25/11/2025

Sylow theorems
Modern Algebra project

1. Quintic equations and the start of symmetries

In 1766, after an invitation from Frederick the Great, took over the post of "Director of Mathematics" in the prussian academy of sciences.

During this time one of the most researched problems in mathematics was the possibility of discovering a quintic formula similar to the quadratic and cubic formulas discovered in the previous decades. One of the mathematicians working on this problem was the prodigy Joseph-Louis Lagrange.

Lagrange started studying mathematics at the age of 14, became one of the first founders of calculus of variations at just 17, became a professor of mathematics at the age of 19, and lastly - based on recommendations of Euler and D'Alembert - got that after from Frederick the great.

During his time in Berlin he started working on quintic equations but in an unusual manner. His main question was "Why the ~~quadratic and~~ cubic formulas worked?"

Lagrange considered a cubic equation with arbitrary coefficents:

$$a_3x^3 + a_2x^2 + a_3x + a_4 = 0$$

with roots labled x_1, x_2, x_3.

Lagrange's idea was constructing a new, seemingly more complicated function:

$$y = (x_1 + wx_2 + w^2x_3)^3 \qquad w \neq \pm 1,\ w^3 = 1$$

Insert visualization

Lagrange now asked "what if the order of the roots were permuted?". Meaning instead of considering the roots in the order of (x_1, x_2, x_3) like the above equation, what if (x_2, x_1, x_3) was chosen instead? What will happen to the value of y?

Lagrange discovered that y has only 2 possible values of y meaning that we can create a much simpler quadratic function with its roots being the 2 possible values of y from the coefficents of the original cubic equation and find its roots and just take their root and one would find a system of linear equations that can be solved to find the original roots.

Great! Why can't we do the same with quintic equations? The problem was whenever Lagrange tried to simplify the quintic equation by constructing a function like y it resulted in a degree 6 or more equation instead of simplifying it.

Maybe add an interactive visualization.

The question now shifted. Given an n-degree equation of n-roots and a function $f(x_1, x_2, \ldots, x_n)$ on the n-roots, how many values will this function have (orbit), and how many roots will stabilize each value?

2. The Orbit - Stabilizer theorem.

In his paper, Réflexions sur la résolution algébrique des équations, Lagrange sat the basics of using symmetries in group theory and even the abstract group theory we know today.

While studying the roots of equations, Lagrange proved the orbit - stabilizer theorem which is a more general statement of the theory which will be known by his name: Lagrange theorem.

What lagrange was doing was acting the group of symmetries on 3 letters on the set of cubic roots $\{x_1, x_2, x_3\}$. Which will be our first definition.

Definition 2.1. : Group Actions

Let $(G, *)$ be a group and X be a non-empty set. A left group action is a function: $\Phi : G \times X \longrightarrow X$

written $g \cdot x$ satisfying the following:

1. $e \cdot x = x$
2. $g_1 \cdot (g_2 \cdot x) = (g_1 * g_2) \cdot x$

Insert examples and visualizations of this, maybe the dihedral group

Alternatively group actions could be thought of as a group homomorphism

$$\Psi : G \longrightarrow Sym(X)$$

$$\forall g \in G, \quad \Psi(g) = \pi_g \quad , \quad \pi_g = g \cdot x$$

An orbit is like the possible values of the function y, or more pricesely the orbit of an element x is all possible elements x can be under the action.

Definition: 2.2: Orbits

Let G be a group acting on X. For any element $x \in X$, the orbit of x under G is the subset of X defined as

$$Orb_G(x) = \{g \cdot x \mid g \in G\}$$

Insert visualization and visual examples

Remarks: - If for some x, $X = Orb_G(x)$, then the action is called transitive.

Lemma 2.1.

The orbits of a group action partitions the set X.

Visualize

Proof:

1. Non-emptiness.

By the identity axiom of group actions

$$e \cdot x = x \in Orb_G(x)$$

2. Disjointness

Assume for the sake of contradiction that $Orb_G(x) \neq Orb_G(y)$ and $Orb_G(x) \cap Orb_G(y) \neq \emptyset$. Then there must be at least one element $z \in Orb_G(x) \cap Orb_G(y)$, and

$$z = g_1 \cdot x = g_2 \cdot y$$
$$\rightarrow (g_1^{-1} * g_1) \cdot x = (g_1^{-1} * g_2) \cdot y$$
$$\rightarrow x = (g_1^{-1} * g_2) \cdot y$$

Which means $x \in Orb_G(y)$, then for any element a in $orb_G(x)$

$$a = b \cdot x = b \cdot (g_1^{-1} * g_2) \cdot y$$
$$= (b * g_1^{-1} * g_2) \cdot y \in Orb_G(y)$$

Meaning $Orb_G(x) \subseteq Orb_G(y)$.

