ALGEBRA 2 HONORS: GALOIS THEORY

DAVID SMYTH

1. Polynomial Equations: High School Approach

1.1. **Solving polynomial equations.** Most of modern algebra was constructed in order to come to grips with the following problem: Given a polynomial

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n,$$

how can we write down a number α such that $f(\alpha) = 0$. For concreteness, let's think of $a_0, \ldots, a_n$ as rational numbers.

Linear Equations $(n = 1)$. I know that you know how to solve a linear equation, but humor me. Consider the equation:

$$f(x) = a_0x + a_1 = 0.$$

We can divide through by a_0, then subtract a_1/a_0 from both sides.

$$\begin{aligned} x + \frac{a_1}{a_0} &= 0 \\ x &= -\frac{a_1}{a_0} \end{aligned}$$

Even though this is completely trivial, I would like to make two observations.

- If a_0 and a_1 are rational numbers, then our solution $-a_1/a_0$ is still a rational number. Similarly, if a_0 and a_1 are real numbers, then our solution is still a real number. You probably remember from high school that simple quadratic polynomials sometimes have complex solutions, but nothing like that happens here. In modern terminology, we can say that *no field extension is necessary* in order to find solutions of linear equations.

- Our first bit of algebra - dividing through by the leading coefficient a_0 - is actually a completely general recipe for reducing an arbitrary polynomial to a *monic* polynomial, i.e. a polynomial with $a_0 = 1$. If we can solve monic polynomials, we can solve all polynomials. Thus, from now on, I'll simply assume our polynomial is monic to begin with.

Quadratic Equations $(n = 2)$.

$$f(x) = x^2 + a_1x + a_2 = 0$$

The basic trick here, probably known to most of you, is to make a substitution which cancels the a_1-term:

$$x = y - \frac{a_1}{2}$$

If we simply plug this in, we get

$$\begin{aligned}\left(y-\frac{a_1}{2}\right)^2+a_1(y-\frac{a_1}{2})+a_2&=0\\ \left(y^2-a_1y+\frac{a_1^2}{4}\right)+\left(a_1y-\frac{a_1^2}{2}\right)+a_2&=0\\ y^2+\left(a_2-\frac{a_1^2}{4}\right)&=0\\ y&=\pm\sqrt{\left(\frac{a_1^2}{4}-a_2\right)}\end{aligned}$$

Now that we have found y, we can go back and find x. We get

$$x=-\frac{a_1}{2}\pm\sqrt{\left(\frac{a_1^2}{4}-a_2\right)}$$

Once again, I would like to make two observations.

- It is no longer true that if our coefficients are rational, then we will necessarily have a rational solution. On the other hand, we can see precisely what we need to "add" in order to get solutions, namely $\sqrt{\left(\frac{a_1^2}{4}-a_2\right)}$. This is the only piece of the solution that might not be defined over the original field of definition. In modern terminology, we'll say that you always have a solution after passing to a degree two extension.
- Once again, the key trick here generalizes in an obvious way. Using a linear substitution of the form $x=y-a_1/n$, we can reduce an arbitrary degree-n equation to an equation satisfying $a_1=0$. We are making progress! Perhaps with enough tricks, we can solve any polynomial.

Cubic Equations $(n=3)$. Applying our previous two tricks, we may assume that we have an equation of the form:

$$f(x)=x^3+a_2x+a_3=0$$

Any ideas? As with the quadratic equations, it's easy to see that if we could cancel the a_2-term we'd be in business - we could simply take cube roots. The new trick is to recognize the possibility of a non-linear substitution:

$$x=y-\frac{a_2}{3y}$$

Let's see what happens when we plug this in:

$$\begin{aligned}\left(y-\frac{a_2}{3y}\right)^3+a_2\left(y-\frac{a_2}{3y}\right)+a_3&=0\\ \left(y^3-a_2y+\frac{a_2^2}{3y}-\frac{a_2^3}{27y^3}\right)+\left(a_2y-\frac{a_2^2}{3y}\right)+a_3&=0.\\ y^3+a_3-\frac{a_2^3}{27y^3}&=0\end{aligned}$$

