Instructor: Leo Goldmakher

Williams College
Department of Mathematics and Statistics

MATH 394 : GALOIS THEORY

Solution Set 1

1.1 In class we observed that complex conjugation 'commutes' with addition and multiplication, in the sense that $\overline{x+y} = \overline{x} + \overline{y}$ and $\overline{xy} = \overline{x} \cdot \overline{y}$ for any $x, y \in \mathbb{C}$.

(a) Prove that complex conjugation commutes with all four operations $+, -, \times, \div$. (We asserted it for the first two operations, but didn't prove it for any of them.)

We will show that complex conjugation commutes with $+$, $-$, $\times$, $\div$. Given any two complex numbers, $a+bi$ and $c+di$ we will show that $\overline{(a+bi) \star (c+di)} = \overline{a+bi} \star \overline{c+di}$ where $\star$ is each of the four basic operations.

$$\overline{(a+bi)+(c+di)} = \overline{(a+c)+(b+d)i} = (a+c)-(b+d)i = (a-bi)+(c-di) = \overline{a+bi} + \overline{c+di}$$

$$\overline{(a+bi)-(c+di)} = \overline{(a-c)+(b-d)i} = (a-c)-(b-d)i = (a-bi)-(c-di) = \overline{a+bi} - \overline{c+di}$$

$$\begin{aligned}\overline{(a+bi)\cdot(c+di)} &= \overline{(ac-bd)+(ad+bc)i} = (ac-bd)-(ad+bc)i \\ &= (a-bi)\cdot(c-di) = \overline{a+bi}\cdot\overline{c+di}\end{aligned}$$

$$\begin{aligned}\overline{\left(\frac{a+bi}{c+di}\right)} &= \frac{\overline{(ac+bd)+(bc-ad)i}}{c^2+d^2} = \frac{(ac+bd)-(bc-ad)i}{c^2+d^2} \\ &= \frac{(a-bi)(c+di)}{(c-di)(c+di)} = \frac{a-bi}{c-di} = \frac{\overline{a+bi}}{\overline{c+di}}\end{aligned}$$

Thus we see that complex conjugation commutes with each of the four basic operations.

(b) Prove that complex conjugation commutes with the functions exp() and sin(). [*Hint: how can one define these functions meaningfully for complex inputs? Taylor series! Don't stress about convergence.*]

We will prove that complex conjugation commutes with exp() and sin(). Notice that we can express each function by its Taylor series expansion,

$$\sin(z) = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots \qquad e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots$$

which is in terms of only the four basic operations. As we have shown that complex conjugation commutes with these, it follows that it commutes with both Taylor series above, and thus with the functions $\sin(z)$ and e^z.

(c) Can you construct a function f and a choice of $z \in \mathbb{C}$ such that $f(\mathbb{R}) \subseteq \mathbb{R}$ but $\overline{f(z)} \neq f(\overline{z})$?

Define $f : \mathbb{C} \to \mathbb{C}$ by

$$f(z) = \begin{cases} 1 & z = i \\ -1 & z \neq i \end{cases}$$

Notice that $\overline{f(i)} = \overline{1} = 1$. However $f(\overline{i}) = f(-i) = -1$. Thus $\overline{f(z)} \neq f(\overline{z})$.

1.2 Show that for any $a, b \in \mathbb{Q}$ such that $\sqrt{b} \notin \mathbb{Q}$, the two numbers $a \pm \sqrt{b}$ are algebraically indistinguishable over $\mathbb{Q}$. [*Hint: start by proving that* $\pm\sqrt{b}$ *are algebraically indistinguishable.*]

First we will show that $\pm\sqrt{b}$ are algebraically indistinguishable over $\mathbb{Q}$ by showing that for any polynomial over $\mathbb{Q}$ for which $\sqrt{b}$ is a root, so is $-\sqrt{b}$. Given some function $f \in \mathbb{Q}[x]$, we can write $f(x) = g(x^2) + xh(x^2)$ for $g, h \in \mathbb{Q}[x]$. If $\sqrt{b}$ is a root of f, then $0 = f(\sqrt{b}) = g(b) + \sqrt{b}h(b)$. If $h(b) \neq 0$, then $\sqrt{b} = -\frac{g(b)}{h(b)} \in \mathbb{Q}$, contradicting our hypothesis. Thus, $h(b) = 0$, whence $g(b) = 0$. It follows that $f(-\sqrt{b}) = g(b) - \sqrt{b}h(b) = 0$, as claimed. The same argument shows that if $-\sqrt{b}$ is a root of f, then $\sqrt{b}$ must be as well.

