

DERIVING SOLUTIONS TO HIGHER ORDER POLYNOMIALS (WITH A MODERN AFTERWORD)

VICTOR WANG

ABSTRACT. Through a high school student's lens, we investigate the classical problem of relating a polynomial's roots to its coefficients and demonstrate how to derive the solution to the cubic and quartic using radicals, as well as the trigonometric solution to the real-rooted cubic. In the afterword, we discuss the history of the problem, as well as the history of how this work came to be.

Deriving solutions to higher order polynomials
(with a modern afterword)

Mathematics HL Internal Assessment

Victor Wang
000XXX-00XX

Session May 20XX

Ms. XXXXX

1 Introduction

1.1 Rationale

As a student of mathematics, I am always impressed by the accomplishments of mathematicians past. And yet, their work feels like an insurmountable peak. How did Euler come up with his famous identity, which so elegantly related five of the most important numbers in mathematics in the same equation? Or what was Gauss' insight, which allowed him to conquer the fundamental theorem of algebra? It scares me to think that, despite my mathematical prowess, my understanding is far from the understanding of the greats from centuries ago.

On the recommendation of a friend, I watched *The Man Who Knew Infinity*, a film about the brilliant mathematician Srinivasa Ramanujan. When I decided to read more about Ramanujan, I became truly in awe at his discoveries. But what set me aflame was learning that at just the age of fifteen, he developed his own technique to solve the general quartic. If he, at that age, could do it, then I, at seventeen, should be able to do so too. Was I really set out for a future in mathematics? I had just found my litmus test.

Beyond simply a personal challenge, the aim of this paper is to practice expressing and communicating these ideas both rigorously and in a way that is accessible to students. It is my hope that this paper is able to be understood by dedicated readers with an advanced high school understanding of algebra. With personal derivations of both old and new solutions, the motivations for each solution can help readers understand how and why each solution was discovered.

1.2 Definitions

Definition 1.1 (Polynomial). “A mathematical expression involving a sum of powers in one or more variables multiplied by coefficients” (Weisstein).

Definition 1.2 (Symmetric polynomial). A polynomial $f(x_1, x_2, \dots, x_n)$ over n variables such that swapping any two x_i and x_j ($1 \leq i < j \leq n$, $i, j \in \mathbb{N}$) returns the same polynomial. By this definition, a polynomial over a single variable is trivially a symmetric polynomial.

Definition 1.3 (Elementary symmetric polynomial). The elementary symmetric polynomials in n variables r_1, r_2, \dots, r_n , written $e_k(r_1, r_2, \dots, r_n)$ for

$k = 0, 1, \dots, n$ are defined by $e_k(r_1, r_2, \dots, r_n) = \sum_{1 \leq j_1 < j_2 < \dots < j_k \leq n} (r_{j_1} r_{j_2} \dots r_{j_k})$.

Example 1.3.1. The first, second, and third elementary symmetric polynomials on r_1, r_2, r_3 are $r_1 + r_2 + r_3$, $r_1 r_2 + r_1 r_3 + r_2 r_3$, and $r_1 r_2 r_3$, respectively.

Theorem 1.4 (Fundamental theorem of symmetric polynomials). *Any symmetric polynomial can be expressed as a polynomial in the elementary symmetric polynomials on those variables.*

Theorem 1.4 is offered without proof, because its difficulty is outside of the scope of this paper. For interested readers, the simplest proof I could find is one by induction described by mathematician Leonid Lerner in his 2011 paper.

Theorem 1.5 (Vieta's formulas). *The coefficient of x^k in $a_n x^n + a_{n-1} x^{n-1} + \dots + a_0 x^0$, which has roots r_1, r_2, \dots, r_n , is equal to $(-1)^{n-k} a_n$ multiplied by the $(n - k)$ th elementary symmetric polynomial $e_{n-k}(r_1, r_2, \dots, r_n)$.*

Proof. This is clear by expanding the polynomial $a_n(x - r_1)(x - r_2) \dots (x - r_n)$ and adding like terms. \square

Example 1.5.1. Expanding the cubic $a_n(x - r)(x - s)(x - t)$ gives $a_n x^3 - a_n(r + s + t)x^2 + a_n(rs + rt + st)x - a_n rst = a_n e_0(r, s, t)x^3 - a_n e_1(r, s, t)x^2 + a_n e_2(r, s, t)x - e_3(r, s, t)$, as desired.

Lemma 1.6. *Knowing the values of the coefficients of a single variable polynomial is sufficient information to find the value of any symmetric polynomial over the original polynomial's roots.*

Proof. This follows directly from applying Theorem 1.5 and Theorem 1.4 in conjunction. \square

2 Quadratic

Here, we will offer an elementary derivation of the quadratic formula and seek to find properties and patterns that we may be able to use in solving the cubic or the quartic.

The solutions to the quadratic $ax^2 + bx + c = 0$ can be found easily by completing the square. We have $ax^2 + bx + c = a((x + \frac{b}{2a})^2 + (\frac{c}{a} - \frac{b^2}{4a^2})) = 0$. That is, $(x + \frac{b}{2a})^2 = \frac{b^2 - 4ac}{4a^2}$. Finally, $x + \frac{b}{2a} = \pm \frac{\sqrt{b^2 - 4ac}}{2a}$, which gives us the well-known formula $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$. With only a few steps, we have proved the following theorem.

Theorem 2.1 (Quadratic formula). *The solutions to the quadratic expression $ax^2 + bx + c = 0$ can be found by*

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

What sorts of insights were obtained from this? How can we use what we have learned here to tackle the cubic and the quartic? For one, we learned the idea of “completing the square”. Applying such an initial substitution to a polynomial to remove the second term is known as converting a polynomial to its depressed form. The most common derivation of the solutions to a cubic involves converting the cubic to its depressed form, and applying another substitution from there. This solution will be explored in Section 3.1.1.

