Notes on Group Theory

by Avinash Sathaye, Professor of Mathematics
January 26, 2015

Contents

1 Preparation.

Be sure to read through the Preliminaries: pages 1-15 from the book. I will assume that you have understood them and can do the exercises. Of course, you are free to ask questions about the material, privately, or in class, and get it clarified.

2 Group axioms and definitions.

Definition: Group We define a group to be a non empty set G together with a **binary operation** $f : G \times G \Rightarrow G$ such that:

1. f is associative. (*What does it mean?)*
2. *There exists an element $e \in G$ such that*

$$f(e, g) = f(g, e) = g \ \forall g \in G.$$

3. *For each $g \in G$ there is some element $g' \in G$ such that*

$$f(g, g') = f(g', g) = e.$$

We shall take the following shortcuts.

1. **Definition: commutative group** A group G is said to be commutative if $f(g, h) = f(h, g) \ \forall g, h \in G$. Such a group is also called **abelian**.
2. We shall rewrite $f(g, h)$ as $g + h$ if the group is abelian and $g \cdot h$ or just gh in general, dropping all reference to "f". On rare occasions, for comparing different operations, we may use other symbols to denote the operation.
3. **Extended product notation.** Suppose we are using the convention to write the product of elements g, h as simply gh. We shall extend this to subsets $S, T \subset G$ as follows.

   By ST we shall denote the set $\{st | s \in S, t \in T\}$.

   We shall extend this to products of several sets as needed.

   If a set is a singleton, then we may simply write the element in place of the set. Thus $\{s\}T$ can be shortened to sT.

   We may also use the inverses of sets, thus $S^{-1} = \{s^{-1} | s \in S\}$. We could also use powers $S^2 = SS$ etc., but we prefer not to take this shortcut!

4. $f(g, g)$ will become g^2 with no mention of "f". This is naturally extended to g^n. The same expression becomes ng in case the group is abelian and we are using "+" as the operation symbol.

5. The element e is proved to be unique and is called the **identity element** of the group. It may be denoted by 0 for an abelian group and by 1 or e in general.

6. The associated element g' is also shown to be unique and is called the inverse of g. It is denoted by the more suggestive notation g^{-1} which is replaced by $-g$ when the group is abelian and the operation symbol is $+$.

2.1 Cyclic groups.

Definition: Subgroup Suppose that G is a group with binary operation f.

- Suppose that $H \subset G$ is a non empty subset which is closed under the binary operation f, i.e. $f(h_1, h_2) \in H$ whenever $h_1, h_2 \in H$.
- Also suppose that for every $h \in H$ the inverse h^{-1} also belongs to H.

Then it is easy to see that the restriction of the operation f to H makes it into a group and H is said to be a subgroup of G.

We shall write this in notation as $H < G$. **Caution!** Note that this notation may lead you to believe that H has to be smaller than G, so H cannot equal G. That is not the implication! Many authors on group theory avoid this notation, perhaps to avoid this confusion!

Definition: Order of an element If G is a group and $g \in G$ then we define:

$$\text{order of } g \ = \ |g| = \min\{n | g^n = e\}.$$

Remember the convention that the minimum of an empty set is ∞.

We shall use the notation:

$$< g >= \{g^n | n \in \mathbb{Z}\}$$

and call it the **cyclic group** generated by g. Note that we are tacitly claiming that it is a subgroup of G, i.e. $< g >< G$.

Alternatively we could define $< g >$ as the smallest subgroup of G which contains g.

Also note that if $n = -m$ a negative integer, then by g^n we mean $(g^{-1})^m$. It is easy to note that $(g^{-1})^m g^m = e$ and hence $(g^{-1})^m = (g^m)^{-1}$.

It can be easily show that the set $< g >$ is an infinite set if $|g| = \infty$ and has exactly d elements if $|g| = d$.

Moreover, if $|g| = d$ is finite, then this set $< g >$ can be explicitly written as $\{e = g^0, g = g^1, g^2, \cdots, g^{d-1}\}$.

Further, for any $n \in \mathbb{Z}$, we can prove that $g^n = g^r$ where r is the remainder when n is divided by d.

Observation. If the elements of H are of finite order, then the condition about the inverses can be omitted, since the inverse of an element with order d is seen to be its $(d-1)$-th power.

Note that the cyclic group generated by a single element $g \in G$ can be also characterized as the smallest subgroup of G containing g. Equivalently, it can be also defined as

$$< g >= \bigcap \{H \mid g \in H < G\}$$

or the intersection of all subgroups of G containing g.

This observation can be strengthened to the

Definition: Group generated by a set . Given a set $S \subset G$, we define a subgroup of G generated by S as:

$$< S >= \bigcap \{H \mid S \subset H < G\}.$$

For calculation purposes, it can also be characterized as a set of finite products $s_1 \cdots s_r$ with $s_i \in < t_i >$ where $t_i \in S$ for $i = 1, \cdots, r$.

Despite the simple definition, the actual calculation of the resulting group can be quite complicated!

To prove these observations, you only need to know that an arbitrary intersection of subgroups of G is also a subgroup of G. This is easy to prove!

2.1.1 Exercises on groups.

1. For any subset $S \subset G$ **recall:**

$$SS = \{s_1 s_2 \mid s_1, s_2 \in S\}.$$

   Prove that $S < G$ iff $SS = S$ and for every $s \in S$, we have $s^{-1} \in S$.

   As before, if every element of S has a finite order, then $S < G$ iff $SS = S$.

   In particular, if G is a finite group, then $S < G$ iff $SS = S$.

2. Given elements $g, h \in G$ the element ghg^{-1} will turn out to be very useful. We shall call it the conjugate of h by g. It may be denoted by the suggestive notation h^g.

   Prove that $|h| = |ghg^{-1}|$.

Deduce that $|xy| = |yx|$ for any $x, y \in G$.

Note that the equations are supposed to work naturally even for infinite orders.

3. **Conjugate subgroups.**

   Prove that if $H < G$ and $g \in G$, then $gHg^{-1} < G$.

   For proof, simply note that $(gHg^{-1})(gHg^{-1}) = gHHg^{-1} = gHg^{-1}$, the last equation follows since $H < G$.

   Also, it is obvious that gHg^{-1} contains the inverses of all its elements, since

   $$\left(ghg^{-1}\right)^{-1} = gh^{-1}g^{-1} \in gHg^{-1}$$

   since $H < G$ implies $h^{-1} \in H$ when $h \in H$.

   The group gHg^{-1} is said to be a conjugate of H by g and is also written in the suggestive notation H^g.

4. This calculation also leads to **the most important concept** in group theory.

   Definition: Normal subgroup. A subgroup $H < G$ is said to be a normal subgroup if $gHg^{-1} = H$ for all $g \in G$.

   We shall use the customary notation

   $$H \lhd G$$

   to indicate that H is a normal subgroup of G.

   Prove that the subgroups $< e >$ and G are always normal. These are also called the **trivial subgroups** for obvious reasons.

   Definition: Simple group. A group G is said to be simple if $< e >$ and G are the only normal subgroups, i.e. G has only trivial normal subgroups.

   One of the biggest achievements of group theory is considered to be the **classification of all finite simple groups**. [1]

[1] Despite the announcement of this great theorem some years back, some doubts still persist about the complete validity of this theorem. The number of mathematicians in the world who may know the complete proof of the theorem is rather small and people are still working on writing the definitive version.

Many group theorists are said to have left group theory after the announcement of the theorem, considering that there was not much left to prove!

5. Prove that if $|g| = d < \infty$ and n is any integer, then:

$$|g^n| = \frac{d}{\text{GCD}(n,d)}.$$

Make and prove an appropriate statement if $|g| = \infty$.

6. Deduce that if $|g| = d < \infty$ and if $h = g^n \in < g >$ then

$$< h > = < g > \text{ iff } \text{GCD}(n,d) = 1.$$

Make and prove an appropriate statement if $|g| = \infty$.

7. We shall say that **an element $\mathbf{g} \in \mathbf{G}$ generates $\mathbf{G}$** if $G = < g >$, i.e. G is a cyclic group consisting of all powers of g.

Let n be a positive integer and define $\mathbb{Z}_n$ by:

$$\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}.$$

Prove that $\mathbb{Z}_n$ is an abelian group under the usual "+" inherited from $\mathbb{Z}$. Prove that it is a cyclic group.

Prove that $\overline{a} \in \mathbb{Z}_n$ is a generator of $\mathbb{Z}_n$ iff $\text{GCD}(a,n) = 1$.

Deduce that the number of distinct generators of $\mathbb{Z}_n$ are exactly $\phi(n)$ the Euler ϕ function evaluated at n.

8. **Commutativity in groups.**

- Prove that for any elements $g, h \in G$ we have $(gh)^{-1} = h^{-1}g^{-1}$. More generally $(g_1 g_2 \cdots g_r)^{-1} = g_r^{-1} \cdots g_2^{-1} g_1^{-1}$.
- We say that elements $g, h \in G$ commute with each other if $gh = hg$.

  Prove that g, h commute with each other iff $ghg^{-1}h^{-1} = e$. Thus the element $ghg^{-1}h^{-1}$ measures how far the elements don't commute!

  We define the **Commutator** of elements $g, h \in G$ to be the element $ghg^{-1}h^{-1}$. We shall denote it by the symbol $[g,h]$.

  The group generated by all the commutators of elements of G is called the **Commutator subgroup** of G. It is often denoted by the convenient symbol G', or by a more suggestive $[G,G]$.

  Caution: Don't forget that elements of $[G,G]$ are products of Commutator elements and cannot, in general, be written as a single Commutator.

  Thus $gh = hg$ iff $[g,h] = e$.

  Note: We could write $[g,h] = h^g h^{-1} = g(g^{-1})^h$. This can be useful later!

- Prove the formula $[g,h]^{-1} = [h,g]$.
- Prove that if a group has the property that $|g| \leq 2$ for each $g \in G$ then G is abelian.

  Hint: Try to show that $[g,h] = e$ for all $g, h \in G$.
- **Challenging problem!**

  Let us say that a group has property P_m if $(gh)^m = g^m h^m$ for all $g, h \in G$.

  Note that G is abelian iff G has property P_m for each $m = 1, 2, 3, \cdots$.

  Prove that G is abelian iff there is some non negative integer n such that G has property P_n, P_{n+1}, P_{n+2}.

  Hint: Note that P_0 and P_1 are trivially true. Prove that P_2 alone implies that G is abelian.

  Thus, the first non trivial case is $n = 3$.
- **Another challenging problem!**

  Suppose we replace our group axioms by the following apparently weaker ones:

  - G is a non empty set with a binary *associative* operation denoted by "$\cdot$".
  - G has an element e such that $a \cdot e = a \ \forall a \in G$.
  - Each element $a \in G$ has an associated element denoted as $h(a)$ such that $a \cdot h(a) = e$.

  Prove that $(G, \cdot)$ is a group! So, the axioms are not weaker at all, despite their appearance!

  Now here is a shocker. Suppose that we keep the first two axioms as they are but replace the third axiom by a similar but different one thus:

  Each element $a \in G$ has an associated element denoted as $h(a)$ such that $h(a) \cdot a = e$.

  Find an example $(G, \cdot)$ where the new axioms are satisfied, but G is not a group!

2.2 Permutation groups.

A more useful way of thinking about a group is this.

Fix some set Ω. **Define** S_Ω to be the set of all bijective functions from Ω to itself. If the set Ω is finite, then these are easily viewed as permutations of the set. The same can be thought for infinite sets with a suitable imagination.

The set S_Ω with the binary operation of composition of the functions forms a natural group. The identity is the identity function and is often denoted as I or Id. The inverse is simply the inverse function.

Later on, we shall prove a theorem which says that all groups can be viewed as subgroups of S_Ω for some Ω.

For a finite set Ω, it is convenient to develop efficient notation and conventions. Since most of our groups will be finite, this will be very useful.

2.2.1 Conventions for a finite permutation group.

Here are the usual simplifications.

1. If Ω has n elements, we agree to list them as simply the natural numbers $1, 2, \cdots, n$ and simplify the notation S_Ω to S_n.

   The permutation functions are then best described by simply listing the images of various elements of Ω in the form of a table. Example:

   $$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

2. While this is efficient, it has two drawbacks. We keep on writing the same first row of n numbers for every element and we probably should economize on the second row.

   For example, this σ above can be quickly described as "exchange 1 and 3, leave the rest unchanged".

   In symbols, we shall agree to write it as $\begin{pmatrix} 1 & 3 \end{pmatrix}$ which tells us about swapping $1, 3$ and we agree that the unmentioned ones are unchanged!

   More generally, we shall agree that $\begin{pmatrix} 2 & 3 & 4 \end{pmatrix}$ shall be the permutation which send 2 to 3, 3 to 4 and 4 to 2. It then leaves 1 unchanged.

   We formalize this thus:

   Cycle notation A sequence $\begin{pmatrix} a_1 & a_2 & \cdots & a_{r-1} & a_r \end{pmatrix}$ is called a cycle of length r (or simply an r-cycle) and intended to permute the elements $\begin{pmatrix} a_1 & a_2 & \cdots & a_{r-1} & a_r \end{pmatrix}$ in a cycle.

3. **Combining cycles** In permutation groups we often omit the composition symbol and simply write cycles next to each other, with the understanding that each cycle represents a function and these are to be composed from right to left.

   Thus $\begin{pmatrix} 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & 4 \end{pmatrix}$ is the same as the permutation

   $$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}.$$

When the cycles have no common elements, then they commute, but not otherwise.

It is a theorem that every permutation can be written as a product of disjoint cycles and the expression is unique, except for the order of the cycles and the fact that entries of a cycle can be permuted in a cyclic manner without changing its meaning.

An example of permuting the entries in a cycle is:

$$\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 2 \end{pmatrix}$$

4. **Orders of permutations.** It is easy to check that the order of an r-cycle is simply r. If a permutation σ is equal to a product $\sigma_1\sigma_2\cdots\sigma_m$ of m **mutually disjoint cycles**, then it is easy to see that

$$\sigma^i = \sigma_1^i\sigma_2^i\cdots\sigma_m^i$$

and indeed, this expression gives a disjoint cycle representation (after breaking up the powers separately into cycles if needed).

Then we can easily see that

$$\sigma^n = Id \text{ iff } \sigma_i^n = Id \ \forall i = 1, \cdots, m.$$

Thus clearly $\sigma^n = Id$ iff $\text{LCM}(|\sigma_1|, |\sigma_2|, \cdots, |\sigma_m|)$ divides n.

In other words $|\sigma| = \text{LCM}(|\sigma_1|, |\sigma_2|, \cdots, |\sigma_m|)$.

Warning: If the cycles are not disjoint, then the formula is not even expected to work! For example,

$$|\begin{pmatrix} 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 3 \end{pmatrix}| = |\begin{pmatrix} 1 & 3 & 2 \end{pmatrix}| = 3$$

but the LCM of the orders of the separate cycles is only 2.

2.2.2 Exercises on Permutations.

1. **Conjugation.** Given two permutations σ and τ we observe the following fact for the conjugate $\theta = \tau\sigma\tau^{-1}$:

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

Thus, the easiest way to describe θ is to take the display of σ and if it has $i \to \sigma(i)$ act by τ on both the entries to write $\tau(i) \to \tau(\sigma(i))$.

The best way to understand this is to work this example:

Let $\sigma = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} 4 & 5 \end{pmatrix}$ and $\tau = \begin{pmatrix} 1 & 4 \end{pmatrix}\begin{pmatrix} 2 & 5 \end{pmatrix}$.