Next we will show that this implies that $a \pm \sqrt{b}$ are algebraically indistinguishable over $\mathbb{Q}$. Suppose for contradiction that there was a polynomial $f \in \mathbb{Q}[x]$ that distinguishes between them. Then exactly one of $f(a + \sqrt{b})$ and $f(a - \sqrt{b})$ is zero. Define $g(z) = f(z - a)$. Then $g(\sqrt{b}) = f(a + \sqrt{b})$ and $g(-\sqrt{b}) = f(a - \sqrt{b})$. Thus exactly one of $g(\sqrt{b})$ and $g(-\sqrt{b})$ is zero, which means that g distinguishes between $\pm\sqrt{b}$. However since $g(z) = f(z - a)$, we see that $g \in \mathbb{Q}[x]$, so this is a contradiction as we previously showed that $\pm\sqrt{b}$ were algebraically indistinguishable.

1.3 Prove that complex conjugates are algebraically indistinguishable over $\mathbb{R}$.

Since $\overline{\overline{z}} = z$, it suffices to show that if z is a root for some $f \in \mathbb{R}[x]$ then $\overline{z}$ is also a root. Notice that for $f \in \mathbb{R}[x]$, since these are constructed out of the basic four operations which we showed commute with complex conjugation, then $\overline{f(z)} = f(\overline{z})$. Thus if $f(z) = 0$, then $f(\overline{z}) = \overline{f(z)} = \overline{0} = 0$ so then $\overline{z}$ is also a root of f. Thus there is no polynomial with real coefficients for which a complex number is a root and its conjugate is not.

1.4 The goal of this problem is to prove the following assertion from lecture:

Claim. *The only choice of rational numbers* a, b, c *satisfying* $a + b\sqrt[3]{2} + c(\sqrt[3]{2})^2 = 0$ *is* $a = b = c = 0$. *(In other words,* $1, \sqrt[3]{2}, (\sqrt[3]{2})^2$ *are linearly independent over* $\mathbb{Q}$*.)*

Define

$$S := \{(a, b, c) \in \mathbb{Z}^3 : a + b\sqrt[3]{2} + c(\sqrt[3]{2})^2 = 0\}.$$

We call the element $(0, 0, 0)$ the *trivial* element of S.

(a) Prove that S forms a group under addition.

Since $S \subseteq \mathbb{Z}^3$, which is a group, it is sufficient to show that for all $x, y \in S$, $x - y \in S$ to show that S is a subgroup. Consider some $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ in S. In $\mathbb{Z}^3$, we know that the additive inverse of y, $-x$ is $(-y_1, -y_2, -y_3)$. So we need to show that

$$x - y = x + (-y) = (x_1, x_2, x_3) + (-y_1, -y_2, -y_3) = (x_1 - y_1, x_2 - y_2, x_3 - y_3) \in S.$$

Since $x, y \in S$, we know both of them are solutions to the initial equation and so

$$(x_1 - y_1) + (x_2 - y_2)\sqrt[3]{2} + (x_3 - y_3)\sqrt[3]{2}^2 = \left(x_1 + x_2\sqrt[3]{2} + x_3\sqrt[3]{2}^2\right) - \left(y_1 + y_2\sqrt[3]{2} + y_3\sqrt[3]{2}^2\right)$$
$$= 0 - 0 = 0$$

so $x - y$ is also a solution to the equation and so $x - y \in S$.

(b) Prove that if $(x, y, z) \in S$ is nontrivial, then $xyz \neq 0$.

To do this we will show that if any of x, y, or z are zero, then $x = y = z = 0$; in showing this we will have shown that either every element is zero or they are all nonzero. If (x, y, z) is not the trivial element, then they are not all zero, and thus they will all be nonzero. The product of three nonzero integers will be nonzero. Observe that given the equation $a + b\alpha = 0$ for some $\alpha \notin \mathbb{Q}$, the only integer solution is $a = b = 0$. We will use this to show that if x, y, or z is zero, then all of them are zero. There are three natural cases to consider here.