It is also helpful to consider what each term actually is in terms of the roots. If we call the roots of the quadratic r and s (that is, $ax^2 + bx + c = a(x - r)(x - s)$), and substitute each term of the quadratic formula with an expression of r and s by Vieta’s formulas, we find that the formula is telling us that $x = \frac{r+s}{2} \pm \frac{\sqrt{r^2 - 2rs + s^2}}{2}$. In other words... the solutions are $x = r$ and $x = s$!

Of course, this result should be as expected; if the solutions to $a(x - r)(x - s) = 0$ were not $x = r$ and $x = s$, I would expect that I have made an error somewhere. From this result, I have made a couple of observations.

Firstly, as per Lemma 1.6, we know that any symmetric polynomial can be found from a statement of the initial polynomial. It is also obvious that any polynomial of the elementary symmetric polynomials (which can be found from the coefficients of the initial polynomial by Vieta’s) will be symmetric. How is it, then, that the quadratic formula results in the *asymmetric* (over r, s) expressions of r and s ?

Interestingly, the symmetric polynomial $r^2 - 2rs + s^2$ is formed first. The asymmetry appears when the square root of that expression is taken, which gives us the two distinct and asymmetric expressions $r - s$ and $s - r$. In a similar fashion, we will attempt to solve the cubic by forming a symmetric polynomial over the three roots, and then applying a cube root to the expression to form three distinct asymmetric expressions. Hopefully that will give us enough information to solve for the individual roots of the cubic.

The second observation I made was again with regards to the idea of symmetry. The quadratic formula (which we determined to be equivalent to $x = \frac{r+s}{2} \pm \frac{r-s}{2}$) expresses the two solutions in one expression by means of a

“centre of symmetry” at $\frac{r+s}{2}$, the arithmetic mean of the two roots. Perhaps we could extend this property to the cubic with three roots (r , s , and t) by starting from a central pivot at the arithmetic mean $\frac{r+s+t}{3}$ of the three roots and adding an expression with three possible values to arrive at an expression for each root.

3 Cubic

3.1 Radical form

3.1.1 By “completing the cube”

The most well-known solution to the cubic involves “completing the cube”. That is, if we wish to solve $ax^3 + bx^2 + cx + d = 0$, we first apply the substitution $y = x + \frac{b}{3a}$. Note then that $ax^3 + bx^2 + cx + d = a(x^3 + \frac{b}{a}x^2 + \frac{c}{a}x + \frac{d}{a}) = a((x + \frac{b}{3a})^3 + (\frac{3ac-b^2}{3a^2})(x + \frac{b}{3a}) + (\frac{27a^2d+2b^3-9abc}{27a^3}))$.

To solve for $y^3 + py + q = 0$, where $p = \frac{3ac-b^2}{3a^2}$ and $q = \frac{27a^2d+2b^3-9abc}{27a^3}$, we must think of an even cleverer substitution. Our motivation in the second substitution involves a desire to reduce the expression to a quadratic. In particular, we may hope that the substitution $y = z + \frac{k}{z}$ for a well-chosen expression k will simplify the expression to one in terms of z^3 , a constant term, and z^{-3} , upon which we would be able to apply the quadratic formula to solve for z^3 .

Expanding $(z + \frac{k}{z})^3 + p(z + \frac{k}{z}) + q$ gives us $z^3 + 3kz + \frac{3k^2}{z} + \frac{k^3}{z^3} + pz + \frac{kp}{z} + q$. Perhaps surprisingly, setting $k = -\frac{p}{3}$ cleanly cancels out the z and z^{-1} terms! That is, the substitution $y = z - \frac{p}{3z}$ leaves us with $z^3 + q - \frac{p^3}{27z^3} = 0$. Multiplying both sides by z^3 , and we obtain $z^6 + qz^3 - \frac{p^3}{27} = 0$, a quadratic in z^3 .

Solving by the quadratic formula, we have $z^3 = \frac{-q \pm \sqrt{q^2 - \frac{4p^3}{27}}}{2}$. It can be shown algebraically that whether we take the positive or negative root of z^3 , we will obtain the same value for $y = z - \frac{p}{3z}$. So, let $z = \sqrt[3]{\frac{-q + \sqrt{q^2 - \frac{4p^3}{27}}}{2}}$. Here, the cube root implies all three possible values (some of which are complex). Substituting z back into y and y back into x , we obtain the roots $x = z - \frac{p}{3z} - \frac{b}{3a}$.

Theorem 3.1 (Cubic formula). *The solutions to the cubic equation $ax^3 + bx^2 + cx + d = 0$ satisfy*

$$x = z - \frac{p}{3z} - \frac{b}{3a},$$

where $p = \frac{3ac - b^2}{3a^2}$, $q = \frac{27a^2d + 2b^3 - 9abc}{27a^3}$, and z is any of the three solutions to $z^3 = \frac{-q + \sqrt{q^2 - \frac{4p^3}{27}}}{2}$.

3.1.2 By Lagrange resolvents

We would like to exploit Lemma 1.6 somehow. Can we find some symmetric expression that we can find from the coefficients of our cubic, then apply the cube root to “break” the symmetry and obtain three asymmetric expressions?

In solving for the quadratic, we found that $r^2 - 2rs + s^2$ was a symmetric polynomial that could be obtained, by Lemma 1.6. Then, the square root operation gave two different expressions: $r - s$ and $s - r$. To the observant reader, $r - s$ is equal to $r + s \operatorname{cis} \pi$! Why should this substitution be useful?

Note that $\operatorname{cis} \pi$ is a primitive second root of unity. That is, $z = \operatorname{cis} \pi$ satisfies $z^2 = 1$ and does not satisfy $z^k = 1$ for any positive integer $k < 2$. This tells us that $(r + s \operatorname{cis} \pi)^2 = (r + s \operatorname{cis} \pi)^2 \operatorname{cis}^2 \pi = (r \operatorname{cis} \pi + s)^2$! Though the expressions $r + s \operatorname{cis} \pi$ and $r \operatorname{cis} \pi + s$ are distinct, they yield the same value when squared. And in this case, that squared value is symmetric over r, s !

