1996

Polynomial equations and solvability: A historical perspective

Laurie Jan Riggs

Follow this and additional works at: https://scholarworks.lib.csusb.edu/etd-project

Part of the Mathematics Commons

Recommended Citation
https://scholarworks.lib.csusb.edu/etd-project/1186

This Thesis is brought to you for free and open access by the John M. Pfau Library at CSUSB ScholarWorks. It has been accepted for inclusion in Theses Digitization Project by an authorized administrator of CSUSB ScholarWorks. For more information, please contact scholarworks@csusb.edu.
POLYNOMIAL EQUATIONS AND SOLVABILITY:

A HISTORICAL PERSPECTIVE

A Project
Presented to the
Faculty of
California State University,
San Bernardino

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts
in
Mathematics

by
Laurie Jan Riggs
June 1996
POLYNOMIAL EQUATIONS AND SOLVABILITY:
A HISTORICAL PERSPECTIVE

A Project
Presented to the
Faculty of
California State University,
San Bernardino

by
Laurie Jan Riggs
June 1996
Approved by:

Dr. James Okon, Chair, Mathematics

Dr. Robert Stein

Dr. Rolland Trapp

6/5/96
Date
ABSTRACT

The search for the solutions of polynomial equations was the catalyst for many profound mathematical discoveries. Some of the greatest mathematical minds worked on finding a solution in radicals for cubic, quartic, and quintic polynomial equations. In pursuit of the formula for the general quintic polynomial Abel and Galois laid a foundation for what would become known as Modern Algebra. In this project I will trace the solutions of polynomial equations starting with the geometric techniques for quadratics and cubics. I will then look at the fifteenth century discoveries of Cardan, Tartaglia, and Ferrari. I will derive the general cubic and quartic equations using permutations of the roots. Finally I will look at Abel, and Galois, and the non-solvability of the general quintic equation.
ACKNOWLEDGMENTS

I would like to thank the professors on my committee, Dr. R. Stein, and Dr. R. Trapp for their help. I am also deeply grateful to my committee chair, Dr. James Okon for his time, energy, and support in producing this project. His interest in mathematical history has inspired me.
# TABLE OF CONTENTS

ABSTRACT ................................................................. iii

ACKNOWLEDGMENTS ...................................................... iv

CHAPTER ONE

Geometric Techniques for Quadratic Polynomials .................... 1
Geometric Techniques for Cubic Polynomials ........................... 2
Algebraic Solutions for Cubic Polynomials ............................ 6

CHAPTER TWO

Solutions to Cubic Polynomials Using Symmetric Functions ........... 14
Solutions to Quartic Polynomials Using Symmetric Functions ........ 28

CHAPTER THREE

Non Solvability of the Quintic Equation .............................. 38

BIBLIOGRAPHY .................................................................. 52
CHAPTER ONE

Geometric Techniques for Quadratic Polynomials

Islamic mathematicians in the time of Omar Khayyam (1050-1123) had geometric techniques for solving quadratic equations. Mathematicians of this era did not use negative numbers so there were different types of methods for solving different types of quadratics. To solve an equation of the type

\[ x^2 + bx = c \]

we use the idea of completing the square. We will look at the following illustration:

Start with square \( ABCD = x^2 \), then add the area \( bx \) to \( x^2 \). To keep it balanced add this in equal portions to both sides, as shown. Now we have the sum of the areas is \( x^2 + bx \).
Complete the square by adding the square $CEFG = \text{area} \frac{b^2}{4}$.

Now we can see that:

\[
\left(x + \frac{b}{2}\right)^2 = \frac{b^2}{4} + c
\]
\[
\left(x + \frac{b}{2}\right) = \sqrt{\frac{b^2}{4} + c}
\]
\[
x = \pm \sqrt{\frac{b^2}{4} + c} - \frac{b}{2}
\]

which is a form of the quadratic formula.

**Geometric Techniques for Cubic Polynomials**

Omar Khayyam (1050-1123) bridged the gap between algebra and geometry. He is regarded and one of the greatest Persian poets and mathematicians. He wrote, "Whoever thinks algebra is a trick in obtaining unknowns has thought in vain. No attention should be paid to the fact that algebra and geometry are different in appearance. Algebras are geometric facts which are proved." While he gave both algebraic and geometric
solutions for quadratic equations, he thought that for general cubic equations algebraic solutions were impossible and gave only geometric solutions. For his geometric solutions he used the idea of intersecting conics to solve cubic equations and gave a complete classification of equations up to degree three with the methods of solving each. Because he used only positive numbers, he had 14 cubic equations that were not reducible to a linear or quadratic form. They are:

\[
\begin{align*}
x^3 &= d \\
x^3 + cx &= d \\
x^3 + d &= cx \\
x^3 &= cx + d \\
x^3 + bx^2 &= d \\
x^3 &= bx^2 + d \\
x^3 + d &= bx^2 \\
x^3 + bx^2 + cx &= d \\
x^3 + bx^2 + d &= cx \\
x^3 + cx + d &= bx^2 \\
x^3 &= bx^2 + cx + d \\
x^3 + bx^2 &= cx + d
\end{align*}
\]
\[ x^3 + cx = b w^2 + d \]
\[ x^3 + d = b x^2 + cx \]

Not all roots to these equations were given, because he did not accept negative roots.

We illustrate Khayyam’s methods with his solution of a cubic equation of the type

\[ x^3 + cx = d. \]

His geometric solution for this type of cubic involved the intersection of a parabola and a circle. He let \( a^2 = c \), \( a^2 h = d \), and then the equation of the parabola is

\[ x^2 = ay. \]

The equation of the circle is

\[ y^2 = x (h - x). \]

As a concrete example, we use this method to solve

\[ x^3 + 4x = 16. \]

In this case \( a^2 = c = 4 \), and \( a^2 h = d = 16 \), thus \( a = 2 \), \( h = 4 \). The equation for the parabola is
The equation for the circle is

\[ y^2 = x(4 - x) . \]
Intersecting these two conics we have:

The x-coordinate of the intersection is 2, which is the desired solution.

**Algebraic Solutions for Cubic Polynomials**

The algebraic solutions to the cubic and quartic equations have a very scandalous history. Included below are two of the main figures in their discovery.

The first is Niccolo Fontana (1499-1557). He is better known by his nickname, Tartaglia, which means the stammerer. When the French attacked Tartaglia’s home town of Brescia in 1512, a soldier slashed him horribly across his face and jaw. He survived the attack, but was left with scars that made it impossible for him to speak clearly. Little is know of his educational background or family, but history portrays him as a gifted mathematician. [EV] There are some questions on the originality of some of the work he published. Quesiti et inventioni diverse, and an Archimedean
translation, both contained another’s work passed off as his own. What Tartaglia did have was an original solution to one type of cubic equation. This type of discovery was not something to share as the ability to solve equations no one else could meant certain victory in public challenges with other mathematicians. These successes could lead to prestigious academic appointments. Earlier, the talented professor Scipione del Ferro had discovered an algebraic solution to a cubic equation of the type,

\[ x^3 + bx = c \]

(a "depressed" cubic, so called because of the absence of a quadratic term.) On his deathbed he passed his secret to one of his students, Antonio Fior, who then sent a challenge to Tartaglia. Now Tartaglia had already come up with the solution to cubic equations of the type

\[ x^3 + ax^2 = c, \]

and when Fior sent him thirty of the depressed cubics to solve he worked frantically to come up with the solution. He succeeded on February 12, 1535, and was able to solve all of Fior’s problems. Meanwhile Fior had not been able to solve any of the problems posed by Tartaglia, and Tartaglia was the victor in the challenge. Word of his conquest reached Cardan who promptly enticed Tartaglia to his home in Milan. Tartaglia had planned to publish his solution to the cubic, but Cardan persuaded him to reveal his
secret with the solemn oath: "I swear to you by the Sacred Gospel, and on my faith as a gentleman, not only never to publish your discoveries, if you tell them to me, but I also promise and pledge my faith as a true Christian to put them down in cipher so that after my death no one shall be able to understand them." [DU] So in Milan, on March 23, 1539, Tartaglia disclosed his secret in a poem:

When the cube and its things near
Add it a new number discrete,
Determine two new numbers different
By that one; this feat
Will be kept as a rule
Their product always equal, the same,
To the cube of a third
Of the number of things named.
Then, generally speaking
The remaining amount
Of the cube roots subtracted
Will be your desired count. [BO]

Tartaglia was furious when in 1545, Cardan published the solution to the cubic in his book *Ars Magna*, the "Great Art." He protested loudly and often, writing many derogatory letters. None were answered by Cardan, but Cardan was defended by his student Ludovico Ferrari, who later bested Tartaglia in a public challenge. Tartaglia gave his side of the controversy with Cardan in his book *Qvesiti ed invenzioni diverse* (1546).

Mathematician, J. Cardan (1501-1576) of Milan could, be considered "the most bizarre
character in the whole history of mathematics."[DU] He was involved in the algebraic solution of the cubic equation, which was the biggest advance in mathematics since the time of the Greeks. He was a professor at Bologna and Milan and a prolific writer. His writings consist of over seven thousand pages on subjects including mathematics, science, philosophy, and religion. Much of what we know about his life comes from his own book De Vita Propria Liber ("The Book of My Life")[CA]. He was a physician who at first was denied practicing medicine in Milan. He believed this was because he was illegitimate. He moved to Sacco, where he married and had three children, but his private life was tragic. His wife died at age 31, one of his sons was beheaded for killing his own wife, and the other son was a criminal. Cardan was also jailed in 1570 for casting the horoscope of Jesus. Excerpts from his own writings document an unusual life-style. He was plagued with insomnia and would consciously inflict pain upon himself because "I considered that pleasure consisted in relief following severe pain."[CA] He was an ardent astrologer and seer of visions. He believed in spirits and claimed to have an attendant spirit with which he conversed.

Personal things aside, the story of Cardan’s work on cubic and quartic equations is intriguing. He and his student, Luigi Ferrari, made astounding progress on what they had gotten out of Tartaglia, and were able to come up with the solutions to the general cubic and quartic equations. Cardan published these results in his famous book Ars Magma,
[CA] and though he did give Tartaglia credit: "Scipio Ferro of Bologna well-nigh thirty years ago discovered this rule and handed it on to Antonio Maria Fior of Venice, whose contest with Niccolo Tartaglia of Brescia gave Niccolo occasion to discover it. He gave it to me in response to my entreaties, though withholding the demonstration. Armed with this assistance, I sought out its demonstration in [various] forms. This was very difficult."[CA] Tartaglia was not satisfied. He said Cardan had broken a sacred oath and was a vile scoundrel. Cardan never did respond to Tartaglia’s accusations.

A sketch in the margin of Ars Magna shows Cardan’s solution involved dividing the cube into six solids. [SM]
The volume of the large cube is $t^3$ and the six solids are:

Two cubes, one with volume $(t - u)^3$ and the other with volume $u^3$.

Four rectangular prisms, two with volume $ut(t - u)$, one with volume $u(t - u)^2$, and the last with volume $u^2(t - u)$.

