GALOISIAN GALOIS THEORY

Lecture Notes in Computational Algebra

ABSTRACT. These notes are an exposition of Galois Theory from the original Lagrangian and Galoisian point of view. A particular effort was made here to better understand the connection between Lagrange's purely combinatorial approach and Galois algebraic extensions of the latter. Moreover, stimulated by the necessities of present day computer explorations, the algorithmic approach has been given priority here over every other aspect of presentation. In particular, you may not find here the clean antiseptic look characteristic of the classical exposition of E. Artin. In contrast these notes should provide a good starting point in attempting constructions in this most difficult computational arena.

1. Symmetric Functions

Unless otherwise specified all fields we shall work with here will be assumed to have zero characteristic. If $\mathcal{F}$ is such a field and $x_1, x_2, \ldots, x_n$ are indeterminates the expression

$$\mathcal{F}[x_1, x_2, \ldots, x_n]$$

will denote the ring of polynomials in $x_1, x_2, \ldots, x_n$ with coefficients in $\mathcal{F}$. In contrast to customary notation, the field of rational functions of $x_1, x_2, \ldots, x_n$ with coefficients in $\mathcal{F}$ will be denoted by

$$Rat[\mathcal{F}; x_1, \ldots, x_n]$$

We may also write

$$Rat[\mathcal{F}; f_1, f_2, \ldots, f_m]$$

to represent all rational expressions in $f_1, f_2, \ldots, f_n$ with coefficients in $\mathcal{F}$, whatever $f_1, f_2, \ldots, f_m$ may be in any particular situation.

As customary, S_n denotes the group of all permutations of $1, 2, \ldots, n$.

If $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n) \in S_n$ and

$$\Phi(x_1, x_2, \ldots, x_n) \in Rat[\mathcal{F}; x_1, \ldots, x_n]$$

we set $x\sigma = x_{\sigma_1}, x_{\sigma_2}, \ldots, x_{\sigma_n}$ and

$$\sigma\Phi = \Phi(x\sigma) = \Phi(x_{\sigma_1}, x_{\sigma_2}, \ldots, x_{\sigma_n})$$

We shall say that $\Phi(x_1, x_2, \ldots, x_n)$ is $k - valued$ if and only if the collection

$$\{\ \sigma\ \Phi\ :\ \sigma \in S_n\ \}$$

has cardinality k. Of course 1-valued functions are usually called *symmetric*. It will be convenient to denote here by $Sym[\mathcal{F}; x_1, x_2, \ldots, x_n]$ and $Ratsym[\mathcal{F}; x_1, x_2, \ldots, x_n]$ the collections of symmetric elements of $\mathcal{F}[x_1, x_2, \ldots, x_n]$ and $Rat[\mathcal{F}; x_1, \ldots, x_n]$ respectively.

The combinatorial study of k-valued function is properly the domain of Lagrange theory and will be carried out in the next section. In this section we will limit ourselves to establishing the results on symmetric functions that are needed in our further developments.

We recall that the symmetric polynomial

$$e_k(x_1, x_2, \ldots, x_n) = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k}$$

is usually referred to as the k^{th} *elementary* symmetric function.

A vector of integers

$$\lambda = (\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_k > 0)$$

is said to be a partition of n and we write $\lambda \vdash n$ if and only if

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k = n$$

We shall also say that k is *the number of parts* of λ. If λ and μ are partitions of the same number with k and h parts respectively, we shall say that λ *dominates* μ and write $\lambda \ge \mu$ if and only if

$$\lambda_1 + \lambda_2 + \cdots + \lambda_s \ge \mu_1 + \mu_2 + \cdots + \mu_s \qquad (\ for\ \ s = 1, 2, \ldots, min(h, k)\ )$$

It easy to see that this partial order is linearly extended by the lexicographic order of integer vectors. Finally, given a partition μ the partition μ' whose parts μ'_s are given by

$$\mu'_s = \#\{i\ :\ \mu_i \ge s\}$$

is usually referred to as the *conjugate* of μ. A simple combinatorial argument shows that we have $\lambda \ge \mu$ if and only if $\lambda' \le \mu'$.

Given a vector

$$p = (p_1, p_2, \ldots, p_n)$$

of non-negative integers, the weakly decreasing rearrangement of the positive components of p will referred to as the *shape* of p and denoted by $\lambda(p)$. Given a partition λ the symmetric polynomial

$$m_\lambda(x) = \sum_{\lambda(p)=\lambda} x^p \qquad (\ x^p = x_1^{p_1} x_1^{p_2} \cdots x_n^{p_n}\ )$$

is usually referred to as the *monomial symmetric function* indexed by λ. The following fact is immediate.

Theorem 1.1

Every symmetric polynomial $\Phi(x_1, x_2, \ldots, x_n)$ has a unique expansion of the form

$$\Phi(x) = \sum_\lambda c_\lambda\, m_\lambda(x) \qquad 1.1$$

where the c_λ are integers if Φ has integer coefficients, and c_λ is in $\mathcal{F}$ if $\Phi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$.

Proof

The symmetry of Φ implies that the coefficients c_p and c_q of any two monomials x^p and x^q appearing in Φ must be the same if p and q have the same shape. Thus 1.1 is obtained by collecting terms of Φ according to shape.

Let now $\mu = (\mu_1, \mu_2, \ldots, \mu_h)$ be a partition and set

$$e_\mu(x) \;=\; e_{\mu_1}(x) e_{\mu_2}(x) \cdots e_{\mu_h}(x) \qquad 1.2$$

Since this polynomial is clearly symmetric in $x_1, x_2, \ldots, x_n$ it must have an expansion of the form

$$e_\mu(x) \;=\; \sum_\lambda m_\lambda(x)\, c_{\lambda\mu} \;\; . \qquad 1.3$$

It develops that the coefficients $c_{\lambda\mu}$ have a suggestive combinatorial interpretation.

Proposition 1.1

If μ has h parts and λ has k parts then $c_{\lambda\mu}$ gives the number of $h \times k$ matrices with $0,1$-entries and row and column sums given by μ and λ respectively.

Proof

There is a one-to-one correspondence between these matrices and the monomials obtained by expanding the product in 1.3. In fact, if $m_1(x), m_2(x), \ldots, m_h(x)$ are monomials coming out of $e_{\mu_1}(x), e_{\mu_2}(x), \ldots, e_{\mu_h}(x)$ respectively, then $m_i(x)$ corresponds to a subset of $1, 2, \ldots, n$ of cardinality μ_i, thus it may be represented by a $0,1$-vector with n components in an obvious manner. Putting together these vectors as the rows of an $h \times n$ matrix M we see that these monomials multiply to $x^p = x_1^{p_1} x_2^{p_2} \cdots x_n^{p_n}$ if and only if the columns of M add up to $p_1, p_2, \ldots, p_n$ respectively. This is our desired correspondence. Since, since monomials of the same shape have the same coefficient, the assertion follows by taking $x^p = x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_k^{\lambda_k}$.

Remark 1.1

Let M be one of the $0,1$-matrices with row and column sums given by μ and λ. Note that the number of $1\,'s$ in the first s columns of M is equal to

$$\lambda_1 + \lambda_2 + \cdots + \lambda_s$$

On the other hand by moving all the $1\,'s$ along their rows until they are bumper to bumper to the left (and all the $0's$ similarly to the right) we obtain a matrix M' whose column sums are the parts of the partition μ' conjugate to μ. This not only gives the inequality

$$\mu'_1 + \mu'_2 + \cdots + \mu'_s \;\geq \lambda_1 + \lambda_2 + \cdots + \lambda_s \qquad (\; for \;\; s = 1, 2, \ldots, min(h,k) \;\;)$$

but also assures that when $\lambda = \mu$ there can only be one matrix with the desired row and column sums.

We can thus obtain

Theorem 1.2

Every homogeneous symmetric polynomial $\Phi(x_1, x_2, \ldots, x_n)$ of degree m has a unique expansion of the form

$$\Phi(x) \;=\; \sum_{\mu \vdash m} d_\mu\, e_\mu(x)$$

where the coefficients d_μ are integers if Φ has integer coefficients and they are elements of $\mathcal{F}$ if $\Phi \in Sym[\mathcal{F}; x_1, x_2, \ldots, x_n]$. In particular, $Sym[\mathcal{F}; x_1, x_2, \ldots, x_n]$ is the polynomial ring generated by the elementary symmetric functions $e_1(x), e_2(x), \ldots, e_n(x)$.

Proof

Let $D = \|d_{\lambda\mu}\|_{\lambda,\mu \vdash m}$ denote the matrix obtained when the partitions λ are in lexicographic order and the partitions μ are in reverse lexicographic order. This done we see that an immediate consequence of the observations made in the Remark above is that D must be unitriangular and therefore invertible over the integers. This shows that the collection $\{e_\mu(x)\}_{\mu \vdash m}$ must also give a basis and that the elements of the basis $\{m_\lambda(x)\}_{\lambda \vdash m}$ have integral linear expansions in terms of the $e_\mu(x)\,'s$. This establishes our assertions.

Corollary 1.2

$$Ratsym[\mathcal{F}; x_1, x_2, \ldots, x_n] \;=\; Rat[\mathcal{F}; e_1, e_2, \ldots, e_n] \qquad 1.4$$

Proof

Note that every element $\Phi \in Rat[\mathcal{F}; x_1, \ldots, x_n]$ can be expressed in the form

$$\Phi(x) = \frac{P(x)}{Q(x)}$$

with $P, Q \in \mathcal{F}[x_1, x_2, \ldots, x_n]$. Multiplying numerator and denominator by the polynomial

$$\prod_{\substack{\sigma \in A_n \\ \sigma \neq id}} \sigma Q$$

we can write

$$\Phi(x) = \frac{P^*(x)}{Q^*(x)}$$

where

$$P^*(x) = P(x) \prod_{\substack{\sigma \in A_n \\ \sigma \neq id}} \sigma Q \qquad , \qquad Q^*(x) = \prod_{\sigma \in A_n} \sigma Q \;\; .$$

Since by its very construction $Q^*(x)$ is already symmetric, we see that Φ will be symmetric if and only if P^* is. Thus our assertion follows immediately from Theorem 1.2.

2. The Euclidean algorithm the Resultant and the Discriminant.

If A, B are polynomials in $\mathcal{F}[t]$, with $degree\, B < degree\, A$, then the *quotient* and the *remainder* anof the *division* of A by B are respectively the unique polynomials Q and R satisfying the requirements

$$\begin{array}{ll} (1) & A \;=\; B\,Q \;+\; R \\ (2) & degree(R) \;<\; degree(B) \end{array} \quad .$$

To construct these two polynomials we can proceed as follows. Set

$$\Delta(A/B) \;=\; \frac{leading\ coeff\ A}{leading\ coeff\ B}\quad t^{degree(A)-degree(B)}$$

Clearly we have

$$degree\big(A-\Delta(A/B)\,B\big) \;<\; degree(A). \tag{2.1}$$

Now set

$$R^{(o)}=A \quad \text{and} \quad R^{(i)}=R^{(i-1)}-\Delta(R^{(i-1)}/B)\,B \qquad (\ i=1,2,\dots\ ) \tag{2.2}$$

Since, by 2.1, degrees are decreasing at least by one at each step of the recursion, after $k \le degree(A)-degree(B)$ steps we shall have $degree(R^{(k)}) < degree(B)$. At this point we stop the recursion. By adding the identities in 2.1 we easily derive that

$$A \;=\; \Big(\sum_{i=0}^{k-1}\Delta(R^{(i)}/B)\Big)\;B \;+\; R^{(k)} \tag{2.3}$$

Thus we may take

$$Q=\sum_{i=0}^{k-1}\Delta(R^{(i)}/B) \quad , \quad R=R^{(k)}\ . \tag{2.4}$$

The important fact to note is that from 2.3 we deduce that if $A, B \in \mathcal{F}[t]$ then $R(t)$ is in $\mathcal{F}[t]$ as well. We shall refer to the process above as the *division algorithm*.

The Euclidean algorithm is the process which yields the greatest common divisor $D(t)$ of two polynomials A and B. It is shown by Berlekamp [] that D may be computed by the following process. Set

$$r_{-2}=A \quad , \quad r_{-1}=B \quad , \quad p_{-2}=0 \quad , \quad p_{-1}=1 \quad , \quad q_{-2}=1 \quad , \quad q_{-1}=0$$

Then compute a_k, r_k, p_k, q_k according to the recursions

$$\begin{array}{lll} (1) & r_{k-2} \;=\; a_k\, r_{k-1} \;+\; r_k & (division) \\ (2) & p_k \;=\; a_k\, p_{k-1} \;+\; p_{k-2} & \qquad (\ k=0,1,2,\dots\ ) \\ (3) & q_k \;=\; a_k\, q_{k-1} \;+\; q_{k-2} & \end{array} \quad .$$

since $degree(r_k)$ dcreases at least by one at each step after $n < degree(B)$ steps we shall have $r_n = 0$. It is shown in [] that these recursion force the following basic identities

$$\begin{array}{lll} (1) & q_n\, p_{n-1} \;-\; p_n\, q_{n-1} \;=\; =(-1)^n & \\ (2) & A \;=\; r_{n-1}\, p_n & \quad . \\ (3) & B \;=\; r_{n-1}\, q_n & \end{array}$$

Since equation (1) here yields that p_n and q_n are relatively prime, we see that (2) and (3) yield that the greatest common divisor D of A and B is necessarily given by r_{n-1}.

As pointed out in [] the advantage of this process over the one that is usually described in most textbooks is that it provides the final answer without excessive storage of partial results. In fact, only 7 results need to be stored at any particular time, a number that is independent of the choice of A and B.

If the roots of a polynomial

$$P(t) \;=\; a_o + a_1 t + a_2 t^2 + \cdots + a_m t^m$$

are $x_1, x_2, \dots, x_n$ then when $a_n = 1$ we may write it in the form

$$P(t) \;=\; (t-x_1)(t-x_2)\cdots(t-x_m) \tag{2.5}$$

Thus, we see that we must have

$$a_{n-k} \;=\; (-1)^k\, e_k(x_1, x_2, \dots, x_n). \tag{2.6}$$

Clearly, if $P(t)$ and

$$Q(t) \;=\; b_o + a_1 t + b_2 t^2 + \cdots + b_m t^m \;=\; (t-y_1)(t-y_2)\cdots(t-y_n) \qquad (\ b_n = 1\ )$$

have a root in common, the expression

$$\prod_{i=1}^{m}\prod_{j=1}^{n}(x_i-y_j)$$

will necessarily vanish. It develops that a multiple of this expression may be written as a polynomial in the coefficients of P and Q. This polynomial is usually referred to as the *Resultant of P and Q* and will be denoted here by $R[P,Q]$. The case $m = 2$, $n = 3$ is sufficient to get accross the idea and avoids excessive notation.

Theorem 2.1

$$R[P,Q]=det\begin{pmatrix} a_o & a_1 & a_2 & 0 & 0 \\ 0 & a_o & a_1 & a_2 & 0 \\ 0 & 0 & a_o & a_1 & a_2 \\ b_o & b_1 & b_2 & b_3 & 0 \\ 0 & b_o & b_1 & b_2 & b_3 \end{pmatrix} \;=\; (-1)^{2\times 3}\, a_2^3\, b_3^2 \prod_{i=1}^{2}\prod_{j=1}^{3}(x_i-y_j) \tag{2.7}$$

Proof

Note that since the division by a_2 and b_3 does not change the roots of P and Q, we can divide both sides of 2.7 by $a_2^3\, b_3^2$ and reduce ourselves to the case $a_2 = b_3 = 1$. This given, note that we have the following matrix multiplication identity:

$$\begin{pmatrix} a_o & a_1 & a_2 & 0 & 0 \\ 0 & a_o & a_1 & a_2 & 0 \\ 0 & 0 & a_o & a_1 & a_2 \\ b_o & b_1 & b_2 & b_3 & 0 \\ 0 & b_o & b_1 & b_2 & b_3 \end{pmatrix} \times \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ y_1 & y_2 & y_3 & 0 & 0 \\ y_1^2 & y_2^2 & y_3^2 & 0 & 0 \\ y_1^3 & y_2^3 & y_3^3 & 1 & 0 \\ y_1^4 & y_2^4 & y_3^4 & 0 & 1 \end{pmatrix} = \begin{pmatrix} P(y_1) & P(y_2) & P(y_3) & 0 & 0 \\ y_1 P(y_1) & y_2 P(y_2) & y_3 P(y_3) & 0 & 0 \\ y_1^2 P(y_1) & y_2^2 P(y_2) & y_3^2 P(y_3) & 0 & 0 \\ 0 & 0 & 1 & b_3 & 0 \\ 0 & 0 & 0 & 0 & b_3 \end{pmatrix}$$

and 2.7 follows immediately by equating determinants of boths sides and cancelling the common factor.