The other inclusion is similar contradicting our assumption that $Orb_G(x) \neq Orb_G(y)$.

3. Orbits cover X

The fact that $\bigcup_{x \in X} Orb_G(x) \subseteq X$ is immediate, and as $x \in Orb_G(x)$ then $X \subseteq \bigcup_{x \in X} orb_G(x)$. □

The last definition we need is the answer of the question "what elements of G stabilizes an element in the set?" Like the roots giving the same value of y.

Definition 2.3.: Stabilizers.

Let G be a group acting on X, the stabilizer of x in G is the subset of G defined as

$$Stab_G(x) = \{g \in G \mid g \cdot x = x\}.$$

Add visualizations and visual examples.

Lemma. 2.2.

The stabilizer is a subgroup of G.

Proof:

By the identity axiom the stabilizer has at least the identity.

By the two-step test notice that if $g_1, g_2 \in Stab_G(x)$

$$(g_1 * g_2) \cdot x = g_1 \cdot (g_2 \cdot x) = g_1 \cdot x = x$$

$$\rightarrow g_1 * g_2 \in Stab_G(x)$$

and

$$g_1 \cdot x = x$$

$$(g_1^{-1} * g_1) \cdot x = g^{-1} \cdot x$$

$$x = g^{-1} \cdot x \rightarrow g^{-1} \in Stab_G(x)$$

Hence, it is a subgroup. □

Remarks: - The stabilizer of an element is Not always normal.
Consider the $\text{stab}_{D_4}(1) = \{R_0, D\}$ subgroup is not normal in D_4 !

Visualize

Now we approach the most fundamental result of Lagrange's paper.

Theorem 2.1.: The orbit-stabilizer theorem

Let G be a finite group acting on a set X. For any element $x \in X$, we have

$$|G| = |\text{Orb}_G(x)| \cdot |\text{Stab}_G(x)|$$

Visualize Visualize
Visualize Visualize
Visualize Visualize

Proof!

1. The map

Let G/S be the set of all left cosets of the stabilizer S, and define $f: G/S \longrightarrow \text{Orb}_G(x)$ with $f(gS) = g \cdot x$

2. The map is well-defined

Let g_1 and g_2 be two representatives of the same coset:

$$g_1 S = g_2 S \longleftrightarrow g_2^{-1} g_1 \in S$$

Meaning

$$(g_2^{-1} g_1) \cdot x = x$$

$$\longrightarrow \quad g_2 (g_2^{-1} g_1) \cdot x = g_2 \cdot x$$

$$\longrightarrow \quad g_1 \cdot x = g_2 \cdot x$$

3. The map is injective

Suppose that $f(g_1 S) = f(g_2 S) \longrightarrow g_1 \cdot x = g_2 x$

$$\longrightarrow (g_2^{-1} g_1) \cdot x = x$$

$$\longrightarrow g_2^{-1} g_1 \in S$$

then $g_1 S = g_2 S$.

4. The map is surjective

Let y be an arbitrary element in $Orb_G(x)$, then $y = g \cdot x$ for some $g \in G$. Then $f(gS) = g \cdot x = y$. Meaning we found a pre-image of each element in $Orb_G(x)$.

Thus we found a bijection from G/S to $Orb_G(x)$, meaning that

$$|G/S| = |Orb_G(x)|$$

$$\rightarrow \quad \frac{|G|}{|Stab_G(x)|} = |Orb_G(x)|$$

$$\rightarrow |G| = |Orb_G(x)| \cdot |Stab_G(x)|$$

□

If you are paying attention you will notice that Lagrange's theorem follows immediately now by setting the set acted on as

$$X = \{aH \mid a \in G\}$$

for some subgroup H, and the action is

$$g \cdot (aH) = (ga)H$$

Extra visualize

Now if you calculated the orbit and the stabilizer of H you will find that

$$Orb_G(H) = \{g \cdot H \mid g \in G\} = \{gH \mid g \in G\} = X$$

$$Stab_G(H) = \{g \in G \mid gH = H \text{ (i.e. } g \in H)\} = |H|$$

Applying the theorem

$$|G| = |G:H| \cdot |H|$$

Which is precisely the statement of Lagrange's theorem.

With the help of Lagrange now we know the neccesary condition for existence of the subgroup.

Later, with the work of Galois and Abel it was proven that an algebraic formula for quntic equations was impossible, marking an end to a decades -or even centuries- long search.