Algebraically,

$$t^3 = 2ut(t - u) + u(t - u)^2 + u^2(t - u) + (t - u)^2 + u^3$$

$$t^3 - u^3 = 2ut(t - u) + u(t - u)^2 + u^2(t - u) + (t - u)^2$$
and this can be simplified to:

\[ t^3 - u^3 = 3ut(t - u) + (t - u)^3. \]

Comparing this to the cubic equation

\[ x^3 + cx = d \]

\[ (t - u)^3 + 3ut(t - u) = t^3 - u^3 \]

we have

\[ x = (t - u) \]

\[ c = 3ut \]

\[ d = t^3 - u^3. \]

If we solve for \( u \) in the second equation we get

\[ u = \frac{c}{3t}. \]

Substitute this value of \( u \) into the third equation to get:

\[ d = t^3 = \frac{c^3}{27t^3}. \]

Clearing the fractions yields

\[ 27dt^3 = 27t^6 - c^3 \]

\[ 27t^6 - 27dt^3 - c^3 = 0. \]
To solve for $t$ we must use the quadratic formula and:

$$
t^3 = \frac{27d \pm \sqrt{(27d)^2 + 4(27)(c^3)}}{54}
$$

$$
= \frac{d}{2} \pm \sqrt{\frac{d^2}{4} + \frac{c^3}{27}}.
$$

Since $t^3 - u^3 = d$,

$$
u^3 = t^3 - d
$$

$$
= \frac{d}{2} \pm \sqrt{\frac{d^2}{4} + \frac{c^3}{27}}
$$

and since $x = t - u$ we can get Cardan's formula for solving cubic equations by combining these two expressions and taking the positive square root.

$$
x = \sqrt[3]{\frac{d}{2} + \sqrt{\frac{d^2}{4} + \frac{c^3}{27}}} - \sqrt[3]{\frac{d}{2} - \sqrt{\frac{d^2}{4} + \frac{c^3}{27}}}.
$$
CHAPTER TWO

Solutions to Cubic Polynomials Using Symmetric Functions

The search for a general formula for the quintic equation continued from 1570 to the 1800's. The key idea for the solutions of the quadratic, cubic and quartic polynomial equations was to express the polynomial as a symmetric function of all the roots of the polynomial. Next, develop a general algorithm to convert symmetric polynomials in the roots to polynomials in the coefficients. Newton had the first formulas for computing symmetric functions of the roots of polynomials in terms of the coefficients of the polynomial, and in the 18th century Lagrange (1768) and Waring (1770, 1782) made important contributions to the subject. We will proceed in this section to give a proof of Waring's Theorem and then use an algorithm [WA] based on the proof of Waring's Theorem to derive the formulas for solutions of the cubic, and quartic equations. The following is a standard proof of Waring's Theorem. The algorithm, and its proof will follow.

Definition: If $a_1, ..., a_n$ are elements in a ring $R$, then the elementary symmetric functions $\sigma_1, ..., \sigma_n$ are defined as follows:

$$\sigma_1 = a_1 + ... + a_n = \sum x_1$$

$$\sigma_2 = \sum_{i<j} a_i a_j = a_1 a_2 + ... + a_{n-1} a_n = \sum x_1 x_2$$
\[ \sigma_n = \alpha_1 \cdots \alpha_n = \sum x_1 \cdots x_n, \]

where \( \sum x_1 \) is the sum of the distinct elements in the set \( \{ x_{\sigma(1)} | \sigma \in S_n \} \).

For example: Let \( n = 3 \), then to find \( \sum x_1 x_2 \) we sum the distinct elements in the set \( \{ x_{\sigma(1)} x_{\sigma(2)} \} = \{ x_1 x_2, x_1 x_3, x_2 x_3 \} \). Thus,

\[ \sum x_1 x_2 = x_1 x_2 + x_1 x_3 + x_2 x_3. \]

**Theorem 1.** Let \( F \leq E \) be fields and let \( \alpha_1, \alpha_2, \ldots, \alpha_n \in E \). If \( f \in F[x_1, x_2, \ldots, x_n] \) is symmetric in \( \alpha_1, \alpha_2, \ldots, \alpha_n \) then there exists a polynomial \( g \in F[x_1, x_2, \ldots, x_n] \) with \( \deg(g) \leq \deg(f) \) such that \( f(\alpha_1, \alpha_2, \ldots, \alpha_n) = g(\sigma_1, \sigma_2, \ldots, \sigma_n) \).

**Proof.** Assume \( \alpha_1, \alpha_2, \ldots, \alpha_n \), are indeterminates over \( F \). Using induction on the degree of \( f \), if the \( \deg(f) = 0 \) the result is clear. Now assume the \( \deg(f) = d > 0 \) and the result is true for polynomials of degree less than \( n \). Let

\[
\begin{align*}
\sigma_1^* &= \alpha_1 + \ldots + \alpha_{n+1} \\
\sigma_2^* &= \sum_{i<j} \alpha_i \alpha_j = \alpha_1 \alpha_2 + \ldots + \alpha_{n-2} \alpha_{n-1} \\
&\vdots \\
\sigma_{n-1}^* &= \alpha_1 \cdot \alpha_2 \cdot \ldots \cdot \alpha_n
\end{align*}
\]

and let \( \sigma_i \) denote the \( i^{th} \) elementary symmetric function in \( \alpha_1, \alpha_2, \ldots, \alpha_n \) for \( i = 1 \ldots n \). By
the inductive hypothesis there exists a polynomial \( g_1 \in F [x_1, x_2, \ldots, x_{n-1}] \) such that

\[
f(\alpha_1, \alpha_2, \ldots, \alpha_{n-1}, 0) = g_1(\sigma_1^*, \sigma_2^*, \ldots, \sigma_{n-1}^*)
\]

and \( \deg(g_1) \leq \deg(f) \). The function

\[
f_1(\alpha_1, \alpha_2, \ldots, \alpha_n) = f(\alpha_1, \alpha_2, \ldots, \alpha_n) - g_1(\sigma_1, \sigma_2, \ldots, \sigma_{n-1})
\]

is symmetric of degree less than or equal to \( d \). Now set \( \alpha_n = 0 \) to obtain

\[
f_1(\alpha_1, \alpha_2, \ldots, 0) = f(\alpha_1, \alpha_2, \ldots, 0) - g_1(\sigma_1^*, \sigma_2^*, \ldots, \sigma_{n-1}^*) = 0.
\]

This implies \( \alpha_n \) is a factor of \( f_1 \) so

\[
f_1(\alpha_1, \alpha_2, \ldots, \alpha_n) = \alpha_1 \alpha_2 \ldots \alpha_n f_2(\alpha_1, \alpha_2, \ldots, \alpha_n) = \sigma_n f_2(\alpha_1, \alpha_2, \ldots, \alpha_n)
\]

for some symmetric polynomial \( f_2 \in F [x_1, \ldots, x_n] \). Now \( \deg(f_2) < \deg(f_1) \leq \deg(f) \) so by our inductive hypothesis, there exists a polynomial \( g_{n-1} \in F [x_1, \ldots, x_n] \) with \( \deg(g_2) \leq \deg(f) \) such that

\[
f_2(\alpha_1, \alpha_2, \ldots, \alpha_n) = g_2(\sigma_1, \sigma_2, \ldots, \sigma_n).
\]

Then,

\[
f(\alpha_1, \alpha_2, \ldots, \alpha_n) = f_1(\alpha_1, \alpha_2, \ldots, \alpha_n) + g_1(\sigma_1, \sigma_2, \ldots, \sigma_{n-1})
\]

\[
= \sigma_n g_2(\sigma_1, \sigma_2, \ldots, \sigma_n) + g_1(\sigma_1, \sigma_2, \ldots, \sigma_{n-1}). 
\]

\( \square \)
Definition: Let \( f(x_1, \ldots, x_n) \in F[x_1, \ldots, x_n] \). Then if \( \sigma \in S_n \), \( \sigma f(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \). We say \( f(x) \) is a symmetric polynomial if \( \sigma f = f \) for all \( \sigma \in S_n \).

Now we will give a method to convert a given symmetric polynomial to a polynomial in the elementary symmetric functions \( \sigma_1, \sigma_2, \ldots, \sigma_n \).

Definition: Let \( ax_1^{\alpha_1} \cdots x_n^{\alpha_n}, bx_1^{\beta_1} \cdots x_n^{\beta_n} \in F[x_1, \ldots, x_n] \). We say \( ax_1^{\alpha_1} \cdots x_n^{\alpha_n} > bx_1^{\beta_1} \cdots x_n^{\beta_n} \) in lexicographical order if the first non vanishing difference \( \alpha_i - \beta_i \) is positive. If \( f, g \in F[x_1, \ldots, x_n] \) we say \( f > g \) if the monomial of highest degree in \( f \) is greater than the monomial of highest degree in \( g \).

**Algorithm 1.** Let \( p(x) \in F[x_1, \ldots, x_n] \) be a symmetric polynomial.

1. Arrange the terms of the given symmetric polynomial lexicographically.

2. We have to find a product \( \sigma(x) \) of elementary symmetric functions which, when ordered lexicographically, has the same leading term as \( p(x) \). This product is then subtracted from the given polynomial and the terms of the resulting polynomial \( p(x) - \sigma(x) \) are lower in lexicographical order than those of \( p(x) \).

3. The leading term is found again and we continue until we have converted the polynomial.

The following lemmas and theorem prove the algorithm. As we shall see in the next lemma the required product in step 2 is

\[
\sigma(x) = a\sigma_1^{\alpha_1-\alpha_2} \cdots \sigma_{n-1}^{\alpha_{n-1}-\alpha_n} \sigma_n^{\alpha_n}.
\]
Lemma 2: Let \( \sigma_i = \sum x_1 \cdots x_i \) be the \( i \)th symmetric function of \( x_1, x_2, \ldots, x_n \) and let \( m > 0 \).

Then

\[
\sigma_i^m = x_1^m x_2^m \cdots x_i^m + p(x)
\]

where the terms of \( p(x) \) are the monomials less than \( x_1^m x_2^m \cdots x_i^m \) in lexicographical order.

Proof: Let \( \sigma_i = a_1 + a_2 + \cdots + a_k \) where \( a_1 = x_1 \cdots x_i \). Then

\[
\sigma_i^m = (a_1 + a_2 + \cdots + a_k)^m = \sum_{m_1 + \cdots + m_k = m} \left( \begin{array}{c} m \\ m_1, m_2, \ldots, m_k \end{array} \right) a_1^{m_1} \cdots a_k^{m_k}.
\]

Assume \( a_1^{m_1} a_2^{m_2} \cdots a_k^{m_k} \geq a_1^m \). Let \( S_1 = \{a_1, \ldots, a_n\} \) be terms with non-zero powers of \( x_1 \). Then we must have \( m_1 + \cdots + m_{n_1} = m \). Thus \( m_{n_1+1} = \cdots = m_k = 0 \). Let \( S_2 = \{a_1, \ldots, a_{n_2}\} \) be the set of elements of \( S_1 \) with non-zero powers of \( x_2 \). Then \( m_1 + \cdots + m_{n_2} = m \) and \( m_{n_2+1} = m_{n_1} = 0 \). Continuing we get \( m = m_1 \) and \( m_i = 0 \) for all \( i \neq 1 \). Thus \( a_1^m = x_1^m \cdots x_i^m \) is the largest term in lexicographical order. \( \square \)

Lemma 3: If \( a_1, a_2, a_3 \in K[x_1, \ldots, x_n] \) are monomials with \( a_1 < a_2 \) then \( a_1 a_3 < a_2 a_3 \).

Proof: Let \( a_1 = x_1^{d_1} \cdots x_n^{d_n}, a_2 = x_1^{e_1} \cdots x_n^{e_n}, a_3 = x_1^{f_1} \cdots x_n^{f_n}, \) and let \( i \) be the least integer such that \( d_i < e_i \). Then if \( j < i, d_j + f_j = e_j + f_j \), and \( d_i + f_i < e_i + f_i \) since \( d_i < e_i \). Thus \( a_1 a_3 < a_2 a_3 \). \( \square \)

Lemma 4: \( \sigma_1^{a_1} \sigma_2^{a_2} \cdots \sigma_n^{a_n} = \sigma_1^{a_1+\cdots+\alpha_n} x_2^{a_2+\cdots+\alpha_n} x_n^{a_n} \) + terms of lower degree.