Clearing denominators, we get:

$$y^6 + a_3 y^3 - \frac{a_2^3}{27} = 0.$$

In most cases, going from a degree 3 equation to a degree 6 equation would not be considered progress. But here (rather miraculously), we can recognize this equation as a quadratic equation in y^3, i.e. we can rewrite it as:

$$(y^3)^2 + a_3(y^3) - \frac{a_2^3}{27} = 0.$$

Thus, we can solve for y^3 using the quadratic formula:

$$y^3 = -\frac{a_3}{2} \pm \sqrt{\frac{a_3^2}{4} + \frac{a_2^3}{27}}.$$

Taking cube roots, we get:

$$y = \sqrt[3]{-\frac{a_3}{2} \pm \sqrt{\left(\frac{a_3}{2}\right)^2 + \left(\frac{a_2}{3}\right)^3}}.$$

Now we should be a little careful here - just as one needs to account for positive and minus square roots, one needs to account for all possible cube roots. Thus, one really has the following solutions:

$$y = \omega^i \sqrt[3]{-\frac{a_3}{2} \pm \sqrt{\left(\frac{a_3}{2}\right)^2 + \left(\frac{a_2}{3}\right)^3}}, i = 0, 1, 2,$$

where $\omega = e^{2\pi i/3}$ is a third root of unity. (The fact that we are suddenly using complex numbers here is a little unsettling - eventually we will understand roots of unity in a purely algebraic way without any complex numbers entering the picture.) Plugging these solutions back into the formula $x = y - \frac{a_2}{3y}$, we obtain solutions of our original equation.

Optional Aside. You might notice something a little fishy here: We apparently have six possible solutions for y, each of which should give a solution for x. But this would give six possible solutions for our original cubic! In fact, there is no contradiction here. With some elementary but tedious algebra one can check that these six different value y, break into three pairs, with each pair giving the same value for x. The final solutions for x are:

$$\sqrt[3]{-\frac{a_3}{2} + \sqrt{\left(\frac{a_3}{2}\right)^2 + \left(\frac{a_2}{3}\right)^3}} - \sqrt[3]{-\frac{a_3}{2} - \sqrt{\left(\frac{a_3}{2}\right)^2 + \left(\frac{a_2}{3}\right)^3}}$$

$$\omega\sqrt[3]{-\frac{a_3}{2} + \sqrt{\left(\frac{a_3}{2}\right)^2 + \left(\frac{a_2}{3}\right)^3}} - \omega^2\sqrt[3]{-\frac{a_3}{2} - \sqrt{\left(\frac{a_3}{2}\right)^2 + \left(\frac{a_2}{3}\right)^3}}$$

$$\omega^2\sqrt[3]{-\frac{a_3}{2} + \sqrt{\left(\frac{a_3}{2}\right)^2 + \left(\frac{a_2}{3}\right)^3}} - \omega\sqrt[3]{-\frac{a_3}{2} - \sqrt{\left(\frac{a_3}{2}\right)^2 + \left(\frac{a_2}{3}\right)^3}}$$

What lessons can we draw from the cubic? The good news is that we are beginning to see what it means to "write down a number α such that $f(\alpha) = 0$." What we would like to do is find some formula for α in terms of the given coefficients $a_0, a_1, \ldots, a_n$. Based on the pattern we've seen so far, we would expect this formula to involve nothing more than the usual operations of addition/subtraction, multiplication/division, and also taking roots.

But there's also some bad news:

- The algebraic complexity of this problem is rapidly ballooning.
- The trick we used with the cubic does not seem to generalize in the way that our first two tricks did. The first two tricks serve to reduce us to the equation.

  $$x^4 + a_2x^2 + a_3x + a_4$$

  But it is not at all clear whether there is a substitution of the form $x = f(y, 1/y)$ which puts this in a simpler form.

