The Alternating Group

R. C. Daileda

1 The Parity of a Permutation

Let $n \geq 2$ be an integer. We have seen that any $\sigma \in S_n$ can be written as a product of transpositions, that is there exist transpositions $\tau_1, \ldots, \tau_m \in S_n$ so that

$$\sigma = \tau_1 \tau_2 \cdots \tau_m. \tag{1}$$

Although such an expression is never unique for a given σ, there is still an important invariant that can be extracted from the factorization (1), namely the parity (even or odd) of m. Specifically, we will say that σ is *even* if there is an expression of the form (1) with m even, and that σ is odd if otherwise. Note that if σ is odd, then (1) holds only when m is odd. However, the converse is not immediately clear. That is, it is not *a priori* evident that a given permutation can't be expressed as a product of both an even and an odd number of transpositions.

Although it is true that this situation is, indeed, impossible, there is no simple, direct proof that this is the case. The easiest proofs involve an auxiliary quantity known as the *sign* of a permutation. The sign is uniquely determined for any given permutation by construction, and is easily related to the parity of m. This thereby shows that the parity of a permutation is uniquely determined as well. The proof that we give below defines the sign of a permutation indirectly through the *action* of S_n on a specific multivariate polynomial. Our first goal is to define this action in general and deduce some of it's properties.

1.1 The Action of S_n on $\mathbb{Z}[X_1, X_2, \ldots, X_n]$

Let $\mathbb{Z}[X_1, X_2, \ldots, X_n]$ denote the set of all polynomials in the variables $X_1, X_2, \ldots, X_n$ with coefficients in $\mathbb{Z}$. Given $f \in \mathbb{Z}[X_1, X_2, \ldots, X_n]$ and $\sigma \in S_n$, define $\sigma f \in \mathbb{Z}[X_1, X_2, \ldots, X_n]$ by

$$(\sigma f)(X_1, X_2, \ldots, X_n) = f(X_{\sigma(1)}, X_{\sigma(2)}, \ldots, X_{\sigma(n)}).$$

That is, σ replaces the ith variable in f with $X_{\sigma(i)}$, for all i. Because σ permutes the variables of f, but leaves the coefficients alone, it is clear that σf also has integer coefficients. Likewise, if $c \in \mathbb{Z}$, then

$$\sigma(cf) = c(\sigma f), \tag{2}$$

because multiplication by c scales *every* coefficient, but has no effect on any of the variables $X_1, X_2, \ldots, X_n$.

Example 1. If $\sigma = (142)(36) \in S_6$ and $f = X_1X_3X_5^2 - 7X_2^3X_4 + 2X_5X_6$, then

$$\sigma f = X_4X_6X_5^2 - 7X_1^3X_2 + 2X_5X_3.$$

And if $c = 3$, then $3f = 3X_1X_3X_5^2 - 21X_2^3X_4 + 6X_5X_6$ so that

$$\sigma(3f) = 3X_4X_6X_5^2 - 21X_1^3X_2 + 6X_5X_3 = 3(\sigma f).$$

If we also have $\tau \in S_n$, then

$$(\tau(\sigma f))\,(X_1, X_2, \dots, X_n) = (\sigma f)(X_{\tau(1)}, \dots, X_{\tau(n)}) = (\sigma f)(Y_1, \dots, Y_n),$$

where $Y_i = X_{\tau(i)}$ for all i. We then have

$$(\sigma f)(Y_1, \dots, Y_n) = f(Y_{\sigma(1)}, \dots, Y_{\sigma(n)}) = f(X_{\tau(\sigma(1))}, \dots, X_{\tau(\sigma(n))}),$$

Since $\tau(\sigma(i)) = (\tau\sigma)(i)$ for all i, we find that

$$(\tau(\sigma f))\,(X_1, X_2, \dots, X_n) = f(X_{(\tau\sigma)(1)}, \dots, X_{(\tau\sigma)(n)}) = ((\tau\sigma)\,f)(X_1, \dots, X_n).$$

That is:

Proposition 1. *For all $\tau, \sigma \in S_n$ and all $f \in \mathbb{Z}[X_1, \dots, X_n]$ one has*

$$(\tau\sigma)f = \tau(\sigma f).$$

1.2 The Semi-Discriminant

We now introduce a special polynomial that will help us determine the parity of a permutation. Define

$$\Delta(X_1, X_2, \dots, X_n) = \prod_{1 \le i < j \le n} (X_j - X_i) \in \mathbb{Z}[X_1, X_2, \dots, X_N]. \tag{3}$$

The product runs over all ordered pairs (i, j) with $1 \le i < j \le n$. The polynomial Δ^2 is known as the *discriminant*. As such we choose to call Δ the *semi-discriminant*.