Proof: Let \( \beta \) be the sum of the terms of lower degree in \( \sigma_1^{a_1} \cdots \sigma_n^{a_n} \). We will proceed
by induction on $n$. If $n = 1$, this is Lemma 2 and we are done. Now assume this is true for $n - 1$. Then we have:

\[
\sigma_1^{a_1} = x_1^{a_1} + \beta_1
\]
\[
\sigma_2^{a_2} = x_1^{a_2} x_2^{a_2} + \beta_2
\]
\[
\vdots
\]
\[
\sigma_{n-1}^{a_{n-1}} = x_1^{a_{n-1}} x_2^{a_{n-1}} \cdots x_{n-1}^{a_{n-1}} + \beta_{n-1}.
\]

By our inductive hypothesis

\[
\prod_{1}^{n-1} \sigma_i^{a_i} = x_1^{\sum a_i} x_2^{\sum a_i} \cdots x_{n-1}^{\sum a_i} + \beta
\]

where $\beta$ is a sum of terms of lower degree in lexicographical order. We have by Lemma 2,

\[
\sigma_n^{a_n} = \sum x_1^{a_n} \cdots x_n^{a_n} + \beta_n.
\]

Multiplying both sides by $\sigma_n^{a_n}$ we get:

\[
\prod_{1}^{n} \sigma_i^{a_i} = \left(x_1^{a_{n-1}} x_2^{a_{n-1}} \cdots x_{n-1}^{a_{n-1}} + \beta\right) \left(x_1^{a_1} \cdots x_n^{a_n} + \beta\right)
\]
\[
= x_1^{a_1+\cdots+a_n} x_2^{a_2+\cdots+a_n} \cdots x_n^{a_n} + \beta x_1^{a_1} \cdots x_n^{a_n} + \beta x_1^{a_1} \cdots x_n^{a_n} + \beta \beta_n.
\]

By Lemma 3, the terms of $\beta_n x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} + \beta x_1^{a_1} \cdots x_n^{a_n} + \beta \beta_n$ are lower in degree than $x_1^{a_1+\cdots+a_n} x_2^{a_2+\cdots+a_n} \cdots x_n^{a_n}$. $\Box$
Theorem 5: Let \( f \in F[x_1, \ldots, x_n] \) such that a monomial in \( f \) of highest degree is \( a_1 x_1^{e_1} x_2^{e_2} \cdots x_n^{e_n} \). Then \( f - a_1 \sigma_1^{e_1 - e_2} \cdots \sigma_{n-1}^{e_{n-1} - e_n} \sigma_n^{e_n} < f \).

Proof: From Lemma 2 we have:

\[
\sigma_1^{e_1 - e_2} \cdots \sigma_n^{e_n} = x_1^{e_1} \cdots x_n^{e_n} + \beta
\]

where \( \beta \) is a sum of monomials of lower degree. By hypothesis, \( f(x) = a_1 x_1^{e_1} \cdots x_n^{e_n} + \gamma \) where \( \gamma \) is a sum of monomials of lower degree. Now,

\[
f - a_1 \sigma_1^{e_1 - e_2} \cdots \sigma_n^{e_n} = \gamma - \beta
\]

so \( f - a_1 \sigma_1^{e_1 - e_2} \cdots \sigma_{n-1}^{e_{n-1} - e_n} \sigma_n^{e_n} < f \), in lexicographic order. \( \square \)

Thus Algorithm 1 will eventually terminate since each iteration produces a polynomial which is smaller in lexicographic order.

We will calculate the discriminant of a cubic polynomial using Algorithm 1.

Let \( p(x) = x^3 + bx^2 + cx + d \)

\[
= (x - \alpha_1)(x - \alpha_2)(x - \alpha_3).
\]

The discriminant of \( p(x) \) is:

\[
\sqrt{D} = (\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)(\alpha_2 - \alpha_3)
\]

\[
D = [(\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)(\alpha_2 - \alpha_3)]^2.
\]
When we expand $D$ we have

$$D = a_1^4a_2^2 + a_1^4a_3^2 + a_1^2a_3^4 + a_2^2a_3^4 + a_2^2a_1^2 - 2a_1^4a_2a_3 - 2a_1a_2a_3^4 - 2a_2^3a_1a_3$$

$$-2a_1^3a_3^2 - 2a_2^3a_3^2 - 2a_1^2a_3^2 + 2a_1a_3^2a_3 + 2a_1^2a_2a_3 + 2a_2a_1^2a_3 + 2a_2^3a_1a_3$$

$$+2a_2a_3^3a_1 + 2a_3^2a_1a_3 - 6a_1^2a_2^2a_3^2.$$ Collecting terms we have

$$D = \sum a_1^4a_2^2 - 2 \sum a_1^4a_2a_3 - 2 \sum a_1^3a_3^2 + 2 \sum a_1^3a_2^2a_3 - 6 \sum a_1^2a_2^2a_3^2.$$ Now use Algorithm 1 to express $D$ as a polynomial in $\sigma_1, \sigma_2, \sigma_3$:

$$D = \sum a_1^4a_2^2 - 2 \sum a_1^4a_2a_3 - 2 \sum a_1^3a_3^2 + 2 \sum a_1^3a_2^2a_3 - 6 \sum a_1^2a_2^2a_3^2.$$ According to Algorithm 1 we subtract $a_1^2\sigma_1^{\alpha_1}a_2^{\alpha_2}a_3^{\alpha_3}$...$\sigma_n^{\alpha_n} = \sigma_1^{4-2\alpha_1} \sigma_2^{2-\alpha_2} \sigma_3^{2-\alpha_3}$.

Using Maple we find:

$$\sigma_1^2\sigma_2^2 = (a_1 + a_2 + a_3)^2 (a_1a_2 + a_1a_3 + a_2a_3)^2$$

$$= a_1^4a_2^2 + a_1^4a_3^2 + 2a_1^2a_3^4 + a_2^4a_3^2 + 2a_1^2a_3^2a_3 + 2a_1a_3^2a_3^3 + a_3^4a_1^2 + a_3^4a_2^2 + 2a_1^4a_2a_3 + 8a_1^2a_2a_3^2 + 8a_1^2a_2a_3^2 + 8a_1^2a_3^2a_3 + 15a_1^2a_2^2a_3^2 + 2a_1a_2^4a_3 + 8a_1a_3^2a_3^2 + 8a_1^2a_3^2a_3a_2 + 8a_1a_2^3a_3^2 + 2a_3^4a_1a_2.$$ Collecting terms:

$$\sigma_1^2\sigma_2^2 = \sum a_1^4a_2^2 + 2 \sum a_1^4a_2a_3 + 2 \sum a_1^3a_3^2 + 8 \sum a_1^3a_2^2a_3 + 15 \sum a_1^2a_2^2a_3^2.$$
We now have
\[
D = 4 \sum \alpha_1^4 \alpha_2^2 - 2 \sum \alpha_1^4 \alpha_2 \alpha_3 - 2 \sum \alpha_1^3 \alpha_2^3 + 2 \sum \alpha_1^3 \alpha_2 \alpha_3 - 6 \sum \alpha_1^2 \alpha_2^2 \alpha_3
\]
\[
-\sigma_1^2 \sigma_2^2 = - \sum \alpha_1^4 \alpha_2^2 - 2 \sum \alpha_1^4 \alpha_2 \alpha_3 - 2 \sum \alpha_1^3 \alpha_2^3 - 8 \sum \alpha_1^3 \alpha_2^2 \alpha_3 - 15 \sum \alpha_1^2 \alpha_2^2 \alpha_3
\]
\[
D - \sigma_1^2 \sigma_2^2 = -4 \sum \alpha_1^4 \alpha_2 \alpha_3 - 4 \sum \alpha_1^3 \alpha_2^3 - 6 \sum \alpha_1^3 \alpha_2 \alpha_3 - 21 \sum \alpha_1^2 \alpha_2^2 \alpha_3.
\]
Next we substitute
\[
-4 \sigma_1^{a_1-a_2} \sigma_2^{a_2-a_3} \ldots \sigma_n^{a_n} = -4 \sigma_1^{2} \sigma_2^{0} \sigma_3
\]
since our leading term is $-4 \alpha_1^4 \alpha_2 \alpha_3$. Then we have:
\[
4 \sigma_1^3 \sigma_3 = 4 (\alpha_1 + \alpha_2 + \alpha_3)^3 \cdot (\alpha_1 \alpha_2 \alpha_3)
\]
\[
= \alpha_1^4 \alpha_2 \alpha_3 + 3 \alpha_1^3 \alpha_2^2 \alpha_3 + 3 \alpha_1^3 \alpha_2 \alpha_3^2 + 3 \alpha_1^2 \alpha_2^3 \alpha_3 + 6 \alpha_1^2 \alpha_2 \alpha_3^2 + 3 \alpha_1 \alpha_2^3 \alpha_3^2 + 3 \alpha_1 \alpha_2^2 \alpha_3^2 + \alpha_1^4 \alpha_2 \alpha_3.
\]
Collecting terms:
\[
4 \sigma_1^3 \sigma_3 = 4 \sum \alpha_1^4 \alpha_2 \alpha_3 + 12 \sum \alpha_1^3 \alpha_2^2 \alpha_3 + 24 \sum \alpha_1^2 \alpha_2 \alpha_3
\]
\[
D - \sigma_1^2 \sigma_2^2 = -4 \sum \alpha_1^4 \alpha_2 \alpha_3 - 4 \sum \alpha_1^3 \alpha_2^3 - 6 \sum \alpha_1^3 \alpha_2 \alpha_3 - 21 \sum \alpha_1^2 \alpha_2^2 \alpha_3.
\]
\[
4 \sigma_1^3 \sigma_3 = 4 \sum \alpha_1^4 \alpha_2 \alpha_3 + 12 \sum \alpha_1^3 \alpha_2^2 \alpha_3 + 24 \sum \alpha_1^2 \alpha_2 \alpha_3
\]
\[
D - \sigma_1^2 \sigma_2^2 + 4 \sigma_1^3 \sigma_3 = -4 \sum \alpha_1^3 \alpha_2^3 + 6 \sum \alpha_1^3 \alpha_2 \alpha_3 + 3 \sum \alpha_1^2 \alpha_2^2 \alpha_3.
\]
Repeating: $a \sigma_1^{a_1-a_2} \sigma_2^{a_2-a_3} \ldots \sigma_n^{a_n} = \sigma_1^{2} \sigma_2^{0}$ and
\[
4 \sigma_2^2 = 4 (\alpha_1 \alpha_2 + \alpha_1 \alpha_3 + \alpha_2 \alpha_3)^2
\]
\[
4\alpha_1^2\alpha_2^2 + 8\alpha_1^2\alpha_2\alpha_3 + 8\alpha_1\alpha_2^2\alpha_3 + 4\alpha_1^2\alpha_3^2 + 8\alpha_1\alpha_3^2\alpha_2 + 4\alpha_2^2\alpha_3^2 \\
= 4 \sum \alpha_1^2\alpha_2^2 + 12 \sum \alpha_1^2\alpha_2\alpha_3 + 24 \sum \alpha_1\alpha_2^2\alpha_3.
\]

Subtracting we get:

\[
D - \sigma_1^2\sigma_2^2 + 4\sigma_1^2\sigma_3 = -4 \sum \alpha_1^2\alpha_2^2 + 6 \sum \alpha_1^2\alpha_2\alpha_3 + 3 \sum \alpha_1^2\alpha_2\alpha_3
\]
\[
4\sigma_2^2 = 4 \sum \alpha_1^2\alpha_2^2 + 12 \sum \alpha_1^2\alpha_2\alpha_3 + 24 \sum \alpha_1^2\alpha_2\alpha_3
\]
\[
D - \sigma_1^2\sigma_2^2 + 4\sigma_1^2\sigma_3 + 4\sigma_2^2 = 18 \sum \alpha_1^2\alpha_2\alpha_3 + 27 \sum \alpha_1^2\alpha_2\alpha_3.
\]

Again:
\[
a\sigma_1^{a_1-a_2}\sigma_2^{a_2-a_3}...\sigma_n^{a_n} = 18\sigma_1\sigma_2\sigma_3
\]
\[
-18\sigma_1\sigma_2\sigma_3 = -18 (\alpha_1 + \alpha_2 + \alpha_3) (\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3) (\alpha_1\alpha_2\alpha_3)
\]
\[
= -18\alpha_1^3\alpha_2\alpha_3 - 18\alpha_1^2\alpha_2^2\alpha_3 - 54\alpha_1^2\alpha_2\alpha_3^2 - 18\alpha_1\alpha_2^2\alpha_3^2
\]
\[
-18\alpha_1\alpha_2\alpha_3^3 - 18\alpha_1^3\alpha_2\alpha_3 - 18\alpha_1^2\alpha_2^2\alpha_3.
\]

Collecting terms:
\[
-18\sigma_1\sigma_2\sigma_3 = -18 \sum \alpha_1^3\alpha_2\alpha_3 - 54 \sum \alpha_1^2\alpha_2\alpha_3^2
\]
\[
D - \sigma_1^2\sigma_2^2 + 4\sigma_1^2\sigma_3 + 4\sigma_2^2 = 18 \sum \alpha_1^3\alpha_2\alpha_3 + 27 \sum \alpha_1^2\alpha_2\alpha_3^2
\]
\[
-18\sigma_1\sigma_2\sigma_3 = -18 \sum \alpha_1^3\alpha_2\alpha_3 - 54 \sum \alpha_1^2\alpha_2\alpha_3^2
\]
\[
D - \sigma_1^2\sigma_2^2 + 4\sigma_1^2\sigma_3 + 4\sigma_2^2 - 18\sigma_1\sigma_2\sigma_3 = -27 \sum \alpha_1^2\alpha_2\alpha_3^2.
\]

Finally:
\[
a\sigma_1^{a_1-a_2}\sigma_2^{a_2-a_3}...\sigma_n^{a_n} = a\sigma_1^0\sigma_2^0\sigma_3^0
\]
Thus,

\[ D = \sigma_1^2 \sigma_2^2 - 4 \sigma_1^3 \sigma_3 - 4 \sigma_2^3 + 18 \sigma_1 \sigma_2 \sigma_3 - 27 \sigma_3^2 \]

\[ = b^2 c^2 - 4b^3 d - 4c^3 + 18bcd - 27d^2 \]

where we have substituted \( \sigma_1 = -b, \sigma_2 = c, \sigma_3 = -d \).