Evidently, there are problems with the high school approach - these algebraic tricks feel unmotivated and to check that they work one has to do many unilluminating calculations. We need a different perspective on this problem. The basic shift in emphasis is to recognize the connection between finding roots and factoring polynomials.

1.2. **From finding roots to factoring.** To see the connection between finding roots and factoring the polynomial, we begin with the following easy lemma. It says that finding a root α of $f(x)$ is the same as factoring $f(x)$ into $(x - \alpha)$ and a lower factor.

Lemma 1.1 (Remainder Theorem)**.** *Let k be a field, and let $f(x) \in k[x]$ be a polynomial. For any $\alpha \in k$, we can write*

$$f(x) = (x - \alpha)g(x) + f(\alpha),$$

where $g(x)$ is a polynomial of degree $n - 1$. In particular if $f(\alpha) = 0$, then $f(x)$ admits a factorization as $(x - \alpha)g(x)$.

Proof. Using the usual division algorithm for polynomials, just divide $f(x)$ by $(x - \alpha)$. We will get

$$f(x) = (x - \alpha)g(x) + c$$

where $c \in k$ is some constant. By plugging α into both sides of this equation, we see that $c = f(\alpha)$. □

If we use this factoring procedure inductively, we get two useful corollaries.

Corollary 1.2. *If $f(x) \in k[x]$ is a degree n polynomial, then f has at most n roots.*

Proof. If f has no roots, then there is nothing to prove, so we may assume that f has a root α. By the Remainder Theorem, we may factor f as

$$f(x) = (x - \alpha)g(x).$$

By induction on the degree of f, we may assume that g has no more than $n - 1$ roots. Since any root of f must be either a root of $(x - \alpha)$ (namely α) or a root of $g(x)$, it follows that $f(x)$ has no more than n roots. □

Corollary 1.3. *If $f(x) \in k[x]$ is a degree n polynomial with n distinct roots $\alpha_1, \dots, \alpha_n \in k$, then f can be factored as:*

$$f(x) = \prod_{i=1}^{n}(x - \alpha_i)$$

Proof. We can factor $f(x)$ as:

$$f(x) = (x - \alpha_1)g(x).$$

Since the roots are distinct $(\alpha_i - \alpha_1) \neq 0$ for all $i = 2, \dots, n$. Thus, $\alpha_2, \alpha_3, \dots, \alpha_n$ must be roots of $g(x)$. By induction on the degree of f, we may assume $g(x) = \prod_{i=2}^{n}(x - \alpha_i)$, and the desired result follows. □

Now let us assume for the time being that $f(x)$ actually has n distinct roots, so that we can factor

$$f(x) = \prod_{i=1}^{n} (x - \alpha_i)$$

Then we can view the roots α_i as variables, and the coefficients of the polynomial as giving equations involving these variables. At least in the case $n = 2$, this idea should be familiar from high school, i.e. if we write

$$(x - \alpha_1)(x - \alpha_2) = x^2 + a_1 x + a_2,$$

where a_1 and a_2 are given to begin with, then we see the the problem of finding the roots of f is just the same as finding two numbers α_1, α_2 such that

$$\begin{aligned} -(\alpha_1 + \alpha_2) &= a_1 \\ \alpha_1 \alpha_2 &= a_2. \end{aligned}$$

In the next lecture, we will generalize this system of equations to higher n.

2. Symmetric Functions

Definition 2.1 (Elementary Symmetric Functions). *Let $k[x_1, \ldots, x_n]$ be a polynomial ring in n variables. For $i = 1, \ldots, n$, we define the following special polynomials $s_i \in k[x_1, \ldots, x_n]$:*

$$\begin{aligned} s_1 &= x_1 + x_2 + \cdots + x_n \\ s_2 &= x_1 x_2 + x_1 x_3 + \ldots + x_{n-1} x_n \\ &\vdots \\ s_k &= \sum_{1 \le i_1 \le i_2 \le \ldots \le i_k \le n} x_{i_1} x_{i_2} \ldots x_{i_k} \\ &\vdots \\ s_n &= x_1 x_2 \cdots x_n \end{aligned}$$

In words, we can say that the k^{th} *symmetric function* is simply the sum of all degree k monomials with no repeated variables.