Example 2. When $n = 3$ we have

$$\Delta(X_1, X_2, X_3) = (X_3 - X_1)(X_3 - X_2)(X_2 - X_1).$$

When $n = 4$ we have

$$\Delta(X_1, X_2, X_3, X_4) = (X_4 - X_3)(X_4 - X_2)(X_4 - X_1)(X_3 - X_2)(X_3 - X_1)(X_2 - X_1).$$

The ordered pairs (i, j) with $1 \le i < j \le n$ that index the product (3) are in bijective correspondence with the two element subsets $\{i, j\}$ of $I_n = \{1, 2, \dots, n\}$, since each such set has a largest element. Let $P \subseteq \mathcal{P}(I_n)$ be the set of all two element subsets of I_n. Let $\sigma \in S_n$ and choose $\{i, j\} \in P$. Since $i \neq j$, we must have $\sigma(i) \neq \sigma(j)$ because σ is a bijection. So $\sigma(\{i, j\}) = \{\sigma(i), \sigma(j)\} \in P$. Thus σ defines function $\sigma : P \to P$. It's surjective since the preimage of $\{i, j\}$ is $\{\sigma^{-1}(i), \sigma^{-1}(j)\}$. And it's therefore also injective by the Pigeonhole Principle, since P is finite. So σ acts to permute the two element subsets of I_n.

Therefore in the product

$$(\sigma\Delta)(X_1, \ldots, X_n) = \Delta(X_{\sigma(1)}, \ldots, X_{\sigma(n)}) = \prod_{1 \leq i < j \leq n} (X_{\sigma(j)} - X_{\sigma(i)}) \tag{4}$$

the pairs of indices $\{\sigma(i), \sigma(j)\} = \sigma(\{i, j\})$ run through every member of P exactly once as $\{i, j\}$ does. However, σ might reverse the ordering $i < j$. So the factors appearing in (3) and (4) are the same, up to the order of the variables in each factor. Since $X_j - X_i = -(X_i - X_j)$, returning the variables in each factor to their original ordering introduces a number of sign changes.[1] This means that $\sigma\Delta = \pm\Delta$, the sign depending on σ, and we define $\epsilon : S_n \to \{\pm 1\}$ by

$$\sigma\Delta = \epsilon(\sigma)\Delta.$$

We call $\epsilon(\sigma)$ the *sign* of $\sigma \in S_n$.

Theorem 1. *The sign $\epsilon : S_n \to \{\pm 1\}$ is a homomorphism.*

Proof. Let $\sigma, \tau \in S_n$. By Proposition 1 we have

$$\epsilon(\tau\sigma)\Delta = (\tau\sigma)\Delta = \tau(\sigma\Delta) = \tau(\epsilon(\sigma)\Delta) = \epsilon(\sigma)(\tau\Delta) = \epsilon(\sigma)\epsilon(\tau)\Delta.$$

Therefore

$$\epsilon(\tau\sigma) = \epsilon(\sigma)\epsilon(\tau) = \epsilon(\tau)\epsilon(\sigma),$$

since the signs belong to $\{\pm 1\}$ and therefore commute with one another. □

In order to relate the parity of a permutation to its sign, we need to determine how ϵ acts on transpositions.

Lemma 1. *For any transposition $(rs) \in S_n$ we have $\epsilon((rs)) = -1$.*

Proof. We may assume $r < s$. The factors in (3) that are altered by (rs) are those whose subscripts include r, s or both. These have the following forms in Δ and $(rs)\Delta$:

[1] The number of sign changes required is called the number of *inversions* in σ, which simply counts how often σ reverses the ordering of a pair of indices. So if $I(\sigma)$ is the number of inversions in σ, we find that $\sigma\Delta = (-1)^{I(\sigma)}\Delta$.

	Δ		$(rs)\Delta$
	$X_s - X_r$	$\rightarrow$	$X_r - X_s$
$s < k$	$X_k - X_s$ $X_k - X_r$	$\rightarrow$	$X_k - X_r$ $X_k - X_s$
$r < k < s$	$X_s - X_k$ $X_k - X_r$	$\rightarrow$	$X_r - X_k$ $X_k - X_s$
$k < r$	$X_s - X_k$ $X_r - X_k$	$\rightarrow$	$X_r - X_k$ $X_s - X_k$

The pairs of factors in the second and final groups are simply interchanged by (rs). The factors in the first and third groups have indices that are inverted (have their order reversed) by (rs) and introduce signs changes in the product defining Δ. But those in the third group occur in pairs, so their sign changes cancel. The only remaining sign change comes from the factor $X_s - X_r$, and we therefore have $(rs)\Delta = -\Delta$. □

Corollary 1. *Let $\sigma \in S_n$. If σ is written as the product of N transpositions, then*

$$\epsilon(\sigma) = (-1)^N.$$

In particular the parity of N is uniquely defined by σ.