This seems like a good lead towards a solution to the cubic. In fact, in Galois theory, these expressions are known as *Lagrange resolvents*. Since we would like to tackle the cubic, we start with the Lagrange resolvent $r + s \operatorname{cis} \frac{2\pi}{3} + t \operatorname{cis} \frac{4\pi}{3}$, where r, s , and t represent the three roots of the cubic. The cube root of the symmetric polynomial $(r + s \operatorname{cis} \frac{2\pi}{3} + t \operatorname{cis} \frac{4\pi}{3})^3$ leads to the three distinct expressions $r + s \operatorname{cis} \frac{2\pi}{3} + t \operatorname{cis} \frac{4\pi}{3}$, $r \operatorname{cis} \frac{2\pi}{3} + s \operatorname{cis} \frac{4\pi}{3} + t$, and $r \operatorname{cis} \frac{4\pi}{3} + s + t \operatorname{cis} \frac{2\pi}{3}$.

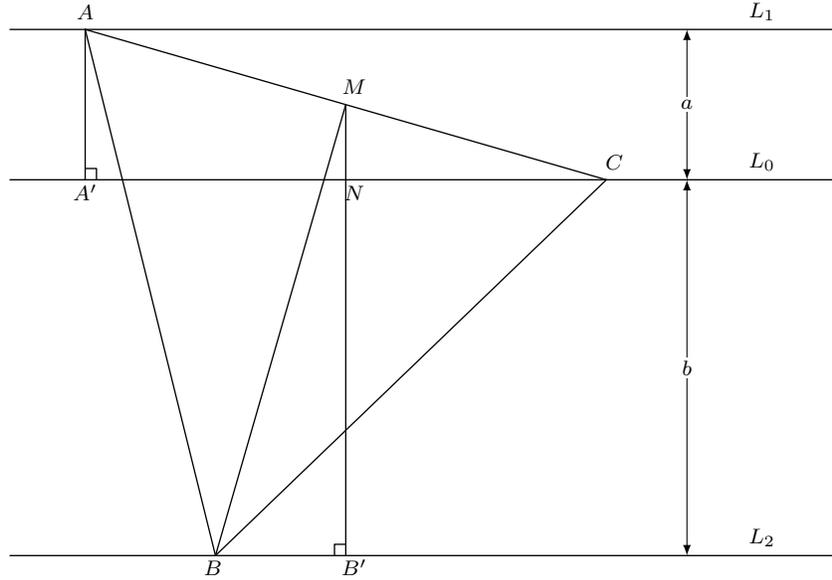
Everything looks good, except... $(r + s \operatorname{cis} \frac{2\pi}{3} + t \operatorname{cis} \frac{4\pi}{3})^3$ is not symmetric. This can be seen by expanding. It is, in fact, distinct from $(s + r \operatorname{cis} \frac{2\pi}{3} + t \operatorname{cis} \frac{4\pi}{3})^3$, so we cannot easily obtain the expression from the information given. Back to the drawing board for us.

3.2 Trigonometric solution

Lemma 3.2. *Given three parallel lines, the middle of which is distance a and b away from the other two, and an equilateral triangle whose vertices lie on the three lines, the sidelength s of the triangle is $\sqrt{\frac{4}{3}(a^2 + ab + b^2)}$.*

Proof. Without loss of generality, let $a \leq b$. Let the middle line be L_0 , the line a away from L_0 be L_1 , and the line b away from L_0 be L_2 . Label the vertices of the triangle A , B , and C , with A on L_1 , B on L_2 , and C on L_0 .

Let M be the midpoint of AC . Let A' be the point on L_0 such that $AA' \perp L_0$ and B' be the point on L_2 such that $MB' \perp L_2$. Let N be the intersection of MB' and L_0 .



Then, $\angle BMB' = \angle BMC - \angle B'MC = 90^\circ - \angle NMC = \angle ACA'$. Because $\angle ACA' = \angle BMB'$ and $\angle AA'C = \angle BB'M = 90^\circ$, $\triangle ACA' \sim \triangle BMB'$. Thus, $A'C = B'M \left(\frac{AC}{BM} \right) = (MN + B'N) \left(\frac{s}{s\sqrt{3}} \right) = \left(\frac{a}{2} + b \right) \left(\frac{2}{\sqrt{3}} \right)$.

By the Pythagorean Theorem, $AA'^2 + A'C^2 = AC^2$, or $a^2 + \frac{4}{3} \left(\frac{a}{2} + b \right)^2 = s^2$. Expanding, $\frac{4}{3}(a^2 + ab + b^2) = s^2$, as desired. \square

Lemma 3.3. *For any angle θ , $\cos(\theta - \frac{2\pi}{3}) + \cos \theta + \cos(\theta + \frac{2\pi}{3}) = 0$.*

Proof. Applying addition and subtraction identities, we find that $\cos(\theta - \frac{2\pi}{3}) + \cos \theta + \cos(\theta + \frac{2\pi}{3}) = \cos \theta \cos \frac{2\pi}{3} + \sin \theta \sin \frac{2\pi}{3} + \cos \theta + \cos \theta \cos \frac{2\pi}{3} - \sin \theta \sin \frac{2\pi}{3} = (1 + 2 \cos \frac{2\pi}{3}) \cos \theta = (1 + 2(-\frac{1}{2})) \cos \theta = 0$, as desired. \square