Afterwards in the early 19th century the main area of focus was groups of symmetries. after the mathematical community noticed that symmetries was indeed a powerfull tool of capturing mathematical elegance and solving hard problems.

Later this will get more abstract and will be known as the group theory we know today. In today's group theory the main area of research is trying to classify all groups up to isomorphism.

A simpler question rose in the contex of symmetries. What are the orders of subgroups that are always guranteed to be found?

3. The dawn of the sufficient condition.

Generally, the converse of Lagrange's theorem doesn't hold. For example $|A_4| = 12$, but it has no subgroups of order 6!

The first gurantee of existence and the first existence theorem in group theory was done by Augustin-Louis Cauchy in his paper in 1845 along with very important and impactful results till this day like the modern cycle notation and other important theorems in the symmetric group.

We will start with a definition that can be seen as a generalization of the notion of a coset.

Definition 3.1: Double cosets

Let G be a group, and let H and K be subgroups of G, the double coset of g with respect to H and K is the set

$$HgK = \{hgk \mid h \in H, k \in K\}.$$

Visualize

Example:

In S_3. Let $H = \{(1), (12)\}$ and $K = \{(1), (13)\}$

The double cosets are

$H(1)K = HK = \{(1), (12), (13), (132)\}$

$H(23)K = \{(123), (23)\}$

$\longrightarrow S_3 = H(1)K \cup H(23)K$

It should be easy to prove that double cosets partition the group.

Now we start building the arsenal we need to tackle the most important theorem of this section.

Visualize and motivate with exomples before discussing the lemma

Lemma 3.1

Let G and H be two subgroups of the symmetric subgroup S_n. If the product of their orders doesn't divide the order of S_n, then there must exist a non-identity element $g \in G$ and a non-identity element $h \in H$ such that g and h are conjugates in S_n

$$g = a h a^{-1}$$

Proof:

Notice that the size of the double coset is given by

$$|G \times H| = \frac{|G| \cdot |H|}{|G \cap x H x^{-1}|}$$

Assume that there is no non-identity element of G is a conjugate of any non-identity element of H. This means that for all $x \in S_n$

$$|G \cap x H x^{-1}| = 1$$

$$\longrightarrow |G \times H| = |G| \cdot |H|$$

And as S_n is partitioned by the double cosets then

$$|S_n| = k(|G| \cdot |H|)$$

then $(|G| \cdot |H|)$ divides the order of S_n.

The contrapositive is the theorem. □

Now what follows is the a similar proof to the original proof by Cauchy. Always refer to the original paper by Cauchy.

We will consider a special action of the group on itself defined as

$$g \cdot a = gag^{-1}$$

The orbit of this action is called the conjugacy class of the element a, and the stabilizer is called the centeralizer of a.

$$Cl(a) = \{gag^{-1} \mid g \in G\} \qquad C_G(a) = \{g \in G \mid gag^{-1} = a\}$$

Now notice that conjugacy classes partition the group, so

$$|G| = \sum |Cl(g)|$$

but it can be that g is in the center, then $|Cl(z)| = 1$, so we can write this equation in terms of centeral element z, and non-centeral elements a_i as

$$|G| = |Z(G)| + \sum |Cl(a_i)|$$

$$\longrightarrow \quad |G| = |Z(G)| + \sum [G : C_G(a_i)]$$

Which is called the class equation.

Theorem 3.1: Cauchy's theorem

Let G be a finite group and p be a prime dividing the order of G, then G has a subgroup of order p.

Proof:

We know that theorem holds for abelian groups. then if p divides the order of the center of the group G, then G has a subgroup of order p.

Otherwise we may assume that p doesn't divide the order of the center. This will make the class equation of G be

$$0 \equiv |Z(G)| + \sum |Cl(a_i)| \pmod{p}$$

Then at least one of the orders in the sum is not divisible by p. and let that conjugacy class be $Cl(a)$, then

$$0 \equiv |C_G(a)| \cdot |Cl(a)| \pmod{p}$$

and as $|Cl(a)|$ is not divisible by p, then $|C_G(a)|$ must be divisible by p.

Therefore we found a subgroup that has the order of p. □

Now that we saw the modern proof of the theorem, we can see the original proof of Cauchy.

Note: Cauchy and Sylow and all mathematicians at the period considered only groups as subgroups of S_n, not abstract groups like we do today. The bridge between the two is Cayley's theorem.

Cauchy's original proof:

add construction method and visualize it.

Let $G \leq S_n$ and $p \mid |G|$, now construct* a group H with $|H| = p^t$ which is the largest order of p dividing $n!$. We know that this group exist by construction.