Now we will find the solutions to the equation \( p(x) = x^3 + bx^2 + cx + d = 0 \). We will assume \( b, c, \) and \( d \) are elements of \( F \), a number field containing the cube roots of unity \( 1, \omega, \omega^2 \). Since \( \omega = -\frac{1}{2} + \frac{1}{2} \sqrt{-3} \) and \( \omega^2 = -\frac{1}{2} - \frac{1}{2} \sqrt{-3} \), it is enough that \( F \) contain \( \sqrt{-3} \). Let \( E \) denote the splitting field of \( p(x) \), then \( E \) is a Galois extension of \( F \), and \( G(E/F) \) is isomorphic to \( S_3 \), the permutation group of the roots, \( \alpha_1, \alpha_2, \alpha_3 \) of \( p(x) \).

\( S_3 \) is a solvable group, and the composition series

\[ S_3 \triangleright A_3 \triangleright \{e\} \]

has cyclic factors \( A_3 \) of order 3, and \( S_3/A_3 \) of order 2. The tower of fields is:

\[ F \triangleleft F_1 \triangleleft F_2. \]

where \( F_1 = F \left( \sqrt{D} \right) \) (See Theorem 16), and \( F_2 \) is the splitting field of \( p(x) \). We define the Lagrange resolvents:

\[ f(1) = \alpha_1 + \alpha_2 + \alpha_3 \]
\[ f(\omega) = \alpha_1 + \omega \alpha_2 + \omega^2 \alpha_3 \]
\[ f(\omega^2) = \alpha_1 + \omega^2 \alpha_2 + \omega \alpha_3. \]

Then,
\[ 3\alpha_1 = f(1) + f(\omega) + f(\omega^2) \]
\[ 3\alpha_2 = f(1) + \omega f(\omega) + \omega f(\omega^2) \]
\[ 3\alpha_3 = f(1) + \omega f(\omega) + \omega^2 f(\omega^2). \]

Thus,
\[ F_2 = F_1\left( f(1), f(\omega), f(\omega^2) \right) = F_1\left( f(\omega), f(\omega^2) \right) \]

since \( f(1) = -b \in F \).

We also have
\[ f(\omega) \cdot f(\omega^2) = \left( \alpha_1 + \omega \alpha_2 + \omega^2 \alpha_3 \right) \left( \alpha_1 + \omega^2 \alpha_2 + \omega \alpha_3 \right) \]
\[ = \alpha_1^2 + \alpha_1 \omega^2 \alpha_2 + \alpha_1 \omega \alpha_3 + \omega \alpha_2 \alpha_1 + \omega^3 \alpha_2^2 + \omega^2 \alpha_2 \alpha_3 + \omega \alpha_3 \alpha_1 + \omega^4 \alpha_3 \alpha_2 + \omega^3 \alpha_3^2 \]

is symmetric hence, \( f(\omega), f(\omega^2) \in F \). Thus \( f(\omega^2) \in F (f(\omega)) \) and \( F_2 = F_1 (f(\omega)) \).

To find the roots of \( p(x) \) we must find \( f(\omega) \) and \( f(\omega^2) \). Now
\[ (f(\omega))^3 = \left( \alpha_1 + \omega \alpha_2 + \omega^2 \alpha_3 \right)^3 \]
\[ = \alpha_1^3 + 3\omega \alpha_2 \alpha_1^2 + 3\omega^2 \alpha_3 \alpha_1^2 + 3\alpha_1 \omega^2 \alpha_2^2 + 6\alpha_1 \omega \alpha_2 \alpha_3 + 3\alpha_1 \omega \alpha_4 \alpha_3^2 + \omega^3 \alpha_2^3 + 3\omega^4 \alpha_2 \alpha_3 + 3\omega^5 \alpha_2 \alpha_3 + \omega^6 \alpha_3^3 \]
= \alpha_1^3 + \alpha_2^3 + \alpha_3^3 + 3\omega (\alpha_1^2\alpha_2 + \alpha_2^2\alpha_3 + \alpha_1\alpha_3^2) + 3\omega^2 (\alpha_1\alpha_2^2 + \alpha_1^2\alpha_3 + \alpha_2\alpha_3^2) + 6\alpha_1\alpha_2\alpha_3.

Now substitute

\begin{align*}
\omega &= \frac{1}{2} + \frac{1}{2}\sqrt{-3} \\
\omega^2 &= \frac{1}{2} - \frac{1}{2}\sqrt{-3}
\end{align*}

to get:

\[(f(\omega))^3 = \sum \alpha_1^3 - \frac{3}{2} \sum \alpha_1^2\alpha_2 + i3\sqrt{3} (\alpha_1^2\alpha_2 + \alpha_2^2\alpha_3 + \alpha_1\alpha_3^2 - \alpha_1\alpha_2^2 - \alpha_1^2\alpha_3 - \alpha_2\alpha_3^2) =
\sum \alpha_1^3 - \frac{3}{2} \sum \alpha_1^2\alpha_2 + 6\alpha_1\alpha_2\alpha_3 + \sqrt{-27D}.
\]

Now we will solve \(\sum \alpha_1^2\alpha_2\). We use Algorithm 1 to express \(\sum \alpha_1^3\) and \(\sum \alpha_1^2\alpha_2\) in terms of \(b, c, d\).

\[\sigma_1\sigma_2 = (\alpha_1 + \alpha_2 + \alpha_3)(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3)\]

\[= \alpha_1^2\alpha_2 + \alpha_1^2\alpha_3 + 3\alpha_1\alpha_2\alpha_3 + \alpha_1\alpha_2^2 + \alpha_2\alpha_3^2 + \alpha_1\alpha_3^2 + \alpha_2\alpha_3^2\]

\[\sum \alpha_1^2\alpha_2 = \sum \alpha_1^2\alpha_2 \]

\[-\sigma_1\sigma_2 = -\sum \alpha_1^2\alpha_2 - 3\sum \alpha_1\alpha_2\alpha_3\]

\[\sum \alpha_1^2\alpha_2 - \sigma_1\sigma_2 = 3\sum \alpha_1\alpha_2\alpha_3.\]

Then using Algorithm 1, \(a_1^{\sigma_1^1}a_2^{\sigma_2^2}a_3^{\sigma_3^3}...a_n^{\sigma_n^n} = 3\sigma_3 = 3\sum \alpha_1\alpha_2\alpha_3\)

\[\sum \alpha_1^2\alpha_2 - \sigma_1\sigma_2 = -3\sum \alpha_1\alpha_2\alpha_3 + 3\sigma_3 = 3\sum \alpha_1\alpha_2\alpha_3\]

26
\[
\sum \alpha_1^2 \alpha_2 - \sigma_1 \sigma_2 + 3 \sigma_3 = 0.
\]

Therefore, \( \sum \alpha_1^2 \alpha_2 = \sigma_1 \sigma_2 - 3 \sigma_3 \). Since \( \sigma_1 = -b \), \( \sigma_2 = c \) and \( \sigma_3 = -d \),

\[
\sum \alpha_1^2 \alpha_2 = 3d - bc.
\]

Similarly, we convert \( \sum \alpha_1^3 \) to a polynomial in \( b, c, d \).

Using Algorithm 1, \( \sigma_1^3 \sigma_2 \sigma_3 \ldots \sigma_n^3 \) = \( \sigma_1^3 \)

\[
\sigma_1^3 = (\alpha_1 + \alpha_2 + \alpha_3)^3
\]

\[
= \alpha_1^3 + 3\alpha_1^2\alpha_2 + 3\alpha_1^2\alpha_3 + 3\alpha_1\alpha_2^2 + 6\alpha_1\alpha_2\alpha_3 + 3\alpha_1\alpha_3^2 + \alpha_2^3 + 3\alpha_2\alpha_3^2 + \alpha_3^3
\]

\[
= \sum \alpha_1^3 + 3 \sum \alpha_1^2 \alpha_2 + 6 \alpha_1 \alpha_2 \alpha_3
\]

then,

\[
\sum \alpha_1^3 = \sum \alpha_1^3
\]

\[
-\sigma_1^3 = -3 \sum \alpha_1^2 \alpha_2 - 6 \alpha_1 \alpha_2 \alpha_3
\]

\[
\sum \alpha_1^3 - \sigma_1^3 = -3 \sum \alpha_1^2 \alpha_2 - 6 \alpha_1 \alpha_2 \alpha_3.
\]

Again, \( \sigma_1^3 \sigma_2 \sigma_3 \ldots \sigma_n^3 \) = \( 3 \sigma_1 \sigma_2 \). We have

\[
3 \sigma_1 \sigma_2 = 3 (\alpha_1 + \alpha_2 + \alpha_3) (\alpha_1 \alpha_2 + \alpha_1 \alpha_3 + \alpha_2 \alpha_3)
\]

\[
= 3 \alpha_1^2 \alpha_2 + 3 \alpha_1^2 \alpha_3 + 9 \alpha_1 \alpha_2 \alpha_3 + 3 \alpha_1 \alpha_2^2 + 3 \alpha_2 \alpha_3^2 + 3 \alpha_1 \alpha_3^2 + 3 \alpha_2 \alpha_3^2
\]

27
\[ \sum \alpha_1^3 - \sigma_1^3 = -3\sigma_1^2\alpha_2 - 6\alpha_1\alpha_2\alpha_3 \]
\[ 3\sigma_1\sigma_2 = 3\sigma_1^2\alpha_2 + 9\alpha_1\alpha_2\alpha_3 \]
\[ \sum \alpha_1^3 - \sigma_1^3 + 3\sigma_1\sigma_2 = 3\alpha_1\alpha_2\alpha_3 \]

and thus \[ \sum \alpha_1^3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3 = -b^3 + 3bc - 3d. \]

Therefore,
\[
(f(\omega))^3 = -b^3 + 3bc - 3d - \frac{3}{2}(3d - bc) - 6d + \frac{1}{2}\sqrt{-27D} \\
= -b^3 + \frac{9}{2}bc - \frac{27}{2}d + \frac{1}{2}\sqrt{-27D}.
\]

Similarly,
\[
(f(\omega^2))^3 = (\alpha_1 + \alpha_2\omega^2 + \alpha_3\omega)^3 \\
= -b^3 + \frac{9}{2}bc - \frac{27}{3}d + \frac{1}{2}\sqrt{-27D}.
\]

Finally we have,
\[
\alpha_1 = \frac{1}{3} \left( -b + \sqrt[3]{-b^3 + \frac{9}{2}bc - \frac{27}{2}d + \frac{1}{2}\sqrt{-27D}} + \sqrt[3]{-b^3 + \frac{9}{2}bc - \frac{27}{2}d - \frac{1}{2}\sqrt{-27D}} \right) \\
\alpha_2 = \frac{1}{3} \left( -b + \omega\sqrt[3]{-b^3 + \frac{9}{2}bc - \frac{27}{2}d + \frac{1}{2}\sqrt{-27D}} + \omega^2\sqrt[3]{-b^3 + \frac{9}{2}bc - \frac{27}{2}d - \frac{1}{2}\sqrt{-27D}} \right) \\
\alpha_3 = \frac{1}{3} \left( -b + \omega^2\sqrt[3]{-b^3 + \frac{9}{2}bc - \frac{27}{2}d + \frac{1}{2}\sqrt{-27D}} + \omega^3\sqrt[3]{-b^3 + \frac{9}{2}bc - \frac{27}{2}d - \frac{1}{2}\sqrt{-27D}} \right)
\]

which is Cardano’s solution to the cubic equation.