Then prove that

$$\sigma^\tau = \tau\sigma\tau^{-1} = \begin{pmatrix} \tau(1) & \tau(2) & \tau(3) \end{pmatrix}\begin{pmatrix} \tau(4) & \tau(5) \end{pmatrix} = \begin{pmatrix} 4 & 5 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \end{pmatrix}.$$

Verify this with the observation as well as with direct computation. **This justifies the notation σ^τ, since it is τ acting on σ.**

Similarly, show that

$$\tau^\sigma = \begin{pmatrix} 2 & 5 \end{pmatrix}\begin{pmatrix} 3 & 4 \end{pmatrix}.$$

2. Use the above to give another proof that a permutation has the same order as that of any of its conjugate.

   It is useful define the **type of a permutation** to be a sequence of pairs $[r, s]$ if the permutation contains s disjoint r-cycles in its cycle representation. To make the definition meaningful, we list the cycle lengths in decreasing order. Also, for convenience, we shall drop the "1-cycles" from consideration! [2]

   Thus our σ above has type $([3, 1], [2, 1])$ while the τ above has type $([2, 2])$.

   As a convention, we define the type of identity to be $([0, 0])$. This is a technical convenience!

   Prove that a permutation has the same type as any of its conjugates. Also, prove that the order of a permutation of type $([r_1, s_1], \cdots, [r_m, s_m])$ is $\text{LCM}(r_1, \cdots, r_m)$.

3. Define the function

   $$M(n) = \max\{|\sigma| \mid \sigma \in S_n\}.$$

   Determine the values of $M(n)$ for $n = 1, 2, \cdots, 10$.

   Challenge: Can you make a formula for $M(n)$ in general?

4. Suppose that σ is of type $([r, 1])$, i.e. is an r-cycle.

   Prove that for every $i = 1, 2, \cdots$ we have that σ^i has type $([a, b])$ where $ab = r$ or $\sigma^i = Id$. In words, this says that a power of a cycle breaks up into a certain number of cycles of equal length.

   Hint: Identify the formula for cycles of small length and then guess. Try cycles of lengths $5, 6, 9, 10, 12$.

[2] **Be aware** that some people might choose to keep them in notation, so that the sum of the terms rs is always equal to n.

5. Prove that a permutation is of order 2 iff it has type $([2, s])$ for some s.

6. **Conjugacy classes.** The conjugacy class of an element $g \in G$ is said to be $\{g^h \mid h \in G\}$. The number of elements in a conjugacy class is going to be an important concept later on.

   Prove that the conjugacy class of a permutation $\sigma \in S_n$ is the set of permutations with the same type as σ.

   For $\sigma = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} 4 & 5 \end{pmatrix} \in S_5$, determine its conjugacy class explicitly.

   Note that the class depends on the group, so for the same σ the class is bigger if we work in $\mathbf{S_6}$.

7. Let $n \geq 4$.

   Prove that the conjugacy class of $\tau = \begin{pmatrix} 1 & 2 \end{pmatrix}\begin{pmatrix} 3 & 4 \end{pmatrix} \in S_n$ has $\frac{n(n-1)(n-2)(n-3)}{8}$ elements.

   Hint: You are counting all permutations of type $([2, 2])$.

2.3 Matrix Groups.

The Euclidean or affine spaces $\Re^n$ as well as the general vector spaces F^n over a field F are extremely important in many branches of mathematics as well as applications.

These spaces naturally give rise to a nice collection of groups which described the "change of coordinate" transformations. These can also be thought of as invertible transformations on the underlying set of points with certain extra restrictions, depending on the kind of geometry being studied.

Rather than get into the details of these spaces, we shall concentrate on the underlying groups as follows.

1. **Field** A field is an abelian group under an addition operation with identity 0 together with a commutative multiplicative group structure on its non zero elements, conveniently denoted by $F^\times = \{a \in F | a \neq 0\}$.

   Further the addition and the multiplication operations satisfy the distributive law:

   $$a(b + c) = ab + ac \ \forall a, b, c \in F.$$

2. **General Linear Group** The set of all $n \times n$-matrices with entries in F and a non zero determinant is denoted by $GL_n(F)$.

   Abstractly, the group can be thought off as the set of invertible linear transformations of an n-dimensional vector space over F. The columns of the matrix, then can be thought of as the images of standard basis

vectors, if we identify the vector space with columns of vectors with n-entries.

The identity of $GL_n(F)$ is the identity matrix I_n.

Many people use a simpler notation $GL(n, F)$ with a similar change for the other notations below.

We shall have use for the **Special Linear Group** $SL_n(F)$ which consists of those matrices in $GL_n(F)$ which have determinant 1. It is easy to see that $SL_n(F) < GL_n(F)$.

An $n \times n$ matrix A is said to be **elementary** if its entries are the same as that of I_n, except for one entry, say $A_{ij} = c$ with $i \neq j$. (This means A_{ij} is allowed to be non zero.) We shall denote such a matrix by $E_{ij}(c)$.

We shall define $E_n(F)$ to be the group generated by the various $E_{ij}(c)$. As before, this can be simply thought of as the smallest subgroup of $GL_n(F)$ which contains all the elementary $n \times n$ matrices. [3]

It is evident that all elementary matrices are in $SL_n(F)$ and it is an important theorem that they generate $SL_n(F)$ in the sense that every element of $SL_n(F)$ is a product of elementary matrices! [4]

3. **Finite fields** Of vital interest in group theory and many applications, including Engineering, is the concept of finite fields. The simplest examples of finite fields are $\mathbb{Z}/n\mathbb{Z}$ or what we called $\mathbb{Z}_n$ earlier.

   The field axioms force the value of n to be a prime number and it is called the characteristic of the field.

   Explicitly, we make the **Definition: Characteristic of a field**

   $$char(F) = \min\{m | m \cdot 1 = 0 \in F\}.$$

   Here, by $m \cdot 1$ we mean the sum $1 + \cdots + 1$ with m terms. If no such m exists (as in the case of the usual real field $\Re$), then we define $char(F) = 0$. [5]

[3] As before, we **caution** the reader that a product of elementary matrices cannot usually be written as a single elementary matrix. It is customary to define a product of elementary matrices to be also an elementary matrix, but don't let it confuse you!

[4] At higher level of group theory, we may replace the field F by a commutative ring and investigate corresponding theorems, which lead to very important and interesting concepts!

[5] You may object, saying that this does not conform with our earlier convention of the minimum of an empty set being infinity! However, the problem is only with our wording! If we define it as the GCD of the set of all n such that $n \cdot 1 = 0$ then in the real field, the GCD comes out to be 0.

It is proved that a finite field F contains a unique $\mathbb{Z}_p$ where p is its characteristic. This $\mathbb{Z}_p$ is said to be its prime field. For characteristic zero, the prime field, the field is always infinite and contains the prime field $\mathbb{Q}$.

Moreover, a finite field F is an n-dimensional vector space over its prime field $\mathbb{Z}_p$ and hence has p^n elements. It is customary to write q for the power p^n and the book denotes the finite field as $\mathbb{F}_q$. [6]

2.3.1 Exercises on Matrices.

1. In $GL_n(F)$ we make a:

   Definition: A diagonal matrix is an $n \times n$ matrix M such that $M_{ij} = 0$ if $1 \leq i, j \leq n$ and $i \neq j$. In other words, its non zero entries are all on the main diagonal.

   Note that if $M \in GL_n(F)$, then none of the M_{ii} are zero.

   A diagonal matrix of the form $d \cdot I_n$ is said to be **a scalar matrix.** where we are using the scalar multiplication by $d \in F$.

   Note that a 2×2 matrix $D = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ with $ab = 1$ is in $SL_2(F)$, since its determinant is $ab = 1$.

   Prove that D is a product of elementary 2×2 matrices. **Hint:** Try to multiply it by a sequence of elementary matrices to convert it to I_2. Recall that this amounts to making elementary row transformations.

   If you recall the so-called LU-decomposition from your Linear Algebra, this calculation will put you on track to proving that every member of $SL_2(F)$ is a product of elementary matrices.

   Usually we define $E_n(F)$ to be the set of all products of elementary matrices in $GL_n(F)$. This is seen to be a group and indeed proved to be equal to $SL_n(F)$.

2. **The Heisenberg Group** We can extend the book definition by setting $H_n(F)$ to be the set of $n \times n$ unit upper triangular matrices, i.e. matrices M which satisfy:

   - $M_{ii} = 1$ for each $i = 1, \cdots, n$.
   - $M_{ij} = 0$ if $i > j$.

[6] Other conventional notations are $GF(n, p)$ or $GF(q)$ where GF is short for Galois Field in honor of E. Galois.

Prove that every element of $H_n(F)$ is a product of upper triangular elementary matrices, i.e. matrices $E_{ij}(c)$ with $i < j$ and $c \in F$. **Hint:** Imitate the Gauss Elimination proof.

In particular, you get that $H_n(F)$ is a subgroup of $E_n(F)$.

2.4 Dihedral groups.

One of the simplest and ubiquitous examples of non abelian groups are the so-called dihedral groups. These are given by: **Definition: Dihedral group** A group generated by elements r, s satisfying the relations

$$s^2 = e = r^n,\ r^s = r^{-1}$$

is said to be a dihedral group of order n. It is denoted by the symbol D_{2n}.

Important. We shall follow the convention in the book and assume that $n \geq 3$ for a dihedral group. The group, as defined, even makes sense for $n = 1, 2$ but it comes out to be a less interesting abelian group in each case.

Caution: Be aware that some books use the notation D_n for the same group.

The group can be explicitly listed as

$$D_{2n} \ = \ \{e, r, r^2, \cdots, r^{n-1}, s, sr, sr^2, \cdots, sr^{n-1}\}.$$

In short, its elements can be listed as $s^i r^j$ where $i = 0, 1$ and $j = 0, 1, \cdots, (n-1)$. It is easy to explicitly describe the product of any two elements

$$s^i r^j s^l r^m = s^p r^q$$

as follows:

- If $l = 0$ then $p = i$ and q equals the remainder of $m + j$ modulo n.
- If $l = 1$, then p is the remainder of $i + l$ modulo 2 and q is the remainder of $m - j$ modulo n.

This can be established by induction.

2.4.1 Exercises on dihedral groups.

As stated above, let D_{2n} be the dihedral group of order $2n$ with $n \geq 3$.

1. Prove that $< r > \lhd D_{2n}$, i.e. $< r >$ is a normal subgroup of the dihedral group D_{2n}.
2. Prove that $< s >$ is **not a normal** subgroup of D_{2n}. (You will need to use $n \geq 3$.)

3. Prove that every element of D_{2n} outside $< r >$ is of order 2. **Hint:** Compute $(sr^i)^2$ from the known formula.

4. For any i, prove that $[s, sr^i] = r^{2i}$. Deduce that r^i commutes with s iff $2i \equiv 0 \mod n$.

5. Determine all elements of D_{2n} which commute with every element of D_{2n}, i.e. find all $z \in D_{2n}$ such that $[z, s^i r^j] = e$ for all $i = 0, 1$ and $j = 0, 1, \cdots, (n-1)$.

   First observe that it is enough to check the conditions:

   $$[z, r] = [z, s] = e.$$

   By calculating the commutators explicitly, show that the only possibility for z is $z = e$ or when $z = r^k$ with $n = 2k$.

   You may wish to prove these formulas first:

   $$[sr^p, r] = r^{-2} \ , \ [r^p, r] = e \ , \ [r^p, s] = r^2 p.$$

2.5 Homomorphisms and Isomorphisms.

We have discussed many ways of looking at the groups. So, it becomes essential to decide when a group is essentially the same as another.

For this, we need two main definitions.

Definition: Homomorphism of group. Given groups G, H and a map $\phi : G \to H$ we say that ϕ is a group homomorphism if it satisfies the condition:

$$\phi(g_1 g_2) = \phi(g_1)\phi(g_2) \ \forall g_1, g_2 \in G.$$

These observations are useful and immediate.

- Here, we have used no group operation symbol in G as well as H. It may have to be used, especially if G, H happen to be the same sets! It is stated in the book with separate operation symbols to clarify this point.

- The total image (usually called the range of ϕ) is easily seen to be a subgroup of H.

  Here are the steps for this:

  – For convenience, write e_G and e_H for the identities in G and H respectively. ¿From $e_G e_G = e_G$, deduce that $\phi(e_G)\phi(e_G) = \phi(e_G)$. This shows that $\phi(e_G) = e_H$ (why?)

- By calculating $\phi(gg^{-1}) = \phi(e_G)$ as $\phi(g)\phi(g^{-1}) = \phi(e_G) = e_H$, deduce that $\phi(g)^{-1} = \phi(g^{-1}) \in \phi(G)$.
- Now it is easy to deduce that $\phi(G)$ is actually a group and a subgroup of H.

- We make a **Definition: Kernel of a homomorphism** The set $\{g \in G \mid \phi(g) = e_H\}$ is said to be the kernel of the homomorphism $\phi : G \to H$.

  We shall use the notation $Ker(\phi)$ to denote this set. The most important property of the kernel is that it is a normal subgroup of G.

  The verification of the subgroup condition is easy. Now if $g \in G$ and $k \in Ker(\phi)$, then note that

  $$\phi(gkg^{-1}) = \phi(g)\phi(k)\phi(g^{-1}) = \phi(g)e_H\phi(g)^{-1} = e_H$$

  and thus $k^g \in Ker(\phi)$ whenever $k \in Ker(\phi)$. This proves the normality.

  Later on, we shall see how creating a suitable homomorphism helps us locate a normal subgroups of a group.

Now we make the next **Definition: Isomorphism of a group** Given groups G, H and a map $\phi : G \to H$ we say that ϕ is a group isomorphism if ϕ is an isomorphism which is both injective and surjective.

Typically, the definition of injectivity requires one to prove that every pair of distinct elements has distinct images. However, the homomorphism hypothesis says that a homomorphism ϕ is injective iff $Ker(\phi) = \{e_G\}$.

Proof. Note that

$$\begin{array}{ll} & \phi \text{ is injective} \\ \text{iff} & \phi(g_1) = \phi(g_2) \text{ implies } g_1 = g_2 \\ \text{iff} & \phi(g_1 g_2^{-1}) = e_H \text{ implies } g_1 g_2^{-1} = e_G \\ \text{iff} & g \in Ker(\phi) \text{ implies } g = e_G \\ \text{iff} & Ker(\phi) = \{e_G\} \end{array}$$

If $G = H$ then we say that an isomorphism from G to H is **an automorphism.**

2.5.1 Exercises on Homomorphisms.

1. Consider two groups $\mathbb{Z}_n, \mathbb{Z}_m$ in our usual notation. The notation $\overline{a}$ used so far is not sufficient since for the same integer a, we need to distinguish between the two barred elements in the different groups.

So, we shall make a better notation $[a]_n$ to denote the $\overline{a} \in \mathbb{Z}_n$ and similarly $[a]_m$ for $\overline{a} \in \mathbb{Z}_m$.

With this in mind, answer the following questions.

(a) For any $[x]_n \in \mathbb{Z}_n$, we have $|[x]_n|$ divides n.

(b) If $F : \mathbb{Z}_n \to \mathbb{Z}_m$ is a homomorphism and $F([1]_n) = [a]_m$ then $|[a]_m|$ divides n.