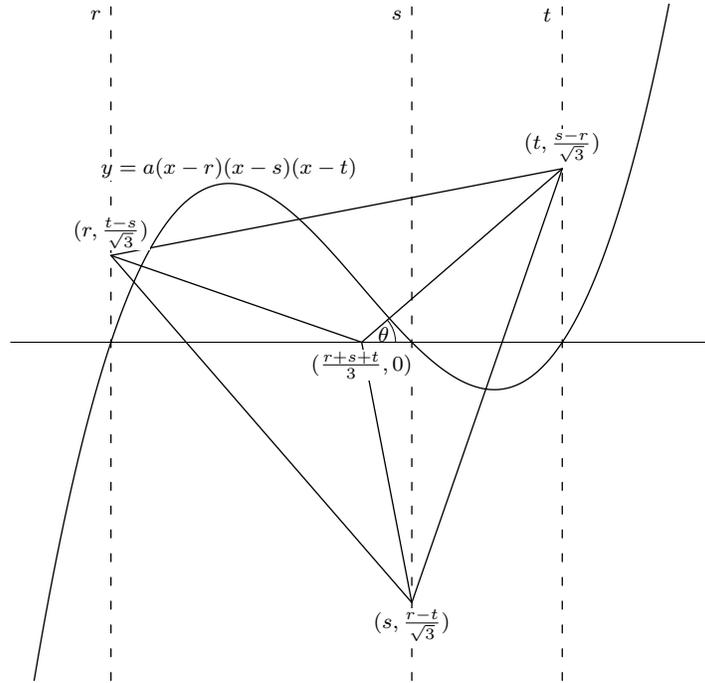
Lemma 3.4. For any angle θ , $\cos(\theta - \frac{2\pi}{3}) \cos \theta + \cos(\theta - \frac{2\pi}{3}) \cos(\theta + \frac{2\pi}{3}) + \cos \theta \cos(\theta + \frac{2\pi}{3}) = -\frac{3}{4}$.

Proof. Applying the addition and subtraction identities, expanding, and simplifying turns the expression into $2 \cos^2 \theta \cos \frac{2\pi}{3} + \cos^2 \theta \cos^2 \frac{2\pi}{3} - \sin^2 \theta \sin^2 \frac{2\pi}{3}$. From here, we apply the Pythagorean trigonometric identity to get rid of the $\sin^2 \theta$ term and obtain the expression $2 \cos^2 \theta \cos \frac{2\pi}{3} + \cos^2 \theta \cos^2 \frac{2\pi}{3} - (1 - \cos^2 \theta) \sin^2 \frac{2\pi}{3} = -\cos^2 \theta + \frac{1}{4} \cos^2 \theta - \frac{3}{4} + \frac{3}{4} \cos^2 \theta = -\frac{3}{4}$. \square

Lemma 3.5. For any angle θ , $\cos(\theta - \frac{2\pi}{3}) \cos \theta \cos(\theta + \frac{2\pi}{3}) = \cos^3 \theta - \frac{3}{4} \cos \theta$.

Proof. Using both addition/subtraction identities and the Pythagorean trigonometric identity, $\cos(\theta - \frac{2\pi}{3}) \cos \theta \cos(\theta + \frac{2\pi}{3}) = (\cos^2 \theta \cos^2 \frac{2\pi}{3} - \sin^2 \theta \sin^2 \frac{2\pi}{3}) \cos \theta = (\frac{1}{4} \cos^2 \theta - \frac{3}{4} \sin^2 \theta) \cos \theta = (\cos^2 \theta - \frac{3}{4}) \cos \theta = \cos^3 \theta - \frac{3}{4} \cos \theta$. \square

With the above proofs under our belt, we can begin to tackle a trigonometric solution to the cubic. In the quadratic formula, we noticed a symmetry of the two roots around their arithmetic mean. We will attempt to solve a cubic $ax^3 + bx^2 + cx + d = a(x - r)(x - s)(x - t) = 0$ by interpreting its three roots to be somehow symmetric around their arithmetic mean. Working with the restriction that r , s , and t are real numbers places the three values on the same axis, and allows for the following geometric interpretation:



Here, we have placed the centre of the three roots on their arithmetic average, at $x = \frac{r+s+t}{3}$. For simplicity, the centre is placed on the Cartesian coordinate $(\frac{r+s+t}{3}, 0)$. The three roots are expressed symmetrically as the three vertices of an equilateral triangle on the Cartesian plane centred around $(\frac{r+s+t}{3}, 0)$ and are placed on the lines $x = r$, $x = s$, and $x = t$.

Without loss of generality, take $r \leq s \leq t$. We can apply Lemma 3.2 to determine the sidelength of the equilateral triangle we have placed on the Cartesian plane. Note that we have an equilateral triangle with vertices on three parallel lines, the middle of which is $s - r$ and $t - s$ away from the other two. By Lemma 3.2, the sidelength of our triangle is $\sqrt{\frac{4}{3}((s-r)^2 + (s-r)(t-s) + (t-s)^2)} = \sqrt{\frac{4}{3}(r^2 + s^2 + t^2 - rs - rt - st)}$. The distance from the centre to a vertex, as in any equilateral triangle, is $\frac{\sqrt{3}}{3}$ times the sidelength, which would be equal to $\frac{2}{3}\sqrt{r^2 + s^2 + t^2 - rs - rt - st}$.

We can use this distance to find the y -coordinate of the vertex on $x = t$. Without loss of generality, let the y -coordinate of the t -vertex be non-negative. By the distance formula, $(t - \frac{r+s+t}{3})^2 + (y - 0)^2 = \frac{4}{9}(r^2 + s^2 + t^2 - rs - rt - st)$ gives us $y^2 = \frac{r^2 - 2rs + s^2}{3}$ or $y = \frac{s-r}{\sqrt{3}}$, since the y -coordinate is non-negative. We can similarly solve for the y -coordinates of the r - and s -vertices, which are $y = \frac{t-s}{\sqrt{3}}$ and $y = \frac{r-t}{\sqrt{3}}$, respectively.