The point of this definition is that the functions s_i precisely encode the relationship between the roots of a polynomial and its coefficients. By some straight-forward high school algebra, you can check:

$$\prod_{i=1}^{n} (x - \alpha_i) = x^n - s_1(\alpha_1, \ldots, \alpha_n) x^{n-1} + s_2(\alpha_1, \ldots, \alpha_n) x^{n-2} - \ldots + (-1)^n s_n(\alpha_1, \ldots, \alpha_n),$$

This means that finding the roots of a given polynomial $f(x) = x^n + a_1 x^{n-1} + \ldots + a_n$ (at least under the assumption that $f(x)$ has n distinct roots - later we'll see this assumption is unnecessary) is precisely equivalent to finding $\alpha_1, \dots, \alpha_n$ which satisfy the following equations.

$$\begin{aligned}\alpha_1 + \ldots + \alpha_n &= a_1 \\ \alpha_1\alpha_2 + \ldots + \alpha_{n-1}\alpha_n &= -a_2 \\ &\vdots \\ \alpha_1\alpha_2 \cdots \alpha_n &= (-1)^n a_n\end{aligned}$$

In these equations, you should think of a_i as given, and α_i as being unknown numbers that you are trying to find.

We started off with a single equation in one variable, and now we have n equations in n variables. How on earth could this be any easier than the original problem? The point is that these are not just any random old equations; the elementary symmetric functions have very special properties that will make them easier to work with than arbitrary functions. For starters, they are symmetric. What exactly does that mean?

Definition 2.2. *Let S_n act on $k[x_1, \dots, x_n]$ by permuting the variables, i.e. $\sigma(f(x_1, \dots, x_n)) = f(x_\sigma(1), \dots, x_{\sigma(n)})$. We say that a function is* symmetric *if $\sigma(f) = f$.*

Example 2.3 ($n = 3$)**.** *If $\sigma = (123)$ is the cyclic permutation of 3 variables, then $\sigma(x_1^3x_2^2x_3) = x_2^3x_3^2x_1$. Evidently, $x_1^3x_2^2x_3$ is not a symmetric function. On the other hand, $x_1^3 + x_2^3 + x_3^3$ is a symmetric function.*

The elementary symmetric functions s_i are all symmetric. While there are many symmetric functions besides the elementary ones, it turns out that they are all generated as polynomial combinations of the elementary symmetric functions. This is an astounding fact!

Theorem 2.4 (Fundamental Theorem of Symmetric Functions)**.** *Let $f(x_1, \dots, x_n)$ be any symmetric polynomial. Then, f can be expressed as polynomial in the symmetric function, i.e $f = g(s_1, \dots, s_n)$ for some polynomial g.*

Example 2.5. *The theorem says that one can express $x_1^3+x_2^3+x_3^3$ as a polynomial in s_1, s_2, s_3. One can easily check that*

$$x_1^3 + x_2^3 + x_3^3 = s_1^3 - 3s_1s_2 + 3s_3$$

Is there a way to derive this formula systematically? Yes, there is - we shall spell out the algorithm in full detail when we prove the fundamental theorem, but it may be useful to sketch the idea informally in the context of this example. To begin with, it's easy to see that s_1^3, s_1s_2, s_3 are the only monomials in s_1, s_2, s_3 that give rise to degree 3 monomials in x_1, x_2, x_3, so these are the only monomials that can appear in our formula. In other words, we must have a formula like

$$x_1^3 + x_2^3 + x_3^3 = as_1^3 + bs_1s_2 + cs_3,$$

for some coefficients a, b, c, and the question is how to figure out these coefficients.

First, focus attention on the x_1^3 term. On the left, it occurs with coefficient 1. On the right, it's easy to see that only s_1^3 contains an x_1^3 term and it occurs with coefficient one. Thus, we must have $a = 1$.