(c) Deduce that if $\mathrm{GCD}(m, n) = 1$ then there is only one possible homomorphism $F : \mathbb{Z}_n \to \mathbb{Z}_m$, namely the zero homomorphism, i.e. $F([x]_n) = [0]_m$ for all $x \in \mathbb{Z}$.

(d) Suppose that $\mathrm{GCD}(m, n) = d > 1$ and write $m = bd$.

Then we can define a homomorphism $F : \mathbb{Z}_n \to \mathbb{Z}_m$ by $F([x]_n) = [bx]_m$.

Hint: The main problem in defining a homomorphism is to verify that it is well defined.

This means to prove that $[x]_n = [y]_n$ implies $[bx]_m = [by]_m$.

The homomorphism condition itself is trivial for all such definitions.

Determine all possible homomorphisms from $\mathbb{Z}_n$ to $\mathbb{Z}_m$ in this situation (i.e. the situation of GCD being bigger than 1).

(e) Prove that if $n \neq m$ then there is no isomorphism from $\mathbb{Z}_n$ to $\mathbb{Z}_m$.

2. Suppose that G is a cyclic group of finite order $n > 1$ generated by $x \in G$. Let t be any integer and consider the map $F_t : G \to G$ defined by $F_t(y) = y^t$ for any $y \in G$.

   Prove that F_t is a homomorphism.

3. Prove that the homomorphism F_t as described above is injective iff $\mathrm{GCD}(t, n) = 1$.

   Also prove that the homomorphism F_t as described above is surjective iff $\mathrm{GCD}(t, n) = 1$.

4. Prove that for any homomorphism $H : G \to G$ there is a unique integer t between 0 and $n - 1$ such that $H = F_t$.

   Moreover, H is an automorphism iff the corresponding t is coprime with n.

   This is often stated by saying that the number of automorphisms of a cyclic group of order n is exactly $\phi(n)$ the Euler ϕ-function evaluated at n.

5. Let G be any group. Prove that $F : G \to G$ defined by $F(g) = g^{-1}$ is a homomorphism iff G is abelian. Moreover, in case it is a homomorphism, it is actually an automorphism.

6. Let G be a group two elements x, y of order 2. Let $n = |xy|$. If G is generated by x, y, then prove that G is isomorphic to D_{2n}.

   Hint: The elements xy, x can be thought of as r, s in the definition of the dihedral group.

7. **Challenging Problem.** Let G be a finite group and σ an automorphism of G. Assume that $\sigma^2 = \sigma \circ \sigma = Id$.

   Also assume that $\sigma(x) = x$ iff $x = e$.

   Prove that G is abelian!

   Hint: Try the following steps.

   (a) Consider the map $x \to x^{-1}\sigma(x)$ from G to G. Using the given conditions prove that this map is injective.

   Since the group G is finite, every element y of the group can be written as $y = x^{-1}\sigma(x)$, i.e. $\sigma(x) = xy$ for some x.

   (b) Using σ deduce that $y\sigma(y) = e$.

   (c) Use known results to finish!

2.6 Group Actions.

If G is a permutation group contained in some S_n, then every elements of the group permutes elements of the set $\{1, 2, \cdots, n\}$. We say that the group G acts on the set $\{1, 2, \cdots, n\}$.

We now generalize this concept by letting an abstract group act on a set. We make a

Definition: Group acting on a set Let G be a group and let A be some set. Suppose that we have a map $T : G \times A \to A$ such that for every fixed $g \in G$ the map $a \to T(g, a)$ is a permutation of the set A.

We also require the following compatibility conditions:

1.
$$T(g_1, T(g_2, a)) = T(g_1 g_2, a) \ \forall g_1, g_2 \in G \text{ and } a \in A.$$

2.
$$T(e, a) = a \ \forall a \in A$$

where $e = e_G$ the identity of G.

Then we say that T is a **group action** and that G acts on A by the action T.

As before, we shorten the notation, letting go of the symbol T in favor of a single dot, so instead of $T(g, a)$ we agree to write $g \cdot a$.

Warning: It is important not to confuse the "$\cdot$" with the group operation.

2.6.1 Properties of group actions.

Here are a few observations about the group actions.

1. Let T be a group action of a group G on a set A. Let the permutation defined by $a \to T(g, a)$ be denoted by σ_g. Then it is easy to see that the map $\sigma : G \to S_A$ given by $g \to \sigma_g$ is a group homomorphism! Indeed, the compatibility conditions are designed to exactly guarantee this.

   Thus, we could have simply defined a group action of G on A as a homomorphism of G in to S_A.

2. It is possible to have a trivial action, defined as $T(g, a) = a$ for all $g \in G$ and $a \in A$. It is not very interesting, of course.

3. It is more interesting to consider actions for which $\sigma_g \neq Id$ unless $g = e$. Such actions are said to be **faithful.** In this case the kernel of the homomorphism $g \to \sigma_g$ reduces to identity $\{e\}$ and the homomorphism is injective.

4. One of the simplest actions is obtained when we let a group act on itself by left multiplication. Thus, we define $T(g, a) = ga$ for all $g, a \in G$.

   Note that this action is always faithful and it is called **the left regular action.** We can obviously define a corresponding **right regular action** by taking $T(g, a) = ag$.

   When the action is faithful, the homomorphism $g \to \sigma_g$ is indeed an isomorphism between G and its image. Thus the group can be thought of as a subgroup of the permutation group S_A up to isomorphism.

   Thus, we have proved the famous **Cayley Theorem** Every group is isomorphic to a subgroup of a permutation group.

5. Given a finite group G with $n > 1$ elements, we can think about the smallest d for which G has an isomorphic copy in some S_d.

   Cayley's Theorem shows that the smallest d is less than or equal to n. There are examples of groups of arbitrary order for which the smallest d is equal to n. (Can you find them?)

Thus, in some sense, the regular representation is the best possible. However, for many groups, we can make isomorphic copies in smaller permutation groups, which helps in their study.

We introduce some new concepts to help us develop a few more interesting useful actions.

2.7 Cosets of a subgroup.

Let H be a subgroup of a group G. We make a **Definition: Coset** For each $g \in G$ the set $gH = \{gh | h \in H\}$ is defined to be the left coset of H and is called the left coset of H defined by g.

It is easy to define a right coset Hg similarly and it is shown to have similar properties. We leave the parallel treatment for the reader to finish.

It is however **important to remember** that in general as sets $gH \neq Hg$ and the condition of when they are equal has important theoretical consequences.

We easily see that $g_1H = g_2H$ iff $g_1^{-1}g_2 \in H$. Moreover

$$z \in g_1H \bigcap g_2H \text{ iff } z = g_1h_1 = g_2h_2$$

for some $h_1, h_2 \in H$. Thus we see that $z \in g_1H \bigcap g_2H$ implies that $g_1^{-1}g_2 \in H$.

Thus, we have proved that two cosets g_1H, g_2H are either disjoint or equal.

For technical convenience Assume that G is a finite group. This guarantees that we have only finitely many left cosets of H. We can handle the case of infinite groups as well, but some adjustments are needed. These will be taken up later.

Choose elements $g_1 = e, g_2, \cdots, g_r, \in G$ such that $A = \{g_iH | i = 1, \cdots, r\}$ is the set of all possible distinct left cosets of H in G.

It is then obvious that G is the disjoint union of all the left cosets of H and moreover any two cosets obviously have the same number of elements, namely $|H|$.

Thus, we see that $|G| = r|H|$ and we have proved the famous

Lagrange Theorem For any finite group G of finite order n, the order $|H|$ of any subgroup H divides n. In particular, the order $|g|$ of any element of G also divides n, since we already know that $|g|$ equals the order of the cyclic subgroup $< g >$.

The number $r = \frac{|G|}{|H|}$ is clearly useful and is called the **index** of H in G. We shall use the notation $[G : H]$

Now we describe the **action of a group on its cosets.** [7]

Define an action T of G on A, the left cosets of H, as follows:

For a given $g \in G$ and $g_iH \in A$ we know that the coset $(gg_i)H$ is a member of A and hence is equal to g_jH for some j. Define $T(g, g_iH) = g_jH$.

For convenience, let us denote the map from G into S_A by the symbol ψ and let us agree to write the permutation $\psi(g)$ by setting $\sigma_g(g_iH) = g_jH$ if $gg_iH = g_jH$. Thus, σ_g is the name of the permutation function corresponding to the element $\psi(g) \in S_A$. [8]

This gives a homomorphism of G into S_r, where r is the index $[G : H]$. Is this faithful?

The kernel is the set of those elements $h \in G$ such that $hgH = gH$ for all $g \in G$. What does this mean?

We see that

$$hgH = gH \textbf{ iff } g^{-1}hgH = H \textbf{ iff } g^{-1}hg \in H.$$

Since the condition $g^{-1}hg \in H$ holds for all g, applying it with $g = e$ we see that $h \in H$.

Indeed the kernel of the homomorphism is a subgroup K of H such that all conjugates of K are again contained in H and it is the largest such subgroup!(This needs a little thought.) It is often called **the core** of H in G.

2.8 Applications of the above theory.

1. **Groups of prime order** Let G be a group of order $p > 1$, a prime number. Then G is a cyclic group isomorphic to $\mathbb{Z}_p$.

   PROOF. Let $x \in G$ such that $x \neq e$. Then $|x| > 1$ and divides the prime p, so $|x| = p$. Thus the $< x > \subset G$ both have the same number of elements and hence $G =< x >$.

   Define a map $\psi : \mathbb{Z}_p \to G$ by $\psi(\overline{a}) = x^a \ \forall a \in \mathbb{Z}$.

   Note that we defined the map by choosing $a \in \mathbb{Z}$ and describing $\psi(\overline{a})$. This has the advantage that it is easy to write, but we need to show that it is well defined. For this, we write:

   We check that ψ is well defined, i.e. if $\overline{a} = \overline{b}$ then $\psi(\overline{a}) = (\overline{b})$ i.e. $x^a = x^b$.

   Now we know that:

   $$\overline{a} = \overline{b} \text{ iff } a = b + mp \text{ for some } m \in \mathbb{Z}.$$

[7] This action will become a very important tool later on. Be sure to study many examples.

[8] In some sense, we don't need a new symbol, we could have simply written, $\psi(g)(g_iH) = gg_iH = g_jH$. If you find it more convenient, you may skip the notation σ_g.

Also we know that:

$$x^a = x^b \text{ iff } a = b + mp \text{ for some } m \in \mathbb{Z}.$$

Hence ψ is well defined. The homomorphism property is easy. The two observations above also show that the map is injective. Hence $|\psi(\mathbb{Z}_p)| = p = |G|$. Therefore the map is also surjective, thus ψ is an isomorphism from $\mathbb{Z}_p$ to G.

This is often paraphrased by saying that for a given prime p, up to isomorphism, there is only one group of order p.

Moreover, note that we never used the primeness while proving the properties of ψ. Hence, the above argument also shows that any cyclic group G of order n is isomorphic to $\mathbb{Z}_n$.

2. **Analyzing the dihedral groups.** Recall that a dihedral group D_{2n} was described as a group with generators r, s such that $|r| = n, |s| = 2$ and $r^s = r^{-1}$. Since $s^2 = e$ we see that $r^s = srs^{-1} = srs = r^{-1}$.

   We shall now study the group action on cosets for the example of a dihedral group.

   - Let $G = D_{2n}$ and $H =< r >$. As usual, assume $n \geq 3$. Note that $|H| = n$ and hence $|G| = 2|H|$.

     Thus we have only two cosets, $eH = H$ and sH. For any $r^i \in H$ we see that $T(r^i, H) = r^iH = H$ and using $s^2 = e$ as well as $srs = r^{-1}$ we see that:

     $$T(r^i, sH) = r^isH = ssr^isH = sr^{-i}H = sH.$$

     Thus each $r^i \in Ker(T)$, i.e. induces the identity permutation on the pair (H, sH).

     The element s in turn, produces the swap of the two cosets, namely $T(s, H) = sH$ and $T(s, sH) = ssH = H$. Further, it now obvious that $T(sr^i, H) = T(s, T(r^i, H)) = sH$ and $T(sr^i, sH) = T(s, T(r^i, sH)) = H$.

   - Now take $G = D_{2n}$ with $n \geq 3$ as above, but take the subgroup to be $K =< s >$. Then we have n cosets $K, rK, \cdots, r^{n-1}K$. For convenience, name them as $K_i = r^{i-1}K$.

     Consider the corresponding action $T(g, K_i) = gr^{i-1}K$.

     First, we look for the kernel. By theory, it is a subgroup of K, so we only need to check $T(s, K_i) = sr^{i-1}K$. But

     $$sr^{i-1}K = sr^{i-1}ssK = r^{1-i}sK = r^jK$$

where j is the remainder of $1-i$ modulo n. Already, for $i=2$ we get $T(s,K_2)=T(s,rK)=r^{-1}K=r^{n-1}K=K_n$. Since $n\geq 3$, we get that this is not the identity permutation and hence the kernel $Ker(\psi)=\{e\}$.

Thus, G is isomorphic to its copy in the permutation group $S_{\{K_1,\cdots,K_n\}}$. We shall explicitly calculate the image by determining the permutations σ_r,σ_s. For convenience and understanding, we shall write these as permutations of $(1,2,\cdots,n)$, i.e. elements of S_n rather that $S_{\{K_1,\cdots,K_n\}}$.

Note that $\sigma_r(K_i)=rr^{i-1}K=r^iK$. Thus as a member of S_n we get $\sigma_r=\begin{pmatrix}1 & 2 & \cdots & n\end{pmatrix}$.

Now, $\sigma_s(K_i)=sr^{i-1}K=r^{1-i}sK=r^{n+1-i}K=K_{n-i}$. Thus $\sigma_s(K_1)=K_{n-1}$ and $\sigma_s(K_{n-1})=K_1$. Thus clearly

$$\sigma_s=\begin{pmatrix}1 & (n-1)\end{pmatrix}\begin{pmatrix}2 & (n-2)\end{pmatrix}\cdots\begin{pmatrix}i & (n-i)\end{pmatrix}\cdots.$$

Note that the middle two cycle may collapse.

For example, for $n=3$, we get $\begin{pmatrix}1 & 3\end{pmatrix}\begin{pmatrix}2 & 2\end{pmatrix}=\begin{pmatrix}1 & 3\end{pmatrix}$.

For $n=4$ on the other hand, we get $\begin{pmatrix}1 & 4\end{pmatrix}\begin{pmatrix}2 & 3\end{pmatrix}$.

The resulting isomorphic copy of G is generated by the two elements σ_r and σ_s.

For concrete values of n, the permutation representation is easier to calculate with.

3. **Existence of a normal group.** Let G be a group of order 15. Assume that G has an element x of order 5 and y an element of order 3.[9]

   Here is a dramatic conclusion.

   We claim that the group $H=<x>$ is necessarily a normal subgroup of G.

   Proof. Consider the cosets of H in G. We must have exactly $3=[G:H]=\frac{|G|}{|H|}$.

   We claim that $y\notin H$, since otherwise the group H of order 5 will contain a subgroup $<y>$ of order 3 which is not a factor of 5.

   Thus $H\neq yH$. By a similar reasoning using $<y^2>=<y>$, we see that the three cosets are $A=\{H,yH,y^2H\}$.

   Thus the coset action map gives

   $$\psi: G\to S_A\cong S_3.$$

[9]Both of these assumptions will be known facts after some more theory. See Cauchy Theorem proved later.

Now $\psi(G)$ the image of G, is a subgroup of a group of order 6 (namely S_A.) Hence ψ cannot be injective. (Otherwise, the order of $\psi(G)$ would be also 15 and cannot divide 6.)

