LECTURE 1 : SYMMETRIES & EQUATIONS, MOTIVATIONS

Abstract algebra's origin story would seem to most naturally fall to the story of polynomial equations and our struggle to solve them. We begin with the quadratic eq^n,

$$f(x) = x^2 + bx + c = \underbrace{(x + b/2)}_{y}{}^2 \underbrace{- b^2/4 + c}_{-d}$$

$$g(y) = y^2 - d = 0 \quad \Rightarrow y^2 = d$$

$$\Rightarrow y = \pm\sqrt{d}$$

$$\Rightarrow x + b/2 = \pm\sqrt{b^2/4 - c}$$

$$\Rightarrow \boxed{x = \frac{-b \pm \sqrt{b^2 - 4c}}{2}}$$

The substitution $y = x + b/2$ depresses the quadratic and allows us to solve using radicals. The discriminant d is useful to categorize the type of solution.

For $b, c \in \mathbb{Q}$ we see no rational solution exists unless d is a square. If $d < 0$ then no real solution exists. Completing the square is rather old and initially it was tied to geometric problems which made lack of real solutions physically plausible.

Def^n/ with two indeterminants X_1 & X_2 we define symmetric functions $S_1 = X_1 + X_2$ and $S_2 = X_1 X_2$.
The general quadratic has form $f(x) = (x - X_1)(x - X_2)$

page 610 in Dummit & Foote.

$$f(x) = (x - X_1)(x - X_2)$$
$$= x^2 - (X_1 + X_2)x + X_1 X_2$$
$$= x^2 - S_1 x + S_2$$

$$d = (X_2 - X_1)^2$$
$$= X_2^2 - 2X_1 X_2 + X_1^2$$
$$= (X_1 + X_2)^2 - 4X_1 X_2$$
$$= S_1^2 - 4S_2$$

Remark: this formulation which focuses abstractly on roots X_1 & X_2 and symmetric functions is due to Lagrange.

CUBIC EQUATION

The solution to $x^3 + ax^2 + bx + c = 0$ is far more complicated than the quadratic case. First we should note the general problem can be reformulated to the depressed cubic

$$f(x) = x^3 + ax^2 + bx + c \qquad x = y - a/3$$

$$= (y - a/3)^3 + a(y - a/3)^2 + b(y - a/3) + c$$

$$= y^3 + py + q = g(y)$$

Notice $f(y - a/3) = g(y)$ thus $g(y_0) = 0 \Rightarrow f(y_0 - a/3) = 0$.

We find y_0 is root of $g(y) \Leftrightarrow x_0 = y_0 - a/3$ is root of $f(x)$.

Remark: $p = \frac{1}{3}(3b - a^2)$, $q = \frac{1}{27}(2a^3 - 9ab + 27c)$

this seems like a relaxing homework exercise.

Suppose α, β, γ are the solutions of $y^3 + py + q = 0$.

Def'n/ The discriminant of $(x - x_1)(x - x_2)(x - x_3)$ is given by $(x_1 - x_2)^2(x_1 - x_3)^2(x_2 - x_3)^2$ then the discriminant $D = (\alpha - \beta)^2(\alpha - \gamma)^2(\beta - \gamma)^2$

It turns out that $D = -4p^3 - 27q^2$

p. 630 - 631 of Dummit & Foote give some details of this derivation which is based on Lagrange's theory of resolvents.

CARDANO'S FORMULAS

To solve $y^3 + py + q = 0$ construct

$$D = -4p^3 - 27q^2$$

$$A = \sqrt[3]{-\frac{27}{2}q + \frac{3}{2}\sqrt{-3D}}$$

$$B = \sqrt[3]{-\frac{27}{2}q - \frac{3}{2}\sqrt{-3D}}$$

subject $AB = -3p$

Th^m/ $\alpha = \frac{A+B}{3}$, $\beta = \frac{\rho^2 A + \rho B}{3}$, $\gamma = \frac{\rho A + \rho^2 B}{3}$

where $\rho = \frac{-1}{2} + \frac{1}{2}\sqrt{-3}$ yield solutions of the depressed cubic $y^3 + py + q = 0$