Remark 2.1

Note that the vanishing of the determinant in 2.7 assures that we can find a non trivial solution to the corresponding homogeneous system. Now a simple computation shows that we have

$$(\, \alpha_o \quad \alpha_1 \quad \alpha_2 \quad -\beta_o \quad -\beta_1 \,) \times \begin{pmatrix} a_o & a_1 & a_2 & 0 & 0 \\ 0 & a_o & a_1 & a_2 & 0 \\ 0 & 0 & a_o & a_1 & a_2 \\ b_o & b_1 & b_2 & b_3 & 0 \\ 0 & b_o & b_1 & b_2 & b_3 \end{pmatrix} = 0$$

if and only if

$$(\alpha_o + \alpha_1 t + \alpha_2 t^2)\, P(t) \;-\; (\beta_o + \beta_1 t)\, Q(t) \;=\; 0$$

and this is equivalent to the statement that P and Q have a non trivial common factor.

In the same vein we see that the expression

$$\Delta(x) \;=\; \prod_{1 \le i < j \le n} (x_i - x_j)$$

vanishes if and only if $P(x)$ as given by 2.5 has multiple roots. Since, $\Delta(x)^2$ is clearly symmetric in $x_1, x_2, \ldots, x_n$, Theorem 1.2 and 2.6 guarantee that the latter polynomial should be expressible as a polynomial in the coefficients $a_o, a_1, \ldots, a_n$ of $P(t)$. In fact, we need only replace Q by the derivative $P'(t) = \frac{d}{dt} P(t)$ in $R[P, Q]$ to obtain

Theorem 2.2

$$R[P, P'] \;=\; a_m^{m(m-1)} \prod_{i=1}^{n} \prod_{j=1}^{n}{}^{(i)}\, (x_i - x_j) = (-1)^{n(n-1)/2}\, a_m^{m(m-1)}\, \Delta(x)^2 \qquad 2.8$$

where the superscript " $^{(i)}$ " in the product is to indicate that the factor corresponding to $j = i$ is to be omitted.

Proof

It is easily seen that the general form of 2.7 may also be written as

$$R[P, Q] \;=\; (-1)^{nm} a_m^n \prod_{i=1}^{n} Q(x_1) Q(x_2) \cdots Q(x_n) \qquad 2.9$$

Now we easily see that

$$P'(x_i) \;=\; a_n \prod_{j=1}^{n}{}^{(i)}\, (x_i - x_j) \qquad 2.10$$

and 2.8 follows by setting $Q = P'$ in 2.9.

The polynomial $R[P, P']$ is usually referred to as the *discriminant* of the equation

$$P(t) = 0 \;\; .$$

We shall terminante with a simple fact that will play a crucial role in the sequel. Recall that a polynomial $B(t)$ with coefficients in a field $\mathcal{F}$ is said to be irreducible over $\mathcal{F}$ if it does not admit a factorization

$$B(t) = P(t) Q(t)$$

into two polynomials $P, Q \in \mathcal{F}[t]$ of strictly lesser degree than B. The Euclidean Algorithm immediately yields that if $B(t)$ is irreducible and $A(t)$ is a polynomial in $\mathcal{F}[t]$ which has a root in common with $A(t)$ then $B(t)$ must be a factor of $A(t)$. The reason for this is that if $A(t)$ and $B(t)$ share a root (a fact which may be verified by computing $R[A, B]$) then the greatest common divisor $D(t)$ of A and B (as yielded by the process descrbed above), is in $\mathcal{F}[t]$ as well and we would be led to a contradiction unless D is equal to a constant multiple of B. The important conclusion we draw from this is that if a polynomial $A \in \mathcal{F}[t]$ shares a root with an irreducible polynomial $B(t) \in \mathcal{F}[t]$ then it must vanish for all the other roots of $B(t)$. This fact has the following immediate extension.

Proposition 2.1

Let $B(t) \in \mathcal{F}[t]$ be irreducible in $\mathcal{F}$ and let $\Phi(t) \in Rat[\mathcal{F}; t]$ vanish at one of the roots of $B(t)$ then $\Phi(t)$ vanishes at all the other roots of $B(t)$.

Proof

By hypothesis $\Phi(t) = P(t)/Q(t)$ with $P, Q \in \mathcal{F}$ using the Euclidean Algorithm we can cancel out (if necessary) the greatest common divisor of P and Q assure that P and Q have no common root. But then $\Phi(t)$ vanishes if and only if $P(t)$ does, and we are thus reduced to the case discussed above.

3. The cubic and the quartic

Formulas giving the general solution of the cubic equation

$$E_3(t) \;=\; (t-x_1)(t-x_2)(t-x3) \;=\; t^3 - e_1t^2 + e_2t - e_3 \;=\; 0 \tag{3.1}$$

where first discovered by Ferreo (sometimes before 1505) rediscoverd by Tartaglia and published by Cardano in 1545. Setting $\omega = e^{2\pi\, i/3}$ the three roots of 3.1 may be written as follows

$$\begin{aligned} x_1 &= \frac{e_1}{3} + \sqrt[3]{-\tfrac{1}{2}q + \sqrt[2]{R}} + \sqrt[3]{-\tfrac{1}{2}q - \sqrt[2]{R}}\ , \\ x_2 &= \frac{e_1}{3} + \omega\sqrt[3]{-\tfrac{1}{2}q + \sqrt[2]{R}} + \omega^2\sqrt[3]{-\tfrac{1}{2}q - \sqrt[2]{R}}\ , \\ x_3 &= \frac{e_1}{3} + \omega^2\sqrt[3]{-\tfrac{1}{2}q + \sqrt[2]{R}} + \omega\sqrt[3]{-\tfrac{1}{2}q - \sqrt[2]{R}}\ , \end{aligned} \tag{3.2}$$

where

$$p \;=\; e_2 - e_1^2/3 \quad , \quad q = -e_3 + \tfrac{1}{3}e_1e_2 - \tfrac{2}{27}\,e_1^3 \quad \textit{and} \quad R \;=\; (q/2)^2 + (p/3)^3\ . \tag{3.3}$$

These formulas are usually derived by the following process, apparently due to Hudde (1650). We start by making the substitution $t = y + \frac{1}{3}e_1$ in 3.1 and transform it to

$$y^3 + py + q = 0$$

This given the further substitution

$$y = z - \frac{p}{3z} \tag{3.4}$$

brings us to the equation

$$z^3 - \frac{p^3}{27z^3} + q \;=\; 0$$

or better yet

$$z^6 + qz^3 - p^3/27 \;=\; 0\ . \tag{3.5}$$

Since this is a quadratic equation for z^3 we immediately derive the two solutions

$$z_1^3 \;=\; -\tfrac{1}{2}\,q \;+\; \sqrt[2]{R} \quad \textit{or} \quad z_2^3 \;=\; -\tfrac{1}{2}\,q \;-\; \sqrt[2]{R}$$

since

$$(-\tfrac{1}{2}\,q \;+\; \sqrt[2]{R})(-\tfrac{1}{2}\,q \;-\; \sqrt[2]{R}) \;=\; -(p/3)^3$$

we may extract cube roots so that

$$z_2 \;=\; -\,\frac{p}{3z_1} \tag{3.6}$$

This is given the six roots of 3.5 are

$$\begin{matrix} z_1 & z_2 \\ \omega z_1 & \omega^2 z_2 \\ \omega^2 z_1 & \omega z_2 \end{matrix}$$

Where each of these pairs multiplies to $-p/3$.

We can now use 3.4 and derive that the three roots of 3.3 may be written in the form

$$\begin{aligned} y_1 &= z_1 + z_2 \\ y_2 &= \omega z_1 + \omega^2 z_2 \\ y_3 &= \omega^2 z_1 + \omega z_2 \end{aligned} \tag{3.7}$$

From which the formulas in 3.2 can be immediately obtained.

The quartic equation was treated in a similar manner. That is "ad hoc" manipulations were used to transform it to equations which could be solved by extraction of roots. To give a brief idea of the process in this case we start with

$$E_4(t) \;=\; (t-x_1)(t-x_2)(t-x_3)(t-x_4) \;=\; t^4 - e_1t^3 + e_2t^2 - e_3t + e_4 \;=\; 0\ . \tag{3.8}$$

Completing the square suggested by the first two terms we can rewrite this equation in the form

$$(t^2 - \tfrac{1}{2}e_1\,t)^2 \;=\; (\tfrac{1}{4}e_1^2 - e_2)\,t^2 \;+\; e_3\,t - e_4\ .$$

We then add $(t^2 - \frac{1}{2}\,e_1t)y + \frac{1}{4}y^2$ to both sides and get

$$(t^2 - \tfrac{1}{2}\,e_1t + \tfrac{1}{2}y)^2 \;=\; (\tfrac{1}{4}e_1^2 - e_2 + y)\,t^2 + (e_3 - \tfrac{1}{2}e_1y)\,t + \tfrac{1}{4}y^2 - e_4\ . \tag{3.9}$$

Next, y is determined so that also the term on the right becomes a perfect square. This requires the coefficients

$$A = \tfrac{1}{4}e_1^2 - e_2 + y \quad , \quad B = e_3 - \tfrac{1}{2}\,e_1\,y \quad , \quad C = \tfrac{1}{4}\,y^2 - e_4$$

of the quadratic on the right hand side of 3.9 satisfy the equation

$$B^2 - 4AC \;=\; 0\ .$$

This leads to a cubic equation for y. To see what are the roots of this equation, we should try to factor $B^2 - 4AC$. Nowdays this is easily done, using any of the available computer algebra packages. In this manner we discover the pleasing fact that $B^2 - 4AC$ factors beautifully in terms of the roots of $E_4(t)$. Namely, we have

$$B^2 - 4AC \;=\; (y - x_2x_4 - x_1x_3)\;(y - x_1x_2 - x_3x_4)\;(y - x_1x_4 - x_2x_3)\ .$$

Thus the three roots of this cubic when expressed in terms of the roots of 3.8 are none other than

$$y_1 \;=\; x_1x_2 + x_3x_4 \quad , \quad y_2 \;=\; x_1x_3 + x_2x_4 \quad , \quad y_3 \;=\; x_1x_4 + x_2x_3$$

We should also note that setting $y = y_1$ we have

$$\begin{aligned} A &= \tfrac{1}{4}e_1^2 - e_2 + y_1 \;=\; \tfrac{1}{4}\,(x_1 + x_2 - x_3 - x_4)^2 \\ B &= e_3 - \tfrac{1}{2}\,e_1\,y_1 \;=\; -\tfrac{1}{2}\,(x_1x_2 - x_3x_4)(x_1 + x_2 - x_3 - x_4) \\ C &= \tfrac{1}{4}\,y_1^2 - e_4 \;=\; \tfrac{1}{4}\,(x_1x_2 - x_3x_4)^2 \end{aligned} \tag{3.10}$$

which yield that the right hand side of 3.9 when $y = y_1$ reduces to

$$\tfrac{1}{4}\,\big(\,-(x_1+x_2-x_3-x_4)\,t+x_1x_2-x_3x_4\,\big)^2\ .$$

This allows us to rewrite 3.9 in the form

$$L^2-R^2\ =\ (L\ -\ R)\ (L\ +\ R)\ =\ 0\ , \tag{3.11}$$

with

$$L\ =\ t^2-\tfrac{1}{2}\,e_1t+\tfrac{1}{2}y_1\qquad\text{and}\qquad R\ =\ \tfrac{1}{2}\,(-(x_1+x_2-x_3-x_4)\,t+x_1x_2-x_3x_4)$$

Now the factorization of A in 3.10 suggests setting

$$z\ =\ x_1+x_2-x_3-x_4\ =\ \sqrt[2]{e_1^2-4e_2+4y_1}\ , \tag{3.12}$$

and then the factorization of B in 3.10 gives that

$$-\tfrac{1}{2}(x_1x_2-x_3x_4)\ =\ (e_3-\tfrac{1}{2}e_1y_1)/z\ .$$

This given, we may write

$$-R\ =\ \tfrac{1}{2}z\,t\ +\ (e_3-\tfrac{1}{2}\,e_1y_1)/z$$

and the equation in 3.11 may yet be rewritten as

$$\big(t^2-\tfrac{1}{2}(e_1+z)\,t+\tfrac{1}{2}y_1-(e_3-\tfrac{1}{2}e_1y_1)/z\big)\,\big(t^2-\tfrac{1}{2}(e_1-z)\,t+\tfrac{1}{2}y_1+(e_3-\tfrac{1}{2}e_1y_1)/z\big)=0\ .$$

Now this is none other than factoring $E_4(t)$ in the form

$$E_4(t)\ =\ \big(t^2-(x_1+x_2)\,t-x_1x_2\big)\ \big(t^2-(x_3+x_4)\,t-x_3x_4\big)\ .$$

In fact, it may be easily verified that

$$\begin{aligned} x_1+x_2\ &=\ \tfrac{1}{2}e_1+\tfrac{1}{2}z\quad,\quad x_1x_2\ =\ \tfrac{1}{2}y_1-(e_3-\tfrac{1}{2}e_1y_1)/z\ ,\\ x_3+x_4\ &=\ \tfrac{1}{2}e_1-\tfrac{1}{2}z\quad,\quad x_3x_4\ =\ \tfrac{1}{2}y_1+(e_3-\tfrac{1}{2}e_1y_1)/z\ . \end{aligned}$$

Thus we may obtain the desired expressions for the pairs x_1, x_2 and x_3, x_4 by solving the two quadratic equations

$$\begin{aligned} t^2-\tfrac{1}{2}(e_1+z)\,t+\tfrac{1}{2}y_1-(e_3-\tfrac{1}{2}e_1y_1)/z\ &=\ 0\ ,\\ t^2-\tfrac{1}{2}(e_1-z)\,t+\tfrac{1}{2}y_1+(e_3-\tfrac{1}{2}e_1y_1)/z\ &=\ 0\ . \end{aligned}$$

This given, in the 17^{th} and 18^{th} centuries it was natural to assume that the solution of the general polynomial equation should be obtainable by similar manipulations and successive root extractions. This was the motivating force [] in Lagrange's investigations in the 1770's that led him to his historic paper *Réflections sur la résolution algébrique des équations* .

To be precise Lagrange was investigating the possibility of finding *closed form* expressions for the roots $x_1, x_2, \ldots, x_n$ which (like those appearing in 3.2) only involved the elementary symmetric functions $e_1, e_2, \ldots, e_n$, roots of unity and radicals. We shall refer to this as *"solving the general equation by radicals"*.

His point of departure was a close examination of the solutions of the cubic and the quartic. Remarkably, he was able to sort out of those simingly ad hoc manipulations a unifying general mechanism of solution. As we shall see Lagrange discovered that in both cases the final formulas could be reached by a sequence of identical, **purely combinatorial,** steps.

This done, he tried to apply this mechanism to the quintic only to discover that the possibility of pushing it through to the production of general formulas for the roots of the quintic appeared to lead to a contradiction!

In fact, he was (and he new he was) within reach of proving the impossibility of solving the quintic equation by radicals.

Around 1799 Ruffini tried to complete Lagranges proof and although he was able to push the argument quite a bit further he nevertheless was left with a hypothesis which he could not remove.

The glory of proving the impossibility of solving the general equation by radicals was bestowed to Abel (for his 1826 paper) (see []) even though he was only concerned with the quintic and, as in Ruffini's work, there were still a number of gaps in his arguments. We shall not deal with Abel's work here since it it departs from the combinatorial approach proposed by Lagrange and later completed by Galois. In fact, the missing step needed to complete Lagrange argument and obtain the unsolvability of the quintic by radicals can be supplied by one single idea of Galois.