The $Ker(\psi) \neq \{e\}$ being a subgroup of H, must be a group of order 5 (since the order divides $5 = |H|$) and hence $Ker(\psi) = H$.

As we already know, kernel of any homomorphism is a normal subgroup, hence our claim is proved.

4. **We can indeed go on to prove** that the group G of order 15 is actually abelian and in fact, a cyclic group generated by xy.

   We shall then consider a challenge to extend this to groups of order pq where $1 < p < q$ are prime numbers.

   This uses other ideas already developed and goes thus:

   (a) First we show that G is abelian:

   Suppose, if possible, that G is not abelian. Note that then x, y don't commute, since otherwise, it is easy to see that all elements of G commute with each other.

   Now that we know that $H \lhd G$, we know that $yxy^{-1} = x^i \in H$ for some $i = 1, 2, 3, 4, 5$. Define the map $\theta : H \to H$ by $\theta(g) = g^y$.

   Since x, y don't commute by assumption, we see that $i \neq 1$ and that θ is an automorphism of H. Thus i is one of $i = 2, 3, 4$.

   Now $\theta^2(x) = \theta(x^i) = x^{i^2}$ and $\theta^3(x) = x^{i^3}$. But $\theta^3(x) = y^3xy^{-3} = x$. Since $y^3 = e$, we must have $i^3 \equiv 1 \pmod 5$. By checking $i = 2, 3, 4$, we see a contradiction in each case!

   (b) Now we have shown G to be abelian and $|x| = 5, |y| = 3$. The intersection of the two subgroups $< x >$ and $< y >$ must be $\{e\}$ since its order must divide the coprime numbers $3, 5$ (by Langrange Theorem).

   We claim that $|xy| = \text{LCM}(3, 5) = 15$. First note that $(xy)^{15} = x^{15}y^{15} = e$, so $|xy||15$.

   Suppose $|xy| = n$.

   Now $(xy)^n = e$ implies $x^n y^n = e$ and this means

$$x^n = y^{-n} \in < x > \bigcap < y > = < e > .$$

   It follows that $x^n = y^{-n} = e$ and hence $\text{LCM}(3, 5)|n$.

   Thus $|xy| = n = 15$ and the cyclic group $< xy >$ must coincide with the group G having the same order.

(c) **Challenge.** The above argument can be repeated for several possible cases as follows.

Let G be a group of order pq where $1 < p < q$ are primes. As we shall show later, by the Cauchy Theorem, G necessarily has elements of orders p, q called y, x respectively.

Find out conditions on primes $\mathbf{p}, \mathbf{q}$ for which you can prove that G must be a cyclic group of order pq.

Hint: Note that for $p = 2$, you always have the example of the dihedral group, so you cannot allow $p = 2$.

Also, for $p = 3$ and $q = 7$, it is possible to have a non abelian group defined by $y^3 = x^7 = e, x^y = x^2$. Check it out and understand why it escapes our method of proof above.

3 Preliminary Theorems in Groups.

3.1 Normalizers, centralizers and stabilizers.

Suppose that a group G acts on a set A and let σ_g denote the permutation of A defined by $a \to g \cdot a$.

We **define** the orbit of a subset $B \subset A$ as follows:

The orbit of B under the action of G is

$$G \cdot B = \{g \cdot b | g \in G, b \in B\}.$$

We are using our convenient notation $G \cdot B$.

As before, if B is a singleton $\{b\}$, then we agree to shorten the notation $G \cdot \{b\}$ to $G \cdot b$.

We now define two important concepts associated with a subset of A.

We define the stabilizer of a set $B \subset A$ to be

$$G_B = \{g \in G | g \cdot B \subset B\}.$$

It is easy to show that G_B is indeed a subgroup of G and is the largest subgroup which induces an action on the subset B.

When $B = \{b\}$ for some element $b \in A$, we may simplify the notation $G_{\{b\}}$ to G_b.

Given a subset $B \subset A$, we can consider another smaller subgroup of G, namely

$$\{g \in G | g \cdot b = b \quad \forall b \in B\}.$$

This is easily seen to be the intersection $\bigcap\{G_b | b \in B\}$. We can call it **the fixer of the subset $\mathbf{B} \subset \mathbf{A}$**.

We may use more suggestive notations as follows:

$$\textbf{Stabilizer } Stab(B;G) = G_B$$

and

$$\textbf{Fixer } Fix(B;G) = \bigcap\{G_b | b \in B\}.$$

Note that in general, the fixer is a smaller subgroup than the stabilizer.

Given any two elements $a, b \in A$ it is easy to see that the orbits $G \cdot a$ and $G \cdot b$ are either equal or disjoint.

To see this, first note that for any $a \in A$, $G \cdot (g \cdot a) = (Gg) \cdot a = G \cdot a$.

Thus, if $G \cdot a$ and $G \cdot b$ have a common element $c \in A$, then $c = g_1 \cdot a = g_2 \cdot b$ for some $g_1, g_2 \in G$. Hence $G \cdot c = G \cdot a$. Similarly, $G \cdot c = G \cdot b$ and hence our claim is proved.

We now make a series of simple calculations leading to a theorem.

Orbit length The number of elements in the orbit containing a is exactly $[G : G_a]$.

PROOF. Consider the left cosets of G_a. First observe that for any $G_a \cdot a = \{a\}$. Now we see that for any coset gG_a we get $gG_a \cdot a = \{g \cdot a\}$. Also, then for different elements $g_1 \cdot a, g_2 \cdot a$ in the orbit of a the corresponding cosets g_1G_a and g_2G_a must be distinct, since their actions on a give different elements $g_1 \cdot a$ and $g_2 \cdot a$.

Orbit Equation Let us choose a collection of m elements of A, say $\{a_i\}$ such that every orbit contains exactly one a_i. Then we have the obvious equation:

$$|A| = \sum_i |G \cdot a_i| = \sum_i [G : G_{a_i}].$$

This equation is seen by writing A as a disjoint union of orbits and adding up the number of elements in each.

Now we apply the above equation to various special actions and derive useful theorems.

Case of $\mathbb{Z}_p$ acting on a subset of G^p

Let G be a finite group and p a prime factor of $|G|$. Let A be the subset of G^p defined as follows:

$$A = \{(g_1, g_2, \cdots, g_p) | g_i \in G \text{ and } g_1 g_2 \cdots g_p = e\}.$$

Let the group $\mathbb{Z}_p$ act on A by the action

$$\overline{1} \cdot (g_1, g_2, \cdots, g_p) = (g_2, g_3, \cdots, g_p, g_1).$$

In other words, $\overline{1}$ rotates the sequence cyclically to the left. In general, $\overline{a}$ acts by rotating a times to the left.

To check the required properties, it is necessary to note that for for any p-tuple $g = (g_1, g_2, \cdots, g_p) \in A$ the resulting p-tuple is also in A, i.e. $g_2 g_3 \cdots g_p g_1$ is also equal to e.

For any such $g \in A$, we can also see that the stabilizer $(\mathbb{Z}_p)_g$ is either $\{\overline{0}\}$, if at least two of the g_i are different and equal to $\mathbb{Z}_p$ in case all components of g are the same as g_1 and thus $g_1^p = e$.

Thus the orbit of a g with at least two different components has exactly p elements (i.e. $[\mathbb{Z}_p :< \overline{0} >]$) or exactly 1 element when $g = (g_1, g_1, \cdots, g_1)$.

It is easy to count that $|A| = |G|^{p-1}$ since the p-tuples of A can be thought of as random $(p-1)$ elements of G followed by the inverse of their product as the last element.

Thus we have an equation

$|G|^{p-1} =$ number of singleton orbits$+p($ number of distinct orbits of length $p)$.

Since the left hand side of this equation as well as the second part of the right hand side are both divisible p, the number of singleton orbits is also a multiple of p. One singleton orbit is $\{(e, e, \cdots, e)\}$ and so there must exist an element $x \neq e$ in G such that $(x, x, \cdots, x) \in A$, i.e. $x \neq e$ but $x^p = e$.

Thus we have proved the

Cauchy Theorem. If a prime p divides $|G|$, then G has an element x of order p. This is actually the elegant **proof by McKay** of the original Cauchy Theorem.

The original proof is an intricate induction proof with a special argument for the abelian case. It is a good exercise in induction!

Food for thought. The proof shows that indeed there are at least $p-1$ elements of order p. Do you see how to get the others?

Where did we use the fact that p was a prime? Could the proof still work if p is an arbitrary integer? Should the theorem be even expected to be true?

The Class Equation An important group action happens when a group G acts on itself by conjugation. Thus here for elements $g, h \in G$ we let $g \cdot h$ be $h^g = ghg^{-1}$.

It is well worth redefining and renaming the various subgroups mentioned above.

For $h \in G$ the stabilizer is defined as

$$C_G(h) = \{g \in G | ghg^{-1} = h\}.$$

Its elements commute with h and the group is called **the centralizer** of h.

The fixer of any subset $S \subset G$ is also called its centralizer

$$C_G(S) = \{g \in G | gsg^{-1} = s \quad \forall s \in S\}.$$

On the other hand the stabilizer of the subset S is denoted by

$$N_G(S) = \{g \in G | gSg^{-1} \subset S\}.$$

This is called **the normalizer** of S.

There is a reason for the term normalizer. If S is a subgroup of G then the normalizer $N_G(S)$ is a group between S and G with the property that S is a normal subgroup of $N_G(S)$. Moreover, it can be shown to be the largest such group.

Under the conjugation action, the orbit of an element h is the set of all possible conjugates ghg^{-1} of h and the number of such conjugates is the index of its centralizer $[G : C_G(h)]$. A singleton orbit clearly consists of an element which commutes with every element of the group. Such elements are said to be central. In turn, it is possible to prove that all central elements form a subgroup of G called its **center** and this group is denoted by the symbol

$$Z(G) = \{g \in G | ghg^{-1} = e \text{ i.e. } gh = hg \quad \forall h \in G\}.$$

The orbit equation now takes the following shape:

$$|G| = \text{ number of singleton orbits} + \sum [G : C_G(h)]$$

where the sum is taken by choosing one representative from each conjugation orbit with at least two elements. In view of the definition of the center, we get the famous **Class Equation:**

$$|G| = |Z(G)| + \sum [G : C_G(h)]$$

the sum being over distinct non singleton orbits as before.

A property of a p-group. Let p be a prime. We define a group P to be a **p-group** if $|P|$ is a power of p. [10]

Theorem on p-groups. If p is a prime and P is a p-group, then the center $Z(P)$ is non trivial, i.e. has at least one non identity element.

Let P act on itself by conjugation and notice that all the indices appearing as lengths of non singleton orbits are powers of p, so at least divisible by p.

Since the left hand side $|P|$ of the class equation is also divisible by p, we get that p divides $|Z(P)|$. Since $e \in Z(P)$ we have $|Z(P)| \neq 0$, hence it contains non identity elements!

3.2 Quotient groups.

Before we proceed further it is useful to understand how the homomorphisms and especially their kernels arise.

We have already shown that the kernel of any homomorphism is a normal subgroup of the domain. We shall now show the converse that a normal subgroup of a group is always the kernel of a certain natural homomorphism.

Existence of a quotient group Let G be any group and H a subgroup. Let A be the collection of all left cosets of H in G.

Assume that H is a *normal* subgroup of G. Then A is naturally a group such that

1. For any g_1H, g_2H we define $g_1Hg_2H = g_1g_2H$.
2. There is a natural homomorphism $\Psi : G \to A$ defined by $\psi(g) = gH$.
3. The homomorphism Ψ is surjective and its kernel is exactly H.

Proof. Here are the steps of the proof.

1. Note that the normality of H implies $gH = Hg$ as sets for any $g \in G$. Thus the definition $g_1Hg_2H = g_1g_2H$ is simply a calculation

$$g_1Hg_2H \;=\; g_1(Hg_2)H = = \; g_1(g_2H)H \;=\; g_1g_2(HH) \;=\; g_1g_2H$$

where the last equation follows from the fact that H is a subgroup and hence $HH = H$.

[10] We can alternatively define a p-group as a group where the order of every element is a power of the prime p. Then we need to use the Cauchy Theorem to show that this definition is equivalent to the above for a finite group. This definition has the advantage that it makes sense even for infinite groups!

2. The fact that A becomes a group under this operation is left for the reader to verify. Let us record that the identity is simply the coset $e_G H = H$ and we shall call it e_A.

   Moreover, $(gH)(g^{-1}H) = (gg^{-})H = H$ implies that $(gH)^{-1} = g^{-1}H$.

3. The fact that Ψ is a homomorphism is evident from its definition.

4. The map Ψ is surjective since any $gH \in A$ is equal to $\Psi(g)$ by definition. Also $\Psi(g) = e_A$ iff $gH = H$ iff $g \in H$. hence $Ker(\Psi) = H$.

Notation. We shall conveniently write G/H for the set of cosets A and call it the quotient group of G by H.

Warning: It is important to remember that the quotient group G/H is defined only when H is a normal subgroup, but the cosets as a set are defined for any subgroup. We don't get the group structure if H is not normal.

3.2.1 Generation of groups.

Let G be a group and H, K its subgroups. We consider the set $HK = \{hk | h \in H, k \in K\}$. We are interested in deciding if it a subgroup and if indeed it is equal to G itself.

Here is a sequence of useful results.

1. HK is a subgroup of G iff $HK = KH$.

   PROOF. We use our convenient subgroup criterion, namely $L \subset G$ is a subgroup iff $LL = L$ and $L^{-1} = L$.

   Suppose $HK = KH$. Then $HKHK = H(KH)K = HHKK$ and hence reduces to HK since the fact that H, K are subgroups says that $HH = H$ and $KK = K$.

   Also, $(HK)^{-1} = K^{-1}H^{-1}$ reduces to KH since H, K are subgroups and hence to HK by hypothesis.

   Thus $HK < G$.

   Now assume that $HK < G$. Then $(HK)^{-1} = K^{-1}H^{-1} = KH$ since H and K are subgroups. But $(HK)^{-1} = HK$ since HK is a subgroup, thus we have proved $HK = KH$.

2. If H, K are subgroups of G and if $H < N_G(K)$ then $HK < G$. Similarly, if $K < N_G(H)$ then we also get $HK < G$.

   PROOF. If $H < N_G(K)$ then $hK = Kh$ for each $h \in H$ and hence $HK = KH$. Therefore we are done by the above result.

   Similar proof holds if $K < N_G(H)$.

In view of this result, we make a

Definition: A group H normalizes a group K (in **G**) if $H < N_G(K)$.

Similarly, we make a

Definition: A group H centralizes a group K (in **G**) if $H < C_G(K)$.

3. **Corollary.** If H, K are subgroups of G and either H or K is normal then HK is a subgroup of G.

   PROOF. The hypothesis means either $N_G(H)$ or $N_G(K)$ equals G and hence we are done by the above result.

4. **Counting HK.** If H, K are finite subgroups of G then we can count the elements of the set HK regardless of whether it is a subgroup and this is sometimes useful.

   We can regard HK as a union of left cosets of K by elements of H.

   Let L be the subgroup $H \bigcap K$ (This is easily seen to be a subgroup).