**Solutions to Quartic Polynomials Using Symmetric Functions**

We will now look at the general solution for the quartic polynomial. Let
\[ f(z) = z^4 + bz^3 + cz^2 + dz + e = 0 \]. Let \( E \) denote the splitting field over \( F \) a number
field containing $b, c, d, e,$ and the cube roots of unity. Since $\alpha_1, \ldots, \alpha_n$ are algebraically independent $G(E/F)$ contains all the permutations of the roots $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ of $f(z)$.

So $G(E/F) \cong S_4$ and $[E : F] = |S_4| = 24$

$S_4$ has a normal subgroup $A_4$, and $A_4$ has a normal subgroup $N$ which contains $N = \{(e), (12)(34), (13)(24), (14)(23)\}$. $N$ is abelian and therefore the subgroup $K = \{(e), (12)(34)\}$ is normal in $N$. Thus $S_4$ has a composition series:

$$S_4 \triangleright A_4 \triangleright N \triangleright K \triangleright \{1\}.$$

We will now compute the fields in the normal tower

$$F < \beta_1 < \beta_2 < \beta_3 < E$$

where $\beta_1, \beta_2, \beta_3$, are the fixed fields of $A_4, N, K$, respectively. $\beta_1 = F(\sqrt{D})$ (see Theorem 16) where

$$D = (\alpha_1 - \alpha_2)^2 (\alpha_1 - \alpha_3)^2 (\alpha_1 - \alpha_4)^2 (\alpha_2 - \alpha_3)^2 (\alpha_2 - \alpha_4)^2 (\alpha_3 - \alpha_4)^2$$

is the discriminant of $f(z)$. Note that $D \subset F_{S_4} = F$. It is not necessary with the quartic to compute $D$ as it falls out later. The general equation of the fourth degree

$$z^4 + bz^3 + cz^2 + dz + e = 0$$

can be transformed into

$$x^4 + px^2 + qx + r = 0$$
by the substitution $z = x - \frac{1}{4}b$.

To compute $\beta_2$, there is an element $\theta_1 = \alpha_1\alpha_2 + \alpha_3\alpha_4$ which is left fixed by $N$ and the permutations $(12), (34), (1324)$, and $(1423)$. Thus $\theta_1 \in B_2$, the fixed field of $N$. All of the permutations of $S_4$ applied to $\theta_1$ yield the following conjugates:

$$
\theta_1 = (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4)
$$

$$
\theta_2 = (\alpha_1 + \alpha_3)(\alpha_2 + \alpha_4)
$$

$$
\theta_3 = (\alpha_1 + \alpha_4)(\alpha_2 + \alpha_3).
$$

Now,

$$
G(E/F(\theta_1)) = N \cup \{(12), (34)\}
$$

$$
G(E/F(\theta_2)) = N \cup \{(13), (24)\}
$$

$$
G(E/F(\theta_3)) = N \cup \{(14), (23)\}.
$$

Therefore,

$$
G(E/F(\theta_1, \theta_2, \theta_3)) = N.
$$

So the elements of $N$ are the only permutations leaving fixed all three numbers $\theta_1$, $\theta_2$, and $\theta_3$. Because of this $\beta_2 = \beta_1(\theta_1, \theta_2, \theta_3)$. These conjugates are the roots of an
equation of the third degree.

\[ \theta^3 - B_1\theta^2 + B_2\theta - B_3 = 0. \]

The \( b_i \) are just the elementary symmetric functions of \( \theta_1, \theta_2, \theta_3 \):

\[
B_1 = \theta_1 + \theta_2 + \theta_3
\]

\[
B_2 = \theta_1\theta_2 + \theta_1\theta_3 + \theta_2\theta_3
\]

\[
B_3 = \theta_1\theta_2\theta_3.
\]

Thus,

\[
B_1 = (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4) + (\alpha_1 + \alpha_3)(\alpha_2 + \alpha_4) + (\alpha_1 + \alpha_4)(\alpha_2 + \alpha_3)
= 2\alpha_1\alpha_3 + 2\alpha_1\alpha_4 + 2\alpha_2\alpha_3 + 2\alpha_2\alpha_4 + 2\alpha_1\alpha_2 + 2\alpha_3\alpha_4
\]

\[
B_1 = 2\sum \alpha_1\alpha_2 = 2p
\]

\[
B_2 = (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4)(\alpha_1 + \alpha_3)(\alpha_2 + \alpha_4) + (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4)(\alpha_1 + \alpha_4)(\alpha_2 + \alpha_3) + (\alpha_1 + \alpha_3)(\alpha_2 + \alpha_4)(\alpha_1 + \alpha_4)(\alpha_2 + \alpha_3)
= \alpha_1^2\alpha_3\alpha_2 + \alpha_1^2\alpha_3\alpha_4 + \alpha_1^2\alpha_2\alpha_4 + \alpha_1^2\alpha_4\alpha_2 + \alpha_1^2\alpha_4^2 + 2\alpha_1\alpha_4\alpha_3\alpha_2 + \alpha_1\alpha_4^2\alpha_3
\]

\[
B_2 = \alpha_2^2\alpha_3\alpha_1 + \alpha_2^2\alpha_3\alpha_4 + \alpha_2^2\alpha_4\alpha_1 + \alpha_2^2\alpha_4\alpha_3 + \alpha_2\alpha_4\alpha_3 + \alpha_3\alpha_2\alpha_3 + \alpha_3\alpha_2^2 + 2\alpha_1\alpha_4\alpha_3\alpha_2 + \alpha_1\alpha_3^2\alpha_4 + \alpha_1^2\alpha_4\alpha_2 + \alpha_1\alpha_4\alpha_3 + \alpha_1\alpha_4\alpha_2 + \alpha_1\alpha_3\alpha_2 + \alpha_1\alpha_3^2 + \alpha_1^2\alpha_3\alpha_2
\]

31
\[
\begin{align*}
\alpha_1^2\alpha_4\alpha_2 + \alpha_1^2\alpha_2\alpha_4 + \alpha_2\alpha_4^2\alpha_1 + \alpha_1\alpha_4^2\alpha_3 + \alpha_2^2\alpha_3\alpha_1 + \alpha_1\alpha_3^2\alpha_2 + \alpha_2^2\alpha_4\alpha_3 + \alpha_2\alpha_3^2\alpha_4 + \\
\alpha_1\alpha_2^2\alpha_4 + \alpha_2\alpha_4^2\alpha_3 + \alpha_2^2\alpha_4^2
\end{align*}
\]

Thus,
\[
B_2 = \sum \alpha_1^2\alpha_2^2 + 3 \sum \alpha_1^2\alpha_2\alpha_3 + 6\alpha_1\alpha_2\alpha_3\alpha_4.
\]

Similarly,
\[
B_3 = \sum \alpha_1^3\alpha_2 + 2 \sum \alpha_1^3\alpha_2\alpha_3 + 2 \sum \alpha_1^3\alpha_2\alpha_3 + 2 \sum \alpha_1^3\alpha_2\alpha_3 + 4 \sum \alpha_1^3\alpha_2\alpha_3
\]

Thus,
\[
B_3 = \sum \alpha_1^3\alpha_2^2\alpha_3 + 2 \sum \alpha_1^3\alpha_2\alpha_3\alpha_4 + 2 \sum \alpha_1^2\alpha_2\alpha_3^2 + 4 \sum \alpha_1^2\alpha_2\alpha_3\alpha_4 + 4 \sum \alpha_1^2\alpha_2\alpha_3\alpha_4
\]

\[
B_2 \text{ and } B_3 \text{ can be expressed in terms of the elementary symmetric functions } \sigma_1, \sigma_2, \sigma_3, \sigma_4, \text{ of the } \alpha_i:
\]
\[
B_2 = \sum \alpha_1^2\alpha_2^2 + 3 \sum \alpha_1^2\alpha_2\alpha_3 + 6\alpha_1\alpha_2\alpha_3\alpha_4 = p^2 - 4r.
\]
Collecting terms:

\[ \delta \alpha_0 + \delta \gamma_0 + \delta \alpha_1 + \delta \gamma_1 + \delta \alpha_2 + \delta \gamma_2 + \delta \alpha_3 + \delta \gamma_3 = \delta \alpha_0 \]

Next we substitute \( \delta \alpha_0 - \delta \gamma_0 \) into:

\[ \delta \alpha_0 + \delta \gamma_0 \rightarrow \delta \alpha_0 - \delta \gamma_0 \]

\[ \delta \alpha_0 - \delta \gamma_0 \]

Collecting terms:

\[ \delta \alpha_0 + \delta \gamma_0 + \delta \alpha_1 + \delta \gamma_1 + \delta \alpha_2 + \delta \gamma_2 + \delta \alpha_3 + \delta \gamma_3 \]

Using algorithm we first substitute \( \delta \alpha_0 - \delta \gamma_0 \) into:
\[ B_2 - \sigma_2^2 = \sum \alpha_1^2 \alpha_2 \alpha_3 \]

\[ \sigma_1 \sigma_3 = \sum \alpha_1^2 \alpha_2 \alpha_3 + 4 \alpha_1 \alpha_2 \alpha_3 \alpha_4 \]

\[ B_2 - \sigma_2^2 - \sigma_1 \sigma_3 = -4 \alpha_1 \alpha_2 \alpha_3 \alpha_4. \]

Finally,

\[ \sigma_1^{\alpha_1 - \alpha_2} \sigma_2^{\alpha_2 - \alpha_3} \ldots \sigma_n^{\alpha_n} = -4 \sigma_4 = -4 \alpha_1 \alpha_2 \alpha_3 \alpha_4 \]

\[ B_2 - \sigma_2^2 - \sigma_1 \sigma_3 = -4 \alpha_1 \alpha_2 \alpha_3 \alpha_4 \]

\[ -4 \sigma_4 = -4 \alpha_1 \alpha_2 \alpha_3 \alpha_4 = -4r \]

\[ B_2 - \sigma_2^2 - \sigma_1 \sigma_3 + 4 \sigma_4 = 0 \]

and thus,

\[ b_2 = p^2 - 4r. \]

For \( B_3 \) recall,

\[ B_3 = \sum \alpha_1^3 \alpha_2 \alpha_3 + 2 \sum \alpha_1^3 \alpha_2 \alpha_3 \alpha_4 + 2 \sum \alpha_1^3 \alpha_2 \alpha_4 + 4 \sum \alpha_1^3 \alpha_2 \alpha_3 \alpha_4. \]