Letting θ (as shown on the diagram) be the angle between the x -axis and the line segment connecting the triangle centre and the t -vertex, we arrive at the following expressions:

$$\begin{aligned} r &= \frac{r+s+t}{3} + \frac{2}{3}\sqrt{r^2 + s^2 + t^2 - rs - rt - st} \cos(\theta + \frac{2\pi}{3}) \\ &= \frac{-b + 2 \cos(\theta + \frac{2\pi}{3})\sqrt{b^2 - 3ac}}{3a} \\ s &= \frac{-b + 2 \cos(\theta - \frac{2\pi}{3})\sqrt{b^2 - 3ac}}{3a} \\ t &= \frac{-b + 2 \cos \theta \sqrt{b^2 - 3ac}}{3a} \end{aligned}$$

Multiplying and expanding these equations together, as well as applying Lemmas 3.3, 3.4, and 3.5, will give us $rst = -\frac{d}{a} = (-\frac{b}{3a})^3 + (-\frac{b}{3a})^2(\frac{2\sqrt{b^2-3ac}}{3a})(0) + (-\frac{b}{3a})(\frac{2\sqrt{b^2-3ac}}{3a})^2(-\frac{3}{4}) + (\frac{2\sqrt{b^2-3ac}}{3a})^3(\cos^3 \theta - \frac{3}{4} \cos \theta)$. Simplifying and rearranging, we have $\frac{9abc - 27a^2d - 2b^3}{27a^3} = \frac{8(b^2 - 3ac)\sqrt{b^2 - 3ac}}{27a^3}(\cos^3 \theta - \frac{3}{4} \cos \theta)$. This tells us that $\cos^3 \theta - \frac{3}{4} \cos \theta = \frac{9abc - 27a^2d - 2b^3}{8(b^2 - 3ac)\sqrt{b^2 - 3ac}}$.

What do we do now? At first glance, this result seems discouraging. To solve for $\cos \theta$, it appears that we would need the cubic formula... However, the keen student will notice that a triple angle identity is hidden in this expression. Multiplying both sides by 4, we obtain $4 \cos^3 \theta - 3 \cos \theta = \frac{9abc - 27a^2d - 2b^3}{2(b^2 - 3ac)\sqrt{b^2 - 3ac}}$. But of course, $4 \cos^3 \theta - 3 \cos \theta$ is none other than $\cos 3\theta$!

That is, θ can be found to equal $\arccos\left(\frac{9abc - 27a^2d - 2b^3}{2(b^2 - 3ac)\sqrt{b^2 - 3ac}}\right)/3$. Here, we double-check to make sure that everything makes sense. Based on our definition of θ (refer to diagram), we should have $0 \leq \theta \leq \frac{\pi}{3}$. The range of the arccosine function is $[0, \pi]$, so the range of our θ is $[0, \frac{\pi}{3}]$, as expected. So we have found another cubic formula!

Theorem 3.6 (Cubic formula). *If the cubic $ax^3 + bx^2 + cx + d = 0$ has three real roots $r \leq s \leq t$, then the roots satisfy*

$$\begin{aligned} r &= \frac{-b + 2 \cos\left(\arccos\left(\frac{9abc - 27a^2d - 2b^3}{2(b^2 - 3ac)\sqrt{b^2 - 3ac}}\right)/3 + \frac{2\pi}{3}\right)\sqrt{b^2 - 3ac}}{3a}, \\ s &= \frac{-b + 2 \cos\left(\arccos\left(\frac{9abc - 27a^2d - 2b^3}{2(b^2 - 3ac)\sqrt{b^2 - 3ac}}\right)/3 - \frac{2\pi}{3}\right)\sqrt{b^2 - 3ac}}{3a}, \\ t &= \frac{-b + 2 \cos\left(\arccos\left(\frac{9abc - 27a^2d - 2b^3}{2(b^2 - 3ac)\sqrt{b^2 - 3ac}}\right)/3\right)\sqrt{b^2 - 3ac}}{3a}. \end{aligned}$$

4 Quartic

4.1 Radical solution

Lemma 4.1. *Every coefficient of polynomial $f(x) = a(x - G_1)(x - G_2)\dots(x - G_k)$ where G_1, G_2, \dots, G_k are all k distinct permutations of some polynomial $g(r_1, r_2, \dots, r_n)$ is a symmetric polynomial over the n variables r_1, r_2, \dots, r_n .*

Proof. Swapping any two of r_1, r_2, \dots, r_n in the factorization of $f(x)$ swaps pairs of the factors $(x - G_1), (x - G_2), \dots, (x - G_k)$ preserving the polynomial $f(x)$. Because $f(x)$ remains unchanged after swapping any two of r_1, r_2, \dots, r_n , the individual coefficients of its expansion remain unchanged as well. By Definition 1.2, they must be symmetric polynomials over r_1, r_2, \dots, r_n . \square

Example 4.1.1. Let $g(r_1, r_2) = r_1^2 - r_2$. Its distinct permutations G_1 and G_2 are $r_1^2 - r_2$ and $r_2^2 - r_1$. The coefficients of the polynomial $f(x) = (x -$

$(r_1^2 - r_2^2)(x - (r_2^2 - r_1)) = x^2 + (-r_1^2 - r_2^2 + r_1 + r_2)x + (r_1^2 r_2^2 - r_1^3 - r_2^3 + r_1 r_2)$ are 1, $-r_1^2 - r_2^2 + r_1 + r_2$, and $r_1^2 r_2^2 - r_1^3 - r_2^3 + r_1 r_2$, all symmetric polynomials on r_1, r_2 .

One of our initial impressions as to how to solve a quartic polynomial might be to somehow express it as a product of two quadratics. That is, if we start with our quartic polynomial $f(x) = ax^4 + bx^3 + cx^2 + dx + e = a(x - r)(x - s)(x - t)(x - u)$, we would like to find some factorization $f(x) = a(x^2 - (r + s)x + rs)(x^2 - (t + u)x + tu)$ and solve for the roots using the quadratic formula on the two quadratic factors.

The key insight here is that the factorization $\frac{f(x)}{a} = (x^2 - (r + s)x + rs)(x^2 - (t + u)x + tu)$ can be expressed as a difference of squares. We know that any product AB can be expressed as $(\frac{A+B}{2})^2 - (\frac{A-B}{2})^2$, so we can similarly say that $\frac{f(x)}{a} = (x^2 - \frac{(r+s+t+u)}{2}x + \frac{(rs+tu)}{2})^2 - (-\frac{(r+s-t-u)}{2}x + \frac{(rs-tu)}{2})^2 = p(x)^2 - q(x)^2$.