Proof. If $\sigma = \tau_1 \cdots \tau_N$ and every $\tau_i \in S_n$ is a transposition, then Theorem 1 and Lemma 1 imply

$$\epsilon(\sigma) = \epsilon(\tau_1) \cdots \epsilon(\tau_N) = (-1)^N.$$

The second part follows from the fact that $(-1)^M = (-1)^N$ if and only if $M \equiv N \pmod 2$. □

Corollary 1 finally tells us that that every member of S_n is either even or odd, but never both: $\sigma \in S_n$ is even if and only if $\epsilon(\sigma) = 1$ and odd if and only if $\epsilon(\sigma) = -1$.

1.3 Exercises

The following exercises provide an alternate proof of the identity $\sigma\Delta = \pm\Delta$ for $\sigma \in S_n$.

Exercise 1. The set $\mathbb{Z}[X_1, X_2, \ldots, X_n]$ is a *ring* under ordinary addition and multiplication of polynomials. Show that the action of S_n on $\mathbb{Z}[X_1, \ldots, X_n]$ respects this ring structure. That is, show that for all $\sigma \in S_n$ and $f, g \in \mathbb{Z}[X_1, \ldots, X_n]$ one has $\sigma(f + g) = \sigma f + \sigma g$ and $\sigma(fg) = (\sigma f)(\sigma g)$.

Exercise 2. Explain why

$$\Delta^2 = (-1)^{\frac{n(n-1)}{2}} \prod_{i \neq j} (X_i - X_j).$$

Use this to show directly that $\sigma(\Delta^2) = \Delta^2$ for any $\sigma \in S_n$.

Exercise 3. Use the results of the preceding exercises to deduce that $(\sigma\Delta - \Delta)(\sigma\Delta + \Delta) = 0$ for any $\sigma \in S_n$. Conclude that $\sigma\Delta = \pm\Delta$.

2 The Alternating Group

The *alternating group* A_n is defined to be the kernel of ϵ:

$$A_n = \ker \epsilon.$$

This means that A_n consists of all even permutations and is normal in S_n. Since Lemma 1 implies that ϵ is an epimorphism, the First Isomorphism Theorem tells us that we have:

Theorem 2. *The sign epimorphism $\epsilon : S_n \to \{\pm 1\}$ induces an isomorphism*

$$S_n/A_n \cong \{\pm 1\}.$$

Corollary 2. *We have $[S_n : A_n] = 2$.*

An important feature of the alternating group is that, unless $n = 4$, it is a simple group. A group G is said to be *simple* if it has no nontrivial proper normal subgroups. For example, Lagrange's Theorem implies that every group of prime order is simple. But this is a somewhat uninteresting result: a group of prime order doesn't have *any* nontrivial proper subgroups. An alternating group, on the other hand, can have a multitude of subgroups, and so the alternating groups A_n furnish a more satisfying example of a class of simple groups.

A_2 is simple because it's the trivial group. We have actually already proven that A_3 is simple, since $|A_3| = 3$ is prime. The subgroup $K = \{\mathrm{Id}, (12)(34), (13)(24), (14)(23)\}$, which is isomorphic to the Klein 4-group, is normal in S_4. Since $K < A_4$, this proves A_4 fails to be simple. The proof that A_n is simple for $n \geq 5$ is a bit more involved, but is purely computational. It involves nothing more than careful manipulations of permutations, 3-cycles in particular.

We require three preparatory lemmas.

Lemma 2. *A_n is generated by 3-cycles.*

Proof. First notice that if i, j, k are distinct, then

$$(ijk) = (ik)(ij) \in A_n,$$

so that A_n contains every 3-cycle. So it suffices to show that every product $\tau_1\tau_2$ of a pair of transpositions is a product of 3-cycles. If τ_1 and τ_2 are not disjoint, the computation above shows that their product is a 3-cycle. On the other hand, if $\tau_1 = (ij)$ and $\tau_2 = (rs)$ with i, j, r, s distinct, then

$$\tau_1\tau_2 = (ij)(ir)(ri)(rs) = (jir)(irs).$$

This completes the proof. □

Lemma 3. *If $n \geq 5$, then all 3-cycles are conjugate in A_n.*

Proof. Because of the identity

$$\sigma(i_1\, i_2\, \cdots\, i_r)\sigma^{-1} = (\sigma(i_1)\, \sigma(i_2)\, \cdots \sigma(i_r)), \tag{5}$$

all cycles of any given length are conjugate in S_n. We must show that when $r = 3$, we can always take σ to be even. So let (ijk) and (rst) be 3-cycles, and choose $\sigma \in S_n$ so that $\sigma(ijk)\sigma^{-1} = (rst)$. If σ is even there's nothing to prove, so suppose σ is odd. Because $n \geq 5$, we can find $a, b \in \{1, 2, \ldots, n\}$ so that a, b, i, j, k are all distinct. Then (ab) commutes with (ijk), $\sigma(ab) \in A_n$ and

$$(\sigma(ab))(ijk)(\sigma(ab))^{-1} = \sigma(ab)(ijk)(ab)\sigma^{-1} = \sigma(ijk)\sigma^{-1} = (rst).$$