Now consider the double coset $G \times H$, the order of the double coset if the two groups share no conjugate element is

$$|G| \cdot |H| = (k \cdot p) \cdot p^t = k \cdot p^{t+1} \mid n!$$

Contradicting the choice of H, meaning the have conjugate elements in common. Meaning that there exists an element in G with

$$|g| = |h| = p^k \qquad k \geq 1$$

Now consider the element $g^{p^{k-1}}$

$$|g^{p^{k-1}}| = \frac{|g|}{\gcd(|g|, p^{k-1})} = \frac{p^k}{p^{k-1}} = p$$ □

This has brought us one step closer to understanding groups! Now we know some groups that must exist, and the question now was how much further can we push this?

The last tool we need to tackle this problem is a powerful theorem called the correspondence theorem.

4. The correspondence theorem

This theorem developed through the decades in group theory. The first person, however, who formalated it in modern abstract generality of today was the brilliant algebraist Emmy Noether.*

Theorem 4.1.: The correspondence theorem.

Let G be a group and $N \trianglelefteq G$, and

$$S = \{H \mid N \subseteq H \leq G\} \quad \text{and} \quad \overline{S} = \{\bar{K} \mid \bar{K} \leq G/N\}$$

, then there is a bijective map

$$\varphi : S \longrightarrow \overline{S}$$

$$\varphi(H) = H/N = \{hN \mid h \in H\}$$

Proof:

1. The map is well defined

We need to prove that $H/N \leq G/N$

Non-empitness: Since $e \in H$, $eN = N \in \varphi(H)$

Closed under the operation: Let $h_1, h_2 \in H$; as h is a supgroup $(h_1 h_2) \in H$, then if $h_1 N, h_2 N \in \varphi(H)$, $(h_1 h_2)N \in \varphi(H)$ too

Closed under inverses: Let $h \in H$. then as H is a subgroup $h^{-1} \in H$. then if $hN \in \varphi(H)$ $(hN)^{-1} = h^{-1}N \in \varphi(H)$ too.

2. The map is injective

Let $\varphi(A) = \varphi(B)$, then if $x \in A$, then $xN \in \varphi(A)$, so there exist some $b \in B$ with $xN = bN$. Meaning that $b^{-1}x \in N$, and as $N \subseteq B$, $b^{-1}x \in B$, as B is a subgroup $b(b^{-1}x) = x \in B$.

Meaning $A \subseteq B$, and the converse is similar.

* I have tried to find Noether's original proof of this theorem, but I have failed to get a hold of the original paper "Abstrakter Aufbau der Idealtheorie...."

3. The map is surjective

Define

$$K = \{ g \in G \mid gN \in \bar{K} \}$$

Since $\bar{K}$ is a subgroup the identity must be included in it ($N \in \bar{K}$), and as $nN = N$ for all $n \in N$, meaning $N \subseteq K$

Let $\pi(g) = gN$, the $K = \pi^{-1}(\bar{K})$

K is a subgroup because if $x, y \in K$, then $\pi(x), \pi(y) \in \bar{K}$ and as $\bar{K}$ is a group $\pi(x)\pi(y) = \pi(xy) \in \bar{K}$, then $xy \in K$.
Similarly, let $x \in K$, then $\pi(x) \in \bar{K}$ and os $\bar{K}$ is a group $(\pi(x))^{-1} \in \bar{K}$ and $\pi(x^{-1}) \in \bar{K} \longrightarrow x^{-1} \in K$.

Lastly by definition K is the preimage of $\bar{K}$. □

Corollary 4.1

For any subgroups $A, B \in S$

$$[B : A] = [\, {}^{B}/_{N} : {}^{A}/_{N} \,]$$

Proof:

Define a map $\psi : {}^{B}/_{A} \longrightarrow {}^{B/N}/_{A/N}$ with

$$\psi(bA) = (bN)\,({}^{A}/_{N})$$

It should be pretty straigth forword to verify that this map is bijective.

5. Sylow theorems

In 1872 the Norwegian highscool teacher Ludvig Sylow published his paper "Théorèmes sur les groupes de substitutions". This paper was a huge step in finding all subgroups of ANY subgroup.

Sylow was deeply inspired by the works of Abel, and it was while reading Abel's manuscipts when Sylow realized that he could generalize Cauchy's works to powers of primes.

Definition 5.1: Sylow p-subgroups

Let G be a finite group of order $p^{\alpha} \cdot m$, where p is a prime and p does not divide m.
A Sylow subgroup is a subgroup P of order p^{α}

From this point onward in our project we may move away a bit from the way mathematicians proved the theorems in the original papers to easier modern proofs; however, the original papers will still be a reference, especially for historical motivation.