To appreciate the beauty of Lagrange's discoveries we should view his results in the original 1771 form. Unfortunately, for clarity we must deviate a bit from Lagrange's terminology. For instance, although Lagrange proved that the order of a subgroup of a group is a divisor of the order of the group, he had to do so in an indirect manner, since the notion of a *group* in its present form really started with Galois. Although using modern terminology distorts somewhat the historical perspective, we will try as much as possible to keep unchanged the contents of Lagrange's discoveries. Our main goal in the next two sections is to present the basic theorems of what is now referred to as *Galois* theory in a sequence that makes the transition from Lagrange to Galois as natural and effortless as possible.

4. Lagrange's "Galois" Theory

Throughout Lagrange's work the roots of an equation

$$E_n(t)\ =\ a_o+a_1t+a_2t^2+\cdots+a_nt^n\ =\ =0$$

are assumed to be independent variables $x_1, x_2, \ldots, x_n$ and $E_n(t)$ is written in the form

$$E_n(t)\ =\ (t-x_1)(t-x_2)\cdots(t-x_n) \tag{4.1}$$

The basic idea that led Lagrange to an *understanding* of the classical solutions of the cubic and the quartic is a careful analysis of the effect that permutations of the roots have on various rational functions of the roots. To make precise what we mean by this we need some notation.

We are given a field $\mathcal{F}$ which remains unchanged throughout, and for a function Φ of the roots which may be in $\mathcal{F}[x_1, x_2, \ldots, x_n]$ or in $Rat[\mathcal{F}; x_1, \ldots, x_n]$ as needed, we set

$$G_\Phi \;=\; \{\, \sigma \in S_n \;:\; \sigma\Phi = \Phi \,\}\ . \tag{4.2}$$

Although Lagrange did not realize (nor did he need) that G_Φ is a group, we shall not ignore this fact here and obtain Lagrange's results by standard present day techniques. We recall that G_Φ is usually referred to as the *stabilizer* of Φ.

The first basic result of Lagrange can be stated as follows

Theorem 4.1

For $\Phi, \Psi \in Rat[\mathcal{F}; x_1, \ldots, x_n]$ we have

$$G_\Psi \subseteq G_\Phi \tag{4.3}$$

If and only if

$$\Phi \in Rat[\mathcal{F}, e_1, e_2, \ldots, e_n, \Psi]$$

Proof

If

$$\theta(y_1, y_2, \ldots, y_n, t) \;\in\; Rat[\mathcal{F}, y_1, y_2, \ldots, y_n, t]$$

and

$$\Phi(x_1, x_2, \ldots, x_n) \;=\; \theta(e_1, e_2, \ldots, e_n, \Psi)\ . \tag{4.4}$$

Then we clearly have 4.3 since every permutation σ leaves $e_1, e_2, \ldots, e_n$ unchanged and if $\sigma \in G_\Psi$ then also Φ does not change. So the condition in 4.3 is trivially necessary.

To show the converse, we resort to the left coset decomposition

$$S_n \;=\; \tau_1\, G_\Psi \;+\; \tau_2\, G_\Psi \;+\cdots+\; \tau_k\, G_\Psi\ . \qquad (\ \tau_1 = identity\ ) \tag{4.5}$$

(which by the way, was a Lagrange invention) and set

$$Q(t) \;=\; \sum_{i=1}^{k} \tau_i\Phi \prod_{j=1}^{k}{}^{(i)}\, (t - \tau_j\Psi)\ .$$

Since any $\sigma \in S_n$ permutes the left cosets of G_Ψ we may write

$$\sigma\, \tau_i \;=\; \tau_{\pi_i} h_i \qquad (\ h_i \in G_\Psi\ )$$

where the map $i \to \pi_i$ is a permutation of $(1, 2, \ldots, k)$. In particular, from 4.3 we deduce that $\sigma\, \Phi = \tau_{\pi_i}\Phi$, and thus we must have

$$\sigma Q(t) \;=\; \sum_{i=1}^{k} \tau_{\pi_i}\Phi \prod_{j=1}^{k}{}^{(\pi_i)}\, (t - \tau_j\Psi) \;=\; Q(t)\ .$$

This implies that the coefficients of $Q(t)$ are in $Ratsym[\mathcal{F}; x_1, x_2, \ldots, x_n]$, so by Theorem 1.2 they are in $Rat[\mathcal{F}, e_1, e_2, \ldots, e_n]$. The same can be said about the polynomial

$$P(t) \;=\; \prod_{i=1}^{k}(t - \tau_i\Psi)\ . \tag{4.6}$$

Now setting $t = \Psi$ in $Q(t)$ we get

$$Q(\Psi) \;=\; \Phi \prod_{j=1}^{k}{}^{(1)}\, (\Psi - \tau_j\Psi) \;=\; \Phi\, P'(\Psi)\ . \tag{4.7}$$

Since, by construction, the values $\tau_i\Psi$ (for $i = 1, 2, \ldots, k$) are all distinct we shall have $P'(\Psi) \neq 0$ and we can divide it out in 4.7 to obtain

$$\Phi \;=\; \frac{Q(\Psi)}{P'(\Psi)} \;=\; \theta(\Psi)$$

with

$$\theta(t) \;=\; \frac{Q(t)}{P'(t)} \;\in\; Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, t]$$

as desired.

Remark 4.1

If G is a group and $H \subseteq G$ is a subgroup then the left coset decomposition

$$G \;=\; \tau_1\, H \;+\; \tau_2\, H \;+\; \cdots \;+\; \tau_m\, H \qquad (\ \tau_1 = identity\ ) \tag{4.8}$$

yields that

If $H \subseteq G$ then the order of H divides the order of G 4.9

In fact, the equation above gives that $|G|/|H| = m$. This result, which appeared for the first time in the work of Lagrange, was formulated there as a statement concerning the *number* of different values taken by rational functions of the roots. More precisely, we can derive from 4.5 that Ψ is a k-valued rational function of the roots if and only if $|S_n|/|G_\Psi| = k$. Similarly, Φ is h-valued if and only if $|S_n|/|G_\Phi| = h$. This gives that when $G_\Psi \subseteq G_\Phi$ we have $k/h = |G_\Phi|/|G_\Psi|$. We shall refer to h and k respectively as the *multiplicities* of Ψ and Φ. So taking $H = G_\Psi$ and $G = G_\Phi$ in 4.8 we get that $k = hm$. We see then that in Lagrange's language the statement in 4.9 becomes

If Φ is a rational function of Ψ then the multiplicity of Φ divides that of Ψ .

Lagrange's proof was based precisely on the coset decomposition. Only he did not have to call it that way. Indeed to get 4.8 for $H = G_\Psi$ and $G = G_\Phi$ all he had to do was *bunch together* the elements of G_Φ that yielded the same value of Ψ.

It will be good here and after for H a subgroup of G to express the fact that $m = |G|/|H|$ by writing $H \subseteq_m G$. This given, these observations can be sharpened into the following corollary of Theorem 4.1.

Theorem 4.2

If for $\Phi, \Psi \in Rat[\mathcal{F}; x_1, \ldots, x_n]$ we have

$$G_\Psi \subseteq_k G_\Phi \tag{4.10}$$

then Ψ satisfies an equation of degree k with coefficients in $Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Phi]$ which is irreducible in $Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Phi]$.

Proof

Let

$$G_\Phi \;=\; \tau_1 G_\Psi \;+\; \tau_2 G_\Psi \;+\; \cdots \;+\; \tau_k G_\Psi \qquad (\ \tau_1 = identity\ ) \tag{4.11}$$

and set

$$Q(t) \;=\; \prod_{i=1}^{k} (t - \tau_i \Psi) \;=\; q_o(x) + q_1(x)\, t + \cdots + q_k(x)\, t^k \ . \tag{4.12}$$

Since any $\sigma \in G_\Phi$ permutes the left cosets of G_Ψ in 4.11 we may write

$$\sigma\, \tau_i \;=\; \tau_{\pi_i} h_i \qquad (\ h_i \in G_\Psi\ )$$

where the map $i \to \pi_i$ is again a permutation of $(1, 2, \ldots, k)$. This gives that for all $\sigma \in G_\Phi$ we have

$$\sigma Q(t) \;=\; \prod_{i=1}^{n} (t - \tau_{\pi_i} \Psi) \;=\; Q(t) \ .$$

Consequently each of the coefficients $q_i(x)$ is left unchanged by the elements of G_Φ. From Theorem 4.1 we then derive that each $q_i(x) \in Rat[\mathcal{F}, e_1, e_2, \ldots, e_n, \Phi]$. On the other hand from 4.12 we get that

$$Q(\Psi) \;=\; q_o(x) + q_1(x)\Psi + \cdots + q_k(x)\Psi^k \;=\; 0 \ .$$

Now suppose, if possible, that $Q(t)$ has a factorisation $Q(t) = Q_1(t)Q_2(t)$ where both polynomials $Q_1(t)$ and $Q_2(t)$ have coefficients in $Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Phi]$ It will then follow that both of them will be invariant under the action of G_Φ. In particular, we must have

$$\tau_i(Q_1(\Psi)) \;=\; Q_1(\tau_i \Psi) \qquad (\ for\ any\quad i = 1, 2, \ldots, k\ )$$

So if Ψ is a root of the equation $Q_1(t) = 0$ then all the other roots of $Q(t)$ must satisfy it as well and $Q_2(t)$ must reduce to a constant in $Rat[\mathcal{F}, e_1, e_2, \ldots, e_n, \Phi]$. The analogous conclusion holds if $Q_2(\Psi) = 0$. Thus $Q(t)$ is irreducible as asserted. This completes our proof.

Remark 4.2

Here and after, if $G_\Psi \subseteq_k G_\Phi$ and we have the left coset decomposition in 4.11, then functions

$$\Psi_1 = \tau_1 \Psi \ ,\ \Psi_2 = \tau_2 \Psi \ ,\ \ldots, \Psi_k = \tau_k \Psi$$

will be referred to as the *conjugates* of Ψ in G_Φ. Note that if Φ is in $Rat[\mathcal{F}, e_1, e_2, \ldots, e_n, \Psi]$, Theorem 4.1 assures that $G_\Psi \in G_\Phi$. So in any case we must have 4.11 for some k. Now suppose that Ψ is a root of the equation $R(t) = 0$ where $R(t)$ is a polynomial of degree h with coefficients in $Rat[\mathcal{F}, e_1, e_2, \ldots, e_n, \Phi]$. Since this polynomial is then invariant under the action of G_Φ all the conjugates of Ψ in G_Φ must also be roots of $R(t) = 0$. This implies that the polynomial $Q(t)$ in 4.12 must be a factor of $R(t)$. However, if $R(t)$ is also irreducible in $Rat[\mathcal{F}, e_1, e_2, \ldots, e_n, \Phi]$, then $R(t)$ and $Q(t)$ can only differ by a factor in $Rat[\mathcal{F}, e_1, e_2, \ldots, e_n, \Phi]$ and we must also $h = k$. This should explain why we call $PS_1, \Psi_2, \ldots, \Psi_k$ the "conjugates" of Ψ.

We have reached a point where to proceed further we need to make more precise what we mean by *solving the general equation by radicals*. To begin with we shall assume that the given field $\mathcal{F}$ (nowdays referred to as the *ground field*) contains all the roots of unity of any order $\leq n$. This given, *solving by radicals* the n^{tic} in 4.1, *in the Lagrange setting* is to mean that we can find a sequence of rational functions $\Phi_i \in Rat[\mathcal{F}; x_1, x_2, \ldots, x_n]$ $(i = 0, 1, 2, \ldots)$ such that

$$\begin{aligned} \Phi_o \;&\in\; Rat[\mathcal{F}; e_1, e_2, \ldots, e_n] \\ \Phi_i \;&=\; \sqrt[p_i]{\theta_i(e_1, e_2, \ldots, e_n, \Phi_{i-1})} \end{aligned} \tag{4.13}$$

where each θ_i is a rational function

$$\theta_i \;=\; \theta_i(y_1, y_2, \ldots, y_n, t) \;\in\; Rat[\mathcal{F}, y_1, y_2, \ldots, y_n, t] \ . \tag{4.14}$$

Finally, we shall require that the end function of this sequence say Φ_d be one of the roots or better yet (as we shall see) a function from which all the roots may be derived by rational operations.

Note first that since for any two integers p and q and for any Φ we have

$$\sqrt[pq]{\Phi} \;=\; \sqrt[p]{\sqrt[q]{\Phi}}$$

there is no loss in requiring that the integers p_i in 4.13 are all primes. Finally we can simplify the convoluted form of the recursion in 4.14 by rewriting it in the form

$$\begin{aligned} &a)\quad \Phi_o \;\in Rat[\mathcal{F}; e_1, e_2, \ldots, e_n] \\ &b)\quad \Phi_i^{p_i} \in Rat[\mathcal{F}, e_1, e_2, \ldots, e_n, \Phi_{i-1}) \end{aligned} \tag{4.15}$$

Remark 4.3

Note further that as long as the coefficients $m_1, m_2, \ldots, m_n$ are all distinct the function

$$v(x) \;=\; v(x_1, x_2, \ldots, x_n) \;=\; m_1 x_1 + m_2 x_2 + \cdots + m_n x_n \tag{4.16}$$

will necessarily be $n!$-valued. Since its stabilizer consists of just the identity permutation, the hypotheses of Theorem 4.1 are satisfied for any rational function of the roots $x_1, x_2, \ldots, x_n$. Thus for any $\Phi \in Rat[\mathcal{F}; x_1, \ldots, x_n]$ we can construct a rational function $\theta_\Phi(t) \in Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, t]$ giving

$$\Phi(x_1, x_2, \ldots, x_n) \;=\; \theta_\Phi\big(v(x_1, x_2, \ldots, x_n)\big) \;. \tag{4.17}$$

We should note here for further reference that the proof of Theorem 4.1 yields that

$$\theta_\Phi \;=\; \frac{Q_\Phi(t)}{P'(t)} \tag{4.18}$$

where

$$Q_\Phi(t) \;=\; \sum_{\sigma\in S_n} \sigma\Phi \prod_{\tau\in S_n}{}^{(\sigma)}\,(t-\tau v) \tag{4.19}$$

and

$$P(t) \;=\; \prod_{\tau\in S_n} (t-\tau v) \;. \tag{4.20}$$

It is important to notice that the denominator of θ_Φ in 4.18 is independent of Φ itself.

5. Lagrange's derivation of the roots of the cubic and the quartic.

Armed with this information, Lagrange could then come up with the following a-priori reconstruction of the solutions of the cubic and the quartic.

a) *The cubic*

Let $\omega = e^{2\pi i/3}$ and note that, since $1, \omega, \omega^2$ are distinct, the expression

$$(x_1 + \omega x_2 + \omega^2 x_3)/3$$

is necessarily a 6-valued function of the roots of $E_3(t)$. Its values are

$$\begin{array}{ll} z_1 = (x_1+\omega x_2+\omega^2 x_3)/3 & z_2 = (x_1+\omega x_3+\omega^2 x_2)/3 \\ z_3 = (x_2+\omega x_3+\omega^2 x_1)/3 & z_4 = (x_2+\omega x_1+\omega^2 x_3)/3 \\ z_5 = (x_3+\omega x_1+\omega^2 x_2)/3 & z_6 = (x_3+\omega x_2+\omega^2 x_1)/3 \end{array} \;. \tag{5.1}$$

Now we see that

$$z_3 = \omega^2 z_1 \;\;,\;\; z_5 = \omega z_1 \qquad \text{and} \qquad z_4 = \omega z_2 \;\;,\;\; z_6 = \omega^2 z_2 \;.$$

Thus

$$\begin{aligned} (z-z_1)(z-z_3)(z-z_5) \;&=\; z^3 - z_1^3 \\ (z-z_2)(z-z_4)(z-z_6) \;&=\; z^3 - z_2^3 \end{aligned}$$

This implies that the 6-degree equation

$$(z-z_1)(z-z_2)(z-z_3)(z-z_4)(z-z_5)(z-z_6) \;=\; 0 \tag{5.2}$$

must take the form

$$(z^3 - z_1^3)(z^3 - z_2^3) \;=\; z^6 - (z_1^3 + z_2^3)\, z^3 \;+\; z_1^3 z_2^3 \;=\; 0 \;. \tag{5.3}$$

Lagrange, on the basis that 5.2 is symmetric in the roots of $E_3(t)$, could now predict that the expressions

$$z_1^3 + z_2^3 \quad \text{and} \quad z_1^3 z_2^3$$

must necessarily be polynomials in e_1, e_2, e_3. And indeed it can directly be checked from 5.1 that

$$z_1^3 + z_2^3 \;=\; -q \quad \text{and} \quad z_1^3 z_2^3 \;=\; -(p/3)^3 \tag{5.4}$$

with p and q given by 3.3. This immediately leads to the bicubic equation

$$z^6 + qz^3 - (p/3)^3 \;=\; 0 \;. \tag{5.5}$$

Now its solution leads to the extraction of the square root of

$$\Phi_o = q^2 + 4(p/3)^3 = 4\,R \;.$$

This expression is essentially the discriminant of $E_3(t)$. More precisely we have

$$\Phi_o(x) \;=\; -\tfrac{1}{27}(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2$$

and its square root may be chosen to be

$$\Phi_1(x) \;=\; i\big(\frac{x_1-x_2}{\sqrt[2]{3}}\big)\big(\frac{x_1-x_3}{\sqrt[2]{3}}\big)\big(\frac{x_2-x_3}{\sqrt[3]{3}}\big) \;.$$

This is a polynomial in the roots of $E_3(t)$ whose stabilizer G_{Φ_1} is the group of even permutations of $S_3 = G_{\Phi_o}$. Now this is in perfect agreement with Theorem 4.2.