E1 $y^3 - 8 = 0$

$p = 0$, $q = -8$ $\Rightarrow$ $D = -4(0)^3 - 27(-8)^2$ $\Rightarrow$ $\sqrt{-3D} = 72$

$$A = \sqrt[3]{\frac{-27}{2}(-8) + \frac{3}{2}(72)} = 3\sqrt[3]{8}$$

$$B = \sqrt[3]{\frac{-27}{2}(-8) - \frac{3}{2}(72)} = \sqrt[3]{4(27) - 3(36)} = 0$$

Then $AB = 0 = -3p$ ✓. Hence,

$$\alpha = \frac{A+B}{3} = \sqrt[3]{8} = 2.$$

$$\beta = \frac{\rho^2 A}{3} = 2e^{4\pi i/3} = 2e^{-2\pi i/3}$$

$$\gamma = \frac{\rho A}{3} = 2e^{2\pi i/3}$$

↔ conjugate sol^n/s.

Notice, $y^3 - 8 = (y-2)(y - 2e^{2\pi i/3})(y - 2e^{4\pi i/3})$

E1 continued

$$y^3 - 8 = (y-2)\left(y - 2e^{2\pi i/3}\right)\left(y - 2e^{-2\pi i/3}\right)$$

$$= (y-2)\left(y^2 - 2\left(e^{2\pi i/3} + e^{-2\pi i/3}\right)y + 4\right)$$

$$= (y-2)\left(y^2 - 4\cos(2\pi/3)y + 4\right)$$

$$= (y-2)(y^2 + 2y + 4)$$

$y^2 + 2y + 4$: irreducible, no real roots

Remark: obviously the usual algebra is way easier than using Cardano's formulas.

E2 Consider $y^3 - 4y = 0$

$p = -4, \quad q = 0$

$$D = -4p^3 - 27q^2 = -4(-4)^3 - 27(0)^2 = 256$$

$$\sqrt{-3D} = 16i\sqrt{3}$$

$$A = \sqrt[3]{\frac{3}{2}\sqrt{-3D}} = \sqrt[3]{24i\sqrt{3}} = \sqrt[3]{24\sqrt{3}}\, e^{\pi i/6}$$

$$B = \sqrt[3]{\frac{-3}{2}\sqrt{-3D}} = \sqrt[3]{-24i\sqrt{3}} = \sqrt[3]{24\sqrt{3}\, e^{-\pi i/2}} = \sqrt[3]{24\sqrt{3}}\, e^{-\pi i/6}$$

$$AB = \sqrt[3]{24\sqrt{3}}\sqrt[3]{24\sqrt{3}}\, e^{\pi i/6} e^{-\pi i/6} = \sqrt[3]{24\cdot 24\cdot 3} = 12$$

This is good, we need $-3p = -3(-4) = 12 = AB$.

$$\alpha = \frac{A+B}{3} = \frac{\sqrt[3]{24\sqrt{3}}}{3}\left(e^{\pi i/6} + e^{-\pi i/6}\right)$$

$$= \frac{2}{3}\sqrt[3]{24\sqrt{3}}\cos(\pi/6)$$

$$= \frac{2}{3}\sqrt[3]{24\sqrt{3}}\,\frac{\sqrt{3}}{2} = 2.$$

E2 continued

$$\beta = \frac{\rho^2 A + \rho B}{3} = \frac{1}{3}\sqrt[3]{24\sqrt{3}}\left(e^{4\pi i/3}e^{\pi i/6} + e^{2\pi i/3}e^{-\pi i/6}\right)$$

$$= \frac{1}{3}\sqrt[3]{24\sqrt{3}}\left(e^{3\pi i/2} + e^{\pi i/2}\right)$$

$$= \frac{1}{3}\sqrt[3]{24\sqrt{3}}\left(-i + i\right)$$

$$= 0.$$

$$\gamma = \frac{\rho A + \rho^2 B}{3} = \frac{1}{3}\sqrt[3]{24\sqrt{3}}\left(e^{2\pi i/3}e^{\pi i/6} + e^{4\pi i/3}e^{-\pi i/6}\right)$$