   Note that L is a subgroup of both H, K. Using this, write

$$K = \bigcup_{i=1}^{m} k_i L \text{ where } m = |K|/|L| = [K : L].$$

   Similarly, write

$$H = \bigcup_{j=1}^{n} h_j L \text{ where } n = |H|/|L| = [H : L].$$

   Then we see that [11]

$$HK = \bigcup_{j=1}^{n} \bigcup_{i=1}^{m} (h_j k_i L)$$

   and hence

$$|HK| = nm|L| = \frac{|H||K||L|}{|L||L|} = \frac{|H||K|}{|L|}.$$

   Thus we have proved:

$$|HK| = \frac{|H||K|}{|H \bigcap K|}.$$

[11] First write $HK = \bigcup h_j K$ noticing that these are exactly all the distinct cosets of this form. Then write $K = \bigcup k_i L$ in the usual coset representation.

3.2.2 Isomorphism Theorems.

Preamble.

If G is a group and $\phi : G \to L$ is a homomorphism, then we know that $\phi(G) < L$ and $Ker(\phi) < G$. We also know that $Ker(\phi)$ is in fact a normal subgroup, say $H \triangleleft G$. Thus the quotient group G/H is defined as the group formed by the left cosets of H in G.

The first important theorem states that

First Isomorphism Theorem. With the preamble in force, we have $G/H \cong \phi(G)$.

PROOF. Define a map $\theta : G/H \to \phi(G)$ by $\theta(gH) = \phi(g)$.

We note that if $g_1H = g_2H$ then $g_1 = g_2h$ for some $h \in H = Ker(\phi)$ and hence $\phi(g_1) = \phi(g_2h) = \phi(g_2)\phi(h) = \phi(g_2)$ since $\phi(h) = e_L$.

This shows that θ is well defined. The fact that it is a homomorphism is now easily checked.

Also, $\theta(gH) = e_L$ iff $g \in Ker(\phi) = H$ iff $gH = H$. Thus $Ker(\theta) = \{H\}$.

Thus θ is injective. It is clearly surjective onto $\phi(G)$ proving the result.

Next we have the

Second Isomorphism Theorem. Let A, B be subgroups of a group G such that $A < N_G(B)$ and thus AB is a subgroup of G.

Then we have:

1. B is a normal subgroup of AB
2. $A \bigcap B$ is a normal subgroup of A and
3. We have:
$$AB/B \cong A/(A \bigcap B).$$

PROOF. The normality of B follows since $A < N_G(B)$. The same fact says that for $a \in A$ and $b \in B$ we have $b^a \in B$. Moreover, if b is also in A, then we also have $b^a \in A$. Thus $b^a \in A \bigcap B$ for all $b \in A \bigcap B$; this proves the second claim.

Finally, let $F : A \to AB/B$ be the map defined by $F(a) = aB$. We show that F is a surjective homomorphism with kernel $A \bigcap B$, thereby proving the result from the first isomorphism theorem.

First $F(a_1a_2) = a_1a_2B = a_1Ba_2B$ since B is normal in AB. Thus F is a homomorphism.

For any $abB \in AB/B$ we see that $abB = aB$ since $b \in B$ and thus F is surjective.

Now for $a \in A$, we see that $F(a) = B$ iff $aB = B$ i.e. $a \in B$. Thus the kernel is exactly $A \bigcap B$.

Hence, we are done by the first isomorphism theorem.

We now derive a very useful **Homomorphism Principle.**

Assume that we have a group G with a normal subgroup H and some group K.

Then any homomorphism F from G/H to K induces a homomorphism F^* from G to K such that $F^*(H) = \{e_K\}$. This is defined as $F^*(g) = F(gH)$.

Conversely, given any homomorphism F^* from G to K satisfying $F^*(H) = \{e_K\}$ we can find a homomorphism F from G/H to K defined by $F(gH) = F^*(g)$.

We describe this by saying that F is a factor of F^*.

Typically, homomorphisms from a quotient group G/H are created as factors of convenient homomorphisms of G.

PROOF. The first part is obvious by the definition of the quotient group. The second part only needs the verification that the map is well defined, i.e. to check that $g_1H = g_2H$ implies $F(g_1H) = F(g_2H)$, i.e. $F^*(g_1) = F^*(g_2)$, but this follows since $g_1 = g_2h$ for some $h \in H$ and hence $F^*(g_1) = F^*(g_2)F^*(h) = F^*(g_2)$ by hypothesis.

Third Isomorphism Theorem. Suppose we have a tower of groups $H < K < G$. Further assume that H, K are normal subgroups of G. Then it is obvious that K/H is a normal subgroup of G/H.

Further, we have

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

PROOF. Define a homomorphism $\psi : G/H \to G/K$ by $\psi(gH) = gK$. We have the natural surjective homomorphism ψ^* from G to G/K defined by $\psi^*(g) = gK$. Clearly $\psi^*(k) = K$ if $k \in K$ and hence by the homomorphism principle, we get that ψ factors this ψ^*.

Now the kernel of ψ is the set of cosets of the form kH where $k \in K$, i,e, K/H. Thus, K/H is a normal subgroup of G/H and by the first isomorphism theorem we get the result.

3.3 Sylow Theorems.

Now we come to some of the most powerful theorems in group theory. We begin with a far reaching generalization of Cauchy Theorem.

First some terminology. Given a finite group G and a prime p let p^r be the highest power of p dividing $|G|$. Any subgroup of G with order p^r is said to be a **p-Sylow subgroup** of G. The set of all such Sylow subgroups shall be denoted by $Syl_p(G)$. The number of distinct Sylow subgroups shall be denoted by the $n_p(G)$. We may drop the G from these notations, if clear from the context.

We begin by proving that $Syl_p(G)$ is a non empty set.

First Sylow Theorem. Given a prime p, amy finite group G has a p-Sylow subgroup.

Note that by definition, if p does not divide d, then the trivial subgroup $\{e_G\}$ satisfies the definition of a p-Sylow subgroup and the theorem is clearly true.

PROOF. First we prove a

LEMMA. Let B be a normal subgroup of a group A and C a subgroup A/B. Let C^* be the subgroup of A defined as the union of the cosets $xB \in C$. Then $|C^*| = |C||B|$.

PROOF of Lemma. Note that C is isomorphic to C^*/B and hence the result follows.

Now we prove the theorem by induction on $d = |G|$. The result being trivial for $d = 1$ we assume the result for all groups with order less than d and prove it for the given G.

Suppose that p^r is the highest power of p which divides d. As already observed, it is enough to consider the case when p divides d and thus $r \geq 1$.

If the center $Z(G)$ of the group contains an element x of order p, then note that $< x > \triangleleft G$ and p^{r-1} is the highest power of p which divides $|G|/| < x > | = |G|/|x|$.

By induction hypothesis $G/ < x >$ contains a p-Sylow subgroup C of order p^{r-1} and the corresponding subgroup consisting of the union of cosets $g < x >$ such that $g < x > \in C$ forms a subgroup of G of order $p^{r-1}p = p^r$, by the LEMMA. It gives the necessary p-Sylow subgroup of G.

Now assume that p does not divide $|Z(G)|$. Let G act on itself by conjugation and write the class equation:

$$|G| = |Z(G)| + \sum [G : C_G(h)]$$

the sum being over distinct non singleton orbits.

Since p divides the left hand side of the equation and does not divide the first term of the right hand side, it must not divide at least one of the terms $[G : C_G(h)]$ on the right hand side. But then p^r must divide $|C_G(h)| = \dfrac{|G|}{[G : C_G(h)]}$.

Clearly $C_G(h)$ is necessarily a group of smaller order that G, and hence it contains a p-Sylow subgroup of order p^r which is automatically the needed subgroup of G.

Having proved the existence of p-Sylow subgroups, we learn to count their number.

Sylow Theorem Part 2. Let p be a prime dividing the order of a group G and let $P \in Syl_p(G)$. Then the number of p-Sylow subgroups of G **conjugate to** P is congruent to 1 modulo p.

PROOF. Let S be the set of p-Sylow subgroups of G conjugate to P. Let P act on this set by conjugation. Since P is a p-group, it follows that every orbit has length equal to a power p^s of p for some integer $s \geq 0$.

All orbits of length bigger than 1 will then have orbit length divisible by p.

We know one orbit with length $p^0 = 1$, namely the orbit of P itself. If we show that there is no other singleton orbit then our orbit equation shows that

$$|S| = 1 + (\text{ sum of orbits lengths which are bigger than } 1\ ) = 1 + mp$$

where m is some non negative integer.

Now suppose if possible, that Q is another p-Sylow subgroup which also leads to a singleton orbit, i.e. for all $x \in P$ we get $xQx^{-1} = Q$. This means P normalizes Q. It follows that PQ is a subgroup of G. Moreover,

$$|PQ| = \frac{|P||Q|}{|P \bigcap Q|}$$

would say that $|PQ| \geq |P|$. On the other hand, since every term in the fraction is a power of p, it follows that $|PQ|$ is also a power of p and by the "Sylow-ness" of P we get $|PQ| \leq |P|$. It follows that $PQ = P$ and then $Q < P$ and consequently $Q = P$ since $|Q| = |P|$. [12]

Now we show:

Sylow Theorem Part 3. If p is a prime dividing $|G|$, then $Syl_p(G) = S$. Thus, any two p-Sylow subgroups are conjugate in G.

In particular $n_p(G) = |S| = [G : N_G(P)]$ where P is some p-Sylow subgroup.

PROOF. We already know that $S \subset Syl_p(G)$.

Suppose if possible, that R is another p-Sylow subgroup outside S. Let R act on S. By the argument used in the above proof, we see that this action would have no singleton orbits!

But then we get that $|S| = 1 + mp$ is a sum of orbit lengths of the action by R, i.e. is a sum of numbers divisible by p. This is clearly a contradiction!

The last part is obvious.

Sylow Theorem Complete. Let G be a finite group and p a prime dividing $|G|$. Then we have the following:

1. $Syl_p(G)$ is non empty.
2. The number $n_p(G)$ of p-Sylow subgroups $= |Syl_p(G)| \equiv 1 \pmod{p}$.

   Moreover, $n_p(G) = [G : N_G(P)]$ for any $P \in Syl_p(G)$ and thus $n_p(G)$ is a factor of $[G : P]$ and is equal to 1 modulo p.

[12] Actually what we have established is that if $Q < G$ is any p-group which is normalized by P, then $Q < P$.

3.4 Applications of the Sylow Theorems.

1. Let G be a finite group and let p be a prime dividing $|G|$. Prove that a $P \in Syl_p(G)$ is normal in G iff $n_p(G) = 1$.

2. If G is a group of order pq where $p < q$ are prime numbers then G is cyclic unless $q \equiv 1 \pmod{p}$.

   Hint: Let P and Q be the Sylow subgroups of orders p, q respectively. Prove that Q is a normal subgroup and P is also normal unless $q \equiv 1 \pmod{p}$.

   If both P, Q are normal, then prove that G is abelian and indeed cyclic.

   Definitive form. Let G be a group of order pq where $p < q$ are primes. Assume that $q \equiv 1 \pmod{p}$. Then either G is cyclic or G is a group generated by x, y with $|x| = p, |y| = q$ and $xyx^{-1} = y^m$ for some m such that $m^p \equiv 1 \pmod{q}$.

   It turns out that indeed there is an integer m such that $q = 1 + mp$ and this is one good choice of m. More on this will be discussed later. The choice of m is not unique, but there is essentially only one choice up to automorphism.

3. Let G be a group of order pq where $p < q$ are prime numbers and assume that $q \equiv 1 \pmod{p}$. Then either G is cyclic or a group which which generated with x, y of orders $|x| = p, |y| = q$ and $xyx^{-1} = y^m$ for some m such that $m^p \equiv 1 \pmod{q}$.

4. Let G be a group of finite order d. Prove that G is not simple in each of the following cases. (i) $d = 12$ (ii) $d = 24$ (iii) $d = 72$ (iv) $d = 312$ (v) $d = 351$

5. Prove that a group of order 56 has a normal p-Sylow subgroup for some $p > 1$.

6. Let $n \geq 2$ be an integer and let G be a group of order $2^n p$, where $p = 2^n - 1$ is a prime number. For example:

   $$12 = 2^2(2^2 - 1)\ ,\ 56 = 2^3(2^3 - 1)\ ,\ 992 = 2^5(2^5 - 1)$$

   Show that G is not a simple group.

7. Suppose that a finite group G has $n_p(G) = 5$ for some prime p. Prove that G has elements of order 5 and 2.

   More generally, assume that a finite group G has $n_p(G) = 2^n + 1$ for some $n \geq 1$. Prove that G has elements of order 2 as well as order q for every prime factor q of $2^n + 1$.

8. Prove that there is no simple group of order 96.

9. Let G be a group of order 2006. Prove that it is abelian or has a cyclic subgroup of index 2.

3.5 Simplicity of $\mathbf{A_n}$.

Now we discuss a large family of known simple groups. These are the so-called alternating groups A_n which are defined as follows.

Let $n \geq 2$ be an integer and let S_n be the usual permutation group acting on the set $\{1, 2, \cdots, n\}$. We have already discussed the cycle decomposition of a permutation into disjoint cycles. Recall that an r-cycle is a cycle which permutes r objects in a cyclic manner. We may call a 2-cycle a transposition.

¿From a sample calculation like

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 5 \end{pmatrix} \begin{pmatrix} 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix}$$

we see that any r-cycle can be written as a product of $(r-1)$-transpositions.

¿From a calculation like

$$\begin{pmatrix} 1 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix}$$

we can see that the above 5-cycle can also be rewritten as a product of 6 transpositions instead of the original 4.

What is surprizing is the

Theorem of the parity of a permutation. Any permutation $\sigma \in S_n$ can be written as a product of a finite number of transpositions. While the expression is not unique, the parity of the number of transpositions (oddness or evenness) depends only on the permutation.

Thus we make a

Definition: Sign of a permutation.

A permutation is said to be **odd** if it can be written as a product of an odd number of permutations and its sign is said to be -1.

A permutation is said to be **even** if it can be written as a product of an even number of permutations and its sign is said to be 1.

PROOF of the Theorem of the parity. Given a permutation $\sigma \in S_n$ and a pair of integers $i \neq j$ in $\{1, 2, \cdots, n\}$ consider a rational number

$$\frac{i-j}{(\sigma(i) - \sigma(j))}.$$

Note that interchanging the order of i, j does not change the value of this rational number. Thus, we can define a well defined function

$$\psi(\sigma, \{i, j\}) = \frac{i-j}{(\sigma(i) - \sigma(j))}.$$

Let T_n be the set of all two element subsets of $\{1, 2, \cdots, n\}$ and note that ψ is a function from $S_n \times T_n$ into $\mathbb{Q}$.

Now define

$$\theta(\sigma) = \prod_{t \in T_n} \psi(\sigma, t).$$

We claim that θ is a function from S_n into the multiplicative group $\{-1, 1\} \subset \mathbb{Q}$ and in fact it **is a group homomorphism.**

It is easy to check that $\theta(\sigma) = -1$ if σ is a transposition thus $\theta(\sigma)$ is 1 if σ is an even permutation and $\theta(\sigma)$ is -1 if σ is an odd permutation; thus proving the theorem.

Note that when we write out the product in $\theta(\sigma)$ the absolute values of the terms in the numerator are $|(i - j)|$ for $\{i, j\} \in T_n$ and the same is true for the denominator. Thus, the ratio is clearly ± 1.