Of course, in view of 5.4, we also have

$$\Phi_o(x) \;=\; (z_1^3 + z_2^3)^2 - 4(z_1^3 z_2^3) \;=\; (z_1^3 - z_2^3)^2$$

and it can be easily verified that

$$\Phi_1(x) \;=\; z_1^3 - z_2^3 \;.$$

In other words $\Phi_1(x)$ is the solution of

$$\Phi_1^2(x) \;=\; \Phi_o \;.$$

Using 5.4 again we deduce that

$$z_1^3 \;=\; -\tfrac{1}{2}q \;+\; \tfrac{1}{2}\Phi_1(x) \;.$$

This given, the final step is the construction of the 6-valued function

$$\Phi_2(x) = z_1 = (x_1 + \omega x_2 + \omega^2 x_3)/3$$

whose stabilizer G_{Φ_2} is trivial and is the solution of

$$\Phi_2(x)^3 \;=\; -\tfrac{1}{2}\,q \;+\; \tfrac{1}{2}\Phi_1(x)\ .$$

In summary, this construction of a 3!-valued function of the roots of $E_3(t)$ has led us to the following scheme:

function		*expression*	*group*
Φ_o	$=$	$-\frac{1}{27}(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2$	S_3
$\downarrow$			$3\downarrow$
Φ_1	$=$	$i(x_1-x_2)(x_1-x_3)(x_2-x_3)/3\sqrt[2]{3}$	$\{id,(1,2,3),(1,3,2)\}$
$\downarrow$			$2\downarrow$
Φ_2	$=$	$(x_1+\omega x_2+\omega^2 x_3)/3$	$\{id\}$

Moreover we have that

$$\Phi_2^3 \;=\; -\tfrac{1}{2}\,q+\tfrac{1}{2}\,\Phi_1 \quad , \quad \Phi_1^2 \;=\; \Phi_o$$

This shows that the solution of the cubic can be obtained the succession of steps

$$\begin{array}{ccccc} \Phi_2^3\in Rat[\mathcal{F},e_1,e_2,e_3,\Phi_1] & \longleftarrow & \Phi_1^2\in Rat[\mathcal{F},e_1,e_2,e_3,\Phi_o] & \longleftarrow & \Phi_o\in Rat[\mathcal{F},e_1,e_2,e_3] \\ G_{\Phi_2}\subseteq G_{\Phi_1} & & G_{\Phi_1}\subseteq G_{\Phi_o} & & G_{\Phi_o}=S_3 \end{array}$$

b) *The quartic*

We can proceed in the same manner as for the cubic and construct a sequence of rational functions of the roots according to the scheme expressed in 4.15, terminating again with a 4!-valued function

$$v(x) \;=\; m_1x_1+m_2x_2+m_3x_3+m_4x_4\ .$$

We may choose here

$$v(x) \;=\; x_1-x_2+i(x_3-x_4)\ . \tag{5.6}$$

As in Section 3, we set $\omega=e^{2\pi i/3}$ and

$$y_1=x_1x_2+x_3x_4 \quad , \quad y_2=x_1x_3+x_2x_4 \quad , \quad y_3=x_1x_4+x_2x_3\ . \tag{5.7}$$

Moreover we let

$$w_1=x_1+x_2-x_3-x_4 \quad , \quad w_2=x_1+x_3-x_2-x_4 \quad , \quad w_3=x_1+x_4-x_2-x_3\ . \tag{5.7}$$

This given, we find that in this case repetitive uses of Lagrange's Theorem 4.1 naturally leads us to the the following scheme:

function		*expression*	*group*
Φ_o	$=$	$(x_1-x_2)^2(x_1-x_3)^2\cdots(x_3-x_4)^2$	S_4
$\downarrow$			$2\downarrow$
Φ_1	$=$	$(x_1-x_2)(x_1-x_3)\cdots(x_3-x_4)$	A_4
$\downarrow$			$3\downarrow$
Φ_2	$=$	$y_1+\omega y_2+\omega^2 y_3$	$\{id,(1,2)(3,4),(1,3)(2,4),(1,4)(2,3)\}$
$\downarrow$			$2\downarrow$
Φ_3	$=$	$w_2w_3+i(y_2-y_3)$	$\{id,(1,2)(3,4)\}$
$\downarrow$			$2\downarrow$
Φ_4	$=$	$x_1-x_2+i(x_3-x_4)$	$\{id\}$

(5.8)

The fact that $G_{\Phi_o}=S_4$ is immediate since Φ_o is a symmetric function of the roots whose expression in terms of e_1,e_2,e_2,e_4 is given by the discriminant formula 2.8. Clearly Φ_1 is invariant only under even permutations of the roots. So G_{Φ_1} is simply the alternating group A_4. To obtain G_{Φ_2}, we simply observe that, since $1,\omega,\omega^2$ are distinct, Φ_2 is invariant only under those permutations of the roots that leave y_1,y_2,y_3 individually invariant. In other words $G_{\Phi_2}=G_{y_1}\cap G_{y_3}\cap G_{y_3}$. This gives us the third entry in the fourth column of 5.8. To obtain G_{Φ_3} we note that for y_2-y_3 not to change we need each of y_1,y_2,y_3 to remain unchanged, thus $G_{y_2-y_3}=G_{\Phi_2}$. Now it is easily seen that

$$G_{w_2w_3} \;=\; \{id,(1,2),(3,4),(1,2)(3,4)\}$$

and thus we must have

$$G_{\Phi_3} \;=\; G_{w_2w_3}\cap G_{\Phi_2} \;=\; \{id,(1,2)(3,4)\}$$

as asserted in 5.8.

We can easily see that

$$\Phi_1^2 \;=\; \Phi_o \quad \text{and} \quad \Phi_4^2 \;=\; \Phi_3 \tag{5.9}$$

We can painlessly check on the computer that

$$\Phi_2^3 \;=\; \tfrac{3}{2}\sqrt[2]{3}\,\Phi_1 \;+\; J\ , \tag{5.10}$$

where

$$J \;=\; \tfrac{1}{2}(2y_1-y_2-y_3)(2y_2-y_1-y_3)(2y_3-y_1-y_2)\ .$$

Note next that we have the coset decomposition

$$G_{\Phi_2} \;=\; G_{\Phi_2} \;+\; (1,3)(2,4)G_{\Phi_2}$$

Thus the *conjugate* of Φ_3 in G_{Φ_2} is

$$\Phi_3' \;=\; (1,3)(2,4)\Phi_3 \;=\; -w_2w_3+i(y_2-y_3)\ .$$

We can now immediately conclude from Theorem 4.2 that the coefficients of the polynomial

$$(t-\Phi_3)(t-\Phi_3') = \big(t-i(y_2-y_3)\big)^2-w_2^2w_3^2 = t^2-4i(y_2-y_3)t-4(y_2-y_3)^2-w_2^2w_3^2 \;=\; \tag{5.11}$$

must be in $Rat[\mathcal{F};e_1,e_2,e_3,e_4,\Phi_2]$. It is interesting to see what they actually turn out to be. For instance we can write

$$w_2^2w_3^2 \;=\; \frac{(w_1w_2w_3)^2}{w_1^2}\ . \tag{5.12}$$

Now from 3.10 we get that

$$w_1^2 \;=\; e_1^2-4e_2+4y_1\ . \tag{5.13}$$

Since $G_{\Phi_2} \subseteq G_{y_1}$, we know that y_1 should be in $Rat[\mathcal{F}; e_1, e_2, e_3, e_4, \Phi_2]$. In fact, we can easily verify that

$$y_1 \;=\; \tfrac{1}{3}\,(e_2 + \Phi_2 + \overline{\Phi}_2)\ . \tag{5.14}$$

In case we might worry that the complex conjugate $\overline{\Phi}_2$ may not be in $Rat[\mathcal{F}; e_1, e_2, e_3, e_4, \Phi_2]$, I will quickly point out that we have the identity

$$\overline{\Phi}_2 \;=\; \frac{e_1^2 - 3e_2}{\Phi_2}\ .$$

Similarly, we can easily derive that

$$i(y_2 - y_3) \;=\; \frac{1}{\sqrt[3]{2}}\,(\Phi_2 - \overline{\Phi}_2)\ .$$

Finally combining 5.11,5.12,5.13,5.14 we get that the function

$$\Phi_3^* \;=\; \Phi_3 - i(y_2 - y_3)$$

satisfies the equation

$$(\Phi_3^*)^2 \;=\; \frac{w_1^2 w_2^2 w_3^2}{e_1^2 - 4e_2 + 4(e_2 + \Phi_2 + \overline{\Phi}_2)/3}\ .$$

Thus we have again obtained the desired $n!$-valued function by constructing a sequence of functions $\Phi_i(x) \in Rat[\mathcal{F}; x_1, x_2, \ldots, x_n]$ satisfying the recurrence

$$\Phi_i(x)^{p_i} \in Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Phi_{i-1}] \qquad (\ \text{with} \quad \Phi_o \in Rat[\mathcal{F}; e_1, e_2, \ldots, e_n]\ ) \tag{5.14}$$

Where, exponents $p_1, p_2, p_3 \ldots$ give the prime factorization of $n!$. Moreover we also have verified that the additional condition

$$G_{\Phi_i} \subseteq G_{\Phi_{i-1}} \tag{5.15}$$

holds throughout.

This given, Lagrange was convinced (and, with hindsight, he was right), that if the quintic should be solvable by radicals, then the $5!$-valued function of its roots must be reachable by a sequence of steps as given in 5.14 and 5.15. However in trying to reproduce the same scheme for the quintic equation he run into insurmountable difficulties. His final efforts were towards constructing a rational function of the roots $x_1, x_2 \ldots, x_5$ that took less than 5 values or one that took exactly five values but was the solution of a binomial quintic. The reason for this search was his need to reduce the construction of the roots to the solution of an equation of degree less than five or to an equation of degree 5 he could solve by taking radicals. Of course he was allowing the possibility, of having to solve an equation of degree greater than five as long as this equation, as in the case of the cubic, could be reduced to the solution of an equation of lesser degree for a power of one of its roots. However, this power would then have to be a k-valued function for some $k < 5$. The best he could do was to produce a 6-valued function. He concludes his work (see []) by saying that although he

had not tried every possibility, the search (without MAPLE or MATHEMATICA) was considerably time consuming... and he had no further time to spend in the search for something

"whose existence is very much in doubt"

He had again the correct suspicion! The non existence of what he was looking for was proved (30 years later) by Ruffini and generalized for $n > 5$ by Cauchy.

Remark 5.1

It is interesting in this connection to observe that there is a neat representation theoretical reason why there is no k-valued function of the roots of $E_n(t)$ when $2 < k < n$ and $n \geq 5$. In fact, the action of S_n on the stabilizer of any k-valued function Φ induces a permutation representation with only one occurrence of the trivial. If the remaining irreducible constituents are all sign representations then Φ is only 2 valued. So for $k > 2$ this representation would have to have a irreducible constituent of degree ≥ 2. But for $n \geq 5$ except for the trivial and the sign representation all the other irreducible representations have dimension $\geq n - 1$. So $k \leq 2$ or $k \geq n$. At any rate we can also give this result an elementary proof.

It is important to see at this point what are the implications of the existence of a sequence of functions Φ_i satisfying 5.14 and 5.15. To this end we terminate this section with a collection of results which shed considerable light on the difficulties encountered by Lagrange in completing his program. We should note that although some of the arguments that follow use the "group" structure of stabilizers, which is one of the main discoveries of Galois, most of the calculations are actually due to Lagrange. This apparent paradox illustrates in a remarkable way how close Lagrange got to discover some of the main points of "Galois" theory.

Theorem 5.1

Let Φ and Ψ be in $Rat[\mathcal{F}; e_1, e_2, \ldots, e_n]$. Suppose that

$$\begin{aligned} &a) \quad G_\Psi \subseteq G_\Phi \quad but \quad G_\Psi \neq G_\Phi \quad , and \\ &b) \quad \Psi^p \in Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Phi] \end{aligned} \tag{5.16}$$

where p is prime. Then the decomposition of G_Φ into left cosets of G_Ψ may be written in the form

$$G_\Phi \;=\; G_\Psi + \gamma G_\Psi + \gamma^2 G_\Psi + \cdots + \gamma^{p-1} G_\Psi \tag{5.17}$$

with γ a p-cycle that commutes with G_Ψ. Moreover, the conjugates of Ψ in G_Φ can be written in the form

$$\Psi_i = \gamma^{i-1}\Psi = \omega^{i-1}\Psi \qquad (\ with \quad \omega = e^{2\pi i/p}\ ) \tag{5.18}$$

Proof

Since by assumption $G_\Psi \neq G_\Phi$ we have

$$G_\Phi \;=\; \tau_1 G_\Psi + \tau_2 G_\Psi + \cdots + \tau_k G_\Psi\ .$$

with $k > 1$. Set $\omega = e^{2\pi i/p}$ and let

$$Q(t) \;=\; \prod_{i=0}^{p-1} (t - \omega^i \Psi)\ . \tag{5.19}$$

From 5.16 b) we deduce that

$$Q(t) \;=\; t^p \;-\; R(x)$$

with $R(x) \in Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Phi]$. In particular for any $\sigma \in G_\Phi$ we must have

$$\sigma Q(t) \;=\; \prod_{i=0}^{p-1} (t - \omega^i \sigma\Psi) \;=\; Q(t) \;=\; \prod_{i=0}^{p-1} (t - \omega^i \Psi) \;.$$

and this gives

$$\sigma\Psi \;=\; \omega^{i(\sigma)}\, \Psi \qquad 5.20$$

for some $0 \le i(\sigma) \le p-1$. This implies that the conjugates of Ψ are all multiples of Ψ by powers of ω. Thus

$$k \le p \;.$$

Now $k > 1$ gives that $\tau_2\Psi \neq \Psi$ so $\; i(\tau_2) \neq 0 \;\; (mod\;\; p)$. But since p is prime, the successive powers $(\omega^{i(\tau_2)})^s$ for $s = 0, 1, \ldots, p-1$ are simply $1, \omega, \ldots, \omega^{p-1}$ in some order. Thus

$$\Psi, \tau_2\Psi, \tau_2^2\Psi \ldots, \tau_2^{p-1}\Psi$$

are all distinct. This gives

$$p \le k$$

Thus k must be equal to p and

$$\tau_2^p = id \;.$$

Moreover, we could have indexed our coset representatives τ_i to be successive powers of an element $\gamma \in G_\Phi$ for which $\gamma\Psi = \omega\Psi$. This gives 5.17. With this choice we have

$$\Psi_i \;=\; \gamma^i\Psi \;=\; \omega^{i-1}\, \Psi \qquad (\; i = 1, 2, \ldots, p \;)$$

Finally, we see that $\gamma\Psi = \omega\Psi$ gives

$$\gamma^{-1} h \gamma\, \Psi \;=\; \Psi \qquad (\; \forall \quad h \in G_\Psi) \;.$$

In other words

$$\gamma^{-1}\, G_\Psi \gamma \;=\; G_\Psi \qquad 5.21$$

This completes our proof.