Using Algorithm 1 we first subtract \( \sigma_1^{\alpha_1 - \alpha_2} \sigma_2^{\alpha_2 - \alpha_3} \ldots \sigma_n^{\alpha_n} = \sigma_1 \sigma_2 \sigma_3 \)

\[ \sigma_1 \sigma_2 \sigma_3 = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4) (\alpha_1 \alpha_2 + \alpha_1 \alpha_3 + \alpha_1 \alpha_4 + \alpha_2 \alpha_3 + \alpha_2 \alpha_4 + \alpha_3 \alpha_4) \]

\[ (\alpha_1 \alpha_2 \alpha_3 + \alpha_1 \alpha_2 \alpha_4 + \alpha_1 \alpha_3 \alpha_4 + \alpha_2 \alpha_3 \alpha_4) \]

\[ = \alpha_1^3 \alpha_2 \alpha_3 + \alpha_1^3 \alpha_2 \alpha_4 + \alpha_1^3 \alpha_3 \alpha_4 + 3 \alpha_1^2 \alpha_2 \alpha_3 \alpha_4 + \alpha_2^2 \alpha_3 \alpha_4 + \alpha_3^2 \alpha_2 \alpha_3 \alpha_4 + 3 \alpha_1 \alpha_2 \alpha_3 \alpha_4 + 3 \alpha_1 \alpha_2 \alpha_3 \alpha_4 + \alpha_1 \alpha_2 \alpha_3 \alpha_4 + \alpha_1 \alpha_2 \alpha_3 \alpha_4 + \alpha_1 \alpha_2 \alpha_3 \alpha_4 + \alpha_1 \alpha_2 \alpha_3 \alpha_4. \]
\[ \begin{align*}
\alpha_1\alpha_4^2\alpha_3^2 &+ \alpha_1^2\alpha_2\alpha_3^2 + 3\alpha_2^2\alpha_3\alpha_4^2 + \alpha_2\alpha_3\alpha_4^2 + \alpha_1\alpha_2\alpha_4^2 + \alpha_1\alpha_3\alpha_2^2 + \alpha_1\alpha_3\alpha_4^2 + \alpha_1\alpha_4\alpha_2^2 + \\
\alpha_3\alpha_2\alpha_3 &+ \alpha_2\alpha_3\alpha_2^2 + \alpha_1\alpha_3\alpha_2 + \alpha_2^2\alpha_3\alpha_4 + \alpha_1\alpha_3\alpha_2^2 + \alpha_2\alpha_3\alpha_2^2 + \alpha_2\alpha_3\alpha_2^2 + \alpha_3^2\alpha_2\alpha_3 + \\
\alpha_1^2\alpha_4^2 &+ \alpha_1^2\alpha_3^2 + \alpha_2^2\alpha_3\alpha_2^2 + \alpha_1\alpha_3\alpha_2 + \alpha_2\alpha_3\alpha_2^2 + \alpha_3\alpha_2\alpha_3 + \\
\alpha_1\alpha_2\alpha_4^2 &+ \alpha_1\alpha_3\alpha_4^2 + 3\alpha_1\alpha_3\alpha_4^2 + 2\alpha_1^2\alpha_3\alpha_2^2 + 3\alpha_1\alpha_3\alpha_2^2 + 2\alpha_1\alpha_3\alpha_4^2 + 3\alpha_1\alpha_3\alpha_4^2 + 2\alpha_1\alpha_3\alpha_4^2 + 8\alpha_1\alpha_3\alpha_4^2 + 2\alpha_1\alpha_3\alpha_4^2.
\end{align*} \]

Collecting terms:

\[ \begin{align*}
\sigma_1\sigma_2\sigma_3 &= \sum \alpha_1^2\alpha_2^2\alpha_3 + 3\sum \alpha_1\alpha_2\alpha_3\alpha_4 + 3\sum \alpha_1\alpha_2\alpha_3^2 + 8\sum \alpha_1^2\alpha_2^2\alpha_3\alpha_4 \\
B_3 &= \sum \alpha_1^2\alpha_2\alpha_3 + 2\sum \alpha_1^3\alpha_2\alpha_3\alpha_4 + 2\sum \alpha_1^2\alpha_2\alpha_3^2 + 4\sum \alpha_1^2\alpha_2\alpha_3\alpha_4 \\
\sigma_1\sigma_2\sigma_3 &= \sum \alpha_1^2\alpha_2^2\alpha_3 + 3\sum \alpha_1^2\alpha_2\alpha_3\alpha_4 + 3\sum \alpha_1^2\alpha_2\alpha_3^2 + 8\sum \alpha_1^2\alpha_2^2\alpha_3\alpha_4 \\
B_3 - \sigma_1\sigma_2\sigma_3 &= -\sum \alpha_1^2\alpha_2\alpha_3\alpha_4 - \sum \alpha_1^2\alpha_2^2\alpha_3^2 - 4\sum \alpha_1^2\alpha_2^2\alpha_3\alpha_4.
\end{align*} \]

Next we subtract \( \sigma_1^2 \sigma_2^2 \sigma_3^2 \ldots \sigma_n^2 = -\sigma_1^2 \sigma_4 \)

\[ \begin{align*}
-\sigma_1^2 \sigma_4 &= -(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4)^2 (\alpha_1\alpha_2\alpha_3\alpha_4) \\
&= -\alpha_1^2\alpha_3\alpha_4\alpha_2 - 2\alpha_2^2\alpha_3\alpha_4 - 2\alpha_1\alpha_2\alpha_3^2\alpha_4 - 2\alpha_1\alpha_3\alpha_2^2\alpha_4 - \alpha_1\alpha_3^2\alpha_3\alpha_4 - 2\alpha_1\alpha_3\alpha_2^2\alpha_4 - \\
&2\alpha_1\alpha_2\alpha_3\alpha_4 - \alpha_1\alpha_3^2\alpha_2\alpha_4 - 2\alpha_1\alpha_3\alpha_4^2\alpha_2 - \alpha_1\alpha_3\alpha_4^2\alpha_2.
\end{align*} \]

Collecting terms:

\[ \begin{align*}
-\sigma_1^2 \sigma_4 &= -\sum \alpha_1^3\alpha_2\alpha_3\alpha_4 - 2\sum \alpha_1^2\alpha_2\alpha_3\alpha_4 \\
B_3 - \sigma_1\sigma_2\sigma_3 &= -\sum \alpha_1^3\alpha_2\alpha_3\alpha_4 - \sum \alpha_1^2\alpha_2^2\alpha_3 + 4\sum \alpha_1^2\alpha_2^2\alpha_3\alpha_4 \\
-\sigma_1^2 \sigma_4 &= -\sum \alpha_1^3\alpha_2\alpha_3\alpha_4 - 2\sum \alpha_1^2\alpha_2\alpha_3\alpha_4 \\
B_3 - \sigma_1\sigma_2\sigma_3 + \sigma_1^2 \sigma_4 &= -\sum \alpha_1^3\alpha_2^2\alpha_3^2 - 2\sum \alpha_1^2\alpha_2\alpha_3\alpha_4.
\end{align*} \]
Using the algorithm: \( \sigma_1^{a_1-a_2} \sigma_2^{a_2-a_3} \ldots \sigma_n^{a_n} = -\sigma_3^2 \)

\[-\sigma_3^2 = -(a_1a_2a_3 + a_1a_2a_4 + a_1a_3a_4 + a_2a_3a_4)^2\]

\[= -\sigma_1^2a_2^2a_3 - 2\sigma_1^2a_2a_3a_4 - 2\sigma_1^2a_3^2a_4 - 2\sigma_1^2a_4^2a_4 - \sigma_1^2\sigma_2^2 - 2\sigma_1^2a_3a_4 - 2a_1a_2a_4 - a_1a_3a_4 - a_2a_3a_4 - a_3a_4^2 - 2a_1a_2a_3a_4a_2 - 2a_1a_2a_4a_2 - a_2a_3a_4^2 - 2a_1a_2a_3^2a_4 - 2a_1a_3a_4^2a_2 - a_2a_3a_4^2.\]

Collecting terms:

\[-\sigma_3^2 = -\sum \sigma_1^2a_2^2a_3^2 - 2\sum \sigma_1^2a_2a_3a_4 = -q^2\]

\[B_3 - \sigma_1\sigma_2\sigma_3 + \sigma_1^2\sigma_4 = -\sum \sigma_1^2a_2a_3 - 2\sum \sigma_1^2a_2a_3a_4 = -q^2\]

\[-\sigma_3^2 = -\sum \sigma_1^2a_2^2a_3^2 - 2\sum \sigma_1^2a_2a_3a_4 = -q^2\]

\[B_3 - \sigma_1\sigma_2\sigma_3 + \sigma_1^2\sigma_4 + \sigma_3^2 = 0.\]

So the equation \( \theta^3 - B_1\theta^2 + B_2\theta - B_3 = 0 \) becomes:

\[\theta^3 - 2\theta^2 + \left(p^2 - 4r\right)\theta + q^2 = 0.\]

This equation is known as the cubic resolvent of the equation of the fourth degree; and according to Cardan the root can be expressed as radicals. Each \( \theta \) is fixed under a group of eight permutations; the three of them together remain fixed under \( N \) hence \((\theta_1, \theta_2, \theta_3) = B_2\).

The field \( B_3 \) arises from \( B_2 \) by adjoining an element which is not invariant under all the substitutions in \( N \) but is only invariant under the identity and \((12)(24)\). One such
element is $\alpha_1 + \alpha_2$. We know that $(\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4) = \theta_1$ and $(\alpha_1 + \alpha_2) + (\alpha_3 + \alpha_4) = 0$

and so $(\alpha_1 + \alpha_2) = - (\alpha_3 + \alpha_4)$. Then by substitution,

$$\alpha_1 + \alpha_2 = \sqrt{-\theta_1} ; \quad \alpha_3 + \alpha_4 = -\sqrt{-\theta_1}$$

and similarly,

$$\alpha_1 + \alpha_3 = \sqrt{-\theta_2} ; \quad \alpha_2 + \alpha_4 = -\sqrt{-\theta_2}$$

$$\alpha_1 + \alpha_4 = \sqrt{-\theta_3} ; \quad \alpha_2 + \alpha_3 = -\sqrt{-\theta_3}.$$

Finally we have the solution for the general equation of the fourth degree:

$$2\alpha_1 = \sqrt{-\theta_1} + \sqrt{-\theta_2} + \sqrt{-\theta_3}$$

$$2\alpha_2 = \sqrt{-\theta_1} - \sqrt{-\theta_2} - \sqrt{-\theta_3}$$

$$2\alpha_3 = -\sqrt{-\theta_1} + \sqrt{-\theta_2} - \sqrt{-\theta_3}$$

$$2\alpha_4 = -\sqrt{-\theta_1} - \sqrt{-\theta_2} + \sqrt{-\theta_3}.$$

Since, for example

$$\alpha_1 = (\alpha_1 + \alpha_2) + (\alpha_1 + \alpha_3) + (\alpha_1 + \alpha_4)$$

$$= 2\alpha_1 + (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4)$$

$$= 2\alpha_1 + 0.$$
CHAPTER THREE

Non-Solvability of the Quintic Equation

Euler found a method for reducing the solution to the quartic equation into that of a cubic equation, and believed he could reduce the quintic equation to that of a quartic. Lagrange made several unsuccessful attempts along the same lines, and mathematicians began to suspect that a general solution to the quintic polynomial might be impossible. The first to give a convincing proof of this fact was Paolo Ruffini. Generally he is not credited with the proof because of gaps in his argument. Niels Abel goes down in history as being the first to provide a solid proof for the non solvability of the quintic equation, and the following is based on his original work.

The following terms for this section are defined as:

Definition: A bijection from a set \( A \) to itself is called a permutation on \( A \).

Definition: Let \( A = \{1, 2, \ldots, n\} \). The set of all permutations of \( A \) is called the symmetric group of degree \( n \) and is denoted \( S_n \).

Definition: An element \( f \) of \( S_n \) is a cycle if there exists a set \( \{i_1, i_2, \ldots, i_n\} \) of distinct integers such that \( f(i_1) = i_2, f(i_2) = i_3, \ldots, f(i_{n-1}) = i_n, f(i_n) = i_1 \), and \( f \) leaves all other elements fixed.

Definition: A cycle of length 2 is called a transposition.
Definition: A finite sequence of fields, $F = F_0 < F_1 < \ldots < F_m = L$, such that $F_{i+1} = F_i [\beta_i]$ with $\beta_i^{n_i} \in F_i$ for some integer $n_i > 0$, is called a radical tower over $F$, and we say $L/K$ is a radical tower.

To begin we need some properties of the symmetric group $S_n$.

Lemma 6: $S_n$ is generated by transpositions. Every permutation in $S_n$, $n > 1$, is a product of 2-cycles.

Proof. Note that the identity (1) can be expressed as $(1) = (12)(21)$, and so is a product of transpositions. Every permutation $\sigma$ can be written in the form

$$(a_1 a_2 \ldots a_r) (b_1 b_2 \ldots b_s) \ldots (e_1 e_2 \ldots e_t)$$

A direct computation shows that each cycle $(a_1 a_2 \ldots a_r)$ is a product $(a_1 a_r) (a_1 a_{r-1}) \ldots (a_1 a_2)$. Thus every permutation is a product of 2-cycles.