□

Lemma 4. *Suppose $n \geq 5$. If a normal subgroup N of A_n contains a 3-cycle, then $N = A_n$.*

Proof. Let $N \lhd A_n$. If N contains a 3-cycle, normality implies N contains all of its conjugates in A_n. This means N contains every 3-cycle, by Lemma 3. Lemma 2 then tells us that $N = A_n$. □

Theorem 3. *If $n \neq 4$, then A_n is simple.*

Proof. It suffices to assume that $n \geq 5$. Let $N \lhd A_n$ be nontrivial. We will show that $N = A_n$ by proving that N contains a 3-cycle and then appealing to Lemma 4. For convenience, set $I_n = \{1, 2, \ldots, n\}$. We will find the 3-cycle we need by considering the number of fixed points of a nonidentity permutation in N.

For any $\sigma \in S_n$, we say that $i \in I_n$ is a *fixed point* of σ if $\sigma(i) = i$. This is equivalent to the statement that in the disjoint cycle decomposition of σ, i belongs to a 1-cycle (or doesn't appear at all, if we omit 1-cycles). Now suppose that $\sigma \in N$ is nontrivial. We claim that unless σ is a 3-cycle, we can always find a nontrivial element of N with more fixed points than σ.

There are two cases to consider. First, suppose that σ is a product of disjoint transpositions (at least two, since σ is nontrivial and even). Consider a pair $(ij), (rs)$ of disjoint transpositions occurring as cycles in σ. Since $n \geq 5$, there is a $t \in I_n \setminus \{i, j, r, s\}$. Let $\tau = (ij)(rt)$ and set $\sigma' = \sigma\tau\sigma\tau^{-1}$. Since N is normal in A_n and τ is even, $\sigma' \in N$. Write $\sigma = (ij)(rs)\gamma$ with γ disjoint from (ij) and (rs), that is i, j, r and s are all fixed points of γ. Then γ commutes with (ij) and (rs), which commute with each other, so that $\tau = \tau^{-1}$ and hence

$$\sigma' = ((ij)(rs)\gamma(ij)(rt))^2 = ((rs)\gamma(rt))^2 = (\gamma(rs)(rt))^2 = (\gamma(tsr))^2.$$

Since $\sigma'(t) = r \neq t$, σ' is nontrivial. Furthermore, we see that σ' fixes i, j and every fixed point of σ, with the possible exception of t. In particular, $\mathrm{id} \neq \sigma' \in N$ has at least one more fixed point than σ. This proves our claim in this case.

Now we suppose that σ has a cycle of length at least 3, but is not simply a 3-cycle itself. Write $\sigma = (ijk\cdots)\gamma$ with γ fixing $i, j, k, \ldots$ (all distinct). If σ has exactly $n-4$ fixed points, it must be that $\sigma = (ijkr)$ is a 4-cycle. But 4-cycles are odd, so this is impossible. It follows that σ has at most $n-5$ fixed points. Then there must exist distinct $r, s \in I_n \setminus \{i, j, k\}$ that are not fixed by σ. Let $\tau = (krs)$. As in the preceding paragraph, let $\sigma' = \sigma^{-1}\tau\sigma\tau^{-1} \in N$. Because τ fixes i and j as well as every fixed point of σ, σ' fixes i and every fixed point of σ. Thus, σ' has one more fixed point than σ. Since $\sigma'(j) = \sigma^{-1}(r) \neq j$, σ' is nontrivial. This establishes our claim.

We now complete the proof of the theorem. Let $\sigma \in N$ be a nontrivial permutation with the maximum number of fixed points. Then it cannot fall into either of the preceding classes. Thus, σ is a 3-cycle, and $N \lhd A_n$ by Lemma 4.

□

Coupled with the fact that the index of A_n in S_n is as small as possible (without being trivial), the simplicity of A_n prevents the existence of other normal subgroups of S_n. This is an easy consequence of the following general group-theoretic lemmas.