$$= \frac{1}{3}\sqrt[3]{24\sqrt{3}}\left(e^{5\pi i/6} + e^{7\pi i/6}\right)$$

$$= \frac{1}{3}\sqrt[3]{24\sqrt{3}}\left(e^{-\pi i/6} + e^{\pi i/6}\right)e^{\pi i}$$

$$= 2(-1).$$

We've found $y^3 - 4y = 0$ has sol'ns $y = 0, 2, -2$ which is hopefully unsurprising, $\beta \quad \alpha \quad \gamma$

$$y^3 - 4y = y(y^2 - 4) = y(y-2)(y+2).$$

E3 A nice example with both $p, q \neq 0$ is given by $y^3 - 7y + 6 = 0$. It turns out that $D = -4p^3 - 27q^2 = 400$. After some calculation, I found $\frac{A+B}{3} = -3$ and I'll let you find α, β, γ for this problem in the homework.

QUARTIC EQUATIONS (Ferrari in 1540's)

Given $f(x) = x^4 + bx^3 + cx^2 + dx + e$ we can substitute $y = x - b/4$ to obtain the depressed quartic $g(y) = y^4 + qy^2 + ry + s$. Following Descartes ~1637 (see Rotman's Advanced Modern Algebra, 2nd Ed. pg. 3) *

$$g(y) = y^4 + qy^2 + ry + s = (y^2 + jy + l)(y^2 - jy + m)$$

This leads to the resolvent cubic

$$(j^2)^3 + 2q(j^2)^2 + (q^2 - 4s)j^2 - r^2 = 0$$

You can solve this and hence solve any quartic. Let us review some historical tidbits (from *)

letters for variables : Viète in 1591
= sign : Recorde in 1557
exponents : Hume in 1636
+ and − : Widman 1486

} Rotman says this is from Cajori, "A History of Mathematical Notation"

//

Special cases of cubic : Scipio del Ferro ≈ 1515
cubic in general : Niccolò Fontana (Tartaglia) 1535
Girolamo Cardano 1545
quartic : Ferrari ≈ 1540's.

Then from 1500's to about 1800 many others searched for a sol'n to the quintic or higher polynomial eq'n.

1770: Lagrange showed reasonable substitutions produced resolvent which must solve higher than 5th order problem for the quintic.

1799 : Ruffini gave proof with gaps that the quintic could not generally be solved by radicals

1815 : Cauchy introduces multiplication of permutations and proved basic facts about symmetric group S_n (introduced cycle notation) (disjoint cycle factorization etc.)

1824 : Abel proved not all quintics are solved by radicals, his proof involved permutations of the roots of the quintic

1829 : GALOIS solved the quintic problem in the sense he gave a method to decide which quintics could be solved by radicals. This was natural extension of work of Lagrange and Abel. He coined the term "group" as subset of S_n closed under composition.

Ideas already present in Galois,

- conjugation
- normal subgroup
- quotient groups
- simple groups

Proved a polynomial equation (over $\mathbb{Q}$) could be solved via radicals iff the Galois group of the polynomial is solvable

Remark: it is funny that Galois Theory requires so much background, yet it is arguably the origin of abstract alg.

Def^n/ given indeterminants $x_1, x_2, \ldots, x_n$ we define

elementary *symmetric* *functions*

$$s_1 = x_1 + x_2 + \cdots + x_n$$
$$s_2 = x_1x_2 + x_1x_3 + \cdots + x_2x_3 + x_2x_4 + \cdots + x_{n-1}x_n$$
$$\vdots$$
$$s_n = x_1x_2 \cdots x_n$$