Next, let's subtract s_1^3 from both sides, to get:

$$x_1^3 + x_2^3 + x_3^3 - (x_1 + x_2 + x_3)^3 = bs_1s_2 + cs_3,$$

If we expand out the left hand side, the x_1^3 term cancels, so let's examine the next lowest order term, i.e. $x_1^2x_2$. On the left, it occurs with coefficient -3. On the right, it's easy to see that only s_1s_2 contains an $x_1^2x_2$ term and it occurs with coefficient 1. Thus, we must have $b = -3$.

Next, let's add $3s_1s_2$ to both sides to get:

$$x_1^3 + x_2^3 + x_3^3 - (x_1 + x_2 + x_3)^3 + 3(x_1 + x_2 + x_3)(x_1x_2 + x_1x_3 + x_2x_3) = cs_3,$$

If we expand out the left hand side, we see that x_1^3 and $x_1^2x_2$ terms cancel - in fact everything cancels except the $x_1x_2x_3$ term which occurs with coefficient 3. Since $s_3 = x_1x_2x_3$, we must have $c = 3$, and we are done.

In the example, I made use of the idea of the "lowest" monomial without actually explaining what I meant. The key technical tool in proving the general theorem is the introduction of an ordering on monomials which makes this concept precise.

Definition 2.6 (Lexicographic Ordering)**.** *The* lexicographic ordering *is a total ordering on all degree m monomials in n variables, which can be defined as follows. Given any monomial, we can write it as $x_{i_1}x_{i_2}\ldots x_{i_m}$ with $i_1 \leq i_2 \leq \ldots \leq i_m$. In other words, we can write it as a product of m variables whose subscripts go from lowest to highest. To compare two monomials, we then just look at the first subscript were two monomials differ. More formally, we say that*

$$x_{i_1}x_{i_2}\ldots x_{i_m} < x_{j_1}x_{j_2}\ldots x_{j_m}$$

if $i_1 = j_1$, $i_2 = j_2$, $\ldots$, $i_{k-1} = j_{k-1}$ and $i_k < j_k$ for some $k \in \{1, \ldots, m\}$.

Example 2.7. *The lexicographic ordering for degree 3 monomials in 3 variables goes like this:*

$$x_1^3 < x_1^2x_2 < x_1^2x_3 < x_1x_2^2 < x_1x_2x_3 < x_1x_3^2 < x_2^3 < x_2^2x_3 < x_2x_3^2 < x_3^3$$

Definition 2.8. *If f is a homogeneous polynomial of degree m (homogeneous means that every monomial in f has the same degree), we let $L(f)$ be the "lowest" monomial of f, i.e. the monomial of f which is least with respect to the lexicographic ordering.*

Example 2.9. *If $f = 2x_1^2x_2 + x_1x_2x_3 + 3x_3^3$, then $L(f) = 2x_1^2x_2$, because $x_1^2x_2 < x_1x_2x_3 < x_3^3$ in the lexicographic ordering.*

If you think about it, you will see that certain monomials cannot occur as $L(f)$ for a symmetric function f. For example, $x_1x_2^2$ could never be the lowest monomial of a symmetric function. Why not? Because if f contains the monomial $x_1x_2^2$, then it must also (by symmetry) contain the monomial $x_1^2x_2$ and $x_1^2x_2 < x_1x_2^2$. More generally, we have the following lemma.

Lemma 2.10. *If f is a symmetric function, and $L(f) = cx_1^{k_1}x_2^{k_1}\ldots x_n^{k_n}$, then $k_1 \geq k_2 \geq \ldots \geq k_n$.*