Remark 5.2

In modern terminology this theorem simply says that 5.16 a) and b) imply that G_Ψ is a normal subgroup G_Φ and the quotient G_Φ/G_Ψ is isomorphic to the group of integers $mod\, p$. To express these two properties we shall here and after write

$$G_\Psi \lhd_p G_\Phi \;. \qquad 5.22$$

Thus the possibility of solving the general equation $E_n(t)$ by radicals as was done with the cubic and the quartic according to the scheme given by 5.14 and 5.15 requires that we should have a sequence of subgroups $G_i \subseteq S_n$ (for $i = 0, 1, \ldots, k$) satisfying the conditions

$$\{id\} = G_m \lhd_{p_m} G_{m-1} \lhd_{p_{m-1}} \cdots G_2 \lhd_{p_2} G_1 \lhd_{p_1} G_o = S_n \qquad 5.23$$

We shall soon see that this is impossible for $n \ge 5$. However, we must first establish some basic facts about S_n and its subgroups.

Proposition 5.1

Let G be a subgroup of S_n which contains all 3-cycles and let H be a subgroup G. Suppose that for some $\gamma \in G$ we have

$$\begin{aligned} &a) \quad G = H + \gamma H + \cdots + \gamma^{p-1} H \;, \\ &b) \quad \gamma H \;=\; H\gamma \;. \end{aligned} \qquad 5.24$$

Then for $n \ge 5$ also H contains all 3-cycles.

Proof

Condition a) in 5.24 gives that we can write

$$(1,2,3) \;=\; \gamma^i a \quad , \quad (3,4,5) \;=\; \gamma^j b$$

for some $0 \le i, j < p$ and $a, b \in H$. On the other hand b) in 5.24 gives that we can write

$$(1,2,3)(3,4,5) \;=\; \gamma^i a \gamma^j b = \gamma^{i+j}\, c \quad , \quad (3,2,1)(5,4,3) \;=\; a^{-1}\gamma^{-i} b^{-1} \gamma^{-j} = c'\gamma^{-i-j}$$

for some $c, c' \in H$. Thus

$$(2,5,3) \;=\; (3,2,1)(5,4,3)(1,2,3)(3,4,5) \;=\; c'\gamma^{-i-j}\gamma^{i+j}\, c \in H \;.$$

Since the same manipulations can be carried out when $1, 2, 3, 4, 5$ are replaced by any 5 distinct integers i_1, i_2, i_3, i_4, i_5, we see that H must contain all 3 cycles as asserted.

Proposition 5.2

If a subgroup $G \subseteq S_n$ contains all 3-cycles then $G = S_n$ or G is the alternating group A_n

Proof

Note that we have

$$\begin{aligned} (1,2)(2,3) \;&=\; (1,2,3) \\ (1,2)(3,4) \;&=\; (1,2)(2,3)(3,2)(2,4) \;=\; (1,3,2)(2,4,3) \end{aligned}$$

Thus the product of two 2-cycles can always be expressed as either a 3-cycle or as a product of two 3-cycles. This implies that every even permutation is a product of 3-cycles. Thus under our hypothesis

$G \supseteq A_n$. Now if G contains a single permutation σ not in A_n (that is an odd permutation), the identity

$$S_n \;=\; A_n \;+\; \sigma A_n$$

yields that $G = S_n$ as desired.

Combining these two propositions we derive that we cannot have 5.23 when $n \geq 5$. In fact the string of inclusions in 5.23 must stop with $G_1 = A_n$ and can go no further! We can thus state

Theorem 5.2 (Ruffini)

For $n \geq 5$ the general equation $E_n(t) = 0$ cannot be solved by radicals by successive adjunctions of rational functions $\Phi_i \in Rat[\mathcal{F}, x_1, x_2, \ldots, x_n]$ satisfying the recursions in 5.14 and 5.15 as was possible for the cubic and the quartic.

We should note that this result doesn't quite put to rest the possibility of solving the general equation by adjoining "radicals". This is because there are still some unjustified hypotheses in the present setting. First of all, in our original definition (see 4.15) of *solution by radicals* we did not have the extra condition 5.15. As we shall see this is a relatively minor point, easily fixed in the Galois setting. Considerably more troubling is that in this definition we require that each newly constructed Φ_i should turn out to be a *rational function in the roots*. What if we allow such steps as taking a p^{th} root of $x_1 x_2 \cdots x_n$? Can we then construct a solution? These doubts would be removed if we could show that the existence of a sequence of steps as in 5.14, 5.15 without the further assumption that $\Phi_i \in Rat[\mathcal{F}, x_1, x_2, \ldots, x_n]$ implies the existence of the same sequence of steps with this assumption **satisfied**. This is in essence what Abel succeeded in showing. We shall not present Abel's arguments here since what must be proved to remove the remaining doubts can be done in an easier way in the Galois setting. We shall thus terminate our treatment of Lagrange's "Galois" Theory with one final result which in some sense reverses Theorem 5.1.

Theorem 5.3

If $\Psi, \Phi \in Rat[\mathcal{F}; x_1, x_2, \ldots, x_n]$ and

$$G_\Phi \;=\; G_\Psi + \gamma G_\Psi + \gamma^2 G_\Psi + \cdots + \gamma^{p-1} G_\Psi \qquad 5.25$$

with $\gamma^p = id$ and

$$\gamma G_\Psi \;=\; G_\Psi \gamma \;. \qquad 5.26$$

Then we can find $\Theta \in Rat[\mathcal{F}; x_1, x_2, \ldots, x_n]$ such that

$$\begin{aligned} &1) \quad G_\Theta \supseteq G_\Psi \quad (\Theta \in Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Psi]), \\ &2) \quad \Theta^p \in Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Phi]. \end{aligned} \qquad 5.27$$

Proof

Set

$$\Theta \;=\; \Psi + \omega\gamma\Psi + (\omega\gamma)^2\Psi + \cdots + (\omega\gamma)^{p-1}\Psi \;. \qquad (\;\omega = e^{2\pi i/3}\;) \qquad 5.28$$

It is easily verified that

$$\gamma\Theta \;=\; \omega^{-1}\Theta \;. \qquad 5.29$$

Moreover, 5.26 gives that each of the conjugates $\Psi_i = \gamma^{i-1}\Psi$ has the same stabilizer as Ψ. Thus from the definition 5.28 we get that

$$G_\Theta \supseteq G_\Psi \qquad 5.30$$

and therefore (again by Theorem 4.1) we must have

$$\Theta \in Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Psi]$$

Finally, 5.30 together with 5.25 and 5.26 implies that the polynomial

$$Q(t) \;=\; \prod_{i=1}^{p}(t - \gamma^{i-1}\Theta) \;=\; \prod_{i=1}^{p}(t - \Theta/\omega^{i-1}) \;=\; t^p - \Theta^p \qquad 5.31$$

is left invariant by every element of G_Φ. Thus from Theorem 4.1 (with $\Psi \to \Phi$ and $\Phi \to \Theta$) we derive that

$$\Theta^p \in Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Phi] \;,$$

as desired.

Remark 5.3

We should point out that the inequality in 5.30 cannot in general be sharpened to an equality. Indeed, in our construction of the roots of the quartic, the function

$$\Phi_3 = w_2 w_3 + i(y_2 - y_3)$$

satisfies the conditions in 5.25 and 5.26 with $\Psi \to \Phi_3$, $\Phi \to \Phi_2$ $p = 2$. In this case we have

$$\gamma = (1,3)(2,4) \quad , \quad \omega = -1$$

and

$$\Theta = \Phi_3 + \omega\gamma\Phi_3 = 2w_2 w_3 \;.$$

As we have seen

$$G_{w_2 w_3} \;=\; \{id, (1,2), (3,4), (1,2)(3,4)\}$$

while

$$G_{\Phi_3} \;=\; \{id, (1,2)(3,4)\} \;.$$

Nevertheless, in a typical application of Theorem 5.3 we may want to obtain Θ by extracting a p^{th} root of a function in $Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Phi]$ then recover Ψ by rational operations involving Θ and possibly other known functions. Now this is provided by the following beautiful identity.

Theorem 5.4 (Lagrange)

If $\Psi, \Phi \in Rat[\mathcal{F}; x_1, x_2, \ldots, x_n]$ and we have 5.25, 5.26 with $\gamma^p = id$ and p a prime. Then

$$\Psi \;=\; (c_o + c_1\Theta + \cdots + c_{p-1}\Theta^{p-1})/p \tag{5.32}$$

where all the coefficients c_i are in $Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Phi]$.

Proof

For convenience set $\Psi_s = \gamma^s \Psi$

$$\Theta_i \;=\; \sum_{s=0}^{p-1} \omega^{si} \Psi_s \ .$$

Then it is easily verified that

$$(1 - \omega^i \gamma)\Theta_i \;=\; 0 \ . \tag{5.33}$$

Moreover, since for p a prime and any $0 < s < p$ we have

$$1 \;+\; \omega^s \;+\; \omega^{2s} + \cdots + \omega^{(p-1)s} \;=\; 0 \ ,$$

we immediately derive that

$$\Psi \;=\; (\Theta_1 + \Theta_2 + \cdots + \Theta_p)/p \ . \tag{5.34}$$

On the other hand, 5.33 gives that the ratios

$$c_i \;=\; \Theta_i/\Theta^i \tag{5.35}$$

are all invariant under γ. Since as we have seen 5.26 gives that all the conjugates Ψ_i of Ψ in G_Φ have G_Ψ as stabilizer, we deduce that these ratios are stable under G_Φ and (again from Thorem 4.1) we must conclude that they are all in $Rat[\mathcal{F}; e_1, e_2, \ldots, e_n, \Phi]$. This given, we see that by substituting 5.35 in 5.34 we get 5.32 with the desired properties.

6. Galoisian "Galois Theory.

Lagrange's pursuits were brought to a conclusion by Abel around 1829 when Galois started his investigations. Galois' starting point may have been the idea that although there may be no general formulas for the roots of $E_n(t)$ for $n \geq 5$ which involved only extraction of roots, it may still be possible to find them for some special equations. This possibility had already emerged in the work of Gauss and Abel but it is not clear to what extent Galois had been aware of their work. We may say that his discoveries stemmed from a natural extension of Lagrange's original methods. To present Galois' results, we shall need some additional notation and definitions.

Our basic ingredients here will be two fields $\mathcal{F}$ and $\mathcal{E}$, with $\mathcal{F}$ a proper subfield of $\mathcal{E}$. The equation to be solved will be written as

$$\tilde{E}_n(t) \;=\; (t - \alpha_1)(t - \alpha_2)\cdots(t - \alpha_n) \;=\; 0 \ . \tag{6.1}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are distinct and, unless explicitely mentioned, will remain unchanged throughout our presentation. We shall also assume hereafter that

$$\begin{aligned} &1) \quad \alpha_1, \alpha_2, \ldots, \alpha_n \;\in\; \mathcal{E} \quad \text{and} \\ &2) \quad e_1(\alpha_1, \alpha_2, \ldots, \alpha_n), e_2(\alpha_1, \alpha_2, \ldots, \alpha_n), \ldots, e_n(\alpha_1, \alpha_2, \ldots, \alpha_n) \;\in\; \mathcal{F} \end{aligned} \tag{6.2}$$

We see that we shall have to work here with functions $\Phi(x_1, x_2, \ldots, x_n)$ of the independent variables $x_1, x_2, \ldots, x_n$ and at the same study their values when when $x_1, x_2, \ldots, x_n$ are replaced by $\alpha_1, \alpha_2, \ldots, \alpha_n$. As clarity requires, these values will be represented by any of the symbols below

$$\Phi(\alpha_1, \alpha_2, \ldots, \alpha_n) \;=\; \Phi(\alpha) \;=\; \tilde{\Phi} \ . \tag{6.3}$$

For a given $\sigma \in S_n$ we shall also use the symbols

$$\sigma\Phi(\alpha_1, \alpha_2, \ldots, \alpha_n) \;=\; \sigma\Phi(\alpha) \;=\; \sigma\tilde{\Phi}$$

to denote the value $\Phi(\alpha_{\sigma_1}, \alpha_{\sigma_2}, \ldots, \alpha_{\sigma_n})$.

This given, one of the fundamental differences between the Galoisian and Lagrangian setups is that although a given $\Phi \in Rat[\mathcal{F}; x_1, x_2, \ldots, x_n]$ and all its images $\sigma\Phi$ are well defined as elements of $Rat[\mathcal{F}; x_1, x_2, \ldots, x_n]$ some of the values $\sigma\tilde{\Phi}$ may make no sense at at all. An example in point is the rational function

$$\Phi(x) \;=\; \frac{1}{x_1^2 + x_2^2 - x_3^2}$$

when $n = 3$ and $\alpha_2^2 + \alpha_3^2 - \alpha_1^2 = 0$.

In summary in the Galois setting, we have to be careful with denominators! We shall avoid the problem by dealing hereafter only with polynomials functions of the roots. As we shall see this is not a serious restriction, and with it, most of the results of Lagrange Theory can be extended to the Galois setting with nearly identical proofs.

Another important difference is that for some $\Phi \in \mathcal{F}[x_1, \ldots, x_n]$ there may be more permutations of the roots of $\tilde{E}_n(t)$ that leave $\Phi(\alpha_1, \alpha_2, \ldots, \alpha_n)$ unchanged than there are in G_Φ. We shall take account of this difference by setting for $\Phi \in \mathcal{F}[x_1, \ldots, x_n]$

$$S_\Phi \;=\; \{\sigma \in S_n \;:\; \Phi(\alpha_{\sigma_1}, \alpha_{\sigma_2}, \ldots, \alpha_{\sigma_n}) = \Phi(\alpha_1, \alpha_2, \ldots, \alpha_n)\;\} \tag{6.4}$$

Note that S_Φ should be considered a property of Φ and not a property of the value $\Phi(\alpha)$. Moreover we should point out that in general this collection of permutations may not even be a group! For instance for the equation

$$\tilde{E}_3(t) \;=\; =\; (t+1)(t+i)(t-i) \;=\; t^3 + t^2 + t + 1\ ,$$

we can take $\mathcal{F}$ and $\mathcal{E}$ to be the fields of rational and complex numbers respectively. Now, if we label the roots by setting

$$-1 = \alpha_1 \quad , \quad i = \alpha_2 \quad , \quad -i = \alpha_3$$

then for $\Phi(x_1, x_2, x_3) \;=\; x_2^2$ we have (in cycle notation)

$$G_\Phi \;=\; \{id\,,\ (1,3)\}\ .$$

On the other hand the permutations of $\alpha_1, \alpha_2, \alpha_3$ that leave $\tilde{\Phi}$ unchanged form the set

$$\{id\,,\ (1,3)\,,\ (2,3)\,,\ (1,2,3)\}$$

This is not a group since $(1,2) = (1,2,3)(2,3)$, yet

$$\Phi(\alpha_1, \alpha_2, \alpha_3) = \alpha_2 = -1 \quad \text{and} \quad \Phi(\alpha_2, \alpha_1, \alpha_3) = \alpha_1^2 = 1\ .$$

To take account of this possibility we shall say that a given $\Phi \in \mathcal{F}[x_1, \ldots, x_n]$ is *Galois* if and only if

$$S_\Phi \;=\; G_\Phi$$

We are now in a position to proceed with our treatment.