Theorem 5.1: Sylow's first theorem

Sylow subgroups of G exists.

Proof:

We will prove this theorem using strong induction. The base case is trivial

Case 1. P divides $Z(G)$

By Cauchy's theorem $Z(G)$ has a supgroup of order p.
That supgroup N is automatically normal

This means we can form the quotient group $\bar{G}$ with

$$|\bar{G}| = \left|\frac{G}{N}\right| = \frac{|G|}{|N|}$$

$$= \frac{p^\alpha \cdot m}{p} = p^{\alpha-1} \cdot m$$

Which is of order less than G which means the theorem holds on it by the inductive hypothesis.

Meaning it has supgroup $\bar{P}$ of order $p^{\alpha-1}$

By the correspondence theorem there exists a supgroup P in G such that

$$\frac{P}{N} = \bar{P}$$

$$\rightarrow |P| = |\bar{P}| \cdot |N| = p^{\alpha-1} \cdot p = p^\alpha$$

Case 2. $p \nmid o$.

Consider the class equation

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

We know that there must be at least one subgroup $H = C_G(a_j)$ for some j that has all the powers of p because

$$|G| = |H| \cdot |G : C_G(a_j)|$$

divisible by p^α ↑ ; must be divisible by p^α ↓ ; not divisible by p^α ↓

and hence H has order less than H by the inductive hypothesis the theorem holds for it and it has a subgroup of order p^α □

Examples:

- S_3 (order 6) has subgroups of orders 2 and 3
  $\{e, (1)\}$, $\{e, (123), (132)\}$

- A_4 (order 12) has subgroups of orders 4 and 3.
  $\{e, (123), (132)\}$, $\{e, (12)(34), (13)(24), (14)(23)\}$.

Remarks: A group of order p^α with $\alpha \geq 1$ is called p-groups.

· For the rest of the section we will denote the set of all sylow subgroups with $Syl(G)$

Now we will investgate properties of sylow subgroups.

Lemma 5.1

Let P be a sylow subgroup. If Q is any p-subgroup of G, then $Q \cap N_G(P) = Q \cap P$.

Proof:

Clearly $Q \cap P \leq Q \cap N_G(P)$

By definition $Q \cap N_G(P) \leq Q$, then it remains to show that $Q \cap N_G(P) \leq P$.

Since $Q \cap N_G(P) = H \leq N_G(P)$, then PH is a subgroup and

$$|PH| = \frac{|P||H|}{|P \cap H|}$$

all of these orders are powers of p making PH a p-subgroup that contains P which implies that $P \leq PH$ which shows that

$$Q \cap N_G(P) \leq P$$

□

Lemma 5.2

Let $S = \{P_1, P_2, \ldots, P_r\}$ be the set of all [illegible] and Q be a p-subgroup, then

$$r \equiv 1 \pmod{p}$$

Proof:

Let Q acts on S by conjugation, so we can write S as the disjoint union of this action

$$S = \mathcal{O}_1 \cup \mathcal{O}_2 \cup \cdots \cup \mathcal{O}_s$$

Notice that

$$|\mathcal{O}_i| = |Q \cap N_G(P_i)|$$

$$\rightarrow |\mathcal{O}_i| = |Q \cap P_i|$$

Now let $Q = P_1$ which gives

$$|\mathcal{O}_1| = 1$$

and

$$|\mathcal{O}_i| = |P_1 : P_1 \cap P_i| > 1 \qquad 2 \le i \le s$$

and since P_1 is a p-group this must be a power of p, so

$$p \mid |\mathcal{O}_i| \qquad 2 \le i \le s$$

then

$$r = 1 + \sum_{i=2}^{s} |\mathcal{O}_i| \equiv 1 \pmod{p} \qquad \square$$

Theorem 5.2: Sylow's second theorem

Any two sylow subgroups are conjugate.

Proof:

Let Q be any p-group and suppose for the sake of contradiction that $Q \neq gP_ig^{-1}$ for all i.

Considering the action of Q on S, then

$$p \mid |\mathcal{O}_i| \quad \text{for all } i$$

Which contradicts the fact that we proved that

$$r \equiv 1 \pmod{p}$$

Hence $Q \le gP_ig^{-1}$ for some i.

If we let Q be a sylow p-subgroup it turns out that for some g

$$gPg^{-1} = Q \qquad \square$$

Theorem 5.3: Sylow's third theorem

The number of all sylow p-subgroups is

$$n_p = |G : N_G(P)| \equiv 1 \pmod{P}$$

Proof:

By the second theorem all sylow subgroups are conjugate. This follows immediately. $\square$