Because $q(x)^2$ is just $p(x)^2 - \frac{f(x)}{a}$, if we knew just a , $f(x)$, and $p(x)$, we could find the factorization $f(x) = a(p(x) + \sqrt{p(x)^2 - \frac{f(x)}{a}})(p(x) - \sqrt{p(x)^2 - \frac{f(x)}{a}})$ and proceed to find the roots with two applications of the quadratic formula.

We already know $f(x)$ and a ; $f(x)$ is the given quartic (which we wish to solve), and a is that quartic's coefficient of x^4 . All that remains to find is $p(x) = x^2 - \frac{(r+s+t+u)}{2}x + \frac{(rs+tu)}{2}$.

By Vieta's, $-(r+s+t+u)$ is equal to $\frac{b}{a}$, by Vieta's formulas, so $-\frac{r+s+t+u}{2} = \frac{b}{2a}$. But what is $\frac{rs+tu}{2}$? It seems we are at an impasse, as we know that we can combine the coefficients of $f(x)$ (which are the elementary symmetric polynomials on r, s, t, u) to form symmetric polynomials, but $\frac{rs+tu}{2}$ is not symmetric; swapping r and t or s and u gives the permutation $\frac{ru+st}{2}$, and swapping r and u or s and t gives $\frac{rt+su}{2}$.

Are we out of luck? Not necessarily. There are a couple of important things to note here. Firstly, it does not matter whether we find $\frac{rs+tu}{2}$, $\frac{rt+su}{2}$, or $\frac{ru+st}{2}$. If we found either of the two other permutations of $\frac{rs+tu}{2}$, we could just as easily substitute that into our formula for $p(x)$ and arrive at a valid factorization of $f(x)$.

Secondly, all of the coefficients of the cubic $(x - \frac{rs+tu}{2})(x - \frac{rt+su}{2})(x - \frac{ru+st}{2})$ are symmetric polynomials on r, s, t , and u , as per Lemma 4.1. Meaning, the coefficients of this cubic can be found from combining the elementary symmetric polynomials of r, s, t, u ; knowing the coefficients of $f(x)$ is enough information to construct this cubic. After constructing the cubic, all we need to do is apply the cubic formula to it to find any one of $\frac{rs+tu}{2}$, $\frac{rt+su}{2}$, and $\frac{ru+st}{2}$.

and complete our formula for $p(x)$.

It is easy to find an expression for $(x - \frac{rs+tu}{2})(x - \frac{rt+su}{2})(x - \frac{ru+st}{2})$ in terms of the coefficients of the original quartic. Expanding our cubic, we find that the coefficient of x^2 is $-\frac{1}{2}(rs + rt + ru + st + su + tu)$, or $-\frac{c}{2a}$, by Vieta's. The coefficient of x is $\frac{1}{4}(r^2st + r^2su + r^2tu + rs^2t + rs^2u + rst^2 + rsu^2 + rt^2u + rtu^2 + s^2tu + st^2u + stu^2) = \frac{1}{4}((r + s + t + u)(rst + rsu + rtu + stu) - 4rstu) = \frac{1}{4}((-\frac{b}{a})(-\frac{d}{a}) - \frac{4e}{a})$, by Vieta's, which simplifies to $\frac{bd-4ae}{4a^2}$. The constant term of the cubic can be expressed as $-\frac{1}{8}((r^3stu + rs^3tu + rst^3u + rstu^3) + (r^2s^2t^2 + r^2s^2u^2 + r^2t^2u^2 + s^2t^2u^2))$. We can see that $r^3stu + rs^3tu + rst^3u + rstu^3 = (r^2 + s^2 + t^2 + u^2)(rstu) = ((r + s + t + u)^2 - 2(rs + rt + ru + st + su + tu))(rstu) = ((-\frac{b}{a})^2 - \frac{2c}{a})(\frac{e}{a})$. In a similar fashion, we can find $r^2s^2t^2 + r^2s^2u^2 + r^2t^2u^2 + s^2t^2u^2$ to equal $(rst+rsu+rtu+stu)^2 - 2(rs+rt+ru+st+su+tu)(rstu) = (-\frac{d}{a})^2 - 2(\frac{c}{a})(\frac{e}{a})$, again by an application of Vieta's formulas. Adding the two parts together, multiplying by $-\frac{1}{8}$, and simplifying tells us that the constant term is equal to $\frac{4ace-ad^2-b^2e}{8a^3}$.

That is, the polynomial $(x - \frac{rs+tu}{2})(x - \frac{rt+su}{2})(x - \frac{ru+st}{2})$ can be expressed in terms of the coefficients of the original quartic as $x^3 - (\frac{c}{2a})x^2 + (\frac{bd-4ae}{4a^2})x + (\frac{4ace-ad^2-b^2e}{8a^3})$. Let k denote one solution to this cubic. Then, we can let $p(x) = x^2 + \frac{b}{2a}x + k$. This gives us $p(x)^2 - \frac{f(x)}{a} = (\frac{8a^2k+b^2-4ac}{4a^2})x^2 + (\frac{bk-d}{a})x + (\frac{ak^2-e}{a})$. Knowing that this expression is the square of a linear polynomial $q(x)$ allows us to easily factor $(\frac{8a^2k+b^2-4ac}{4a^2})x^2 + (\frac{bk-d}{a})x + (\frac{ak^2-e}{a}) = (\frac{\sqrt{8a^2k+b^2-4ac}}{2a}(x + \frac{2abk-2ad}{8a^2k+b^2-4ac}))^2$, by the quadratic formula.