Theorem 6.1

For every subgroup $G \subseteq S_n$ we have a Galois $\Phi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$ such that

$$S_\Phi = G\ . \tag{6.5}$$

Proof

We start with the case $G = S_n$. Here can take again a linear function

$$v(x_1, x_2, \ldots, x_n) \;=\; m_1 x_1 + m_2 x_2 + \cdots + m_n x_n \tag{6.6}$$

as in 4.16, but we must be a bit more careful in choosing the coefficients m_i. To this end note that if $m_1, m_2, \ldots, m_n$ are chosen to be integers in the interval $[0, M]$, then for any given pair of distinct permutations $\sigma, \tau \in S_n$ the equation

$$m_1\alpha_{\sigma_1} + m_2\alpha_{\sigma_2} + \cdots + m_n\alpha_{\sigma_n} \;=\; m_1\alpha_{\tau_1} + m_2\alpha_{\tau_2} + \cdots + m_n\alpha_{\tau_n}$$

can have at most $(M+1)^{n-1}$ distinct solution vectors $(m_1, m_2, \ldots, m_n)$. Thus to assure that

$$\sigma v(\alpha) \neq \tau v(\alpha) \qquad (\ \forall \quad i \neq j\ )$$

we need to avoid at most $\binom{n!}{2} \times (M+1)^{n-1}$ vectors. However, when $M + 1 > \binom{n!}{2}$ there will remain some n^{tuples} for us to choose and satisfy our requirement that $S_v = S_n$. Having made one such choice of $(m_1, m_2, \ldots, m_n)$, the desired Φ for any given subgroup G can be readily produced. In fact, we can show that we can set

$$\Phi \;=\; \prod_{\sigma \in G} (N - \sigma v(x)) \tag{6.7}$$

where N is a suitably chosen integer. To see this note first that the form of 6.7 guarantees that whatever N we choose we shall have at least $S_\Phi \supseteq G$. Now let

$$S_n \;=\; \tau_1 G \;+\; \tau_2 G \;+\; \cdots \;+\; \tau_k G$$

be the decompostion of S_n into left cosets of G. This given, our choice of $m_1, m_2, \ldots, m_n$ assures that the polynomials

$$\tilde{P}_i(t) \;=\; \tau_i \prod_{\sigma \in G} (t - \sigma v(\alpha)) \;=\; \prod_{\sigma \in G} (t - \tau_i \sigma v(\alpha))$$

have no roots in common. Since they all have degree $|G|$, the equation

$$\tilde{P}_i(t) = \tilde{P}_j(t)$$

for $i \neq j$ can then have at most $|G|$ solutions. Thus if we want an integer N which gives $\tilde{P}_i(N) \neq \tilde{P}_j(N)$ for all $i \neq j$ we need avoid at most $\binom{k}{2} \times |G|$ values. Clearly we can find such an N in the interval $[0, M]$ as soon as $M > \binom{k}{2} \times |G|$. This completes our argument.

Remark 6.1

All the constructions, proofs and definitions in this section will use an $n!$-valued Galois function

$$v(x) \;=\; m_1 x_1 + m_2 x_2 + \cdots + m_n x_n$$

which must remain unchanged throughout the rest of the section. We must therefore make sure that some of the objects we introduce, such as for instance the "Galois Group" of our equation $\tilde{E}_n(t)$ do not depend on the choice of $m_1, m_2, \ldots, m_n$. This is one of the prices we have to pay for not

following the abstract approach. However, we believe that this will be well compensated by the additional insights that our insistence on explicit constructions will provide.

We begin by showing that the values of every polynomial $\Phi(x)$ are in fact polynomials in the values of $v(x)$. More precisely we have

Proposition 6.1

For any $\Phi \in \mathcal{F}[x_1, \dots, x_n]$ we can construct a polynomial $\Theta_\Phi(t) \in \mathcal{F}[t]$ such that for any $\gamma \in S_n$ we have

$$\gamma\tilde{\Phi} \;=\; \Theta_\Phi\big(\gamma\tilde{v}\big)\ . \tag{6.8}$$

Proof

As in section 4 (see 4.19 and 4.20) we let

$$Q(t) \;=\; \sum_{\sigma\in S_n} \sigma\Phi \prod_{\tau\in S_n}{}^{(\sigma)}\,(t-\tau v)\quad,\quad P(t) \;=\; \prod_{\tau\in S_n}(t-\tau v)\ . \tag{6.9}$$

Since both Q and P are by construction S_n-invariant, their coefficients are polynomials in $\mathcal{F}[e_1, e_2, \dots, e_n]$. The hypothesis in 6.3 2) then yields that the polynomials

$$\tilde{Q}(t)\ ,\ \tilde{P}(t)\ ,\ \tilde{P}'(t)$$

have coefficients in $\mathcal{F}$. Setting $t = \gamma\tilde{v}$ in

$$\tilde{Q}(t) \;=\; \sum_{\sigma\in S_n} \sigma\tilde{\Phi} \prod_{\tau\in S_n}{}^{(\sigma)}\,(t-\tau\tilde{v})$$

gives

$$\tilde{Q}(\gamma\tilde{v}) \;=\; \gamma\tilde{\Phi}\,\tilde{P}'(\gamma\tilde{v})\ . \tag{6.10}$$

Note further that since by construction v takes $n!$ distinct values, the polynomials $\tilde{P}$ and $\tilde{P}'$ have no common root. We can thus apply the Berlekamp algorithm and construct a pair of polynomials $p(t), q(t) \in \mathcal{F}[t]$ such that

$$p(t)\tilde{P}'(t) \;+\; q(t)\tilde{P}(t) \;=\; 1$$

Setting $t = \gamma\tilde{v}$ in this equation yields

$$p(\gamma\tilde{v})\tilde{P}'(\gamma\tilde{v}) \;=\; 1\ .$$

Multiplying both sides of 6.10 by $p(\gamma\tilde{v})$ and using this equation we finally get

$$p(\gamma\tilde{v})\tilde{Q}(\gamma\tilde{v}) \;=\; \gamma\tilde{\Phi}\ ,$$

and this gives 6.8 with

$$\Theta_\Phi(t) \;=\; p(t)\tilde{Q}(t)\ .$$

It develops that the polynomial

$$\tilde{P}(t) \;=\; \prod_{\tau\in S_n}(t-\tau\tilde{v}) \tag{6.11}$$

plays a crucial role in our development. It may or may not be reducible in $\mathcal{F}[t]$. If it is, we can write its factorization into irreducibles in the form

$$\tilde{P}(t) \;=\; \prod_{\tau\in T_1}(t-\tau\tilde{v})\ \prod_{\tau\in T_2}(t-\tau\tilde{v})\cdots\prod_{\tau\in T_m}(t-\tau\tilde{v}) \;=\; \tilde{P}_1(t)\tilde{P}_2(t)\cdots\tilde{P}_m(t)\ ,$$

where $T_1, T_2, \cdots, T_m$ are disjoint subsets, and T_1 is the subset that contains the identity permutation. Now we have the following crucial fact.

Proposition 6.2

The subset T_1 is a group

Proof

Since by definition T_1 contains the identity, we only need to show that if $\eta, \xi \in T_1$ then their product $\eta\xi$ is also in T_1. To this end, we use Proposition 6.1 and construct the polynomial $\Theta(t) \in \mathcal{F}[t]$ that gives 6.8 for $\Phi = \xi v$. We shall thus have

$$\gamma\xi\tilde{v} \;=\; \Theta(\gamma\tilde{v}) \qquad (\ \forall\ \ \gamma \in S_n)\ . \tag{6.12}$$

Now note that by hypothesis

$$\tilde{P}_1(\xi\tilde{v}) \;=\; \prod_{\tau\in T_1}(\xi\tilde{v}-\tau\tilde{v}) \;=\; 0.$$

In particular, using 6.12 for $\gamma = id$ we can rewrite this in the form

$$\tilde{P}_1\big(\Theta(\tilde{v})\big) \;=\; 0\ .$$

But this says that the polynomials $\tilde{P}_1(t)$ and $\tilde{P}_1\big(\Theta(t)\big)$ have a root in common. Since they are both in $\mathcal{F}[t]$ and $\tilde{P}_1(t)$ is irreducible in $\mathcal{F}[t]$, we deduce that $\tilde{P}_1\big(\Theta(t)\big)$ must vanish for all the other roots of $\tilde{P}_1(t)$. In particular we must have

$$\tilde{P}_1\big(\Theta(\eta\tilde{v})\big) \;=\; 0$$

Now using 6.12 with $\gamma = \eta$ this may yet be rewritten as

$$\tilde{P}_1(\eta\xi\tilde{v}) \;=\; 0\ .$$

But this implies that the permutation $\eta\xi$ lies in T_1 as well.

Q.E.D.

We shall hereafter denote T_1 by G and refer to it as the *Galois Group* of $\tilde{E}_n(t)$. The polynomial $\tilde{P}_1(t)$ itself will be referred to as a *Galois resolvent* of $\tilde{E}_n(t)$. For instance when $n = 3$ and

$$\tilde{E}_3(t) \;=\; t^3 + t^2 + t + 1\ ,$$

we may take $\mathcal{F}$ to be the field of rational numbers and $v = x_2 - x_1$. This gives

$$P(t) \;=\; \big(t^2-(x_2-x_1)^2\big)\big(t^2-(x_2-x_3)^2\big)\big(t^2-(x_3-x_1)^2\big)\ .$$

Now this can be rewritten as

$$P(t) \;=\; t^6+(6e_2-2e_1^2)+(e_1^2-3e2)^2t^2-(x_1-x_2)^2(x_1-x_3)^2(x_2-x_3)^2$$

and substituting $e_1 = 1$, $e_2 = 1$ and $e_3 = -1$ we get (using formula 2.7)

$$\tilde{P}(t) \;=\; t^6+4t^4+4t^2+16\ .$$

Its irreducible factorization is

$$\tilde{P}(t) \;=\; (t^2-2t+2)(t^2+2t+2)(t^2+4)\ .$$

Now the roots of t^3+t^2+t+1 are $-1, i, -i$ so if we label them $\alpha_1, \alpha_2, \alpha_3$ respectively, then the Galois resolvent is

$$\tilde{P}_1(t) \;=\; t^2-2t+2 \;=\; \big(t-(x_2-x_1)\big)\big(t-(x_3-x_1)\big)\ ,$$

and the Galois group reduces to

$$G = \{id, (2,3)\}\ . \tag{6.13}$$

Note that if we had chosen $v = x_2 - x_3$ then the Galois resolvent would have been

$$t^2+4 \;=\; \big(t-(x_2-x_3)\big)\big(t+(x_2-x_3)\big)$$

and the Galois group would still be as in 6.13. It is easy to see from this example that as a subgroup of S_3, G does depend on our labeling of the roots. Nevertheless we are going to show that as a group of permutations of the set $-1, i, -i$, G only depends on the equation $\tilde{E}_n(t)$ and the given field $\mathcal{F}$.

To this end, we need to introduce two classes of subgroups of S_n. We shall set

$$\mathcal{A} \;=\; \{\, H \subseteq S_n \,:\, \Phi \in \mathcal{F}[x_1,\ldots,x_n] \,\&\, S_\Phi \supseteq H \implies \tilde{\Phi} \in \mathcal{F} \,\} \tag{6.14}$$

and

$$\mathcal{B} \;=\; \{\, H \subseteq S_n \,:\, \Phi \in \mathcal{F}[x_1,\ldots,x_n] \,\&\, \tilde{\Phi} \in \mathcal{F} \implies S_\Phi \supseteq H \,\}\ . \tag{6.15}$$

In words, a subgroup H of S_n belongs to $\mathcal{A}$ if and only if any polynomial $\Phi \in \mathcal{F}[x_1,\ldots,x_n]$ whose value is invariant under H has its value $\tilde{\Phi}$ in $\mathcal{F}$. In the same vein we can say that a subgroup H of S_n belongs to $\mathcal{B}$ if and only if any polynomial $\Phi \in \mathcal{F}[x_1,\ldots,x_n]$ with its value $\tilde{\Phi}$ in $\mathcal{F}$ must remain invariant (by value) under all elements of H. It is immediate from the definitions 6.14 and 6.15 that for any two groups $H, K \subseteq S_n$ we have

$$H \in \mathcal{A}\ \&\ K \supseteq H \implies K \in \mathcal{A} \qquad \text{and} \qquad H \in \mathcal{B}\ \&\ K \subseteq H \implies K \in \mathcal{B} \tag{6.17}$$

In words, $\mathcal{A}$ and $\mathcal{B}$ are respectively upper and lower ideals of subgroups of S_n (under containement). Remarkably, we have the following basic fact

Theorem 6.2

$\mathcal{A}$ and $\mathcal{B}$ are both principal ideals with G as their unique extremal element. That is

$$\mathcal{A} \cap \mathcal{B} \;=\; \{G\} \tag{6.18}$$

Proof

We start by proving that

$$G \in \mathcal{A} \cap \mathcal{B}\ . \tag{6.19}$$

Given a $\Phi \in \mathcal{F}[x_1,\ldots,x_n]$ and using Proposition 6.1 we may write

$$\gamma\tilde{\Phi} \;=\; \Theta_\Phi(\gamma\tilde{v}) \qquad (\ \forall \gamma \in S_n\ ) \tag{6.20}$$

Thus if $S_\Phi \supseteq G$ we have

$$\tilde{\Phi} \;=\; \frac{1}{|G|}\sum_{\gamma\in G}\Theta_\Phi(\gamma\tilde{v})$$

Now the right hand side of this expression is a symmetric polynomial (*) in the roots of $\tilde{P}_1(t)$. This shows that $\tilde{\Phi}$ may equally be expressed as a polynomial (*) in the coefficients of $\tilde{P}_1(t)$ which themselves are in $\mathcal{F}$. This implies that $\tilde{\Phi} \in \mathcal{F}$ and that

$$G \in \mathcal{A}\ .$$

Conversely, if $\tilde{\Phi} = a \in \mathcal{F}$ then using the same polynomial Θ_Φ we may rewrite this as

$$\Theta_\Phi(\tilde{v}) - a \;=\; 0\ .$$

But this says that the polynomial

$$\Theta_\Phi(t) - a \in \mathcal{F}[t]$$

has a root in common with $\tilde{P}_1(t)$. Thus it must vanish at all the other roots of $\tilde{P}_1(t)$. That is we must have

$$\gamma\tilde{\Phi} \;=\; \Theta_\Phi(\gamma\tilde{v}) \;=\; a \;=\; \tilde{\Phi} \qquad (\ \forall\ \ \gamma \in G\ )\ .$$

This implies that $S_\Phi \supseteq G$ and that

$$G \in \mathcal{B} \tag{6.21}$$

To complete our argument we must show that G is contained in all the other elements of $\mathcal{A}$ and that G contains all the other elements of $\mathcal{B}$.

Note that if $G_1 \in \mathcal{A}$, the polynomial

$$Q(t) \;=\; \prod_{\gamma\in G_1}(t-\gamma\tilde{v})\ .$$

(*) with coefficients in $\mathcal{F}$

(whose coefficients are necessarily invariant under G_1) must belong to $\mathcal{F}[t]$. Since it has the root $\tilde{v}$ in common with $\tilde{P}_1(t)$ and $\tilde{P}_1(t)$ is irreducible $Q(t)$ must be divisible by $\tilde{P}_1(t)$. This gives that

$$G \subseteq G_1 \ .$$

Conversely, let $G_1 \in \mathcal{B}$. Consider the polynomial

$$\Phi(x) \ = \ \tilde{P}_1\big(v(x)\big) \in \mathcal{F}[x_1, \ldots, x_n] \ .$$

Since its value $\tilde{\Phi} = \tilde{P}_1(\tilde{v}) = 0$ is clearly in $\mathcal{F}$ it must remain invariant under every element of G_1. That is we must have

$$\tilde{P}_1(\gamma\tilde{v}) = 0$$

for all $\gamma \in G_1$. This shows that

$$G_1 \subseteq G$$

and completes our proof.

Remark 6.2

Since the definitions 6.14 and 6.15 of the classes $\mathcal{A}$ and $\mathcal{B}$ only involve the given field $\mathcal{F}$ and the roots of the equation $\tilde{E}_n(t)$, we see that one of the consequences of Theorem 6.2 is that also G itself only depends on $\mathcal{F}$ and $\tilde{E}_n(t)$. When in our developments we keep the given equation fixed and only vary the field, for simplicity, we shall use the notation $\mathcal{A}(\mathcal{F})$, $\mathcal{B}(\mathcal{F})$, $G(\mathcal{F})$ and leave the dependence on $\tilde{E}_n(t)$ implicit. In all other cases we will indicate this dependence with a subscript. We should also keep in mind that one of the immediate consequence of our definition of a Galois group is that if $\mathcal{F} \subseteq \mathcal{F}_1$ are two fields then we necessarily must have $G_{\tilde{E}_n(t)}(\mathcal{F}_1) \subseteq G_{\tilde{E}_n(t)}(\mathcal{F})$.