Lemma 5. *Let G be a group, $N \lhd G$ and $H \leq G$. Then $H \cap N \lhd H$ and $[H : H \cap N]$ divides $[G : N]$.*

Proof. Let $H \to G/N$ be the homomorphism given by the composition of inclusion and the canonical epimorphism. Its kernel is $H \cap N$, making this a normal subgroup of H, and the First Isomorphism Theorem implies $H/(H \cap N)$ is isomorphic to a subgroup of G/N. The result follows at once. □

Corollary 3. *Let G be a group and $H, N \leq G$ with $[G : N] = 2$ (so that $N \lhd G$). Then $H \leq N$ or $[H : H \cap N] = 2$.*

Lemma 6. *Let G be a group with a simple subgroup N of index 2. If $H \lhd G$ and H is nontrivial, then $N \leq H$, or $|H| = 2$ and $H \leq Z(G)$.*

Proof. By Lemma 5, $H \cap N \lhd N$. As N is simple, we must have $H \cap N = \{e\}$ or $H \cap N = N$. In the second case, $N \leq H$ and we are done. In the first case, Corollary 3 implies that

$$2 = [H : H \cap N] = [H : \{e\}] = |H|.$$

It is an easy exercise to show that a normal subgroup of order two must be contained in $Z(G)$, and this completes the proof. □

Lemma 7. *For $n \geq 3$, $Z(S_n) = \{\mathrm{id}\}$.*

Proof. Let $\sigma \in S_n$, $\sigma \neq \mathrm{id}$. If σ has a fixed point i, choose $j \neq i$ not fixed by σ and set $\tau = (ij)$. Then $\tau\sigma\tau^{-1}$ fixes j and hence $\tau\sigma\tau^{-1} \neq \sigma$. If σ has no fixed points, then $\sigma(1) = i \neq 1$. Choose $j \notin \{1, i\}$ (possible since $n \geq 3$) and set $\tau = (ij)$. Then $\tau\sigma\tau^{-1}(1) = j \neq i = \sigma(1)$ so that $\tau\sigma\tau^{-1} \neq \sigma$. In either case, we see that if $\sigma \neq \mathrm{id}$, then $\sigma \notin Z(G)$, which proves the result.

□

Theorem 4. *If $n \neq 4$, the only nontrivial proper normal subgroup of S_n is A_n.*

Proof. This now follows from Lemmas 6 and 7. □

It is easy to see that the conclusion of Theorem 4 fails when $n = 4$. Indeed, we have already observed that the nontrivial proper normal subgroup K of A_4 is also normal in S_4.

Another consequence of Theorem 3 concerns commutators and solvability. Recall that given a group G its *commutator subgroup* is

$$G' = \langle xyx^{-1}y^{-1} \mid x, y \in G\rangle.$$

The elements $[x, y] = xyx^{-1}y^{-1}$ are called *commutators*. Note that $[x, y] = e$ if and only if $xy = yx$, so that G' is trivial if and only if G is abelian. Any conjugate of a commutator is also a commutator, which implies that G' is a normal subgroup of G. It has the property that for any $H \lhd G$, G/H is abelian if and only if $G' \leq H$. This is simply because $[x, y]H = [xH, yH]$ for all $x, y \in G$, so that H contains every commutator in G (and hence all of G') if and only if $[xH, yH] = H$ for all $xH, yH \in G/H$, which we have already observed is equivalent to G/H being abelian. Therefore G/G' is the largest abelian quotient of G.

Theorem 5. *For all $n \geq 2$, $S'_n = A_n$.*

Proof. By the First Isomorphism Theorem, the sign epimorphism ϵ yields an isomorphism $S_n/A_n \cong \{\pm 1\}$. Since $\{\pm 1\}$ is abelian, $S'_n \leq A_n$. For $n \neq 4$, S'_n is nontrivial, normal in S_n, and A_n is simple. It follows that $S'_n = A_n$.

To treat the case $n = 4$, we replace the final step in the argument above with a somewhat more direct argument (which applies to *any* $n \geq 3$). Given distinct i, j, k, we have

$$[(ij), (jk)] = (ij)(jk)(ij)(jk) = (ijk)^2 = (ikj).$$

By Lemma 2, we conclude that A_n is generated by commutators, and hence $A_n \leq S'_n$. We already know $S'_n \leq A_n$, so $S'_n = A_n$.

□

We can also determine the commutator subgroup of A_n.

Theorem 6. *A'_n is trivial for $n \leq 3$, $[A_4 : A'_4] = 3$, and $A'_n = A_n$ for $n \geq 5$.*

Proof. The first case is trivial, since A_2 and A_3 are abelian. The last case is just as easy, since when $n \geq 5$, A_n is simple and nonabelian. To deal with A_4, recall that in this case there is a normal subgroup K of order 4. Hence A_4/K has order 3, and is therefore abelian. This in turn implies that $A'_4 \leq K$. It's easy to check that K has no nontrivial proper subgroups that are normal in A_4, which means that we must have $A'_4 = K$. □

Our proof that A_4 is not simple was perhaps somewhat unsatisfying. Without motivation, we simply produced a nontrivial proper normal subgroup. Theorem 6 now explains where it came from: it's A'_4. So the reason A_4 fails to be simple is because its commutator subgroup is nontrivial and proper! Notice that when $n = 3$ we also have $[A_3 : A'_3] = 3$. So one could restate Theorem 5 as follows: $A'_n = A_n$, unless $n = 3$ or 4, when $[A_n : A'_n] = 3$.