Also note that for convenience the members of T_n can be chosen in any convennient order. Thus we could let $i = 1, \cdots, n$ and then let $j = i+1, \cdots n$. So

$$\theta(\sigma) = \prod_{i=1}^{n} \prod_{j=i+1}^{n} \left(\psi(\sigma, \{i, j\})\right).$$

Now, given any permutation $\tau \in S_n$ we see that another way of listing members of T_n might be to consider $\{\tau(i), \tau(j)\}$ as i varies from $1, \cdots, n$ and for a fixed i, the index j varies from $i + 1$ to n.

Thus we also have

$$\theta(\sigma) = \prod_{i=1}^{n} \prod_{j=i+1}^{n} \left(\psi(\sigma, \{\tau(i), \tau(j)\})\right).$$

Then we can write

$$\theta(\sigma)\theta(\tau) = \left(\prod_{i=1}^{n} \prod_{j=i+1}^{n} \left(\psi(\sigma, \{\tau(i), \tau(j)\})\right)\right) \left(\prod_{i=1}^{n} \prod_{j=i+1}^{n} \left(\psi(\tau, \{i, j\})\right)\right).$$

A simple cancellation between the numerator of the first product with the denominator of the second product shows that

$$\theta(\sigma)\theta(\tau) = \prod_{i=1}^{n} \prod_{j=i+1}^{n} \psi(\sigma \circ \tau, \{i, \jmath\}) = \theta(\sigma \circ \tau).$$

Thus, we have proved that θ is a group homomorphism from S_n into the multiplicative group $\{1, -1\}$.

The kernel of the homomorphism is then a well defined normal subgroup of S_n and we define:

Definition: Alternating group A_n is the set of permutations σ in S_n such that $\theta(\sigma) = 1$. Equivalently it is an even permutation. If its type is $([r_1, s_1], \cdots, [r_m, s_m])$, then it is easy to check that

$$\theta(\sigma) = \prod_{i=1}^{m} \left((-1)^{(r_i-1)s_i}\right).$$

We have seen above that every permutation in S_n can be written as a product of transpositions $\begin{pmatrix} i & j \end{pmatrix}$ where $1 \leq i < j \leq n$. We now refine this to:

Theorem on generation of $\mathbf{S_n}$ and $\mathbf{A_n}$. We have the following generations.

1. Let $n \geq 2$. Then every permutation in S_n can be written as product of permutations of the form $\begin{pmatrix} 1 & i \end{pmatrix}$ where $1 < i \leq n$.

   PROOF.
   $$\begin{pmatrix} i & j \end{pmatrix} = \begin{pmatrix} 1 & i \end{pmatrix}\begin{pmatrix} 1 & j \end{pmatrix}\begin{pmatrix} 1 & i \end{pmatrix}.$$

2. Let $n \geq 2$. Then every permutation in S_n can be written as product of permutations of the form $\begin{pmatrix} (i-1) & i \end{pmatrix}$ where $1 < i \leq n$.

   Idea of the proof. Induct from these calculations.

   $$\begin{pmatrix} (i-2) & i \end{pmatrix} = \begin{pmatrix} (i-2) & (i-1) \end{pmatrix}\begin{pmatrix} (i-1) & i \end{pmatrix}\begin{pmatrix} (i-2) & (i-1) \end{pmatrix}.$$

   $$\begin{pmatrix} (i-3) & i \end{pmatrix} = \begin{pmatrix} (i-3) & (i-2) \end{pmatrix}\begin{pmatrix} (i-2) & i \end{pmatrix}\begin{pmatrix} (i-3) & (i-2) \end{pmatrix}.$$

3. Let $n \geq 2$. Then every permutation in S_n can be generated with $\begin{pmatrix} 1 & 2 \end{pmatrix}$ and $\begin{pmatrix} 1 & 2 & \cdots & n \end{pmatrix}$.

   Idea of the proof. Conjugate the first one repeatedly by the second. Then use earlier result!

4. Let $n \geq 3$. Then every element of A_n is a product of 3-cycles, not necessarily disjoint, of course!

   PROOF. It is enough to show that a product of two transpositions is a product of 3-cycles.

   Here are the calculations:

   $$\begin{pmatrix} a & b \end{pmatrix}\begin{pmatrix} c & d \end{pmatrix} = \begin{pmatrix} a & c & b \end{pmatrix}\begin{pmatrix} a & c & d \end{pmatrix}.$$

   $$\begin{pmatrix} a & b \end{pmatrix}\begin{pmatrix} a & c \end{pmatrix} = \begin{pmatrix} a & c & b \end{pmatrix}.$$

   Note. This gives another proof that A_n is a normal subgroup of S_n, since under conjugation, a 3-cycle always goes to a 3-cycle!

5. Let $n \geq 5$. Any two 3-cycles are conjugate to each other by elements of A_n.

   PROOF. Let two 3-cycles be $u = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}$ and $v = \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix}$. Make two permutations of $\{1, 2, \cdots, n\}$ as

   $$x = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & \cdots \end{pmatrix} \text{ and } y = \begin{pmatrix} y_1 & y_2 & y_3 & y_4 & y_5 & \cdots \end{pmatrix}.$$

   If the permutation σ defined by $\sigma(x_i) = y_i$ is even then it is the even permutation which conjugates u into v and our proof is done.

   If σ is odd, set $z = \begin{pmatrix} y_1 & y_2 & y_3 & y_5 & y_4 & \cdots \end{pmatrix}$ and note that the permutation τ defined by $\tau(x_i) = z_i$ is then an even permutation and it also conjugates u into v.

   Thus, u, v are conjugate by elements of A_n.

   Think! See what can go wrong if $n = 2, 3, 4$.

6. Let $n \geq 5$. Let H be any *normal* subgroup of A_n. If H contains a 3-cycle, then $H = A_n$.

   PROOF. If it contains one 3-cycle, then it contains all 3-cycles, since these are conjugate in A_n and H is normal in A_n. Then H contains all the necessary generators of A_n and we are done!

We are now set to prove the main

Theorem. $\mathbf{A_n}$ is simple if $\mathbf{n \geq 5}$.

PROOF. Let $n \geq 5$ and let $H < A_n$ be a normal subgroup. By a sequence of calculations, we shall show that if $H \neq \{Id\}$, then H contains all 3-cycles and hence $H = A_n$.

Setup.

Thus, assume that H is not $\{Id\}$ and let T be the set of all integers r such that some element of H has a cycle of length $r \geq 2$ in its disjoint cycle representation.

Let $s = \max\{r | r \in T\}$.

We shall show that $3 \in T$ and indeed a 3-cycle belongs to H. Then we are done by the above generation results.

1. **Main principle.** If $h \in H$ and $g \in A_n$ then $[h, g] = h(gh^{-1}g^{-1}) \in H$, since $H \lhd A_n$.

   This is obvious since the two displayed parts are clearly both in H. Typically, this is used by calculating $[h, g]$ as $(hgh^{-1})g^{-1}$ after fixing $h \in H$ and choosing convenient $g \in G$.

2. Suppose that $s \geq 4$. Then H contains a 3-cycle.

   Without loss of generality we can assume that

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots \end{pmatrix} \cdots \in H.$$

   Set $\tau = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \in A_n$. Then

$$[\sigma, \tau] = \begin{pmatrix} 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 2 \end{pmatrix}.$$

   This is in H by the main principle, hence we are done.

3. Suppose that $s = 3$ then H contains a 3-cycle.

   PROOF. Without loss of generality we have $\sigma \in H$ of the form $\sigma = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \cdots$. If σ is just a 3-cycle, then we are done. Hence we may assume:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 & 5 & \cdots \end{pmatrix} \cdots .$$

   Let $\tau = \begin{pmatrix} 1 & 2 & 4 \end{pmatrix} \in A_n$.

   Then we see

$$[\sigma, \tau] = \begin{pmatrix} 2 & 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 3 & 5 & 2 \end{pmatrix} \in H.$$

   Thus $s \geq 5$ and we have a contradiction!

4. We now claim that $s \geq 3$.

   PROOF. Suppose if possible $s = 2$. Then we can assume without loss of generality that

$$\sigma = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 4 \end{pmatrix} \cdots \in H.$$

   Take $\tau = \begin{pmatrix} 1 & 3 & 5 \end{pmatrix} \in A_n$. Here, we seriously need $n \geq 5$. There are two possible cases.

   Case 1. $\sigma = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 4 \end{pmatrix}$. In this case, we see:

$$[\sigma, \tau] = \begin{pmatrix} 2 & 4 & 5 \end{pmatrix} \begin{pmatrix} 1 & 5 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 4 & 5 & 3 \end{pmatrix} \in H.$$

   Thus we have $s \geq 5$. Thus $s = 2$ is not possible!

   Case 2. $\sigma = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \end{pmatrix} \cdots$. In this case, we see:

$$[\sigma, \tau] = \begin{pmatrix} 2 & 4 & 6 \end{pmatrix} \begin{pmatrix} 1 & 5 & 3 \end{pmatrix} \in H.$$

   But then $s \geq 3$, another contradiction!

5. **Done!**

   We have shown that $s \geq 3$ and for each of the possibilities $s = 3$, $s \geq 4$ we have shown that H contains a 3-cycle as promised!

Remark. We now finish off the discussion by describing the situation for $n < 5$.

1. If $n = 2$ then A_2 is the identity subgroup and is trivially simple.

2. If $n = 3$, then A_3 is the cyclic group of order 3 generated by $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$. It is clearly simple, being of prime order.

3. If $n = 4$, then A_4 contains a famous normal subgroup, namely K, the Klein 4-group defined as

$$K = \{Id, \begin{pmatrix} 1 & 2 \end{pmatrix}\begin{pmatrix} 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 3 \end{pmatrix}\begin{pmatrix} 2 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 4 \end{pmatrix}\begin{pmatrix} 2 & 3 \end{pmatrix}\}.$$

   Note that in our above notation, here $s = 2$ is possible, because we don't have a fifth element to create our τ. This K is a proper normal subgroup of A_4 and is abelian. Hence, every subgroup of K is normal in K, though not in A_4. The subgroup K is indeed the unique proper normal subgroup of A_4.

   It can be shown that K is in fact the commutator subgroup $[A_4, A_4]$.

3.6 Building new groups from old.

Given two groups H, K we now investigate various possible ways of putting them together to form a new group. This naturally helps in analyzing a given group as developed from smaller groups.

The first construction is the simplest product.

Direct product: Given groups H, K by their **direct product** we mean the set $H \times K$ together with the group operation defined by

$$(h_1, k_1)(h_2, k_2) = (h_1h_2, k_1k_2).$$

Note that to avoid cluttering our notation, we are not identifying the symbols for the different group operations and thus h_1h_2 is defined as the product in H while k_1k_2 is defined as the product in K.

This direct product is also called the external direct product in contrast with the internal direct product defined below.

We note the following easy facts about the direct product.

1. If H, K are finite then $|H \times K| = |H||K|$.

2. If L is a subgroup of $H \times K$ then

$$L_H = \{h \in H | (h,k) \in L \text{ for some } k \in K\}$$

is a subgroup of H. Similarly

$$L_K = \{k \in H | (h,k) \in L \text{ for some } h \in H\}$$

is a subgroup of K.

The group $L^* = \{(h,k) | h \in L_H, k \in L_K\}$ is a subgroup of $H \times K$. It evidently contains L but can be bigger!

3. $G = H \times K$ contains two subgroups H_1, K_1 isomorphic to H and K respectively. namely $H_1 = \{(h,1) | h \in H\}$ and $K_1 = \{(1,k) | k \in K\}$. Note that we are using the symbol "1" to denote the identity in both the groups. This is an **abuse of notation,** but helps us keep the notation short.

4. We shall name the natural projection map $\pi_H : G \to H$ defined by $\pi_H(h,k) = h$. Similarly $\pi_K(h,k) = k$.

Note that then $L_H = \pi_H(L)$ and $L_K = \pi_K(L)$. Both projections are group homomorphisms with $Ker(\pi_H) = K_1$ and $Ker(\pi_K) = H_1$. Thus $G/H_1 \cong K$ and $G/K_1 \cong H$.

5. **Internal direct product.** Note that H_1, K_1 are subgroups of G satisfying these properties:

- $H_1 \bigcap K_1 = \{1\}$. **Convention:** Note that this 1 is really the identity $(1,1) \in G$. We may often just write 1 for $\{1\}$ in order to denote the appropriate group containing just the appropriate identity, or the so-called identity subgroup.
- $G = H_1 K_1$.
- $[H_1, K_1] = 1$. This cryptic notation is saying that elements of H_1, K_1 commute with each other and hence their commutators reduce to the identity subgroup.

We shall say that the group G is an internal direct product of H_1, K_1 if the above conditions hold.

It can be deduced from the given conditions that every element $g \in G$ is **uniquely** a product of some $h_1 \in H_1$ and $k_1 \in K_1$ so that $g = h_1 k_1$. As a result, it is also easy to prove that an internal direct product is in turn isomorphic to an external direct product. We simply take the map $H_1 \times K_1 \to G$ defined by $(h_1, k_1) \to h_1 k_1$.

Example. Let $G = \mathbb{Z}_{15}$ and $H_1 =< [3]_{15} >$ and $K_1 =< [5]_{15} >$. Check that G is an internal direct product of H_1, K_1. [13]

Note, in turn, that $H_1 \cong \mathbb{Z}_5$ and $K_1 \cong \mathbb{Z}_3$ and this is simply the reaffirmation that $\mathbb{Z}_{15} \cong \mathbb{Z}_5 \times \mathbb{Z}_3$.

6. **Semidirect product.** Now consider the example of the dihedral group D_{2n} with $n \geq 3$.

   Let $H =< r >$ where r is the usual element of order n and $K =< s >$ where s is the element of order 2 satisfying $srs = r^{-1}$. Then $H \bigcap K = 1$ and $D_{2n} = HK$ but $[H, K] \neq 1$. The group H, however is a normal subgroup and D_{2n}/H is isomorphic to $\mathbb{Z}_2 \cong K$.

   We formalize a product concept to model this behavior.

   Let H, K be groups and $\phi : K \to Aut(H)$ a group homomorphism. **Definition: The semidirect product of H with K induced by** ϕ is denoted by $H \rtimes_\phi K$ and is defined thus:

   - $H \rtimes_\phi K = H \times K$, i.e. as a set it coincides with the direct product.
   - The product (i.e. the group operation) is defined by

     $$(h_1, k_1)(h_2, k_2) = (h_1 \phi(k_1)(h_2), k_1 k_2).$$

     This is sometimes written by using the suggestive notation $h_2^{\phi(k_1)}$ for $\phi(k_1)(h_2)$. Thus we write:

     $$(h_1, k_1)(h_2, k_2) = (h_1 (h_2)^{\phi(k_1)}, k_1 k_2).$$

     The idea is that $\phi(k_1)$ is an automorphism of H and hence we let it act on h_2. [14]
   - Just as in the case of the direct product, we have subgroups H_1, K_1 isomorphic to H, K respectively. This time, only H_1 is expected to be a normal subgroup and we have only one projection map $\pi_H : H \rtimes_\phi K \to K$. [15]
   - **Convention.** We can drop the reference to ϕ in the notation if it is clear from the context.

[13] Remember that we need to use the additive notation here.

[14] The book shortens the notation to $k_1 \cdot h_2$. Some people write it as $h_2^{k_1}$ suppressing the notation ϕ. No matter how you shorten it, it is hiding many calculations and hence needs to be well defined and understood!

[15] The other projection map does exist, but won't be a group homomorphism! You should check in D_{2n}.

Example. Thus in the D_{2n} example, we can take $\phi(s)$ to be the automorphism of $H =< r >$ which takes r to r^{-1}. For convenience, denote this automorphism of H by τ.