In trying to extend Lagrange's Theorems 4.1 & 4.2 to the Galois setting we should be tempted to let the Galois group G play the role of S_n in the arguments. However, given a polynomial $\Psi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$ we may not be in a position to write down the coset decomposition in 4.5 with S_n replaced by G for the simple reason that we may not have $G_\Psi \subseteq G$. Using $G_\Psi \cap G$ in place of G_Ψ doesn't get us anywhere for the simple reason that the intersection $G_\Psi \cap G$ may only consist of the identity permutation. Nor we can use S_Ψ instead of G_Ψ for it may be too big and as we have seen it may not even be a group. It develops that the optimal choice turns out to be the intersection $S_\Psi \cap G$ which here and after will be denoted by $\tilde{G}_\Psi$ and referred to as the *Galois stabilizer* of Ψ. In fact, $\tilde{G}_\Psi$ is neither too small nor to big and remarkably it can be easily shown that

Theorem 6.3

For any polynomial $\Psi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$ the Galois stabilizer

$$\tilde{G}_\Psi \ = \ S_\Psi \cap G$$

is a group

Proof

Since by construction $\tilde{G}_\Psi$ contains the identity, we need only show that

$$\eta, \xi \in \tilde{G}_\Psi \implies \eta^{-1}\xi \in \tilde{G}_\Psi \ . \tag{6.22}$$

However, this is immediate. In fact, $\eta, \xi \in \tilde{G}_\Psi$ implies that

$$\xi\tilde{\Psi} - \eta\tilde{\Psi} = 0 \ \in \mathcal{F} \ .$$

But since $G \in \mathcal{B}$ we must necessarily have

$$S_{\xi\Psi - \eta\Psi} \supseteq G$$

In other words

$$\gamma\xi\tilde{\Psi} - \gamma\eta\tilde{\Psi} \ = \ 0 \ . \qquad (\ \forall \quad \gamma \in G \ )$$

In particular this must hold true for $\gamma = \eta^{-1}$. That is

$$\eta^{-1}\xi\tilde{\Psi} \ = \ \tilde{\Psi} \ ,$$

which gives 6.22 as desired.

We are now in a position to state and prove the Galois version of Theorems 4.1 and 4.2.

Theorem 6.4

If $\Psi \& \Phi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$ and

$$\tilde{G}_\Psi \subseteq \tilde{G}_\Phi \tag{6.23}$$

then we can construct a polynomial $\Theta(t) \in \mathcal{F}[t]$ which gives

$$\gamma\tilde{\Phi} \ = \ \Theta(\gamma\tilde{\Psi}) \qquad (\ \forall \quad \gamma \in G \ ) \tag{6.24}$$

Proof

Let

$$G \ = \ \tau_1 \tilde{G}_\Psi \ + \ \tau_2 \tilde{G}_\Psi \ + \ \cdots \ + \ \tau_k \tilde{G}_\Psi \qquad (\ \tau_1 = \mathit{identity} \ ) \tag{6.25}$$

and set

$$A(t) \ = \ \sum_{i=1}^{k} \tau_i\tilde{\Phi} \prod_{j=1}^{k}{}^{(i)} (t - \tau_j\tilde{\Psi}) \quad , \quad B(t) \ = \ \prod_{i=1}^{k} (t - \tau_i \ \tilde{\Psi}) \ . \tag{6.26}$$

Now from 6.23 and 6.25 we get that $A(t)$ and $B(t)$ are G-invariant, and $G \in \mathcal{A}$ gives that

$$A(t) \ \& \ B(t) \in \mathcal{F}[t] \ .$$

Since by construction $B(t)$ has distinct roots we can find two polynomials $p(t), q(t) \in \mathcal{F}[t]$ such that

$$p(t)B'(t) + q(t)B(t) \ = \ 1 \ .$$

Setting $t = \gamma\tilde{\Psi}$ gives

$$p(\gamma\tilde{\Psi})B'(\gamma\tilde{\Psi}) \;=\; 1 \qquad 6.27$$

On the other hand we see from 6.26 that if $\gamma = \tau_i h$ with $h \in \tilde{G}_\Psi$ then

$$A(\gamma\tilde{\Psi}) \;=\; A(\tau_i\tilde{\Psi}) \;=\; \tau_i\tilde{\Phi}\, B'(\tau_i\tilde{\Psi}) \ .$$

Using 6.23 we may rewrite this as

$$A(\gamma\tilde{\Psi}) \;=\; \gamma\tilde{\Phi}\, B'(\gamma\tilde{\Psi}) \ .$$

Multiplying both sides by $p(\gamma\tilde{\Psi})$ and using 6.27 we finally get

$$\gamma\tilde{\Phi} \;=\; p(\gamma\tilde{\Psi})A(\gamma\tilde{\Psi})$$

which is 6.24 with

$$\Theta(t) \;=\; p(t)A(t) \ .$$

This Theorem has an immediate consequence which can be helpful in the construction of the Galois group of an equation.

Corollary 6.1

Let G be the Galois group of $\tilde{E}_n(t) \in \mathcal{F}[t]$ and suppose that for some $\Psi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$ we have

(1) $S_\Psi = G_\Psi$,

(2) $\tilde{\Psi} \in \mathcal{F}$.

Then we must necessarily have

$$G \;\subseteq\; G_\Psi \ .$$

Proof

In view of Theorem 6.2 we need only show that $G_\Psi \in \mathcal{A}(\mathcal{F})$. To this end let $\Phi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$ and let $S_\Phi \supseteq G_\Psi$. Condition (1) then assures that $\tilde{G}_\Phi \supseteq \tilde{G}_\Psi$. We can thus use Theorem 6.4 and derive that for some $\theta(t) \in \mathcal{F}[t]$ we have $\tilde{\Phi} = \theta(\tilde{\Psi})$. But then condition (2) yields us that $\tilde{\Phi} \in \mathcal{F}$ as desired.

Before we can proceed any further we need to establish the following basic fact

Proposition 6.3

Let $B(t) \in \mathcal{F}[t]$ be a polynomial of degree k which is irreducible in $\mathcal{F}[t]$, and let

$$B(\tilde{\Psi}) \;=\; 0$$

for some $\Psi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$. Then the values

$$1\ ,\ \tilde{\Psi}\ ,\ \tilde{\Psi}^2\ ,\ \ldots\ ,\tilde{\Psi}^{k-1} \qquad 6.28$$

form a basis of a vector space V over $\mathcal{F}$ which is also a field

Proof

Suppose that for some $c_o, c_1, c_2, \ldots, c_{k-1} \in \mathcal{F}$ not all vanishing we had

$$c_o + c_1\tilde{\Psi} + c_2\tilde{\Psi}^2 + \cdots + c_{k-1}\tilde{\Psi}^{k-1} \;=\; 0$$

then the polynomial $R(t) = c_o + c_1 t + c_2 t^2 + \cdots + c_{k-1}t^{k-1}$ would have a root in common with $B(t)$ and the greatest common divisor of $R(t)$ and $B(t)$ would yield a non trivial factorization of $B(t)$ in $\mathcal{F}[T]$ contradicting the irreducibility of $B(t)$. This shows that the elements in 6.28 are independent over $\mathcal{F}$.

To complete the proof we need to show that every non vanishing element v of V has an inverse in V. Now such an element would be given by a linear combination

$$v \;=\; c_o + c_1\tilde{\Psi} + c_2\tilde{\Psi}^2 + \cdots + c_{k-1}\tilde{\Psi}^{k-1} \qquad (\ c_i \in \mathcal{F}\ )$$

with some $c_i \neq 0$. For the same reasons as above, the polynomial $R(t) = c_o + c_1 t + c_2 t^2 + \cdots + c_{k-1}t^{k-1}$ cannot have any root in common with $B(t)$. Thus we can use the Berlekamp algorithm and construct two polynomials $p(t), q(t) \in \mathcal{F}[t]$ giving

$$p(t)R(t) + q(t)B(t) \;=\; 1 \ .$$

Setting $t = \tilde{\Psi}$ we get

$$p(\tilde{\Psi})\, R(\tilde{\Psi}) \;=\; 1 \ ,$$

which shows that $R(\tilde{\Psi})$ is invertible and that its inverse in V is given by $p(\tilde{\Psi})$.

Q.E.D.

The vector space V will here and after be denoted by $\mathcal{F}[\tilde{\Psi}]$ and referred to as the ***Extension of $\mathcal{F}$*** by $\tilde{\Psi}$. The integer k giving the dimension of $\mathcal{F}[\tilde{\Psi}]$ will be called the ***degree*** of the extension. We shall also say that $\mathcal{F}[\tilde{\Psi}]$ is obtained by adjoining $\tilde{\Psi}$ to $\mathcal{F}$. The following theorem provides the crucial tools needed in the applications of Galois theory to the theory of equations.

Theorem 6.5

Let $\Psi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$, set

$$G \;=\; \tau_1\tilde{G}_\Psi \;+\; \tau_2\tilde{G}_\Psi + \cdots + \tau_k\tilde{G}_\Psi \qquad (\ \tau_1 = id\ ) \tag{6.29}$$

and let $\tilde{\Psi}_1 = \tau_1\tilde{\Psi}, \tilde{\Psi}_2 = \tau_2\tilde{\Psi}, \ldots, \tilde{\Psi}_k = \tau_k\tilde{\Psi}$ denote the conjugates of $\tilde{\Psi}$ in G. Then

(i) The polynomial $B(t) = \prod_{i=1}^k (t - \tilde{\Psi}_i)$ is irreducible in $\mathcal{F}[t]$.

(ii) By adjoining $\tilde{\Psi}$ to $\mathcal{F}$ the Galois group of the equation $\tilde{E}_n(t) = 0$ is reduced to $\tilde{G}_\Psi$.

(iii) The Galois group of $B(t)$ is the subgroup $\Gamma_{\tilde{\Psi}}$ of S_k corresponding to the action of G on the left cosets of $\tilde{G}_\Psi$. In particular we have the isomorphism

$$\Gamma_{\tilde{\Psi}} \;\cong G \,/\, \bigcap_{i=1}^{k} \tau_i\tilde{G}_\Psi\tau_i^{-1}\ . \tag{6.30}$$

(iv) Set $\mathcal{F}_1 = \mathcal{F}[\tilde{\Psi}]$ and let $Aut_{\mathcal{F}}(\mathcal{F}_1)$ denote the group of automorphisms of $\mathcal{F}_1$ that leave $\mathcal{F}$ elementwise fixed. This given, $Aut_{\mathcal{F}}(\mathcal{F}_1)$ can be identified with the set $\{\tau_i\tilde{G}_\Psi : \tau_i\tilde{G}_\Psi = \tilde{G}_\Psi\tau_i\ \}$. In particular we have the isomorphism

$$Aut_{\mathcal{F}}(\mathcal{F}_1) \;\cong \Big(\sum_{\tau_i\tilde{G}_\Psi = \tilde{G}_\Psi\tau_i} \tau_i\tilde{G}_\Psi\Big)/\tilde{G}_\Psi \tag{6.31}$$

Proof

Proof of (i)

Suppose that a polynomial $B_1(t) \in \mathcal{F}[t]$ divides $B(t)$ and shares the root $\tau_i\tilde{\Psi}$ with $B(t)$. Then the polynomial $B_1(\tau_i\Psi) \in \mathcal{F}[x_1, x_2, \ldots, x_n]$ has the value $B_1(\tau_i\tilde{\Psi}) = 0 \in \mathcal{F}$ and $G \in \mathcal{B}$ gives that $S_{B_1(\tau_i\Psi)} \supseteq G$. In other words we must have $B_1(\gamma\tau_i\tilde{\Psi}) = 0$ for all $\gamma \in G$. Since the action of G on $\tilde{\Psi}_1, \tilde{\Psi}_2, \ldots, \tilde{\Psi}_k$ is transitive, we see that $B_1(t)$ has to vanish at all the roots of $B(t)$, which forces $B_1(t) = B(t)$. Thus $B(t)$ can't have a proper factor in $\mathcal{F}[t]$.

Proof of (ii)

Since $B(\tilde{\Psi}) = 0$ we can use Proposition 6.3 to construct the field $\mathcal{F}_1 = \mathcal{F}[\tilde{\Psi}]$. This given we need only verify that

$$\mathcal{A}(\mathcal{F}[\tilde{\Psi}]) \cap \mathcal{B}(\mathcal{F}[\tilde{\Psi}]) \;=\; \{\tilde{G}_\Psi\}\ .$$

Now this is immediate. In fact, if for some $\Phi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$ we have $S_\Phi \supseteq \tilde{G}_\Psi$ then we must have $\tilde{G}_\Phi \supseteq \tilde{G}_\Psi$ as well and Theorem 6.4 gives that for some $\Theta(t) \in \mathcal{F}[t]$

$$\tilde{\Phi} = \Theta(\tilde{\Psi})\ \in \mathcal{F}[\Psi]\ .$$

Thus $\tilde{G}_\Psi$ is in $\mathcal{A}(\mathcal{F}[\tilde{\Psi}])$.

Conversely let $\Phi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$ and $\tilde{\Phi} \in \mathcal{F}[\tilde{\Psi}]$. This means that for some $\Theta(t) \in \mathcal{F}[t]$ we have

$$\tilde{\Phi} = \Theta(\tilde{\Psi})\ .$$

In other words for the polynomial $\Xi(x) = \Phi(x) - \Theta(\Psi(x)) \in \mathcal{F}[x_1, x_2, \ldots, x_n]$ we have

$$\tilde{\Xi} = 0\ \in \mathcal{F}$$

so from $G \in \mathcal{B}(\mathcal{F})$ we get that we must have

$$\gamma\tilde{\Phi} - \Theta(\gamma\tilde{\Psi}) = \gamma\tilde{\Xi} \;=\; 0 \qquad (\ \forall\ \ \gamma \in G\ )$$

However, for $\gamma \in \tilde{G}_\Psi$ this yields that

$$\gamma\tilde{\Phi} \;=\; \Theta(\tilde{\Psi}) \;=\; \tilde{\Phi}\ .$$

In other words $\tilde{\Phi} \in \mathcal{F}[\Psi]$ implies that $S_\Phi \supseteq \tilde{G}_\Psi$, which gives $\tilde{G}_\Psi \in \mathcal{B}(\mathcal{F}[\tilde{\Psi}])$ as desired. This completes the proof of (ii).

Proof of (iii)

Let $\Gamma_{\tilde{\Psi}}$ be the image of G in the symmetric group S_k given by the permutation action of G on the left cosets of $\tilde{G}_\Psi$. We want to show that

$$\mathcal{A}_B(\mathcal{F}) \cap \mathcal{B}_B(\mathcal{F}) \;=\; \{\Gamma_{\tilde{\Psi}}\}\ . \tag{6.31}$$

To this end let $\Phi \in \mathcal{F}[y_1, y_2, \ldots, y_k]$ and suppose that for any $\gamma \in \Gamma_{\tilde{\Psi}}$ we have

$$\Phi(\tilde{\Psi}_{\gamma_1}, \tilde{\Psi}_{\gamma_2}, \cdots, \tilde{\Psi}_{\gamma_k}) \;=\; \Phi(\tilde{\Psi}_1, \tilde{\Psi}_2, \ldots, \tilde{\Psi}_k)\ . \tag{6.32}$$

Since each $g \in G$ induces a permutation of $\tilde{\Psi}_1, \tilde{\Psi}_2, \ldots, \tilde{\Psi}_k$ by an element $\gamma \in \Gamma_{\tilde{\Psi}}$, 6.32 implies that

$$S_{\Phi(\Psi_1, \Psi_2, \ldots, \Psi_k)} \;\supseteq\; G\ .$$

But then $G \in \mathcal{A}_{\tilde{E}_n}(\mathcal{F})$ gives that $\Phi(\tilde{\Psi}_1, \tilde{\Psi}_2, \ldots, \tilde{\Psi}_k) \in \mathcal{F}$. Thus $\Gamma_{\tilde{\Psi}} \in \mathcal{A}_B(\mathcal{F})$.

Conversely, say $\Phi(\tilde{\Psi}_1, \tilde{\Psi}_2, \ldots, \tilde{\Psi}_k) = a \in \mathcal{F}$. Then $G \in \mathcal{B}_{\mathcal{E}_n}(\mathcal{F})$ gives that $g\Phi(\tilde{\Psi}_1, \tilde{\Psi}_2, \ldots, \tilde{\Psi}_k) = \Phi(\tilde{\Psi}_1, \tilde{\Psi}_2, \ldots, \tilde{\Psi}_k)$ for all $g \in G$. But if $\gamma = \gamma(g)$ is the image of g in S_k, this simply says that

$$\Phi(\tilde{\Psi}_{\gamma_1}, \tilde{\Psi}_{\gamma_2}, \cdots, \tilde{\Psi}_{\gamma_1}) \;=\; \Phi(\tilde{\Psi}_1, \tilde{\Psi}_2, \ldots, \tilde{\Psi}_k)\ .$$

This gives $\Gamma_{\tilde{\Psi}} \in \mathcal{B}_B(\mathcal{F})$. To complete the proof of (iii) we need only observe that the Galois stabilizer of a conjugate $\tilde{\Psi}_i$ is simply the conjugate subgroup $\tau_i\tilde{G}_\Psi\tau_i^{-1}$, thus the only elements of g that leave invariant all the conjugates of $\tilde{\Psi}$ are those that belong to the intersection

$$\cap_{i=1}^k\ \tau_i\tilde{G}_\Psi\tau_i^{-1}\ .$$

This yields 6.30.