Proof. Let f be a symmetric function with $L(f) = x_1^{k_1}x_2^{k_1}\ldots x_n^{k_n}$, and suppose the statement of the lemma fails, i.e. suppose that the k_i's are not ordered from largest to smallest. Let $\sigma \in S_n$ be a permutation that orders the k_i's correctly, i.e. such that

$$k_{\sigma(1)} \geq k_{\sigma(2)} \geq \ldots \geq k_{\sigma(n)}.$$

Since f is symmetric, f must contain the monomial $x_1^{k_{\sigma(1)}}x_2^{k_{\sigma(2)}}\ldots x_n^{k_{\sigma(n)}}$. By the definition of the lexicographic ordering, we have $x_1^{k_{\sigma(1)}}x_2^{k_{\sigma(2)}}\ldots x_n^{k_{\sigma(n)}} < x_1^{k_1}x_2^{k_1}\ldots x_n^{k_n}$. But this is a contradiction, since we started by assuming that $x_1^{k_1}x_2^{k_1}\ldots x_n^{k_n}$ was the lowest monomial in f. □

Now we are ready to prove the fundamental theorem of symmetric functions. The idea, as demonstrated in Example 2.5, is to focus on the lowest monomial of our symmetric function, and then find a monomial in the elementary symmetric functions which matches it. By successively subtracting off appropriate multiples of monomials in the symmetric functions, we can work our way up the lexicographic ordering until there are monomials left! At that point, we have expressed f as a polynomial in the symmetric functions.

Proof. First, we reduce to the case that f is homogenous. We claim that if we know the fundamental theorem for homogenous symmetric functions, then we know the fundamental theorem for all symmetric functions. To see this, let f be any symmetric function and write $f = f_1 + \ldots + f_m$, where each f_i is homogenous of degree i. If f is symmetric, then each f_i must be as well (because the action of S_n preserves the degree of each monomial of f). If we know the fundamental theorem for homogenous symmetric functions, then we can write each f_i as a polynomial in elementary symmetric function. But then we clearly get a representation of f as a polynomial in elementary symmetric functions as desired.

Now let f be a homogeneous symmetric function and let $L(f) = cx^{k_1} \ldots x^{k_n}$. By Lemma 2.10, we know that $k_1 \geq k_2 \geq \ldots \geq k_n$. We claim that there exists a monomial in the symmetric functions, say $cs_1^{l_1}s_2^{l_2}\ldots s_n^{l_n}$ such that

$$L(f) = L(cs_1^{l_1}s_2^{l_2}\ldots s_n^{l_n}).$$

To check this, we need to investigate the lowest monomials of the elementary symmetric functions. By the definition of the elementary symmetric functions, one easily checks that:

$$L(s_i) = x_1x_2\ldots x_i.$$

It follows that

$$L(cs_1^{l_1}s_2^{l_2}\ldots s_n^{l_n}) = cx_1^{l_1+l_2+\ldots+l_n}x_2^{l_1+l_2+\ldots+l_{n-1}}\ldots x_n^{l_n}.$$

Thus, in order to get $L(f) = L(cs_1^{l_1}s_2^{l_2}\ldots s_n^{l_n})$, we simply need to find non-negative integers $l_1, l_2, \ldots, l_n$ such that

$$\begin{aligned} l_1 + \ldots + l_n &= k_1 \\ l_1 + \ldots + l_{n-1} &= k_2 \\ &\vdots \\ l_n &= k_n. \end{aligned}$$

Happily, the condition $k_1 \geq k_2 \geq \ldots \geq k_n$ guarantees that we can do this. Indeed, we simply set $l_n = k_n$ and $l_i = k_i - k_{i+1}$ for $i = 1, \ldots, n-1$. With this choice of l_i, we have $L(f) = L(cs_1^{l_1}s_2^{l_2}\ldots s_n^{l_n})$ as desired.

Now we are basically done. If we let $f' := f - cs_1^{l_1}s_2^{l_2}\ldots s_n^{l_n}$, then f' is a symmetric function with $L(f') > L(f)$. Thus, we can simply replace f by f' and repeat this procedure. As we do this, we will subtract off multiples of monomials of the symmetric functions to get a sequence of functions $f, f', f'', \ldots$ with higher and higher lowest monomials. The only way this process can terminate is to have $f^k = 0$ for some k. At that point, we have an equation expressing f as a sum of monomials of elementary symmetric functions, i.e. a polynomial in the elementary symmetric functions. □

3. Lagrange's Solution to the Quartic

Lagrange actually proved the fundamental theorem on symmetric functions in the course of developing a more systematic approach to solving polynomial equations. In this lecture, we'll see how Lagrange made use of the theorem to give a solution to the general quartic equation. Lagrange's solution begins with the following observation, which tells you how to turn an arbitrary function into a symmetric function.