Then the image $\phi(< s >)$ is simply $< \tau >$. Thus we have

$$D_{2n} \cong \mathbb{Z}_n \rtimes_\phi DZ_2$$

where, we have identified $< r >$ with $\mathbb{Z}_n$, $< s >$ with $\mathbb{Z}_2$ and $Aut(H)$ with $\mathbb{Z}_n^\times$.

Remark.

1. We may think of the direct product as a special case of a semidirect product by taking the trivial homomorphism ϕ, namely taking $\phi(k)$ to be the identity automorphism for all $k \in K$.

2. If $H \lhd G$ is a normal subgroup and $K < G$ is any subgroup such that $H \bigcap K = 1$ then we know that HK is a subgroup of G. We can identify it as a semidirect product by defining $\phi(k)$ to be the conjugation automorphism (inner automorphism) $\phi(k)(h) = khk^{-1}$.

   It can be easily seen that then $H \rtimes_\phi K \cong HK$ with the identification $(h, k) \to hk$. We only show the homomorphism property.

   We have:
   $$(h_1, k_1)(h_2, k_2) = (h_1(k_1 h_2 k_1^{-1}), k_1 k_2)$$
   and the homomorphism condition requires that

   $$(h_1 k_1)(h_2 k_2) = (h_1(k_1 h_2 k_1^{-1}))(k_1 k_2)$$

   which is evident! The rest of the claim follows.

   We can simply call this an **internal semidirect product** and write HK as $H \rtimes K$. Note that the map ϕ is not mentioned, if it corresponds to matching elements with the inner automorphisms induced by them.

3.7 Some examples and exercises.

Here are some examples with details left for the reader and some exercises to be completed.

1. **Another view of an abelian group.** Let G be a group and use the usual notation G' to denote its commutator subgroup $[G, G]$.

   Prove that G/G' is an abelian group. (*Remember that the notation implies that G' is a normal subgroup and this also needs a proof!*)

Prove that if G/H is abelian for some normal subgroup H, then $G' \triangleleft H$. This is sometimes described as "G' is the smallest normal subgroup with abelian quotient."

2. Prove that if A, B are subgroups of G such that G/A and G/B are abelian, then $G/(A \bigcap B)$ is also abelian.

3. Let K be a cyclic group of order $n > 1$.

   Deduce from old exercises that $Aut(K)$ is isomorphic to $\mathbb{Z}_n^\times$ - the multiplicative group $\{[a]_n | \text{GCD}(a, n) = 1\}$ of order $\phi(n)$.

   In particular, note that $Aut(K)$ is abelian.

4. Let K be a cyclic normal subgroup of G. Prove that $G' \subset C_G(K)$. **Hint:** You need to prove that conjugation by any element of G' is the identity automorphism of K. Use knowledge of the $Aut(K)$ above.

5. Prove that a group of order p^2 where p is a prime is either cyclic or isomorphic to $\mathbb{Z}_p \times \mathbb{Z}_p$. Thus, in either case, it is abelian.

6. Let G be a group of order p^3 where p is a prime. Recall the proof that G has a non trivial center and deduce that there is an element $z \in G$ of order p such that $[z, G] = \{e\}$.

   Conclude that the commutator subgroup $G' \subset < z >$.

7. Assume that the elements $x, y \in G$ commute with $[x, y]$. Prove the formula:
$$(xy)^n = x^n y^n [y, x]^{\frac{n(n-1)}{2}}.$$
   Hint: Use induction on n.

8. Let G be a group of order p^3 where p is an odd prime. Prove that the map $t \to t^p$ is a homomorphism of G into $Z(G)$.

   Why does p have to be odd? What changes when $p = 2$?

9. Combine the above results to describe all groups of order p^3 where p is an odd prime.

   Hint: Start with z in the center as above and separately consider the cases when $G/ < z >$ is either cyclic or isomorphic to $\mathbb{Z}_p \times \mathbb{Z}_p$.

4 Further Theorems in Groups.

We now turn our attention to some topics of importance in applications of Group Theory. We may not get to the actual applications for some time, though!

4.1 Fundamentals of p-groups.

Having studied the existence p-Sylow groups, we now turn our attention to the structure of the p-Sylow subgroups themselves.

First some definitions.

Definition: Maximal subgroup.

A subgroup $H < G$ is said to be maximal if $H \neq G$ and there is no subgroup K strictly between H and G.

Definition: Characteristic subgroup.

A subgroup $H < G$ is said to be characteristic if for every automorphism $\sigma \in Aut(g)$ we have $\sigma(H) = H$. We can alternatively describe this as " every automorphism stabilizes H ".

We write H *char* G, if H is a characteristic subgroup of G.

Note that a normal subgroup only requires that it is stabilized by all inner automorphisms, so a characteristic subgroup is necessarily normal.

Here are some **facts about characteristic subgroups** worthy of remembering:

- $\{e\}$ and G are trivial characteristic subgroups.

  The commutator subgroup G' is also characteristic. The center $Z(G)$ is also characteristic.

- If G has a unique p-Sylow subgroup, then it is clearly characteristic.

- If K *char* H and $H \triangleleft G$, then $K \triangleleft G$. For proof, we simply note that the inner automorphism given by conjugation by an element of G induces an automorphism of H since H is normal in G and hence an automorphism of K since K is characteristic in H.

Here are some **important properties of p-groups.** Let p be a prime. Let P be a p-group. The Sylow theorems don't say anything further about it, since it is its own p-Sylow subgroup. So we need further analysis to understand its properties.

1. P has non trivial center i.e. $|Z(P)| > 1$.

   This is already established by considering the conjugation action of P on itself and invoking the class equation:

   $$|P| = |Z(P)| + \textbf{sum of lengths of non singleton orbits} .$$

   Since the LHS and the second term of RHS are divisible by p, so is $|Z(P)|$ and since $Z(P)$ contains at least the identity element, $|Z(p)|$ is at least p, giving the result!

2. Every normal subgroup H of P also contains non trivial central elements.

   This is done by repeating the above argument, but this time letting P act on H by conjugation. (This is meaningful, since H is normal!) Thus we get:

   $$|H| = |Z(P) \bigcap H| + \textbf{sum of lengths of non singleton orbits}$$

   and thus, as before $|Z(P) \bigcap H| > 1$.

3. Every normal subgroup H of P of order p^s contains subgroups of order p^i, for $i = 0, 1, \cdots, s$ which are **normal in P.**

   To see this, note that we already know the result for $i = 0, 1$ since H contains the identity as well as a nontrivial central element (which we may assume to be of order p by raising to a power if necessary).

   By induction, assume that H_i is a subgroup of H with $|H_i| = p^i$ and $0 \leq i < s$. We will show that there is $H_{i+1} < H$ such that $H_{i+1} \triangleleft P$ and $|H_{i+1}| = p^{i+1}$.

   For proof, simply consider $K = H/H_i < P/H_i$ and note that $|H/H_i| = p^{s-i} > 1$. Hence, we can choose an element $z \in H$ such that its image $\overline{z}$ in H/H_i is a non trivial central element of P/H_i. Further arrange that $|\overline{z}| = p$ by raising it to a suitable power if necessary.

   With a little work, we see that $H_{i+1} =< H_i, z >$, satisfies all the needed conditions.

4. In particular, P contains **normal subgroups** of all orders dividing $|P|$.

   We simply apply the above to $H = P$. The main point is that a sequence of elements $z_1, z_2, \cdots, z_j$ generates a normal subgroup if each z_i is central modulo the subgroup generated by $z_1, \cdots, z_{i-1}$. It is this fact that leads to the idea of studying upper central series.

5. For any *proper subgroup* H of P, we have that $N_P(H)$ is strictly bigger than H.

   We make an induction on $|P|$. If $|P| = 1$ or p or even p^2, then the group is abelian and there is nothing to prove.

   We have two cases.

   - **Case 1.** Suppose $Z(P)$ contains an element $x \notin H$. Then $x \in N_P(H) \setminus H$ and the proof is finished!

- **Case 2.** Suppose $Z(P) \subset H$. Then we go modulo $Z(P)$ and note that $|P/Z(P)| < |P|$, so induction hypothesis can be applied to the group $H/Z(P)$.
  Write $\overline{P} = P/Z(P)$ and $\overline{H} = H/Z(P)$.
  Thus, there is $\overline{x} \in \overline{P}$ such that $\overline{x} \notin \overline{H}$ but $\overline{x} \in N_{\overline{P}}(\overline{H})$.
  It is easy to check that x is then in $N_P(H) \setminus H$.

6. Every maximal subgroup Q of P is normal. Moreover, such a maximal Q has index p.

   This is immediate from the above, since the normalizer $N_P(Q)$ can only be P by maximality of Q, so Q is normal! Moreover, the p-group P/Q has no proper subgroups and hence must be the smallest non trivial p-group, namely the cyclic one of order p.

4.2 Fundamental Theorem of finite abelian groups.

The fundamental theorem of finitely generated abelian groups states that any finitely generated abelian group A is isomorphic to a direct product of cyclic abelian groups. Note that a finite cyclic group is isomorphic to $\mathbb{Z}_n$ for some n, whereas an infinite cyclic group is isomorphic to $\mathbb{Z}$.

There are further uniqueness statements associated with this theorem.

1. The number of infinite cyclic groups appearing in the direct product only depends on the group and is called its **free rank** or **Betti number**.

2. The product representation of the finite cyclic groups is not unique, unless one makes some restrictions. One way is to require that all cyclic factors be p-groups for various primes p and then the representation is unique except for order.

   It is customary to use a more natural decomposition in terms of the so-called invariant factors, but since it can be uniquely deduced from the p-group factorization, we postpone its discussion until it comes up naturally in a more general context (where p-group concept does not generalize well).

In this section, we only prove the version for finite groups, so the free rank is 0 and no copies of $\mathbb{Z}$ appear!

We make a

Definition: Exponent of a group. A positive integer m is said to be the exponent of a group G if $g^m = e$ for each $g \in G$ and that m is the least such integer.

If such an integer does not exist, then we define it to be ∞. It is clearly the LCM of the orders of elements of G.

Note that a finite group always has a finite exponent and it is a factor of $|G|$, in view of Lagrange's Theorem.

Let us note that in the additive notation of abelian groups, we rewrite the condition as $mg = 0$.

Preamble for the Fundamental Theorem of finite abelian groups. Let A be a finite abelian group of order n and exponent $m|n$. We shall assume that $p_1, \cdots, p_r$ are the different prime factors of n

Thus, the theorem we aim to prove is that:

$$A = \prod_{i=1}^{r} A_i \text{ where } A_i \text{ is a } p_i\text{-group.}$$

Here we are using the internal direct product.

Further, if we let $n_i = \text{ord}_{p_i}(n)$ then A_i is the unique normal p_i-Sylow subgroup of A of order $p_i^{n_i}$ and itself is congruent to a product of p_i-groups, $A_{i_1}, \cdots$.

1. **Reduction to the p-group case.** We prove a

   Lemma. Suppose that the exponent m of A has a factorization $m = ab$ where $\text{GCD}(a, b) = 1$, then

   $$A = H \times K$$

   where H and K have exponents a, b respectively. Here, we are actually claiming that A is an internal direct product of H, K.

   Proof. Define $H = \{x \in A | x^a = e\}$ and $K = \{x \in A | x^b = e\}$. Using the abelian nature of A we see that H, K are subgroups. [16]

   Since a, b are coprime, there exist integers u, v such that $au + bv = 1$.

   If $x \in H \bigcap K$ then from $x^a = x^b = e$ we note that

   $$x = x^1 = x^{au+bv} = (x^a)^u (x^b)^v = e.$$

   Thus $H \bigcap K = \{e\}$.

   Moreover, given any $x \in A$, note that $x^b \in H$ since clearly $(x^b)^a = x^m = e$. Similarly, $x^a \in K$ since $(x^a)^b = x^m = e$.

   Thus, by the same calculation as above $x = (x^b)^v (x^a)^u \in HK$. We have thus proved that A is the internal direct product of H and K as

[16] Simplest precise argument might be to note that H is simply the kernel of the a-power homomorphism F_a from A into A. Similarly for K.

asserted. Clearly, the exponent of H is some a_1 dividing a and the exponent of K is some b_1 dividing b.

¿From $A = HK$ it follows that $x^{a_1 b_1} = e$ for all $x \in A$ and hence $m = ab$ divides $a_1 b_1$. From coprimeness of a, b it follows that $a_1 = a$ and $b_1 = b$.

The reduction. Applying the above lemma repeatedly, it is easy to see that the group A is isomorphic to the direct product of groups A_i whose orders are powers of a single prime p_i. This reduces the proof of the theorem to the case where we only need to consider a to be a p-group for some prime p.

2. **Case of a p-group.** Now we **simplify our notation** by assuming that p is a prime and A is a finite abelian p-group of order n and exponent m.

   Important change of representation. We will find it more convenient to think of our abelian group in additive notation.

   We now list the changes in our language caused by this conversion.

   - The identity of the group A is now 0 and the exponent m is the smallest positive integer such that $mx = 0$ for all $x \in A$.
   - We shall use induction on n, the case of $n = 1, p$ being trivial.
   - **The elementary abelian case.** An abelian p-group A is said to be elementary abelian, if every non zero element (non identity element) has order p, i.e. $pA = \{0\}$.

     Another way of saying this is to note that equivalently, the exponent is p.

     Drawing on our knowledge from elementary algebra and linear algebra, we note that such a group is naturally a *vector space* over the field $\mathbb{Z}_p$. We use the usual additive group structure and the scalar multiplication $[r]_p x$ when $r \in \mathbb{Z}_p$ is simply defined as rx. The condition of being elementary abelian makes this well defined.

     Now, the usual basis of the vector space gives the generators necessary to make the direct product (or direct sum in conformity with the additive notation).

     Thus, if $x_1, \cdots, x_t$ is a basis of A, then

     $$A = \prod_{i=1}^{t} < x_i > .$$

     This is easy to see from the vector space theory. Thus our theorem is now considered proved for elementary abelian p-groups.

We shall prove that A is a product of cyclic p-groups(or the so-called direct sum in conformity with the additive notation.)

3. **General case of the p-group.** Let $H = pA = \{px | x \in A\}$ and $K = \{x \in A | px = 0\}$.

   Note that the homomorphism $x \to px$ has kernel K and image H, so we get that $|A| = |K||H|$.

   It is easy to see that $|H| < |A|$ since otherwise we will get an infinite sequence of elements $x_1, \cdot, x_s, \cdots$ such that $x_i = px_{i+1}$, in contradiction with the finiteness of A.

   Thus, we may assume H to be a direct product of cyclic p-groups $< h_i >$, for $i = 1, \cdots, r$, say. For each h_i we can find $g_i \in A$ such that $pg_i = h_i$.

   Then we can easily see that the group generated by $\{g_1, \cdots, g_r\}$ is the internal direct product of $\{< g_i >\}$ and we shall denote it by A_0.

   Suppose that $|h_i| = p^{n_i}$ for each i. Then it is easy to see from the direct product condition that $|H| = \prod_{i=1}^{r}(p^{n_i})$ and $|A_0| = \prod_{i=1}^{r}(p^{n_i+1}) = |H|p^r$.

   Moreover, clearly,

$$H \bigcap K =< p^{n_1-1}h_1 > \times \cdots \times < p^{n_r-1}h_r > .$$

   Thus $|H \bigcap K| = p \cdots p = p^r$.