Def^n/ the general polynomial of degree n is

$$(x - x_1)(x - x_2)\cdots(x - x_n)$$

whose roots are the indeterminants $x_1, x_2, \ldots, x_n$

The coefficients of the general polynomial are given by the elementary symmetric functions upto a sign. Ultimately we'd like to find formulas for $x_1, x_2, \ldots, x_n$ in terms of the coefficients of the given polynomial. Let's play with $n=3$;

$$(x-x_1)(x-x_2)(x-x_3) = \left(x^2 - (x_1+x_2)x + x_1x_2\right)(x - x_3)$$
$$= x^3 - (x_1+x_2+x_3)x^2 + (x_1x_2 + x_1x_3 + x_2x_3)x - x_1x_2x_3$$
$$= x^3 - s_1x^2 + s_2x - s_3$$

Generally,

$$(x-x_1)(x-x_2)\cdots(x-x_n) = x^n - s_1x^{n-1} + s_2x^{n-2} + \cdots + (-1)^n s_n$$

I'll skip the derivation of the cubic formula, my goal here is to share a few highlights of Lagrange's program to investigate finding an analog to the quadratic eq^n for order $n \geq 5$. If you read the details you learn that Lagrange reproduced results of earlier mathematicians with a systematic method focused on exploiting symmetries of the formulas with respect to <u>permutations</u> of roots.

SYMMETRIES AND GROUPS

By the mid 19th century Cayley had abstracted the group concept beyond permutations and about the same time non-Euclidean geometry was discovered. The context of Euclidean geometry with well-defined distance, angle and parallelism gave way to wild examples where ∥-lines intersect and the familar theorems of Euclidean geometry were twisted or lost all together. It was as a responce to this situation that Felix Klein put forth a program to generalize the idea of Galois Theory: roughly analysis of structure preserved by a group to geometry.

ERLANGEN PROGRAM (1872, FELIX KLEIN)

characterize geometries based on group theory and projective geometry

less structure ↑

PROJECTIVE GEOMETRY

AFFINE GEOMETRY

EUCLIDEAN GEOMETRY

↓ more structure

In Euclidean Geometry, only Euclidean transformations were permissable, In Affine geometry just affine trans., In projective geometry just projective trans.

In each geometry, particular objects are preserved under the transformations of the geometry, but perhaps not for less structured geometries.

Remark: look up "Klein geometries" to see how this ERLANGEN PROGRAM was implimented via a pair (G, H) where G is a group and H a subgroup. The space $\mathbb{X} = G/H$ allowed transformations by G (the symmetry group of $\mathbb{X}$). For example,

1.) Affine Space $A(n) \cong \mathbb{R}^n$

$$G = Aff(n) \cong \mathbb{R}^n \rtimes GL(n, \mathbb{R})$$
$$H = GL(n, \mathbb{R})$$

2.) Euclidean Space $E(n)$

semi-direct product

$$G = Euc(n) \cong \mathbb{R}^n \rtimes O(n, \mathbb{R})$$
$$H = O(n, \mathbb{R})$$

The idea of preserving structure through certain preferred mappings is a pattern we see again and again in math. Klein's program has been reinvented multiple times in every generation which followed.

We say an object has a "symmetry" if there is some set S which maps to itself under some collection of transformations

Def'n/ Let S be a nonempty set, then a bijection on S is a permutation of S.

Euclidean Geometry

We'll approach this topic using analytic geometry for the n-dim'l space $\mathbb{R}^n$ where we define

$$X \cdot Y = X_1 Y_1 + X_2 Y_2 + \cdots + X_n Y_n \quad \text{(dot-product)}$$

$$\|X\| = \sqrt{X \cdot X} \quad \text{(vector length or norm)}$$

$$d(X,Y) = \|Y - X\| \quad \text{(distance from } X \text{ to } Y\text{)}$$

The set of all possible permutations of $\mathbb{R}^n$ is huge. Only a select subset is of interest to Euclidean Geom.

