G : C_G(x)]$ for some $x \notin Z(G)$, then $p \mid C_G(x)$ so we can apply the hypothesis to $C_G(x)$. Case 2: If $p \mid [G : C_G(x)]$ for all $x \notin Z(G)$, then by equation 2 we have $|Z(G)| \equiv 0 \mod p$, so there is an element of order p in $Z(G) < G$. □

Corollary 11.3. *G is a p-group if and only if every element has order a power of p.*

Proof. Lagrange's theorem implies one direction and Cauchy's theorem implies the other. □

Corollary 11.4. *Every p-group has nontrivial center (i.e., $Z(G) \neq \{1\}$).*

Proof. For any $x \in G$, $C_G(x)$ is a subgroup of G, hence is also a p-group. Consider the class equation modulo p:

$$0 = |G| = |Z(G)| + \sum [G : C_G(x)] = |Z(G)| .$$

In particular $|Z(G)| \neq 1$. □

Corollary 11.5. *Every group of order p^2 is abelian.*

Proof. By the previous corollary, the center is a nontrivial normal subgroup of G. If $Z(G) \neq G$, then $G/Z(G) \simeq \mathbb{Z}/p\mathbb{Z}$. Choose $a \in G$ such that its image generates $G/Z(G)$. Given $g_1, g_2 \in G$ we can write them as $g_1 = a^{m_1} z_1$ and $g_2 = a^{m_2} z_2$ with $z_i \in Z(G)$. Now an easy computation shows that $g_1 g_2 = g_2 g_1$. Hence G is abelian. □

Corollary 11.6. *Every group of order p^n has normal subgroups of order p^m for $1 \leqslant m \leqslant n$.*

Proof. By induction on n. Since $p||Z(G)|$ (as seen above) Cauchy's theorem shows that there is an element a of order p in $Z(G)$. Then $\langle a \rangle \lhd G$ is a normal subgroup. We can apply the induction hypothesis to $G/\langle a \rangle$ to get normal subgroups of this quotient of index $p, p^2, \ldots, p^{n-1}$. By the correspondence theorem these correspond to normal subgroups of G containing $\langle a \rangle$ of the same indices. □

Corollary 11.7. *The Jordan-Hölder factors of any p-group are all isomorphic to $\mathbb{Z}/p\mathbb{Z}$.*

Proof. The previous corollary shows there is a subnormal series all of whose quotients are isomorphic to $\mathbb{Z}/p\mathbb{Z}$. By the correspondence theorem this is a composition series ($\mathbb{Z}/p\mathbb{Z}$ has no normal subgroups). □

12. Sylow theorems

The so-called Sylow theorems were established by Ludwig Sylow in the 1870's. They are structural theorems concerning the p-subgroups of a finite group. Sylow's motivation came from Galois theory.

Definition 12.1. *Let G be a finite group and p a prime divisor of $|G|$. A subgroup $H < G$ is called a* Sylow-p-subgroup *if it is a p-group and $p \nmid (G : H)$.*

This means that the order of H is the largest power of p dividing the order of G.

Theorem 12.2. *Let G be a group of order $p^r m$ with p prime and m relatively prime to p. Then*

1. *There is a Sylow-p-subgroup of G;*
2. *Any two Sylow-p-subgroups of G are conjugate;*

(3) The number s_p of Sylow-p-subgroups satisfies

$$s_p \equiv 1 \mod p \quad \textit{and} \quad s_p \mid m\,;$$

(4) Every p-subgroup of G is contained in a Sylow-p-subgroup.

Example 12.3.
D_{25} has order 2×5^2. The subgroup $H = \langle \rho \rangle$ of rotations has order 5^2, so it is a Sylow-5 subgroup. This is the only Sylow-5-subgroup. The subgroup $\langle \rho^5 \rangle$ has order 5, so it is not a Sylow-p-subgroup (for any p). Any reflection generates a Sylow-2-subgroup, and conversely any subgroup of order 2 is generated by a reflection. Note that all of the Sylow-2-subgroups are conjugate. Moreover, the number of these is $s_2 = 25$, which is congruent to 1 mod 2 and divides 25.

For another example consider upper-triangular matrices in $\mathrm{GL}_n(\mathbb{F}_p)$.

Proof of Sylow (1). Let $X = \{A \subset G \ : \ |A| = p^r\}$ be the set of subsets of G of size p^r and define an action of G on X by $g \star A = \{ga \ : \ a \in A\}$. Given A, consider $H := \mathrm{Stab}(A) = \{g \in G \ : \ gA = A\} < G$. First we claim that $|H| \leqslant |A| = p^r$. This is because for any $a \in A$ there is an injective map of sets $H \to A$ given by $h \mapsto ha$. Now the orbit-stabiliser theorem gives $p^r m = (G : 1) = (G : H)(H : 1) = |Orb(A)|(H : 1)$ so it suffices to find some A such that $|Orb(A)|$ is not divisible by p. Since the orbits in X partition X we have

$$\sum |\,\mathrm{Orb}(A_i)| = |X| = \binom{p^r m}{p^r} \not\equiv 0 \mod p\,,$$

so there must be an orbit of size not divisible by p. □

Lemma 12.4. *Let P be a Sylow-p-subgroup and $H < G$ a p-group. If $H < N_G(P) = \{g \in G \ : \ gPg^{-1}\}$, then $H < P$.*