   Now, if $K \subset H$ then

$$|K| = |H \bigcap K| = p^r$$

   and we see that $|A_0| = |H|p^r = |H||K| = |A|$. Then $A_0 = A$ and the direct product structure of A_0 finishes the proof.

   In case $K \notin H$ we need to enlarge A_0. Actually, in this case, we invoke the induction in a different attack.

   Consider the group A/H which is clearly elementary abelian by definition of H. Let $x \in K \setminus H$ and note that $\overline{x} \in A/H$ is a non zero element. We can make a basis for the $\mathbb{Z}_p$-vector space A/H as $\overline{x} = \overline{y_1}, \cdots, \overline{y_s}$.

   We claim that

$$A =< x > \times < y_2, \cdots, y_s > .$$

   If we have an element in the intersection of $< x >$ and $< y_2, \cdots, y_s >$, then it must be also in H, since it goes to zero modulo H. But $< x >$

$\bigcap H = \{0\}$ and hence the intersection of the two groups is the zero group.

Let B be the group generated by $< x >$ and $< y_2, \cdots, y_s >$. It remains to show that every element $g \in A$ can be written as an element of B. Clearly, we can write $g = u_1 + pg_1$ where $u \in B$ and $pg_1 \in H$. Applying a similar process to g_1 and collecting terms, we can write $g = u_1 + pu_2 + p^2 g_2$. Continuing, we see write

$$g = u_1 + pu_2 + \cdots + p^i u_{i+1} g_i.$$

For large enough i, we will have p^i divisible by the exponent of A and hence $p^i g_{i+1} = 0$. Thus $g = u_1 + pu_2 + \cdots + p^{i-1} u_i \in B$.

4. **Uniqueness considerations.** We have now shown that for any abelian p-group, we have a sequence of non negative integers $d_i, i = 1, \cdots, t$ such that A is a direct product of d_1 copies of $\mathbb{Z}_p$, d_2 copies of $\mathbb{Z}_{p^2}$ and so on until d_t copies of $\mathbb{Z}_{p^t}$, where $d_t \neq 0$.

   It can be shown, say by counting elements of various order in A that the sequence of numbers $d_1, \cdots, d_t$ is completely determined by the group A.

   This will be done in exercises. We only illustrate two simple results along these lines.

   - The value of t corresponding to the highest nonzero d_t is simply determined by the exponent being p^t. This follows from the fact that the exponent of such a direct product is clearly the exponent of $\mathbb{Z}_{p^t}$.
   - Each copy of $\mathbb{Z}_{p^t}$ contains exactly $p^t - p^{t-1}$ elements of order p^t. Then the product of d_t copies of these clearly contains $(p^t - p^{t-1})^{d_t}$ such elements of order p^t. Thus, the number d_t is uniquely determined as the log to the base $p^t - p^{t-1}$ of the number of elements of order p^t.
   - Continuing this way, we can successively determine each of $d_t, d_{t-1}, \cdots d_1$.
   - Even though the above is simple, it is also kind of tedious to keep track of. Thus, a better strategy is a suitable induction. We do this next.

5. **Formula for the number of elements of a given order.**

5 Extra Topics.

The edits are just finished. It may still need minor proof reading. 9:50 AM 10/5

5.1 Symmetric Functions

We discuss an important application of the ideas of group actions and orbits to the study of symmetric functions which are essential in theory of equations and Galois Theory.

We start with:

Definition: symmetric function A function $f(X_1, \cdots, X_n)$ of several variables $X_1, \cdots, X_n$ is said to be a symmetric function if it is invariant under any rearrangement of its arguments. Of course, the notion is of interest only when $n > 1$.

For example, $X_1 + X_2, X_1X_2$ and $\frac{a+b(X_1+X_2)}{X_1^2+X_2^2+cX_1X_2}$ are all symmetric functions of X_1, X_2 where a, b, c are parameters and not involved in the permutation.

We begin with some convenient notations:

1. As usual, let $\mathbb{N}$ be the set of all non negative integers. We will denote a monomial $X_1^{i_1} \cdots X_n^{i_n}$ by X^i where $i = (i_1, \cdots, i_n)$ and each $i_j \in \mathbb{N}$. We will also use the symbol X to denote $(X_1, \cdots, X_n)$ when convenient.

   Let $M = \{X^i \mid 0 \leq i_j \leq n \forall 1 \leq j \leq n\}$.

   Then, any polynomial in $h(X) \in K[X] = [X_1, \cdots X_n]$ can be expressed as $h(X) \sum_{i \in A} h_i X^i$, where the coefficients $h_i \in K$ and A is a finite subset of $\mathbb{N}^n$.

   If A is empty, then the polynomial is the zero polynomial.

   For any non zero polynomial $h(X) \in K[X]$ we define its support $Supp(h(X))$ by $\{i \mid h_i \neq 0\}$

2. We define an order among the monomial in $K[X]$ by $X^i >_M X^j$ if there is an s such that $1 \leq s \leq n$ such that $i_u = j_u$ if $u < s$ and $i_s > j_s$. This is one of the lexicographical orders and it totally orders the set of all monomials M in X with $1 = X^{(0,\cdots,0)}$ as the smallest and there is no largest.

   However, if we restrict the total degree to an integer D, then the largest would be $X_1^D = X^{(D,\cdots,0)}$.

3. For any non zero $h(X) \in K[X]$ we define its leading form $Lf(h(X)) = h_i X^i$ if X^i is the largest monomial in $Supp(h(X))$.

The corresponding exponent $i = (i_1, \cdots, i_n) \in M$ may be denoted as $\deg(h(X); M)$ and we may clearly extend the $>_M$ to the exponents by defining $i >_M j$ if and only if $X^i >_M X^j$. We may call the largest i in the support of $h(X)$ as the $M - degree$ of $h(X)$.

4. We define the action of the permutation group S_n on M by $\sigma(X^i) = X^{\sigma(i)}$ where $\sigma \in S_n$ and $\sigma(i) = (i_{\sigma(1)}, \cdots, i_{\sigma(n)})$. We extend this action by linearity on $K[X]$ by $\sigma(\sum_{i \in A} c_i X^i) = \sum_{i \in A} c_i X^{\sigma(i)}$.

5. We, next define some basic symmetric functions. If $X^i \in M$, then we define

   - $Sym(i) = \sum_{\sigma \in S_n} \sigma(X^i)$.
   - Let $f(i)$ denote the order of subgroup $Fix_{S_n}(i)$. Then each term in $Sym(i)$ s repeated exactly $f(i)$ times.

     We define $Sym^*(i)$ to be the corresponding sum where we let σ vary only over the coset representatives of $Fix_{S_n}(i)$. If $f(i)$ is not divisible by the characteristic of K, then we could simply note that $Sym^*(i) = \frac{1}{f(i)} Sym(i)$.
   - We note that the largest monomial in the support of $Sym(i)$ as well as $Sym^*(i)$ is j such that $j_1 \geq j_2 \geq \cdots j_n$.
   - Let Z be a new indeterminate. If we expand the polynomial $\prod_{i=1}^{n}(Z + X_j)$ as $Z^n + e_1 Z^{n-1} + \cdots e_n$, then it is easy to see that $e_1 = Sym^*(1, 0, \cdots, 0), e_2 = Sym^*(1, 1, \cdots, 0), \cdots, e_n = Sym^*(1, 1, \cdots, 1)$ and these are the usual basic symmetric functions.
   - Another important set of symmetric functions, often called Newtonian symmetric functions are $s_j = Sym^*(j, 0, \cdots, 0)$.

6. We now prove the

   Fundamental Theorem of Symmetric Functions which states that the set of symmetric polynomials in X is a subring of $K[X]$ generated by $e_1, \cdots, e_n$.

   Proof. Let $E = K[e_1, \cdots, e_n]$ be the ring of polynomials generated by $e_1, \cdots, e_n$.

   Let $h(X) \in K[X]$ be a symmetric polynomials. We wish to show that $h(X) \in E$.

   Since the zero polynomial belongs E, it is enough to assume $h(X)$ to be non zero.

We shall denote the usual total degree in $(X$, $cdots, X_n)$ as $\deg_X(h(X))$.

We make induction on $(\deg_X(h(X)), \deg(h(X); M))$. We note that for the smallest $\deg_X(h(X)) = 0$, we have $0 \neq h(X) \in K \subset E$ and hence the theorem holds.

If $\deg_X(h(X)) = 1$, then for $u = (1, 0, \cdots, 0)$ we have $Sym^*(u) = e_1$ and by symmetric property of $h(X)$ all linear terms must have the same coefficient as h_u, so $h(X) = h_u Sym^*(u) + c$ for some $c \in K$. Thus, $h(X) = h_u e_1 + c \in E$.

Now let $Lf(h(X)) = h_v X^v$ and thus $\deg(h(X); M) = v$. Assume that the theorem holds for all smaller $M - degree$ polynomials as well as for smaller total degree polynomials.

We first note that for all permutations $\sigma \in S_n$ we must have that $h_{\sigma(v)} = h_v$ and thus $\deg(h(X) - h_v Sym^*(v); M) <_M \deg(h(X); M)$.

So, it is enough to prove that $Sym^*(v) \in E$. We note that $Sym^*(v) = Sym^*(\sigma(v))$ for all $\sigma \in S_n$ and so, without loss of generality, we may assume that $v_1 \geq v_2 \geq \cdots \geq v_n$. [17]

We shall continue to use the induction hypothesis and consider $\mu_v(X) = \prod_{i=1}^n e_j^{d_j} \in E$ where $d_n = v_n$ and by decreasing induction, we set $d_j = v_j - \sum_{r=j+1}^n d_r$ for $j = n-1, n-2, \cdots, 1$.

We note that

$$Lf(\mu_v(X)) = X^{(d_1,0,\cdots,0)} X^{(d_2,d_2,0,\cdots,0)} \cdots X^{(d_n,d_n,\cdots,d_n)} = X^v.$$

Hence, $\deg(Sym^*(v) - \mu_v(X); M) < \deg(Sym^*(v); M)$ and so by induction, the difference $Sym^*(v) - \mu_v(X)$ is in E.

Therefore, $(Sym^*(v) - \mu_v(X)) + \mu_v(X) = Sym^*(v) \in E$.

Thus, our Theorem is established.

7. We immediately get the **Corollary:** The field of symmetric rational functions in X is equal to the field of rational functions in $e_1, \cdots e_n$.

   Proof. Let $\frac{f(X)}{g(X)}$ be a symmetric rational function. Consider the polynomial $G(X) = g(X) \prod_{\sigma \in S_n \setminus Id} \sigma(g(X))$. Then it is evident that $g(X)G(X)$ is symmetric and thus, our rational function is equal to $\frac{f(X)G(X)}{g(X)G(X)}$ whose denominator is a symmetric function. It follows that the numerator is also symmetric and hence by our theorem both the numerator and denominator in this expression are in E.

   This proves the Corollary.

[17] This condition is really equivalent to the assumption that X^v is indeed $Lf(Sym^*(v))$.

8. We now consider the extension of fields $K(e_1, \cdots, e_n) \subset K(X_1, \cdots, X_n)$. Both these fields are said to be "pure transcendental of dimension n" which means they are generated by n algebraically independent elements. It is easily seen that they are therefore isomorphic even though one is a proper subfield of the other. The field derived from E may be conveniently described as $Fix_{S_n}(k(X))$.

   Naturally, we may replace the group S_n by other subgroups of S_n. The question of determination of the fixed field is interesting and unsolved.

   First, there is the well known E. Noether's problem which asked if the fixed field is always rational. She proved it for small n ($n \leq 4$). R. G. Swan found a counterexample for $n = 47$. The problem was tantalizing because, if true, it would have given an easy proof of a fundamental question of Galois Theory; namely, is every finite group representable as a Galois group of some polynomial over $\mathbb{Q}$?

   An old Algebra book of Jacobsen posed an exercise to show that the fixed field of the alternating group A_n is always pure transcendental. This field, at least has an explicit description, namely it is generated by the square root of the Z-discriminant of $\prod_{i=1}^{n}(Z + X_i)$ over the field generated by E. After several years, the exercise was removed, since so far, only small values of n have been successfully tackled.

9. While the $\{e_i\}$ is a good finite set of generators for E, often an infinite set of generators is more useful in computations. These are the "Newtonian elementary functions" and we now give the necessary formulas. Their drawback is that they don't generate the same ring in positive characteristic, but they work fine in characteristic zero.

   We now prove the basic identities due to Newton.

   Define s_i for $i \in \mathbb{N}$ by $s_i = \sum_{j=1}^{n} X_j^i$.

   We have **Theorem: Newton**

   Let
   $$P(X) = \prod_{i=1}^{n}(X - X_i) = X^n + p_1 X^{n-1} + \cdots + p_n.$$

   Note: We have $p_i = (-1)^i e_i$.

   Then we have

   (a) For $1 \leq m \leq n-1$ we have:

   $$s_m + p_1 s_{m-1} + \cdots + p_{m-1} s_1 + p_m m.$$

(b) For $m \geq n$ we have:

$$s_m + p_1 s_{m-1} + \cdots + p_{n-1} s_{m+1-n} + p_n s_{m-n}$$

Proof:

By logarithmic differentiation, we have

$$\frac{P'(X)}{P(X)} = \frac{nX^{n-1} + (n-1)p_1 X^{n-2} + \cdots + p_{n-1}}{X^n + p_1 X^{n-1} + \cdots + p_{n-1}X + p_n} = \sum_{i=1}^{n} \frac{1}{X - X_i}.$$

We put $X = T^{-1}$, and first note that

$$\frac{P'(X)}{P(X)} = \frac{nT^{-(n-1))} + (n-1)p_1 T^{-(n-2)} + \cdots + p_{n-1}}{T^{-n} + p_1 T^{-(n-1)} + \cdots + p_n}.$$

Further simplification gives:

$$\frac{P'(X)}{P(X)} = T\frac{n + (n-1)p_1 T + \cdots + p_{n-1}T^{(n-1)}}{1 + p_1 T + \cdots + p_n T^n}.$$

Using Geometric series expansion and a similar substitution, we deduce:

$$\frac{1}{X - X_i} = \frac{1}{T^{-1}(1 - TX_i)} = T\sum_{j=0}^{\infty} X_i^j T^j.$$

Putting it together, we get:

$$T\frac{n + (n-1)p_1 T + \cdots + p_{n-1}T^{(n-1)}}{1 + p_1 T + \cdots + p_n T^n} = T\left(\sum_{i=1}^{n}\sum_{j=0}^{\infty} X_i^j T^j\right).$$

The RHS is now seen to be $T\sum_{j=0}^{\infty} s_j T^j$.

Canceling T from both sides and multiplying both sides by $1 + p_1 T + \cdots p_n T^n$ and collecting coefficients of powers of T on the right, we get:

$$n + (n-1)p_1 T + \cdots + p_{n-1}T^{(n-1)} = (1 + p_1 T + \cdots p_n T^n)\left(\sum_{i=0}^{\infty} s_i T^i\right).$$

The necessary identities can now be read off by comparing the coefficients of powers of T.

Thus, if $m < n$ by comparing coefficients of T^m on both sides, we see that:

$$(n-m)p_m = s_m + p_1 s_{m-1} + \cdots + p_m s_0.$$

Since, $s_0 = n$ we cancel np_m from both sides and moving all terms to one side, we get

$$s_m + p_1 s_{m-1} + \cdots + m p_m.$$

If $m \geq n$, then the LHS has 0 coefficient and the result follows just from the Cauchy product on the right.

To be continued ...