Now the subgroup of even permutations in $S_n$ is called the alternating group and is denoted $A_n$.

Lemma 7: Every element in $A_n$, $n \geq 3$, can be expressed as a 3-cycle or a product of 3-cycles.

Proof. Every element in $A_n$ can be written as a product of an even number of 2-cycles. Now we show that every product of two 2-cycles can be written as a 3-cycle. In general if the 2-cycles are disjoint then $(ab)(cd) = (acb)(acd)$. If the 2-cycles are not disjoint then $(ab)(bc) = (abc)$. This shows every element in $A_n$ can be expressed
as a 3-cycle or a product of 3-cycles. □

In the next lemma we use the fact that every element in $A_n$ can be expressed as a 3-cycle or a product of three cycles to show that each element of $A_n$ is a product of $m$-cycles.

Lemma 8: Let $m$ be an integer, $3 \leq m \leq n$. Then each $\sigma \in A_n$ can be written as a product of $m$-cycles.

Proof. An $m$-cycle is a product of $m-1$ transpositions, so if $m$ is odd an $m$-cycle is in $A_n$. But then the identity $(a_1a_2a_3) = (a_2a_3a_4...a_m)(a_ma_{m-1}...a_4a_3a_2a_1)$ shows that every 3-cycle is in the group generated by the $m$-cycles, so Lemma 8 follows from Lemma 7. □

Now the following lemma was used by Abel in his proof, but it is due to Cauchy.

Lemma 9: Let $K$ be a field and $L = K(x_1,...,x_n)$, let $p$ be the largest prime $\leq n$, and let $f \in L$. Let $S = \{\sigma f \mid \sigma \in S_n\}$ then $|S| \geq p$ or $|S| \leq 2$.

Proof. Let $\sigma \in S_n$ be a $p$-cycle, and $\langle \sigma \rangle$ be the subgroup generated by $\sigma$. We then define $H = \{\phi \in \langle \sigma \rangle \mid \phi f = f\}$. Since $p$ is a prime, either $H = \langle \sigma \rangle$ or $H = \langle e \rangle$. Thus, either $f, \sigma f, ... \sigma^{p-1}f$ are all distinct, or $\sigma f = f$. If $|S| < p$, we have $\sigma f = f$ for all $p$-cycles. Lemma 8 implies $f$ is fixed by $A_n$. Since $A_n$ has index 2 in $S_n$, $S_n = A_n \cup \tau A_n$ for some transposition $\tau$. Then for all $\sigma \in S_n$, either $\sigma \in A_n$ or $\sigma = \tau \sigma$ for some $\sigma \in A_n$. If $\sigma \in A_n$ then $\sigma f = f$. If $\sigma = \tau \sigma$ for some $\sigma \in A_n$, then $\sigma f = \tau f$. Thus
Now we give the first part of Abels proof; that if \( L \) is the splitting field of the polynomial \( x^n - s_1x^{n-1} + \ldots + (-1)^n s_n \) and \( L \) is contained in a radical tower over \( K \) then \( L/K \) is a radical tower. This is what was left out of Ruffini’s proof on the non-solvability of the quintic equation, and the reason the solution is attributed to Abel.

We will assume the base field \( K \) contains all the roots of unity, \( K = k(s_1, s_2, \ldots, s_n) \).

Lemma 10. Let \( F \) be a field containing the \( q \)th roots of unity. Let \( x^q - a \in F[x] \) be irreducible, \( q \) a prime, and let \( \alpha \) be a root. Let \( \gamma \in F(\alpha) \) such that \( \gamma \notin F \). Then there exists \( \beta \in F(\alpha) \) such that \( F(\beta) = F(\alpha) \), \( \beta^q \in F \) and \( \gamma = b_0 + \beta + b_2\beta^2 + \ldots + b_{q-1}\beta^{q-1} \) with \( b_0, b_2, \ldots, b_{q-1} \in \bar{F} \).

Proof. Since \( F \) contains the \( q \)th roots of unity, \( \{1, \ldots, \alpha^{q-1}\} \) is a basis for \( F(\alpha) \) over \( F \).

Let \( \gamma = a_0 + a_1\alpha + \ldots + a_{q-1}\alpha^{q-1} \) and let \( 1 \leq k < q \) be the smallest integer such that \( a_k \neq 0 \).

Let \( \beta = a_k\alpha^k \). So then \( \beta^q \in F \) and for \( k \leq m < q \) there exist integers \( r \) and \( s_m \) such that \( 0 \leq s_m < q \) and \( rq + s_m k = m \). Then \( \alpha^m = (\alpha^q)^r (\alpha^k)^s = \alpha^r \left( \frac{\beta}{a_k} \right)^s = c_s\beta^s \) and \( c_s \in F \).

Substitution yields \( \gamma = a_0 + a_1c_1\beta^{q_1} + \ldots + a_{q-1}c_{q-1}\beta^{q_{q-1}} = b_0 + \beta + b_2\beta^2 + \ldots + b_{q-1}\beta^{q-1} \).

Further, \( s = 1 \) implies \( m = rq + k \), which then implies \( m = k \). Thus, \( a_k\alpha^k = 1 \cdot \beta \) so no other power of \( \alpha \) produces a multiple of \( \beta \). Therefore \( b_1 = 1 \). Also, \( F(\beta) \subseteq F(\alpha) \) since \( \beta = a_k\alpha^k \in F(\alpha) \) and \( 1 = rq + sk \) for some \( r \), \( s \) implies \( \alpha^s = c_s\beta^s \) for some \( s \).

Therefore \( \alpha \in F(\beta) \) and \( F(\beta) = F(\alpha) \). \( \square \)
We will need the next two technical results in the proof of Lemma 14.

Lemma 11. Let $q$ be prime and $\zeta$ be a primitive $q^{th}$ root of unity. Then for all $i$

$$1 + \zeta^i + ... + \zeta^{(q-1)i} = \begin{cases} 0 & \text{if } q \text{ does not divide } i \\ q & \text{if } q \text{ does divide } i \end{cases}$$

Proof. Assume $q$ divides $i$, then $i = mq$ and this implies $\zeta^i = (\zeta^q)^m = 1^m = 1$. So $1 + \zeta^i + ... + \zeta^{(q-1)i} = 1 + 1 + ... + 1 = q$. Now assume that $q$ does not divide $i$. Then $\zeta^i \neq 1$, and

$$1 + \zeta^i + (\zeta^i)^2 + ... (\zeta^i)^{q-1} = \frac{1 - (\zeta^i)^q}{1 - \zeta^i} = \frac{0}{1 - \zeta^i} = 0 . \square$$

Lemma 12. Either $x^p - a$ is irreducible, or $a$ is a $p^{th}$ power in $K$, so that there exists in $K$ a factorization $x^p - a = x^p - \beta^p = (x - \beta)(x^{p-1} + \beta x^{p-2} + ... + \beta^{p-1})$.

Proof. Suppose that $x^p - a$ is reducible. Then, $x^p - a = \varphi(x) \cdot \psi(x)$. In its splitting field $x^p - a$ factors as follows:

$$x^p - a = \prod_{r=0}^{p-1} (x - \zeta^r \theta)$$

where $\theta^p = a$, and $\zeta$ is a $p^{th}$ root of unity. Therefore, the factor $\varphi(x)$ must be a product of certain factors $x - \zeta^r \theta$, and the constant term $\pm b$ of $\varphi(x)$ must have the form $\pm \zeta' \theta^\mu$, where $\zeta'$ is a $p^{th}$ root of unity: $b = \zeta' \theta^\mu$, and $b^p = \theta^{\mu p} = \alpha^\mu$. Since $0 < \mu < p$ we have $(\mu, p) = 1$ and so for suitable integers $\vartheta$ and $\sigma$ we have $\vartheta \mu + \sigma p = 1$, and $a = a^{\vartheta \mu} a^{\sigma p} = b^{\vartheta \mu} b^{\sigma p}$; hence $a$ is a $p^{th}$ power. $\square$

We will now show that if $y$ belongs to a normal extension of a field $F$ then the minimal polynomial of $y$ factors in $L[x]$. 

42
Lemma 13. Let $L \supset F$ be fields and $y \in L$ be algebraic over $F$. Let $g(x) \in F[x]$ be the minimal polynomial of $y$. Then $g(x)$ factors into linear factors in $L[x]$.

Proof. Let $y_1, y_2, \ldots, y_n$ be distinct conjugates of $y$ acted upon by $G(L/F)$. Let $f(x) = (x - y_1)(x - y_2) \cdots (x - y_n)$. Since, for all $\varphi \in G(L/F)$ there exists $\sigma \in S_n$ such that $\varphi(f(x)) = (x - y_{\varphi(1)}) \cdots (x - y_{\varphi(n)}) = f(x)$, the coefficients of $f(x)$ are invariant under the symmetric group, and thus are elements of $F_{G(L/F)} = F$. The minimal polynomial $g(x)$ of $y$ is irreducible over $K$ and divides $f(x)$. Thus $g(x)$ factors into linear factors in $L[x]$.

Now we come to the main part of Abel’s proof, which says basically that if a radical extension containing $K$ is intersected with the splitting field $L$, the resulting field is a radical extension.

Lemma 14. Let $K < E$ be fields, and $\alpha$ be an element of $E$ such that $x^q - \alpha \in E[x]$ is irreducible. Let $\alpha$ be a root of $x^q - \alpha$ and $M = E(\alpha) \cap L$, $M_0 = E \cap L$. If $M \neq M_0$ then $M/M_0$ is a radical extension.

Proof. Let $y \in M$, $y \notin M_0$. By Lemma 10 we can find a $\beta \in E(\alpha)$ such that $E(\alpha) = E(\beta)$, $\beta^q = b \in E$ and

$$y = b_0 + \beta + b_2 \beta^2 + \cdots + b_{q-1} \beta^{q-1}$$
where \( b_i \in E \). Let \( g(x) \in K[x] \) be the minimal polynomial for \( y \) over \( K \), and then set

\[
G(x) = g \left( b_0 + x + b_2x^2 + \ldots + b_{q-1}x^{q-1} \right)
\]

Now \( G(x) \) is contained in \( E[x] \) and \( \beta \) is a root. Since \( \beta \notin E \), \( x^q - b \in E[x] \) is irreducible by Lemma 12, and \( I = \{ f(x) \in E[x] \mid f(\beta) = 0 \} = (a(x)) \) for some monic \( a(x) \in E[x] \). Since \( x^q - b \in I \), \( x^q - b = f(x) \cdot a(x) \) which implies \( a(x) = x^q - b \) or \( a(x) = 1 \). Clearly, \( 1 \notin I \) so \( a(x) = x^q - b \). Also, \( G(\beta) = 0 \) implies \( G(x) \in I = (x^q - b) \) and so \( x^q - b \) divides \( G(x) \). If \( \zeta \) is a primitive \( q^{th} \) root of unity then \( G(x) = a(x) \cdot (x^q - b) \) and \( G(\zeta^i \beta) = a \left( \zeta^i \beta \right)^q \cdot \left( (\zeta^i \beta)^q - b \right) = a \left( \zeta^i \beta \right) \cdot (b - b) = 0 \) and so,

\[
y = y_1 = b_0 + \beta + b_2 \beta^2 + \ldots + b_{q-1} \beta^{q-1}
\]

\[
y_2 = b_0 + \zeta \beta + b_2 \zeta^2 \beta^2 + \ldots + b_{q-1} \zeta^{q-1} \beta^{q-1}
\]

\[
\vdots
\]

\[
y_i = b_0 + \beta \zeta^{i-1} + b_2 \zeta^{2(i-1)} \beta^2 + \ldots + b_{q-1} \zeta^{(i-1)(q-1)} \beta^{q-1}
\]

\[
\vdots
\]

\[
y_q = b_0 + \zeta^{q-1} \beta + b_2 \zeta^{2(q-1)} \beta^2 + \ldots + b_{q-1} \zeta^{(q-1)(q-1)} \beta^{q-1}
\]

are all roots of \( g(x) \). By Lemma 13, \( y_1, \ldots, y_q \) are all in \( L \). We then multiply the \( i^{th} \) equation by \( \zeta^{1-i} \) to get:

\[
\zeta^i y = b_0 + \beta \zeta + b_2 \zeta^2 \beta^2 + \ldots + b_{q-1} \zeta^{q-1} \beta^{q-1}
\]

\[
\zeta^0 y = b_0 + \beta + b_2 \beta^2 + \ldots + b_{q-1} \beta^{q-1}
\]
\[ \zeta^{-1} y_2 = \zeta^{-1} b_0 + \beta + b_2 \zeta^2 + \ldots + b_{q-1} \zeta^{q-2} \beta^{q-1} \]

\[ \vdots \]

\[ \zeta^{1-q} y_q = \zeta^{1-q} b_0 + \beta + b_2 \zeta^{q-1} \beta^2 + \ldots + b_{q-1} \zeta^{(q-2)(q-1)} \beta^{q-1}. \]

Add the resulting equations and we have

\[ \sum_{i=1}^{q} \zeta^{1-i} y_i = b_0 \sum_{i=1}^{q} \zeta^{1-i} + q \beta + b_2 \beta^2 \sum_{i=1}^{q} \zeta^i + \ldots + b_{q-1} \beta^{q-1} \sum_{i=1}^{q} \zeta^{(q-2)(i-1)}. \]

Then by Lemma 11, \( \sum_{i=1}^{q} \zeta^{1-i} y_i = q \beta \). Thus \( \beta = \frac{1}{q} \sum_{i=1}^{q} \zeta^{1-i} y_i \in L, \beta \in L \cap E(\alpha) = M, \) and \( \beta^q = b \in L \cap E = M_0 \). So, \( M_0(\beta) \subseteq M \).