Our next result requires a definition. We say a finite group G is *solvable* if there is a *subnormal series*

$$\{e\} = G_r \lhd G_{r-1} \lhd G_{r-2} \lhd \cdots G_1 \lhd G_0 = G,$$

where G_i/G_{i+1} is abelian for all i (also called an *abelian series*). If one forms the *derived series* for G, by setting $G^{(1)} = G'$ and $G^{(i+1)} = (G^{(i)})'$, it is not difficult to show that G is solvable if and only if there is an r so that $G^{(r)} = \{e\}$.

Every abelian group is clearly solvable, as is D_n for every n (why?). The series $\{\mathrm{id}\} \lhd K \lhd A_4$ shows that A_4 is solvable, too. A deep result of Feit-Thompson states that, in fact, every group of odd order is solvable. On the other hand, the symmetric and alternating groups provide examples of families of groups that are not solvable.

Theorem 7. *For $n \geq 5$, S_n and A_n are not solvable.*

Proof. Since $S_n' = A_n$ and $A_n' = A_n$ by Theorems 5 and 6, the derived series of S_n (and A_n) terminates in an infinite string of A_n's. Thus, S_n and A_n are not solvable. □

We now provide two applications of the theorems we have proven about A_n so far. The first application is to subgroups of index n. For any $i \in I_n = \{1, 2, \ldots, n\}$, we begin by setting

$$H_i = \{\sigma \in S_n \mid \sigma(i) = i\}.$$

It is straightforward to check that $H_i \leq S_n$ for all i. Furthermore, if ι is any bijection between $I_n \setminus \{i\}$ and I_{n-1}, then $\sigma \mapsto \iota\sigma\iota^{-1}$ yields an isomorphism $\kappa : H_i \to S_{n-1}$. Thus $|H_i| = (n-1)!$, so that

$$[S_n : H_i] = \frac{n!}{(n-1)!} = n.$$

Since there are odd permutations fixing i, Corollary 3 implies that $[H_i : H_i \cap A_n] = 2$. Therefore

$$[S_n : A_n][A_n : H_i \cap A_n] = [S_n : H_i \cap A_n] = [S_n : H_i][H_i : H_i \cap A_n] = 2n,$$

and we find that

$$[A_n : H_i \cap A_n] = n.$$

So by Lagrange's Theorem we have

$$|H_i \cap A_n| = \frac{|A_n|}{[A_n : H_i \cap A_n]} = \frac{n!/2}{n} = \frac{(n-1)!}{2}.$$

The subgroups H_i are not normal in S_n, because they are all conjugate to one another. Indeed, if $i \neq j$, then conjugation by (ij) maps H_i onto H_j. If $n \geq 4$, we can choose r, s distinct from i, j and instead conjugate by $(ij)(rs) \in A_n$ to achieve the same result. This then implies that $H_i \cap A_n$ is conjugate to $H_j \cap A_n$ in A_n, as well. We have the same conclusion when $n = 3$, too, simply because $H_i \cap A_3$ is trivial for $i \in I_3$.

One can show that κ carries transpositions to transpositions, and hence that $\kappa(H_i \cap A_n) \leq A_{n-1}$. Since both groups have the same size, it must actually be the case that $\kappa(H_i \cap A_n) = A_{n-1}$. That is,

$$H_i \cap A_n \cong A_{n-1}. \tag{6}$$

We will prove that this statement is true for *any* index n subgroup of A_n.

Now suppose $H \leq A_n$ with $[A_n : H] = n$, and assume $n \neq 4$. We begin by reintroducing a familiar construction. Recall that if we let A_n act on the left coset space A_n/H by left translation, we get a homomorphism

$$T : A_n \to \mathrm{Perm}(A_n/H).$$

Because A_n transitively permutes A_n/H, T is not trivial. But A_n is simple, so if T isn't trivial it must be injective. Hence the image has index

$$\frac{n!}{n!/2} = 2$$

in $\mathrm{Perm}(A_n/H)$. Let $\beta : A_n/H \to I_n$ be a bijection with $\beta(H) = 1$. Then $\gamma \mapsto \beta\gamma\beta^{-1}$ defines an isomorphism $U : \mathrm{Perm}(A_n/H) \to S_n$. The image of $\alpha = U \circ T$ has index 2 in S_n, so by Theorem 4 it must be A_n. This means that α is an automorphism of A_n.