Theorem 4.2 (Quartic formula). *The four roots of the quartic $ax^4 + bx^3 + cx^2 + dx + e = 0$ can be found to equal the solutions of the quadratics*

$$x^2 + \frac{b + \sqrt{8a^2k + b^2 - 4ac}}{2a}x + (k + \frac{(bk - d)\sqrt{8a^2k + b^2 - 4ac}}{8a^2k + b^2 - 4ac}) = 0,$$

$$x^2 + \frac{b - \sqrt{8a^2k + b^2 - 4ac}}{2a}x + (k - \frac{(bk - d)\sqrt{8a^2k + b^2 - 4ac}}{8a^2k + b^2 - 4ac}) = 0,$$

where k is some solution to $k^3 - (\frac{c}{2a})k^2 + (\frac{bd-4ae}{4a^2})k + (\frac{4ace-ad^2-b^2e}{8a^3}) = 0$, solvable by the cubic formula.

Interestingly, I could not find this particular method of solving the quartic online. Although the solution is convoluted, I remain proud of the idea that I may have discovered a novel solution to the quartic. There are, of course, solutions that arrive at the same answer with different steps.

5 Quintic

5.1 Radical solution

Interestingly, when solving for the cubic and quartic in radicals, we found that we required the quadratic and cubic formula, respectively. Why should this be the case?

If we look back to Section 3.1.2, we attempted a radical solution by hoping that the cube of a Lagrange resolvent was symmetric. Unfortunately for us, there were two asymmetric possibilities for such a value. At this point, perhaps a light bulb is going off in your head. Would it be possible, in a similar way to our solution to the quartic, to solve for one of the two asymmetric possibilities with the quadratic formula?

In the case of three roots (r , s , and t), there were two distinct possibilities when taking the cube of a Lagrange resolvent: $(r + s \operatorname{cis} \frac{2\pi}{3} + t \operatorname{cis} \frac{4\pi}{3})^3$ and $(s + r \operatorname{cis} \frac{2\pi}{3} + t \operatorname{cis} \frac{4\pi}{3})^3$. In the case of four roots (r , s , t , and u), we can find that there are three distinct possibilities when taking the fourth power of a Lagrange resolvent: $(r + s \operatorname{cis} \frac{\pi}{2} + t \operatorname{cis} \pi + u \operatorname{cis} \frac{3\pi}{4})^4$, $(s + r \operatorname{cis} \frac{\pi}{2} + t \operatorname{cis} \pi + u \operatorname{cis} \frac{3\pi}{4})^4$, and $(u + s \operatorname{cis} \frac{\pi}{2} + t \operatorname{cis} \pi + r \operatorname{cis} \frac{3\pi}{4})^4$.

The cubic reduced to a quadratic when there were two distinct values when taking the cube of a Lagrange resolvent. The quartic reduced to a cubic when there were three distinct values when taking the fourth power of a Lagrange resolvent.

We can check and see that there are actually six different values when taking the fifth power of a Lagrange resolvent over five roots. From this, we could suspect that a quintic must “reduce” to a sextic, a polynomial that would “reduce” to a polynomial of an even higher degree, *ad nauseum*. It does not appear anywhere near possible to be solved. Indeed, this is the basic idea behind the proof of the insolvability of quintic over radicals, as per Galois theory.

6 Conclusion

This investigation was an opportunity to delve into what higher-level mathematics is really like. In particular, I enjoyed the rigour involved in perfecting each mathematical statement and proof. This exploration greatly challenged my mathematical ability, and has offered me greater appreciation

towards what is known as “pure mathematics”, where mathematicians pursue math not for any practical purpose, but for simply the sake of it.

Throughout the investigation, I was surprised when my initial guesses about Lagrange resolvents were right on the mark. In Section 3.1.2, there was some disappointment in not finding a symmetric polynomial, but Lagrange resolvents ended up being an important part of Galois theory. I have tried to structure this paper as a narrative, so that places where I myself got stuck were conveyed to the reader. This hopefully helps with the aim to communicate in an accessible manner to the reader, helping them understand my thoughts at each step.

Although the results of this investigation were mainly already well-known, I hope that by my own personal derivation of each formula, the motivations and thoughts behind each approach are more clearly understood by the reader. Additionally, as far as I know, my solution to the quartic is unique from the solutions in the public domain.

Overall, this was a very enjoyable experience. I am glad to be able to communicate my knowledge to others through this investigation. Perhaps this paper will encourage others to pursue mathematics more deeply and more passionately.

7 Afterword

The study of the relationship between a polynomial's roots and its coefficients has a long and rich history. Some particularly romantic highlights include challenges to solve cubic equations in a rivalry between da Coi and Tartaglia, a possibly broken promise between Cardano and Tartaglia, and Galois sending his work on insolvability to a friend before his untimely death in a duel. Among other things, the study has led to the development of the theory of symmetric functions from the elementary symmetric polynomials that arise in Vieta's formulas, complex numbers from trying to understand the "inadmissible" solutions that arose in the work of Cardano, del Ferro, and Tartaglia on the cubic, as well as Galois theory and abstract algebra as the language to express Galois' work.

This work was completed to satisfy a requirement for the International Baccalaureate program, which I was enrolled in during my last two years of high school. The IBO gives examples of previous Mathematics IAs on their website. At the time, I noticed a bias towards applied topics in their examples; none of the example papers given on their website would be considered pure mathematics. In addition to the arrogance and pretentiousness of a 17-year-old, I was drawn to the notion of mathematics outside of immediate applications because I had learned to suspect that there was value in the mathematical art beyond simply being "useful".

I read the sample papers given on the IBO website, and was frustrated by the poor judgment applied to one of the student papers, arising from what I perceived as a lack of mathematical understanding. In Example 21 on the IBO website, the student models the rate of regularization of irregular verbs using exponential decay. They use methods to compute some sort of half-life corresponding to a word, and then conclude that the word will fall out of usage after twice the computed half-life. In the assessment of this work, the instructor wrote: "natural logarithms commensurate with the level of the course clearly understood and correctly used".

I submitted a proposal for my Mathematics IA to my teacher. Other topics I had considered were magic squares of squares and finding continuous extensions of discrete patterns (e.g. defining the factorial function for non-integers). My teacher expressed her disappointment that I had submitted my proposal late, whereas I felt unfairly judged knowing I had put work into the proposal and that many of my classmates had only written their proposals the night before the deadline.