Def^n/ If $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ preserves the distance between points then φ is an <u>isometry</u>. That is

φ an isometry $\iff \|\varphi(P) - \varphi(Q)\| = \|P - Q\| \quad \forall P, Q \in \mathbb{R}^n$

I'll leave the proof of the Th^m below to the reader,

Th^m/ Let $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be an isometry

(1.) φ preserves dot-products $\iff \varphi(0) = 0$

That is, $\varphi(X) \cdot \varphi(Y) = X \cdot Y \ \forall X, Y \in \mathbb{R}^n \iff \varphi(0) = 0$.

(2.) If $\theta = \cos^{-1}\left(\frac{X \cdot Y}{\|X\| \, \|Y\|}\right) = \angle(X,Y)$ for $X, Y \neq 0$

then $\angle(P,Q) = \angle(\varphi(P), \varphi(Q)) \quad \forall P, Q \in \mathbb{R}^n$

It is easy to show translations and orthogonal transformations are isometries of Euclidean space.

Def^n/ $T_a : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ with $T_a(X) = X + a$ is the <u>translation</u> by $a \in \mathbb{R}^n$. Likewise, if $R \in \mathbb{R}^{n \times n}$ and $R^T R = I$ then $R \in O(n)$ and $L_R(X) = RX$ defines the orthogonal transformation $L_R : \mathbb{R}^n \longrightarrow \mathbb{R}^n$

The set of all bijections of $\mathbb{R}^n$ forms a group with respect to composition $\mathrm{Perm}(\mathbb{R}^n)$

Th^m $\mathrm{Isom}(\mathbb{R}^n) = \{\varphi : \mathbb{R}^n \to \mathbb{R}^n \mid \varphi \text{ an isometry}\}$ is a subgroup of $\mathrm{Perm}(\mathbb{R}^n)$. Moreover, $\mathrm{Trans}(\mathbb{R}^n) = \{T_a \mid a \in \mathbb{R}^n\}$ and $\mathrm{Orth}(\mathbb{R}^n)$, the set of all orthogonal transformations are both subgroups of $\mathrm{Isom}(\mathbb{R}^n)$; $\mathrm{Trans}(\mathbb{R}^n), \mathrm{Orth}(\mathbb{R}^n) \leq \mathrm{Isom}(\mathbb{R}^n)$. Furthermore, for each $\varphi \in \mathrm{Isom}(\mathbb{R}^n)$ there exist $T_a \in \mathrm{Trans}(\mathbb{R}^n)$ and $L_R \in \mathrm{Orth}(\mathbb{R}^n)$ for which $\varphi = L_R \circ T_a$

Proof: the Th^m above says translations and orthogonal transformations are all we need to construct any distance-preserving map. One step in the argument is to establish the following characterization of $\mathrm{Orth}(\mathbb{R}^n)$

Lemma: If $\varphi \in \mathrm{Isom}(\mathbb{R}^n)$ then $\varphi \in \mathrm{Orth}(\mathbb{R}^n) \Leftrightarrow \varphi(0) = 0$.

Proof of this Lemma is a good exercise.

Remark: I'll probably make parts of the proof homework. The more subtle aspect I'll probably leave to the reader, it is shown in ARTIN OR O'neill's "Elementary Differential Geometry" "ALGEBRA" (CHAPTER 6)

∃ a video where I go through the details.

Isometries of $\mathbb{R}^2$ and $\mathbb{R}^3$ are especially interesting as they are integral to the formulation of planar or solid Euclidean geometry. We can demonstrate isometries send circles to circles, lines to lines, and a given geometric shape to that shape once more.

Def^n/ Let $\Omega \subseteq \mathbb{R}^n$. The symmetry group of Ω is defined via

$$\Sigma(\Omega) = \{\varphi \in \mathrm{Isom}(\mathbb{R}^n) \mid \varphi(\Omega) = \Omega\}$$

E4 Let $\Omega_n \subseteq \mathbb{R}^2$ be the regular, unit-side-length polygon with $n \geq 3$ sides then $D_n = \Sigma(\Omega_n)$.
(See LECTURE 5 of my Math 421 notes for more exposition on this if you wish)