Proof. $H < N_G(P)$ and $P \lhd N_G(P)$, so we may apply the isomorphism theorem to get that $(HP : P) = (H : H \cap P) =$ a power of p, since H is a p-group. On the other hand $(HP : P)$ divides $(G : P)$ which is prime to p, so $(HP : P) = 1$. □

Proof of the rest of Sylow theorem. Let X be the set of Sylow-p-subgroups of G and consider the action of G on X by conjugation. Let $O \subset X$ be an orbit. We will show that $O = X$, $|O| \equiv 1 \mod p$ and $|O| \mid m$.

For any Sylow-p-subgroup $P < G$ we may further decompose $O = \bigcup O_i$ as a union of orbits of O under the action of P. Note that all O_i have order $|O_i| = [P : \mathrm{Stab}(O_i)]$, which is a power of p. We claim that there is at most 1 orbit of size 1, which must be $O_i = \{P\}$ and this only occurs when $P \in O$. To see this note that $O_i = \{Q\}$ if and only if $P \subset N_G(Q)$,which implies $P = Q$ by the lemma. Taking $P \in O$ we find that $|O| \equiv 1 \mod p$. If there were some $P \in X \backslash O$, then we would have no orbits of size 1 and thus conclude $|O| \equiv 0 \mod p$. This contradiction shows that $O = X$.

Now $|O| = (G : N_G(P)) = \frac{(G:1)}{(N_G(P):P)(P:1)} = \frac{m}{(N_G(P):P)} \mid m$.

Finally, we show that every p-subgroup of G is contained in a Sylow-p-subgroup. For this suppose $H < G$ is a p-subgroup and consider the action of H on X by conjugation. Since the orbits have size a power of p and $|X| \equiv 1 \mod p$ we must have at least one orbit of size 1. This means that $H < N_G(P)$ for some P and hence that $H < P$ by the lemma. □

13. The alternating group

Definition 13.1. *Let $\sigma \in S_n$ be a permutation. We say that σ is* even *or* odd *correspondingly as the set*

$$\{(i,j) \ : \ i<j \text{ and } \sigma(i)>\sigma(j)\}$$

has an even or odd number of elements. The signature *of a permutation is ± 1 correspondingly as it is even or odd.*

Example 13.2. (132) *is an even permutation in S_3 since the set in the definition consists of the pairs $\{(2,1),(3,1)\}$.*

Lemma 13.3. *The signature defines a surjective homomorphism* $\operatorname{sgn} : S_n \to \{\pm 1\}$.

Proof. Surjectivity is clear since the identity is even and (12) is odd. Define an action of S_n on the polynomial ring $\mathbb{Z}[x_1,\dots,x_n]$ by $\sigma \star x_i = x_{\sigma(i)}$ and extending linearly. Consider the polynomial $P(x_1,\dots,x_n) = \prod_{1\leqslant i<j\leqslant n}(x_j - x_i)$. Then $\sigma(P) = \operatorname{sgn}(\sigma)P$. Since this is a group action we have

$$\operatorname{sgn}(\sigma\tau)P = (\sigma\tau)\star P = \sigma \star \operatorname{sgn}(\tau)P = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)P$$

so that $\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$. □

Definition 13.4. *The* Alternating group *A_n is the kernel of* $\operatorname{sgn} : S_n \to \{\pm 1\}$.

Note that $|A_n| = |S_n|/2 = n!/2$ and $A_n \lhd S_n$ is a normal subgroup.

Lemma 13.5. *Every $\sigma \in S_n$ can be written as a product of disjoint cycles.*

Proof. The supports of each cycle will be an orbit of $\{1,\dots,b\}$ under the action of σ. □

This gives a partition of $\{1,\dots,n\}$.

Lemma 13.6. *Two permutations are conjugate if and only if they give the same partition of $\{1,\dots,n\}$.*

Proof. Exercise. Use that for $\sigma = (i_1 i_2 \dots i_n)$ we have $g\sigma g^{-1} = (g(i_1)g(i_2)\cdots g(i_n))$ □

Example 13.7. *The permutation $\sigma \in S_5$ given by $i \mapsto 4i \mod 5$ has orbits* $\operatorname{Orb}(1) = \{1,4\}, \operatorname{Orb}(2) = \{2,3\}$ *and* $\operatorname{Orb}(5) = \{5\}$. *We can write this as $(14)(23)(5)$. This is conjugate to the permutation $\tau(23)(45)$. Specifically, $g\sigma g^{-1} = \tau$ for $g = (15)$.*

Corollary 13.8. *Every $\sigma \in S_n$ can be written as a product of transpositions.*