Lemma 3.1. *Let $f_1 \in k[\alpha_1, \ldots, \alpha_n]$ be any polynomial, and let $f_1, \ldots, f_k$ be the orbit of f_1 under the action of S_n, i.e. the set of all functions you get by acting on f_1 with elements of S_n. If $s(x_1, \ldots, x_k)$ is any symmetric function in k variables, then $s(f_1, \ldots, f_k)$ is a symmetric function in $\alpha_1, \ldots, \alpha_n$.*

Example 3.2. *Let $f_1 = (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4) \in k[\alpha_1, \alpha_2, \alpha_3, \alpha_4]$. One can easily check that the orbit of f_1 is:*

$$\begin{aligned} f_1 &= (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4) \\ f_2 &= (\alpha_1 + \alpha_3)(\alpha_2 + \alpha_4) \\ f_3 &= (\alpha_1 + \alpha_4)(\alpha_2 + \alpha_3) \end{aligned}$$

According to the theorem, any symmetric function of f_1, f_2, f_3 must give a symmetric function of the α_i. In particular, one can verify by inspection that $f_1 + f_2 + f_3, f_1 f_2 + f_1 f_3 + f_2 f_3, f_1 f_2 f_3$ are all symmetric in $\alpha_1, \ldots, \alpha_4$. It follows of course that they can be expressed as polynomials in the elementary symmetric functions, and indeed one can easily check:

$$\begin{aligned} f_1 + f_2 + f_3 &= 2s_2(\alpha_1, \ldots, \alpha_4) \\ f_1 f_2 + f_1 f_3 + f_2 f_3 &= (Exercise!) \\ f_1 f_2 f_3 &= (Exercise!) \end{aligned}$$

How does Lagrange use this to solve the quartic polynomial? Given an equation

$$f(x) = x^4 + a_2 x + a_3 + a_4 = 0,$$

Lagrange starts by assuming that $f(x)$ has 4 distinct roots $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. (It is not true of course that all degree 4 polynomials have 4 distinct roots - this extra assumption is a weakness of Lagrange's method, and later in the course we will have the technical machinery to get around it.) As we discussed in the last lecture, these roots must satisfy the four equations:

$$\begin{aligned} s_1(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 &= -a_1 = 0. \\ s_2(\alpha_1, \alpha_2, \alpha_3, \alpha_4) &= a_2 \\ s_3(\alpha_1, \alpha_2, \alpha_3, \alpha_4) &= -a_3 \\ s_4(\alpha_1, \alpha_2, \alpha_3, \alpha_4) &= a_4 \end{aligned}$$

Thus, we are given three complex numbers a_2, a_3, a_4, and we want to find four complex numbers $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ satisfying the above equations. Though we don't know what $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are, Lagrange begins by defining f_1, f_2, f_3 to be the expressions in Example 3.2. In other words, we have used four unknown complex numbers to define three more unknown complex numbers. The key point, however, is that we have a way to find formulas for f_1, f_2, f_3 in terms of a_2, a_3, a_4. Here comes Lagrange's stroke of genius - consider the cubic polynomial:

$$(x - f_1)(x - f_2)(x - f_3) = x^3 - (f_1 + f_2 + f_3)x^2 + (f_1 f_2 + f_1 f_3 + f_2 f_3)x - f_1 f_1 f_3$$

As we argued in Example 3.2, the coefficients of this polynomial are symmetric functions in $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. By the fundamental theorem of symmetric functions, we can therefore write the coefficients of this function as polynomials in a_2, a_3, a_4. But this means that we can use our solution for the cubic to solve for f_1, f_2, f_3 in terms of a_2, a_3, a_4! Needless to say, actually writing down the formula would be a rather tedious affair, but the key point is that we know one exists.