(Remark: $D_n = \langle a, b \mid |a| = n, |b| = 2, bab = a^{-1} \rangle$ is almost always a better way to calculate in D_n)

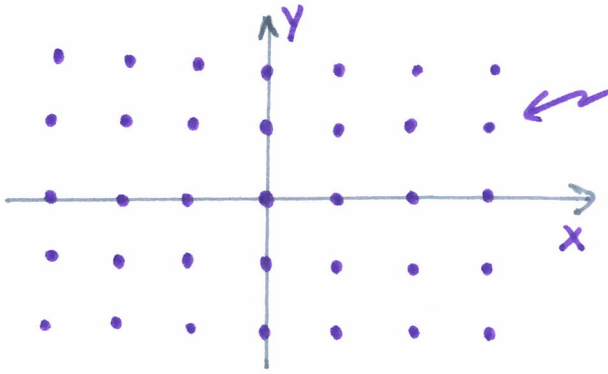
Th^m (Corollary 6.4.10) (ARTIN, p. 167 of "ALGEBRA" 2nd Ed)
Given G a finite subgroup of $\mathrm{Isom}(\mathbb{R}^2)$, if coordinates are suitably chosen, G becomes the cyclic group or dihedral group for appropriate order.

ARTIN goes on to categorize the discrete (like $\mathbb{Z}_n$ or D_n) but infinite subgroups of $\mathrm{Isom}(\mathbb{R}^2)$ and he shows $\exists$ 17 planar cyrstallographic groups (see the picture from p. 174)

Remark: the term discrete has a topological/analytical origin, or, in terms of manifold theory we can say H is discrete if $\dim_{\mathbb{R}}(H) = 0$. The points in H can be separated into disjoint neighborhoods. Certainly H finite $\Rightarrow$ H discrete, but $\exists$ infinite discrete sets.

E5 Consider $\mathbb{Z} \subseteq \mathbb{R}$ then $\mathbb{Z}$ gives discrete subgroup of $\mathbb{R}$ under $+$.

E6 Consider $\mathbb{Z}^2 \subseteq \mathbb{R}^2$, once more $\mathbb{Z}^2$ gives discrete subgroup of $\mathbb{R}^2$ w.r.t. $+$.



$\leftarrow \mathbb{Z} \times \mathbb{Z}$, the unit-lattice

In some sense, $\mathbb{Z}$ and $\mathbb{Z}^2$ are subgroups of $\mathrm{Isom}(\mathbb{R}^2)$, but, to be precise, they are isomorphic to particular discrete subgroups of $\mathrm{Isom}(\mathbb{R}^2)$. I'll try to write some homework to explore the content of this claim.

Remark: $\dim(\mathrm{Isom}(\mathbb{R}^n)) = \dim(\mathrm{Orth}(\mathbb{R}^n)) + \dim(\mathrm{Trans}(\mathbb{R}^n))$
so, for example, $\dim(\mathrm{Isom}(\mathbb{R}^2)) = 1+2 = 3$ whereas $\dim(\mathrm{Isom}(\mathbb{R}^3)) = 3+3 = 6$.

In conclusion, $Isom(\mathbb{R}^n)$ has a wealth of stories to tell, you can read whole books on the details. Let us conclude with a few cryptic comments about what comes next,

- Klein Geometries or Homogeneous Spaces
- CARTAN Geometries, enter the connection which describes how a given model geometry is locally distorted, for examples
  - Riemannian manifold is deformation of Euclidean Space
  - Lorentzian manifold is deformation of Minkowski Space
  - Conformal manifold is deformation of Conformal Sphere
  - Manifold with affine connection is deformation of Affine Space

In each case we use a group of allowed transformations to describe the structure of the space. For Klein Geometries there is a global symmetry whereas the introduction of curvature for the connection allows the way the group acts on space to vary from point to point. This is made precise with the mathematics of Principle Fiber Bundles. Anyway, there is much more to say, but next we go a different path and focus on how using the construction of a group action allows a more robust description of how groups act on sets (obviously I'm not testing manifold theoretic jibber jabber in here ☺)