There are some inaccuracies in the paper and I made or implied false claims about my level of understanding in it (as a result of being at the time a teenager who wanted to appear smart). I will try to elucidate more of what my background and thought process was and correct the false claims in this afterword.

In the rationale, I use the wording “student of mathematics” instead of “mathematics student”, which was thought to be a typo by one classmate. I did mean “student of mathematics”, I think because I recognized then that mathematics itself was a teacher I was learning from. I learned proofs of Euler’s formula and the fundamental theorem of algebra in my undergraduate studies, but I’m not sure I learned the history of the results and which ideas in the proofs were specifically attributed to Euler and Gauss.

I cited “a friend” as the source of my interest in *The Man Who Knew Infinity* and Ramanujan. It was actually my high school counselor who asked me to watch it. He often made good attempts at connecting with me by sharing depictions of mathematicians in pop culture, although I suspected he worried I would become something like a cold unfeeling logical robot devoid of emotions.

I was frustrated by my inadequacy to satisfactorily define what a polynomial was in precise language. I ended up citing *MathWorld* for a definition. I also could not prove the fundamental theorem of symmetric polynomials. I cited a proof in an article of Lerner because it seemed simple, but I did not understand it. I later learned in university a proof using a lexicographic monomial order, which I think is sufficiently elementary.

I was familiar with Vieta’s formulas through my participation in mathematics competitions, although I did not know they were attributed to Vieta until another student referred to them as Vieta’s formulas. I knew that the direct argument gave sufficient justification, but I still felt it lacked somehow in precision.

I don’t remember when I first learned the quadratic formula, but by the time I was 17, I also knew that some students struggled to remember it and did not know how to derive it from first principles. When studying the quadratic formula, I gained some intuition for how a formula for the roots of a polynomial could employ radicals to recover expressions that were not symmetric polynomials in the roots.

I tried to find a cubic formula using radicals, but I was not able to. I read off from the internet the first solution to the cubic in this article. I was not too far with my approach considering the expressions $(r + s \operatorname{cis} \frac{2\pi}{3} + t \operatorname{cis} \frac{4\pi}{3})^3$

and $(r + t \operatorname{cis} \frac{2\pi}{3} + s \operatorname{cis} \frac{4\pi}{3})^3$. The pair can be computed with a quadratic. Then, taking cube roots, one could recover (up to a permutation of variables) $r + \operatorname{cis} \frac{2\pi}{3} + t \operatorname{cis} \frac{4\pi}{3}$ and $r + t \operatorname{cis} \frac{2\pi}{3} + s \operatorname{cis} \frac{4\pi}{3}$. The caveat here is that the two cube roots chosen must be “coupled”, which can be done by requiring that their product is $r^2 + s^2 + t^2 - rs - rt - st$. Then, one uses the known value for $r + s + t$ and solves a system of linear equations. The solution to the cubic using radicals was first published by Cardano in 1545, based on earlier work of del Ferro and Tartaglia.

I recall that I had already briefly seen a picture of an equilateral triangle overlaid on top of the graph of a cubic before I started working on the trigonometric solution to the cubic, so in that sense I had a head start. The intuition I gained from studying the quadratic was useful in understanding how an equilateral triangle could help in determining the roots of a real-rooted cubic. Deriving the solution was a lot of fun, in part because I was able to apply my knowledge of plane geometry (which appears often in mathematics competitions) to the study of algebra. I got stuck once I arrived at the expression $\cos^3 \theta - \frac{3}{4} \cos \theta$; for this I simply looked it up and learned that it was $\frac{1}{4} \cos 3\theta$. For the figures, I think I used *GeoGebra*, which allows the user to export their diagram to TikZ code. The trigonometric solution to the real-rooted cubic was first published by Vieta in 1593.

For the solution to the quartic using radicals, I was very happy with the success of the ideas I tried: writing a quartic as a product of two quadratics, writing the product as a difference of squares, and finding $\frac{rs+tu}{2}$ (up to a permutation of variables) using a cubic. I wrote up the lemma used in the section on the quartic to justify the reduction to the cubic, which I now think was a nice find, considering I didn’t have any notion of group theory or abstract algebra at the time. I knew that there was a small edge case when $bk - d = 0$ (in which case $q(x)^2$ is either zero or a monomial), but I ignored that issue in the presentation of the final formula, since the procedure leading up to the formula still worked. I considered my solution to the quartic new because most or all of the solutions to the cubic or quartic using radicals I could find online involved the preliminary step of depressing the cubic or quartic. The solution to the quartic using radicals was first discovered by Ferrari in 1540, but could not be immediately published because it required a solution to the cubic.

Unfortunately, my analysis of why the quintic could not be solved using radicals was essentially me making things up. (Whoops!) I extrapolated from a small pattern I had observed with few data points, then claimed that

the obstruction I found with applying these ideas to polynomials of degree ≥ 5 was the “basic idea” of insolvability “as per Galois theory”. I did not understand Galois theory as a high school student. I was happy to learn that some of the expressions I played around with were part of the literature, but my claim that Lagrange resolvents were “an important part of Galois theory” was false. The insolvability of the general quintic and general polynomials of higher degree was proven by Abel in 1824; an incomplete proof was given by Ruffini in 1799. In 1830, Galois’ work provided a description of which specific polynomials of degree ≥ 5 could or could not be solved using radicals.

For my Mathematics IA, I wanted to imitate the professional and precise style of modern mathematicians. It was my first LaTeX document! I wanted my work to be aesthetically beautiful, and spent time tweaking and making the diagrams pretty. I was lucky to be able to investigate and learn more about the ideas of mathematicians across time. I do think my work on my Mathematics IA partly influenced my decision to continue studying mathematics in university. As a result, my study of mathematics has become a large influence on my life and philosophy, guiding my understanding of ideals such as beauty, justice, and truth, and how to recognize and incorporate these ideals in my relationships with other people in our shared strange existence.

For my work I was awarded a grade of 17 out of 20, coinciding with the grade given to the paper in Example 21 on the IBO website.

Love,
Victor