This is an extremely interesting construction! The elements of A_n are the even permutations of I_n. By taking a subgroup of this collection of permutations with a particular size (of index n), and letting A_n act on the coset space, the simplicity of A_n yields a realization of A_n as the even permutations on a different set *through an entirely different mechanism.* We obtain two different "copies" of A_n, connected by an isomorphism. But there's only one A_n, up to the names of what's being permuted, so we've managed to cook up an automorphism of A_n. More on that later.

For any $\sigma \in A_n$, $\alpha(\sigma) = U(T(\sigma)) = \beta T(\sigma)\beta^{-1}$. From this it follows that $\alpha(\sigma) \in H_1 \cap A_n$ if and only if $T(\sigma)(H) = H$. But $T(\sigma)(H) = \sigma H$, by definition. We find that $\alpha(\sigma) \in H_1 \cap A_n$ if and only if $\sigma \in H$. That is, $\alpha(H) = H_1 \cap A_n$. Because α is an automorphism of A_n, this proves that $H \cong H_1 \cap A_n$. Referring back to (6), we see that we have succeeded in establishing the following result.

Theorem 8. *If $n \neq 4$, then every subgroup of A_n of index n is isomorphic to A_{n-1}.*

As with any group, A_n has a number of *inner automorphisms*, which are those that are given by conjugation by a fixed even permutation. And, as with any normal subgroup, conjugation by any element of S_n is also an automorphism of A_n (curiously, these automorphisms don't get a name). Does A_n have any other automorphisms? It turns out the answer is "no," unless $n = 6$. Although it's not particularly difficult to prove the "no" part of this result, it would take us too far afield. However, if we assume familiarity with the Sylow theorems, we have the tools in hand to treat the $n = 6$ case.

Let H be a simple group of order 60. Let $P \leq H$ be a 5-Sylow subgroup. By the orbit-stabilizer theorem, the number of conjugates of P is equal to the index of its normalizer, which must divide $[H : P] = 12$. But the number of conjugates of P is also equal to the number of 5-Sylow subgroups of H, which is $\equiv 1 \pmod 5$. Since P isn't normal in H, it must have more than 1 conjugate. The only way these conditions can simultaneously be satisfied is if there are exactly six 5-Sylow subgroups of H.

The 5-Sylow subgroups of H are permuted by conjugation, and mapping each element of H to the permutation it induces gives rise to a homomorphism

$$H \hookrightarrow S_6.$$

It is injective because H is simple and the action is nontrivial (H acts transitively on its 5-Sylow subgroups). We may therefore assume $H \leq S_6$. Since H is simple, Corollary 3 tells us that, in fact, $H \leq A_6$. We find that

$$[A_6 : H] = \frac{6!/2}{60} = 6,$$

so that by Theorem 8, $H \cong A_5$. Although it's not the result we're after, we pause to record what we've now proven.

Theorem 9. *A_5 is the only simple group of order 60, up to isomorphism.*

Because there is no 5-Sylow subgroup of H left inert by conjugation (it would be normal in H, otherwise), when viewed as a subgroup of A_6, $H \neq H_i \cap A_6$ for any i. Consider once again the automorphism α of A_6 arising from the action of A_6 on A_6/H. We have seen that $\alpha(H) = H_1 \cap A_6$. This proves that α is not given by conjugation, because the only images of $H_1 \cap A_6$ under conjugation are the subgroups $H_i \cap A_6$, and H is not one of these. This is what we were trying to prove.

Theorem 10. *There exists an automorphism of A_6 that is not given by conjugation in S_6.*

3 Remarks

Remark 1. One can easily prove that A_n is a normal subgroup of S_n directly from the definition of "even permutation," without the need for any of the machinery of Section 1. Likewise, by definition, if $\sigma, \tau \in S_n$ are both odd, then $\sigma\tau^{-1} \in A_n$, and $\sigma A_n = \tau A_n$. Hence, without any aid from ϵ, we can conclude there are *at most* two cosets of A_n in S_n: the coset of the even permutations and the coset of the odd permutations. But doesn't this mean, automatically, that there are *exactly* two cosets? If so, the index equation of Corollary 2, and hence Corollary 1, follow immediately. Could we have missed something so obvious?

No, we didn't miss anything. Although our elementary argument *appears* to have proven that the even and odd permutations in S_n fall into two cosets of A_n, it was predicated on the assumption that *odd permutations exist.* We actually didn't prove that until we established Lemma 1! Since the index equation $[S_n : A_n] = 2$ implies that $S_n \setminus A_n \neq \varnothing$, it also implies the existence of odd permutations. This means that Lemma 1 is equivalent to the equality $[S_n : A_n] = 2$. So, in some sense, one can view all of Section 1 simply as a proof that $[S_n : A_n] = 2$.