Now we must show \( M_0(\beta) \supseteq M \). Let \( \gamma \in M \). Then, since \( \gamma \in E(\alpha) = E(\beta) \), there exists \( c_i \in E \) such that \( \gamma = c_0 + c_1 \beta + c_2 \beta^2 + \ldots + c_{q-1} \beta^{q-1} \). We will show that the coefficients \( c_i \) must be elements of \( E \cap L = M_0 \). Now let \( \gamma_i = c_0 + c_1 \zeta^{i-1} \beta + \ldots + c_{q-1} \zeta^{(i-1)(q-1)} \beta^{q-1} \) and once again multiply \( \gamma_i \) by \( \zeta^{k(1-i)} \) and sum up over all the \( i \)'s ranging from 1 to \( q \):

\[ \sum_{i=1}^{q} \zeta^{k(1-i)} \gamma_i = \sum_{i=1}^{q} \sum_{t=0}^{q-1} c_t \zeta^{(t-k)(i-1)} \beta^t = \sum_{t=0}^{q-1} c_t \beta^t \sum_{i=1}^{q} \zeta^{(t-k)(i-1)} \]

By Lemma 11, the inner sum is 0 unless \( t = k \). Thus \( c_k \beta^k = \sum_{i=1}^{q} \zeta^{k(1-i)} \gamma_i \in M \). Since \( \beta \in M, c_k \in M \cap E = M_0 \). So \( M = M_0(\beta) \). \( \square \)

Now we come to the main theorem:
Theorem 15. If $L/K$ is contained in a radical tower then $L/K$ is a radical tower.

Proof. Suppose $E/K$ is a radical tower and $L \subseteq E$. Then $K = E_0 \subset E_1 \subset E_2 \subset \ldots \subset E_m = E$ where $E_i = E_{i-1} \left( \sqrt[p_i]{a_i} \right)$, $q_i$ a prime, and $a_i \in E_{i-1}$. Now, $K = E_0 \cap L \subseteq E_1 \cap L \subseteq \ldots \subseteq E_{m-1} \cap L \subseteq L$ is a tower. If $E_{i+1} \cap L \neq E_i \cap L$ then by Lemma 13, $E_{i+1} \cap L/E_i \cap L$ is a radical extension. After eliminating equalities we have $L/K$ is a radical tower. \[\Box\]

Theorem 16. Let $K = K_0 < K_1 < \ldots < K_n < L$ be a radical tower, let $D$ be the discriminant of $x^n - \sigma_1 x^{n-1} + \ldots + (-1)^n \sigma_n$, and $\delta = \sqrt{D}$. Then $K_1 = K_0 \left( \sqrt{D} \right)$.

Proof. Assume $K_1 = K(\alpha)$ where $\alpha = \sqrt[4]{a}$ with $a \in K_0 = K$. Also $\tau \in S_n$ be a transposition and $\alpha \notin K$. Then $\tau(\alpha)^p = a = \alpha^p$. Dividing by $\alpha^p$, $(\tau(\alpha)/\alpha)^p = 1$ so that $\frac{\tau(\alpha)}{\alpha} = \zeta$ where $\zeta^p = 1$. Thus $\tau(\alpha) = \zeta \alpha$. Apply $\tau$ to both sides of the equation $\tau(\alpha) = \zeta \alpha$ and we have $\alpha = \tau(\zeta \alpha) = \zeta \tau(\alpha) = \zeta^2 \alpha$. Since $p$ is prime $\zeta = 1$ or $p = 2$. So, either $\tau(\alpha) \neq \alpha$ for some transposition $\tau$ and $p = 2$, or $\alpha$ is fixed by all transpositions. If $\alpha$ is fixed by all transpositions, then $\alpha$ is fixed by $S_n$ and is an element of $K$, which is contrary to the assumption $\alpha \notin K$. Therefore $\tau(\alpha) = \pm \alpha$ for all transpositions $\tau$. This implies that $\sigma(\alpha) = \pm \alpha$ for all $\sigma \in S_n$. Every 3-cycle is a square since $(abc) = (abc)^2$. Thus $\alpha$ is fixed by 3-cycles and therefore by all of $A_n$. Since $\tau(\alpha) = -\alpha$ for at least one transposition, and $S_n = A_n \cup \tau A_n$, $\tau(\alpha) = -\alpha$ for all transpositions. Since $\tau(\delta) = -\delta$ for all transpositions $\tau$, $\alpha/\delta$ is
fixed by all transpositions, and so by all elements in $S_n$. Thus $\alpha/\delta = b \in K$ and so $a = \alpha^2 = b^2\delta^2 = b^2D$. This shows $K_1 = K(b\sqrt{D}) = K(\sqrt{D})$. □

We will now show that $K(D)$ has no radical extension in the splitting field of the general polynomial with degree greater than or equal to five.

Theorem 17. Let $K_1$ be as in Theorem 10, and let $L$ be the splitting field of the general $n^{th}$ degree polynomial. Then $K_1$ has no radical extension in $L$ if $n \geq 5$.

Proof. Suppose $c \in K_1^*$ and $K_2 = K_1(\sqrt[3]{c})$. Set $\gamma = \sqrt[3]{c}$. We have shown that $A_n$ leaves $K_1$ fixed. Let $\rho$ be a three-cycle, and apply $\rho$ to both sides of the equation $\gamma^3 = c$. We then see that $\rho(\gamma)^3 = c$. Therefore $\rho(\gamma) = \zeta^q \gamma$ where $\zeta^q = 1$. Apply $\rho^3$ to the equation $\rho(\gamma) = \zeta^q \gamma$ and we have $\gamma = \rho^3(\gamma) = \zeta^3 \gamma$. Thus, either $\rho(\gamma) = \gamma$ for all three-cycles $\rho$ or $\rho(\gamma) \neq \gamma$ for some three-cycle and $q = 3$. In the first case $\gamma$ is fixed by $A_n$ and is in $K_1$ which is contrary to our assumption, so we conclude that $q = 3$. If $n \geq 5$, $A_n$ is also generated by five-cycles (Lemma 3). The same argument shows that $q = 5$ thus we have a contradiction. Therefore $K_1$ has no radical extension in $L$.

Theorem 18. $L/K$ is not a radical tower when $n \geq 5$ and thus the general quintic equation $p(x) = x^5 + bx^4 + cx^3 + dx^2 + ex + f = 0$ is not solvable by radicals.

Proof. Suppose

$$K = K_0 \subset K_1 \subset K_2 \subset \ldots \subset K_n = L$$

is a radical tower. Then there is a prime $p$ and element $a \in K^*$ such that $K_1 = K(\sqrt[p]{a})$. 

47
We know from Theorem 15 that \( p = 2 \) and that \( a = b^2 D \) where \( b \in K^* \) and \( D \) is the discriminate of \( p(x) \). Then, from Theorem 16, \( K \) has no radical extension in \( L \). Thus we have no radical tower, and the general quintic equation is not solvable by radicals.\( \square \)

Niels Abel (1802-1829) was born on the southwestern coast of Norway. He and his older brother, Hans Mathias, were sent to the Cathedral School in Oslo when Niels was 14. Niels was an exceptional student and pursued the works of Newton, Lagrange, Gauss, and Euler. In 1819 his instructor wrote of him: "With the most excellent genius he combines an insatiable interest and desire for mathematics, so that if he lives he probably will become a great mathematician" (Ore [1957]). While Abel excelled at school, he had trouble within his family. His brother had dropped to the bottom of his class, and was sent home as "feeble-minded". His father then died in 1820. He found a home for his sister and took his younger brother Peder with him to the University of Oslo in 1821.

Just before his family began to fall apart Abel believed he had discovered the solution to the quintic equation. When he was asked for more details he then discovered his own error. In 1824 he published a memoir, "On the Algebraic Resolution of Equations," in which he gave the first valid proof that no such solution is possible. He based his proof on "simply expressing its root as algebraic functions of the coefficients," and then applying this to theorems on symmetric functions.
At the University of Oslo he quickly mastered the advanced mathematics and began his own research. He worked on polynomial equations and also elliptic functions and integration. In 1824 he received a grant to travel and meet with other mathematicians. He went to Berlin but was unable to see Gauss as he was "absolutely unapproachable". He then traveled to Paris to meet with Cauchy, but found his reception less than warm and wrote: "Every beginner has a great deal of difficulty in getting noticed here. I have just finished an extensive treatise on a certain class of transcendental functions . . . but Mr. Cauchy scarcely deigns to glance at it."[EV]

By 1827 Abel had run out of money, and did not have enough to eat. He returned to Norway in May and got a temporary appointment at the University of Oslo which barely meet his expenses. His health began to deteriorate because of tuberculosis but he continued to send his work to a journal in Berlin. He died in on April 6, 1829, just two days before word came of his appointment as professor in Berlin.

Another brilliant mathematician with a tragically short life was Evariste Galois (1811-1832). He was born in Paris, the second of three children. His parents, though well-educated, were not mathematicians. His father was the mayor of their village, Bourg-la-Reine. He was taught by his mother until age twelve. When he did enter school his class work was mediocre and he was considered unusual, or eccentric by his teachers.
It is said he was fascinated by Legendre’s Geometry. He also studied Lagrange’s work on the theory of equations, and the writings of Abel. Because of his mathematical abilities he felt he was ready to enter the prestigious Ecole Polytechnique, but did not pass the entrance exams. He then began to study under Louis-Paul-Emile Richard, and published his first paper on continued fractions at age 17. Richard saved some of his early work, and one piece from 1828 shows that Galois, like Abel, thought he could solve the quintic equation \([BO]\). In May of 1829 Galois submitted his first paper on the theory of equations to the Paris Academy. It was refereed by Cauchy, but Cauchy apparently lost it and it never appeared in print. He tried again to gain entrance into the Ecole Polytechnique but failed. It was around this time that his father committed suicide. Galois then entered the Ecole Normale to prepare for teaching, and continued his research. He submitted another paper to the Academy in a competition and it was received by Fourier, but he died shortly after, and this manuscript, too, was lost.

France at this time was in a state of political unrest. Galois joined the National Guard which supported revolution. Because of a letter he had written against the head of the Ecole Normale, Galois was expelled. He was also arrested twice and served a five week, and then a six month sentence for his revolutionary activities. He was able to work on his theory of equations while he was in prison, but he also made a suicide attempt during this time. One month after his release from prison he was killed in a duel with
Pescheux d'Herbinville. On the eve of the duel he spent most of the night writing to a friend outlining his discoveries and hoping "some men will find it profitable to sort out this mess." [EV] Evariste Galois was 20 years old when he died.
BIBLIOGRAPHY