- $x = 0$
  In this case, we have that $x + y\sqrt[3]{2} + z\sqrt[3]{2}^2 = y\sqrt[3]{2} + z\sqrt[3]{2}^2 = \sqrt[3]{2}(y + z\sqrt[3]{2}) = 0$ dividing by $\sqrt[3]{2}$, we see that $y + z\sqrt[3]{2} = 0$. Since $\sqrt[3]{2}$ is irrational, by the lemma $y = z = 0$.
- $y = 0$
  In this case, we have that $x + y\sqrt[3]{2} + z\sqrt[3]{2}^2 = x + z\sqrt[3]{2}^2 = 0$. Since $\sqrt[3]{2}^2$ is irrational, by the lemma $x = z = 0$.
- $z = 0$
  Here we have that $x + y\sqrt[3]{2} + z\sqrt[3]{2}^2 = x + y\sqrt[3]{2} = 0$. Since $\sqrt[3]{2}$ is irrational, by the lemma $x = y = 0$.

In each case we saw that $x = y = z = 0$, thus if any one is zero then they are all zero. Thus if (x, y, z) is not trivial, then x, y, and z are all nonzero, or else they would all be zero and thus the trivial element. If they are all nonzero, then $xyz \neq 0$.

(c) Prove that S only contains the trivial element.

Suppose $(x, y, z) \in S$ is nontrivial, i.e.

$$x + y\sqrt[3]{2} + z\sqrt[3]{2}^2 = 0. \tag{$\dagger$}$$

Multiplying through by $\sqrt[3]{2}$, we deduce

$$(2z, x, y) \in S.$$

If instead we square both sides of ($\dagger$), we find

$$(x^2 + 4yz, 2z^2 + 2xy, y^2 + 2xz) \in S.$$

Since S is a group and $x, y, z \in \mathbb{Z}$, we can take linear combinations of these three elements:

$$(0, xy - 2z^2, y^2 - xz) = x(x, y, z) + 2y(2z, x, y) - (x^2 + 4yz, 2xy + 2z^2, 2xz + y^2) \in S.$$

Part (b) implies this element must be the trivial element, whence

$$xy = 2z^2 \qquad \text{and} \qquad xz = y^2.$$

Multiplying the former by z and the latter by y, we conclude

$$y^3 = 2z^3.$$

Since $z \neq 0$ we deduce $y/z = \sqrt[3]{2}$. This is impossible, since the left hand side is rational while the right hand side is not. Thus our choice of nontrivial $(x, y, z) \in S$ must not be possible, i.e. S is the trivial group.

(d) Prove the claim. (*Careful! There's something to check here.*)

Note that S consists of triples of *integers*, while the claim concerns triples of *rationals*. Fortunately, it's easy to get from one to the other. Indeed, suppose

$$a + b\sqrt[3]{2} + c\sqrt[3]{2}^2 = 0 \qquad (\ddagger)$$

for some $a, b, c \in \mathbb{Q}$. Pick positive integers m, n, ℓ such that $ma, nb, \ell c \in \mathbb{Z}$. Multiplying $(\ddagger)$ through by $mn\ell \neq 0$ yields $(amn\ell, bmn\ell, cmn\ell) \in S$. But we proved that S is trivial, whence $amn\ell = bmn\ell = cmn\ell = 0$. It follows that $a = b = c = 0$, as claimed.

1.5 Given a group G and two subsets $A, B \subseteq G$, we define the *commutator of A and B*, denoted $[A, B]$, to be the group *generated* by all the elements $\{[a, b] : a \in A, b \in B\}$, where $[a, b] := aba^{-1}b^{-1}$. This problem does not assume prior experience with these objects, so **please do not look up any information on commutators.** However, you may look up the basic notions from group theory: subgroup, normal subgroup, quotient group, isomorphism, even and odd permutations.

(a) Prove that for any permutations $\sigma, \tau \in S_n$, the permutation $[\sigma, \tau]$ is even.

Given $\sigma, \tau \in S_n$, we know we can write each as a product of transpositions, say

$$\sigma = \sigma_1\sigma_2\cdots\sigma_m \qquad \text{and} \qquad \tau = \tau_1\tau_2\cdots\tau_\ell$$

with all the σ_i and τ_i denoting transpositions. It's straightforward to verify that

$$\sigma^{-1} = \sigma_m\sigma_{m-1}\cdots\sigma_1 \qquad \text{and} \qquad \tau^{-1} = \tau_\ell\tau_{\ell-1}\cdots\tau_1.$$

Thus $[\sigma, \tau] = \sigma\tau\sigma^{-1}\tau^{-1}$ can be written as a product of $2(m + \ell)$ transpositions, hence is an even permutation.