Once we have solved for f_1, f_2, f_3, it is easy to see that we can solve for $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. Indeed, using the equation

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 0,$$

we see that

$$f_1 = (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4) = -(\alpha_1 + \alpha_2)^2,$$

and we conclude that

$$\begin{aligned} \alpha_1 + \alpha_2 &= \sqrt{-f_1}, \\ \alpha_3 + \alpha_4 &= -\sqrt{-f_1}. \end{aligned}$$

Similarly, we have

$$\begin{aligned} \alpha_1 + \alpha_3 &= \sqrt{-f_2}, \\ \alpha_2 + \alpha_4 &= -\sqrt{-f_2}, \\ \alpha_1 + \alpha_4 &= \sqrt{-f_3}, \\ \alpha_2 + \alpha_3 &= -\sqrt{-f_3}. \end{aligned}$$

From here, one can solve for each of the α_i individually. For example,

$$\alpha_1 = \frac{(\alpha_1 + \alpha_2) + (\alpha_1 + \alpha_3) - (\alpha_2 + \alpha_3)}{2} = \frac{\sqrt{-f_1} + \sqrt{-f_2} + \sqrt{-f_3}}{2}.$$

We have solved the quartic!

Now, let us step back from the particulars of this example, and consider Lagrange's overall strategy. If we are trying to solve a polynomial of degree n, Lagrange's strategy is to find a function of the roots $f_1(\alpha_1, \ldots, \alpha_n)$ whose orbit $f_1, \ldots, f_k$ under S_n is less than n. Assuming we know how to solve equations of degree less than n, we can solve for $f_1, \ldots, f_k$ by considering the polynomial

$$\prod_{i=1}^{k} (x - f_i).$$

As in the example of the quartic, the fundamental theorem of symmetric functions implies that the coefficients of this polynomial will be polynomials in the original coefficients $a_1, \ldots, a_n$, so we can solve.

Of course, as someone pointed out in class, if we take the orbit of f_1 to be too small, e.g. if we simply take f_1 to be a symmetric function, then knowing the value of f_1 won't help us much in our quest to solve for the α_i's. The idea, therefore, is to find a function which has a small enough orbit that you can solve for it, but is also a useful bridge to solving for the α_i's. In fact, Lagrange spent the last years of his life looking for a function $f(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$ which would allow him to solve the quintic. Alas, we now know that solving the quintic is impossible.

Let us now step back and consider the overall idea of Galois Theory. The problem with polynomials is that they are really not very transparent. When we solved the cubic in Lecture 1, we found that we could essentially reduce the cubic equation to a quadratic equation. In some sense, therefore, there was a quadratic equation "hidden" inside the cubic equation. Similarly, Lagrange found a cubic equation "hidden" inside a quartic equation. We need to switch to a kind of mathematical structure in which this hidden structure becomes more transparent.

In the following weeks, we will see how to associate to any polynomial $f(x) \in K[x]$ a certain algebraic object L/K, called a field extension. While the field extension contains, in some sense, all the relevant information about f , it is much easier to deal with. For example, the fact that solving the cubic involves a hidden quadratic equation, will appear as the fact the associated field extension L/K has an intermediate extension associated to a quadratic equation, i.e. we have a picture like this:

$$\begin{tikzcd}
L \arrow[dd, dash, "x^3+a_2x+a_3"'] \arrow[dr, dash] & \\
& K' \arrow[dl, dash, "x^2+a_3x-(a_2^3)/27"] \\
K &
\end{tikzcd}$$

Thus, the problem of solving polynomials will be reduced to understanding the structure of field extensions, especially the problem of understanding all the intermediate field extensions of a given field extension. This problem, in turn, will turn out to be solvable in terms of pure group theory. Associated to a field extension L/K, we will define a certain group $G(L/K)$ called the Galois group of the extension. Subfields will correspond to subgroups/quotients of the Galois group, just as Lagrange's intermediate cubic equation was found by finding a homomorphism $S_4 \to S_3$. This will give us a very pretty answer to our original problem!