Proof. It suffices to treat the case when $\sigma = (i_1\, i_2\, \cdots, i_r)$ is a cycle in which case we see that

$$(i_1\, i_2\, \cdots, i_r) = (i_1\, i_r)\cdots(i_{r-2}\, i_r)(i_{r-1}\, i_r)\,.$$

□

Corollary 13.9. *Every $\sigma \in A_n$ can be written as a product of 3-cycles.*

Proof. Write σ as a product of transpositions. Since the signature of a transposition is -1 there are an even number of transpositions in the product. It therefore suffices to show that any product of two transpositions can be written as a product of 3-cycles. For this consider:

$$(i\,j)(k\,\ell) = \begin{cases} (ikj) & j = k\,, \\ (ij\ell)(ik\ell) & i,j,k,\ell \text{ distinct}\,, \\ 1 & (i\,j) = (k\,\ell)\,. \end{cases}$$

□

The following gives our first example of a nonabelian simple group.

Theorem 13.10 (Galois)**.** *For $n \geqslant 5$, A_n has no proper normal subgroups.*

A group is called solvable if all of its Jordan-Hölder factors are abelian. The terminology comes from Galois theory and the relationship to solvability of a polynomial by radicals. We will discuss this connection later in the course.

Corollary 13.11. *For $n \geqslant 5$, the Jordan-Hölder factors of S_n are $\mathbb{Z}/2$ and A_n. In particular, A_n and S_n are not solvable.*

One proves the theorem in two steps, which are given in the following lemmas.

Lemma 13.12. *If $n \geqslant 5$, $N \lhd A_n$ and N contains a 3-cycle, then $N = A_n$.*

Lemma 13.13. *If $N \lhd A_n$ and $N \neq \{1\}$, then N contains a 3-cycle*

Proof of Lemma 13.12. Let $\gamma \in N$ be a 3-cycle and let $\sigma \in S_n$ be any 3-cycle. It suffices to show $\sigma \in A_n$. By Lemma 13.6 there is a $g \in S_n$ such that $g\gamma g^{-1} = \sigma$. If $g \in A_n$, then $\sigma \in A_n$. If $g \notin A_n$, then choose a transposition (ij) disjoint from σ (which is possible since $n \geqslant 5$). Then $tg \in A_n$ and $tg\gamma(tg)^{-1} = t\sigma t = \sigma \in N$. □

Before giving the proof of Lemma 13.13 let us introduce the notion of fixed subsets. Given $g \in G \circlearrowright X$ a group acting on a set X, let $X^g = \{x \in X \ : \ g \star x = x\}$ be the subset of elements fixed by g. For example, if $\sigma = (i_1 i_2 \cdots i_r)$ is an r-cycle in S_n, then $\{1, \ldots, n\}^\sigma = \{i \ : \ i \neq i_j \text{ for any } j = 1, \ldots, r\}$. Note that for an r-cycle σ we have $|\{1, \ldots, n\}^\sigma| = n - r$, except for a cycle of length 1. Note also that any permutation fixing $n-2$ elements is a transposition and any permutation fixing exactly $n-3$ elements must be a 3-cycle.

Proof of Lemma 13.13. Let $\sigma \in N$, $\sigma \neq 1$. We show that if $\{1, \ldots, n\}^\sigma$ has size less than $n-3$, then we can construct $1 \neq \sigma' \in N$ which fixes more elements of $\{1, \ldots, n\}$. Repeating if necessary we eventually construct an element of N which fixes exactly 3 elements and is therefore a 3-cycle.

Write σ as a product of disjoint cycles. We consider two cases. In the first the product contains a cycle $(i_1 i_2 i \cdots i_r)$ of length at least 3. In this case we can find i_4, i_5 distinct from i_1, i_2, i_3 that are not fixed by σ. Let $\gamma = (i_5 i_4 i_3)$ and $\sigma_1 = \gamma\sigma\gamma^{-1} = (i_1 i_2 i_4 \ldots) \cdots \in N$ (since $N \lhd A_n$). Then $\sigma' = \sigma_1\sigma^{-1} \in N$. Now $\sigma' \neq 1$ since $\sigma_1 \neq \sigma$. Everything fixed by σ is fixed by σ' since if $\sigma(i) = i$, then $i \neq i_3, i_4, i_5$ and so i is fixed by σ and γ. Also $\sigma'(i_1) = \sigma_1\sigma^{-1}(i_1) = \sigma_1(i_2) = i_1$, so σ' fixes more than σ does.

In the second case σ is a product of disjoint transpositions. Say $\sigma = (i_1\, i_2)(i_3\, i_4)\cdots$. Choose any i_5 distinct from $i_1, \ldots, i_4$ and let $\tau = (i_5\, i_4\, i_3) \in A_n$. Now set $\sigma' = \tau\sigma\tau^{-1}\sigma^{-1}$. Now check that

- if $i \notin \{i_1, \ldots, i_5\}$ and $\sigma(i) = i$, then $\sigma'(i) = i$.
- $\sigma'(i_4) = i_5$, so $\sigma' \neq 1$
- $\sigma'(i_1) = i_1$

Hence σ' fixes more elements than σ, but not all. □