(b) Show that $[S_3, S_3] \simeq \mathbb{Z}_3$, and that $[\mathbb{Z}_3, \mathbb{Z}_3]$ is trivial. [*Hint: If you use part (a) you get to be lazy!*]

Recall that A_3 is the set of all even permutations of S_3; note that $A_3 \leq S_3$.

Lemma. $A_3 = \{(), (1\ 2\ 3), (1\ 3\ 2)\}$
Proof. Since $A_3 \leq S_3$, Lagrange's theorem implies that the order of A_3 is a divisor of 6. Also, A_3 isn't all of S_3 (because, e.g., $(1\ 2) \in S_3 \setminus A_3$), so $|A_3| \leq 3$. On the other hand, we can find three elements of A_3: we see that $(1\ 2\ 3) = (1\ 3)(1\ 2) \in A_3$, whence A_3 must also contain $(1\ 2\ 3)^2 = (1\ 3\ 2)$ as well as the identity $()$. □

From part (a), all of the generators of $[S_3, S_3]$ are in A_3, whence $[S_3, S_3] \leq A_3$. Finally, observe that $[(2\ 3), (1\ 2)] = (1\ 2\ 3)$, which is a generator of A_3. Thus, $[S_3, S_3] = A_3$. Since $\mathbb{Z}_3$ is the unique group of order 3 (up to isomorphism), we deduce that $[S_3, S_3] \simeq \mathbb{Z}_3$ as claimed.

Since $\mathbb{Z}_3$ is cyclic, it must be abelian, whence the commutator of any two elements is trivial. This shows that $[\mathbb{Z}_3, \mathbb{Z}_3]$ is trivial.

(c) Given a finite group G, consider the set $\{N \trianglelefteq G : G/N \text{ is abelian}\}$. (Recall that $N \trianglelefteq G$ means that N is a normal subgroup of G.) Prove that the smallest element of this set is $[G, G]$.

Given a finite group G, let $A := \{[a, b] : a, b \in G\}$. By definition, $[G, G]$ is the collection of all finite products of elements of A and their inverses. But observe that $[a, b]^{-1} = [b, a]$; this means that $[G, G]$ is simply the set of all finite products of elements of A, or in symbols

$$[G, G] = \Big\{ \prod_i [a_i, b_i] : a_i, b_i \in G \Big\}.$$

A short computation shows that $g[a, b]g^{-1} = [gag^{-1}, gbg^{-1}]$ whence

$$g\big(\prod_i [a_i, b_i]\big)g^{-1} = \prod_i g[a_i, b_i]g^{-1} = \prod_i [ga_ig^{-1}, gb_ig^{-1}] \in [G, G].$$

This shows that $[G, G] \trianglelefteq G$.

Next we show that a normal subgroup $N \trianglelefteq G$ has an abelian quotient group G/N if and only if $[G, G] \subseteq N$. Observe that for any $a, b \in G$ we have

$$\big[[a], [b]\big] = \big[[a, b]\big]$$

where $[a]$ denotes the element $aN \in G/N$. It follows that G/N is abelian if and only if $[e] = \big[[a, b]\big]$ for every $a, b \in G$. Furthermore,

$$\big[[a, b]\big] = [e] \iff [a, b] \in N.$$

Thus we see that G/N is abelian if and only if $[a, b] \in N$ for all $a, b \in G$. Since by definition $[G, G]$ is the smallest group containing all elements of the form $[a, b]$, we deduce that

$$G/N \text{ is abelian} \implies [G, G] \leq N.$$

The reverse implication is straightforward, so we see that

$$G/N \text{ is abelian} \iff [G, G] \leq N.$$

In particular, we see that $G/[G, G]$ is abelian.

Summing up, we've proved

- $[G, G] \trianglelefteq G$,
- $G/[G, G]$ is abelian, and
- G/N is abelian if and only if $[G, G] \leq N$.

This proves the claim.

1.6 The following questions refer to the write-up on Arnold's theorem.

(a) Solve part (a) of Exercise 1 from the write-up.