Remark 2. To every finite group one can associate a unique sequence of simple groups, akin to a prime factorization. Let G be a group. A *composition series* for G is a finite sequence of subgroups G_i of G,

$$G_r = \{e\} \lhd G_{r-1} \lhd G_{r-2} \lhd \cdots \lhd G_1 \lhd G_0 = G, \tag{7}$$

so that G_i/G_{i+1} is simple for all i. By the Correspondence Principle, this means that there are no proper normal subgroups of G_i properly containing G_{i+1}. So, if $G_{i+1} < H \lhd G_i$, then $H = G_{i+1}$ or $H = G_i$. A composition series is therefore a *maximal* subnormal series for G: there is no way to make it longer by inserting more subgroups.

This reformulation actually yields a quick proof that every finite group G has a composition series. Start by taking G_1 to be the largest possible proper normal subgroup of G (everything's finite, so this is no problem). Then let G_2 be the largest possible proper normal subgroup of G_1. Continue in this manner until $G_r = \{e\}$ (since the G_i are finite and shrinking, this must happen eventually). Done.

The somewhat amazing fact is that the *composition factors* G_i/G_{i+1} of (7) are invariants of G. No matter how we build a composition series for G (the algorithm of the preceding paragraph is only one option), we will always get the same factor groups. This somewhat vague statement is made precise in the well-known Jordan-Hölder Theorem.

Theorem 11 (Jordan, Hölder)**.** *Let G be a finite group and suppose*

$$\begin{aligned} G_r &= \{e\} \lhd G_{r-1} \lhd G_{r-2} \lhd \cdots \lhd G_1 \lhd G_0 = G, \\ G'_s &= \{e\} \lhd G'_{s-1} \lhd G'_{s-2} \lhd \cdots \lhd G'_1 \lhd G'_0 = G, \end{aligned}$$

are both composition series for G. Then $r = s$ and there is $\sigma \in S_r$ so that[2]

$$G_{\sigma(i)-1}/G_{\sigma(i)} \cong G'_{i-1}/G'_i$$

for all i.

The proof of the Jordan-Hölder Theorem is a somewhat elaborate application of the fundamental Isomorphism Theorems, utilizing Zassenhaus' Butterfly Lemma (mentioned only because it has such a great name!) [1]. What the theorem tells us is that, in a certain sense, the finite simple groups are the "building blocks" of every finite group.[3] Given this significant role, it is natural to ask if it is possible to describe all of the finite simple groups. It is an astonishing fact that the answer is "yes." After decades of work and tens of thousands of pages of published mathematics, the classification of the finite simple groups was finally completed in 2004. No small feat indeed!

The Classification Theorem states that, with 26 exceptions (the *sporadic groups*), the finite simple groups fall into three infinite families. The first of these is the family of (cyclic) groups of prime order. The second is the family of alternating groups!

Remark 3. The group-theoretic notion of solvability is intimately related to the solvability of polynomial equations by radicals. Roughly speaking, a polynomial is solvable by radicals if it is possible to express all of its roots in terms of arithmetic involving only elements in the field of the coefficients and (perhaps nested) nth roots. For example, the quadratic formula shows that every quadratic polynomial can be solved by radicals. And the polynomial $x^8 - 10x^4 + 1$ is solvable by radicals since its roots are

$$\epsilon_1\sqrt{\epsilon_2\sqrt{2} + \epsilon_3\sqrt{3}}, \;\; \epsilon_i \in \{\pm 1\}.$$

Although the expressions for the roots are more complicated than in the quadratic case, every polynomial of degree 3 or 4 is also solvable by radicals. In other words, there is a "cubic formula" and a "quartic formula."

[2] We have shifted i down by 1 to facilitate the the application of σ.

[3] Although this is the party line, there is no general way to reconstruct a group G from its composition factors. So although the composition factors are indeed invariants of G, knowledge of them alone doesn't usually tell you what G is.

The quest to find similar results for polynomials of higher degree led ultimately to Abel's Theorem: there is no general solution by radicals for a polynomial of degree 5 or more. This is somewhat striking, as it asserts the *nonexistence* of a certain type of formula. It turns out that Abel's Theorem is deeply connected to the theory of finite groups!

Galois was able to show that to any polynomial $p(x)$ one can associate a finite group G, a certain subgroup of the permutations of its roots. This is the so-called *Galois group* of $p(x)$. The amazing fact is that $p(x)$ is solvable by radicals if and only if G is solvable. By showing that G is *not* solvable, one can demonstrate that $p(x)$ *cannot* be solved by radicals!

Because the Galois group of the "generic" degree n polynomial is S_n, Galois theory tells us that the general polynomial of degree $n \geq 5$ cannot be solved by radicals. Put another way, the quadratic, cubic and quartic formulae *cannot* be generalized to any higher degree.

References

[1] Lang, S., *Algebra*, Springer, 2008.