Let $g(z) := z^{1/3}$. Suppose γ is a loop in $\mathbb{C}$ which starts and ends at some point $p \in \mathbb{C}$. Then the image of γ under g starts at one of the cube roots of p. At any point z in the loop γ we have three choices for $g(z)$, but observe that as soon we choose the initial position for the image, there is a unique choice of $g(z)$ that renders the image continuous for every z in the loop γ.

Where does the image of γ terminate? Certainly it must terminate at a cube root of p. To keep track of this we introduce a function W from the space of all loops based at p to $\{1, \omega, \omega^2\}$, defined by $W(\gamma) = \omega^k$ iff the image of γ terminates at $\omega^k\alpha$, where α denotes the initial position of the image.[a] Below we shall make use of the following result:

Lemma. $W(\gamma^{-1})W(\gamma) = 1$.

Taking on faith the Lemma, we consider the image under g of a commutator loop $[\gamma_1, \gamma_2]$, where both loops γ_i are based at p. Let's say the initial position of the image is α. After traversing γ_1 we end up at $W(\gamma_1)\alpha$. From there we traverse γ_2, which leaves us at $W(\gamma_2)W(\gamma_1)\alpha$; then γ_1^{-1}, depositing us at $W(\gamma_1^{-1})W(\gamma_2)W(\gamma_1)\alpha$; and finally γ_2^{-1}, leaving us at $W(\gamma_2^{-1})W(\gamma_1^{-1})W(\gamma_2)W(\gamma_1)\alpha$. Since multiplication in $\mathbb{C}$ is commutative, we see that the image of the commutator loop terminates at

$$W(\gamma_2^{-1})W(\gamma_1^{-1})W(\gamma_2)W(\gamma_1)\alpha = W(\gamma_2^{-1})W(\gamma_2)W(\gamma_1^{-1})W(\gamma_1)\alpha = \alpha,$$

the initial position of the image. We therefore conclude that the image of any commutator loop under g must itself be a loop.

It only remains to prove the Lemma. We parametrize γ by writing

$$\gamma(t) = r(t)e^{i\theta(t)},$$

where t varies from 0 to 1, $\theta(t) \in \mathbb{R}$, and $r(t) \geq 0$. Since γ is a loop, we see that $r(0)e^{i\theta(0)} = r(1)e^{i\theta(1)}$. From this we deduce that $e^{i\left(\theta(1)-\theta(0)\right)} \in \mathbb{R}_{\geq 0}$, which immediately implies that $\theta(1) = \theta(0) + 2\pi n$ for some $n \in \mathbb{Z}$, and thence that $r(0) = r(1)$. As a consequence, we see that

$$W(\gamma) = \omega^n.$$

Given our parametrization of γ, it's not too hard to derive a parametrization of γ^{-1}:

$$\gamma^{-1}(t) = r(1-t)e^{i\theta(1-t)}$$

where t varies from 0 to 1. Since we know from above that $\theta(1) = \theta(0) + 2\pi n$, we deduce $W(\gamma^{-1}) = \omega^{-n}$. This concludes the proof of the Lemma. □

[a] Although not necessary for this problem, you can learn more about this type of function by looking up *winding number*.

(b) Solve exercise 4 from the write-up.

At the start of Arnold's proof, one picks a specific separable quintic p. One then forms several special loops based at p; for the proof to go through, we need a quintic formula which applies to all of the (infinitely many!) points on these loops! In other words, Arnold's approach implies the non-existence of any finite formula built out of the coefficients, radicals, and continuous functions which applies to every polynomial lying in these special loops.

(c) In view of Arnold's proof, explain what problem **1.5**(b) implies about the cubic formula.

In Arnold's proof, the fact that $[S_3, S_3]$ was nontrivial implied that any general cubic formula must contain nested radicals. If $[[S_3, S_3], [S_3, S_3]]$ were also nontrivial, this would imply:

> *Hypothetical Result.* Any general cubic formula built solely out of radicals, continuous functions, and the coefficients of a cubic polynomial must contain doubly-nested radicals.

However, since $[[S_3, S_3], [S_3, S_3]]$ is trivial by **1.5**(b), Arnold's approach doesn't imply the HR! (Note that Arnold's does not disprove the HR either—the only information it yields is that any general cubic formula requires at least one level of nesting of radicals.)