Proof of (iv)

We assume here that $\tilde{E}_n(t) \in \mathcal{F}[t]$ as before and that $\mathcal{E} = \mathcal{F}[\alpha_1, \alpha_2, \ldots, \alpha_n]$. Note that since every element of $\mathcal{F}_1$ is of the form

$$\theta(\tilde{\Psi}) \;=\; c_o + c_1\tilde{\Psi} + \cdots + c_{k-1}\tilde{\Psi}^{k-1} \qquad (\ c_i \in \mathcal{F}\ ) \tag{6.33}$$

with k the index of $\tilde{G}_\Psi$ in G, to find the image by a $g \in Aut_{\mathcal{F}}(\mathcal{F}_1)$ of any element of $\mathcal{F}_1$ we need only know $g\tilde{\Psi}$. This is because, every c_i remaining unchanged by g we must necessarily have $g\theta(\tilde{\Psi}) = \theta(g\tilde{\Psi})$. Moreover, since $B(t) = \prod_{i=1}^{k}(t-\tilde{\Psi}_i) \in \mathcal{F}[t]$, the identity $B(\tilde{\Psi}) = 0$ forces $B(g\tilde{\Psi}) = 0$ as well. In particular, we deduce that $g\tilde{\Psi} = \tilde{\Psi}_i = \tau_i\tilde{\Psi}$ for some $i = 1, 2, \ldots, k$. In addition $g\tilde{\Psi} \in \mathcal{F}_1$ yields that we must have $\tilde{\Psi}_i = \tau_i\tilde{\Psi} = \theta(\tilde{\Psi})$ with $\theta(t) \in \mathcal{F}$. But this gives that $\tilde{G}_\Psi \subseteq \tilde{G}_{\Psi_i}$ and since $\tilde{G}_{\Psi_i} = \tau_i\tilde{G}_\Psi\tau_i^{-1}$ the latter inclusion can hold true if and only if

$$\tilde{G}_{\Psi_i} = \tilde{G}_\Psi \ .$$

Thus we see that the elements of $Aut_{\mathcal{F}}(\mathcal{F}_1)$ can be simply identified with the left cosets $\tau_i\tilde{G}_\Psi$ such that $\tau_i\tilde{G}_\Psi = \tilde{G}_\Psi\tau_i$. This gives 6.31 and completes the proof of the Theorem.

We should note that the argument used in the proof of part (i) of this Theorem leads to the following basic property of the Galois group of an equation.

Proposition 6.4

The polynomial $B(t) = \prod_{i=1}^{n}(t-\alpha_i) \in \mathcal{F}[t]$ is irreducible in $\mathcal{F}[t]$ if and only if its Galois group $G = G_B(\mathcal{F})$ acts transitively on $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Proof

Let $B_1(t) \in \mathcal{F}[t]$ be an irreducible factor of $B(t)$. If α_i is any of the roots of $B_1(t)$ then $B_1(\alpha_i) = 0 \in \mathcal{F}$ and the fact that $G \in \mathcal{A}_B(\mathcal{F})$ gives that $B_1(\alpha_{\gamma_i}) = 0$ for all $\gamma \in G$. But then the transitivity of G yields that $B_1(t)$ can't be a proper factor of $B(t)$. Conversely suppose that G is intransitive. This means that the orbit of any of the roots, say α_1 can't consist of all the roots of $B(t)$. Denote this orbit by $Orb(\alpha_1)$ and set

$$B_1(t) = \prod_{\alpha_i \in Orb(\alpha_1)} (t-\alpha_i) \ .$$

Since $B_1(t)$, by construction is invariant under G then $G \in \mathcal{A}_B(\mathcal{F})$ gives that $B_1(t) \in \mathcal{F}[t]$. Moreover, also by construction, $B_1(t)$ does not contain all then roots of $B(t)$. Thus $B_1(t)$ is necessarily a proper factor of $B(t)$ and the latter must therefore be reducible in $\mathcal{F}[t]$.

Remark 6.3

We should note that parts (i), (ii) and (iii) of Theorem 6.5 are Galois' fundamental breakthroughs in the theory of equations. In the original Galois context groups came first and fields were only accessories. In later interpretations and additions to Galois' work started by Kronecker [], brought to completion by Dedekind [] and Weber and popularized by E. Artin [],[], this viewpoint has been reversed and Galois theory was made to become part of the theory of fields. In particular, part (iv) of Theorem 6.5 in its present interpretation is a later addition.

Since our presentation of Galois theory differs from most available textbooks on the subject, perhaps a few words might be needed to connect this writing to present day literature. For instance in Artin's monograph [] an extension field $\mathcal{F}_1$ of a field $\mathcal{F}$ is called a *normal* extension if *the group of automorphisms of $\mathcal{F}_1$ that leave $\mathcal{F}$ fixed has $\mathcal{F}$ for its fixed field*. We should note that in [] a field $\mathcal{F}$ is

said to be *fixed* by an automorphism g if g fixes *every* element of $\mathcal{F}$. We see then that one of our fields $\mathcal{F}_1 = \mathcal{F}[\Psi]$ is a normal extension of $\mathcal{F}$ if and only if the only elements of $\mathcal{F}_1$ that remain fixed under every element of $Aut_{\mathcal{F}}(\mathcal{F}_1)$ (as given by 6.31) are the elements of $\mathcal{F}$ itself.

This given, we can easily convert Theorem 6.5 into the collection of results that in [] is referred to as the *Fundamental Theorem of Galois Theory*.

Theorem 6.6

Let G be the Galois group of $\tilde{E}_n(t) = \prod_{i=1}^{n}(t-\alpha_i)$ in $\mathcal{F}$ and let $\mathcal{E} = \mathcal{F}[\alpha_1, \alpha_2, \ldots, \alpha_n]$. Then

(i) Each subgroup $G_1 \subseteq G$ is the Galois group of $E_n(t)$ with respect to an intermediate field $\mathcal{F}_1 = \mathcal{F}[\tilde{\Psi}]$. Different groups G_1, G_2 corresponding to different fields $\mathcal{F}_1, \mathcal{F}_2$.

(ii) The subgroup G_1 is a normal subgroup of G if and only if the corresponding field $\mathcal{F}_1$ is a normal extension of $\mathcal{F}$. In that case the group of automorphisms of $\mathcal{F}_1$ that leave $\mathcal{F}$ fixed is isomorphic to the quotient group G/G_1.

(iii) For each $G_1 \subseteq G$ the dimension of $\mathcal{F}_1$ over $\mathcal{F}$ is $|G|/|G_1|$ and the dimension of $\mathcal{E}$ over $\mathcal{F}_1$ is $|G_1|$.

Proof

proof of (i)

We have seen (Theorem 6.1) that given any subgroup $G_1 \subseteq G$ we can find $\Psi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$ such that $\tilde{G}_\Psi = G_1$. From Theorem 6.5 we get that the Galois group of $\tilde{E}_n$ in $\mathcal{F}_1 = \mathcal{F}[\Psi]$ is G_1. If for $G_1, G_2 \subseteq G$ we have $\tilde{G}_{\Psi_1} = G_1$, $\tilde{G}_{\Psi_2} = G_2$ then $\mathcal{F}[\tilde{\Psi}_1] \subseteq \mathcal{F}[\tilde{\Psi}_2]$ gives $\tilde{\Psi}_1\theta(\tilde{\Psi}_2)$ with $\theta(t) \in \mathcal{F}[t]$ thus also $\tilde{G}_{\Psi_2} \subseteq \tilde{G}_{\Psi_1}$ so the equality $\mathcal{F}[\Psi_1] = \mathcal{F}[\Psi_2]$ forces the equality $G_1 = \tilde{G}_{\Psi_1} = \tilde{G}_{\Psi_2} = G_2$. This proves (i).

proof of (ii)

Suppose that $G_1 = \tilde{G}_\Psi$ is not a normal subgroup of G. Then by relabeling the elements τ_i appearing in 6.29 so that 6.34 may be rewritten as

$$Aut_{\mathcal{F}}(\mathcal{F}_1) \ = \ \sum_{i=1}^{s} \tau_i\tilde{G}_\Psi \qquad (\ for \quad s < k\ )$$

Now note that the polynomial

$$B_1(t) \ = \ \prod_{i=1}^{s}(t-\tilde{\Psi}_i) \in \mathcal{F}_1[t]$$

cannot be in $\mathcal{F}[t]$. This is clear since otherwise the irreducibility of $B(t) = \prod_{i=1}^{k}(t-\tilde{\Psi}_i)$ (Theorem 6.5 (i)) would be contraddicted. In particular one of the elementary symmetric functions $e_i(y_1, y_2, \ldots, y_s)$ must take a value $e_i(\tilde{\Psi}_1, \tilde{\Psi}_2, \ldots, \tilde{\Psi}_s)$ not in $\mathcal{F}$. However, since the latter is invariant under permutations of $\tilde{\Psi}_1, \tilde{\Psi}_2, \ldots, \tilde{\Psi}_s$ we see (from 6.31) that we have a element of $\mathcal{F}_1$ that remains invariant under all elements of $Aut_{\mathcal{F}}(\mathcal{F}_1)$. So $\mathcal{F}_1$ in this case is not a normal extension of $\mathcal{F}$. Conversely suppose that $G_1 = \tilde{G}_\Psi$ is a normal subgroup of G then 6.31 gives that

$$Aut_{\mathcal{F}}(\mathcal{F}_1) \ = \ \Big(\sum_{i=1}^{k} \tau_i\tilde{G}_\Psi\Big)/\tilde{G}_\Psi \ = \ G/G_1 \ . \qquad 6.34$$

Since every element of $\mathcal{F}_1$ is already fixed by $\tilde{G}_\Psi$, we see from 6.34 that an element $f \in \mathcal{F}_1$ is fixed by $Aut_{\mathcal{F}}(\mathcal{F}_1)$ if and only if it is fixed by every element of G, but then $G \in \mathcal{A}_{\tilde{E}_n}(\mathcal{F})$ gives that f must lie in $\mathcal{F}$. This gives that $\mathcal{F}$ is the fixed field of $Aut_{\mathcal{F}}(\mathcal{F}_1)$ and completes the proof of (ii).

proof of (iii)

Given that $G_1 = \tilde{G}_\Psi$ and given that we have 6.29, Proposition 6.3 combined with (i) of Theorem 6.5 yields that the dimension of $\mathcal{F}_1$ as a vector space over $\mathcal{F}$ is precisely given by $k = |G|/|G_1|$. It develops that the last assertion of the Theorem is an immediate consequence of part (ii) of Theorem 6.5. In fact, to construct the Galois group of $\tilde{E}_n$ relative to $\mathcal{F}_1$ we can also resort to the original definition. That is we break up the polynomial $\tilde{P}(t)$ given in 6.11 into its irreducible factors in $\mathcal{F}_1[t]$ and take the collection of permutations which give the irreducible factor $P_{1,1}(t)$ that has $\tilde{v}$ as a root. (*)

However, part (ii) of Theorem 6.5 gives that this procedure must deliver the polynomial

$$P_{1,1}(t) \;=\; \prod_{\tau\in\tilde{G}_\Psi} (t - \tau\tilde{v}) \;. \qquad 6.35$$

Since $P_{1,1}(t)$ is in $\mathcal{F}_1[t]$ we can use Proposition 6.3 and deduce that the extension $\mathcal{F}_1[\tilde{v}]$ must be of dimension $|\tilde{G}_\Psi|$ as a vector space over $\mathcal{F}_1$. But then (iii) follows from the fact that $\mathcal{F}_1[\tilde{v}]$ and $\mathcal{F}[\alpha_1, \alpha_2, \ldots, \alpha_n]$ are one and the same. In fact, we trivially have $\mathcal{F}_1[\tilde{v}] \subseteq \mathcal{F}[\alpha_1, \alpha_2, \ldots, \alpha_n]$ and the reverse containement $\mathcal{F}_1[\tilde{v}] \supseteq \mathcal{F}[\alpha_1, \alpha_2, \ldots, \alpha_n]$ follows from Theorem 6.1. This completes our proof.

Remark 6.4

We should note that under the definition of some texts (see [] ,[]) one of our extensions $\mathcal{F}_1 = \mathcal{F}[\tilde{\Psi}]$ would be called *normal* if and only if any irreducible polynomial $Q(t) \in \mathcal{F}[t]$ that has a root in $\mathcal{F}_1$ has all the other roots in $\mathcal{F}_1$. Now we can easily show that, again this happens if and only if $\tilde{G}_\Psi$ is a normal subgroup of the Galois group $G = G_{\tilde{E}_n(t)}(\mathcal{F})$. Let us use the same notation as in the proof of Theorem 6.5 and let us set $Q(t) = \prod_{i=1}^{m}(t - \beta_i)$. Note first that if the "normality" condition in [] and [] is satisfied then one of the polynomials all whose roots, by this condition, would have to lie in $\mathcal{F}_1$ is the polynomial $B(t)$ itself. But as we have seen this is equivalent to $\tilde{G}_\Psi$ being a normal subgroup of G.

Conversely suppose that $\tilde{G}_\Psi$ is a normal subgroup of $G = G_{\tilde{E}_n(t)}(\mathcal{F})$. Then $\beta_1 \in \mathcal{F}_1$, simply means that $\beta_1 = \theta(\tilde{\Psi})$ with $\theta(t) \in \mathcal{F}[t]$. In other words we have

$$\beta_1 = \Phi(\alpha_1, \alpha_2, \ldots, \alpha_n) = \theta(\Psi(\alpha_1, \alpha_2, \ldots, \alpha_n)) \;.$$

But then we must also have

$$\gamma\beta_1 = \gamma\tilde{\Phi} = \theta(\gamma\tilde{\Psi}).$$

Letting Δ denote a set of representatives for the left cosets of $\tilde{G}_\Phi$ in G construct the polynomial

$$Q_1(t) \;=\; \prod_{\tau\in\Delta} (t - \tau\tilde{\Phi}) \in \mathcal{F}[t]$$

(*) Note that $P_{1,1}(t)$ must also be an irreducible factor of the polynomial $\tilde{P}_1(t)$ which gave us the Galois group G of $\tilde{E}_n$ in $\mathcal{F}$.

Since $Q_1(t)$t shares a root with $Q(t)$, the irreducibility of the latter forces all the roots of $Q(t)$ to be roots of $Q_1(t)$. In other words, every root β_i has an expression of the form

$$\beta_i = \tau_{j_i}\tilde{\Phi} = \theta(\tau_{j_i}\tilde{\Psi}) \;.$$

Now, since the normality of $\tilde{G}_\Psi$ forces all the conjugates of $\tilde{\Psi}$ in G to belong to $\mathcal{F}[\tilde{\Psi}]$ we must conclude that each $\beta_i \in \mathcal{F}[\tilde{\Psi}]$ as desired.

We should note that $Q(t)$ being a factor of $Q_1(t)$ yields that $m \le k = degree\ B(t)$.

7. Solving cyclic equations

In these notes we shall say that an n^{th} degree polynomial $\tilde{E}_n(t) = \prod_{i=1}^{n}(t - \alpha_i) \in \mathcal{F}[t]$ and the corresponding equation $\tilde{E}_n(t) = 0$ *cyclic* in $\mathcal{F}[t]$ if its Galois group $G = G_{\tilde{E}_n(t)}(\mathcal{F})$ is the cyclic group on n letters. More precisely, $\tilde{E}_n(t)$ is cyclic with respect to $\mathcal{F}[t]$ if and only if by a suitable labeling of the roots $\alpha_1, \alpha_2, \ldots, \alpha_n$ we have

$$G = G(\mathcal{F}) = \{id, \gamma, \gamma^2, \ldots, \gamma^{n-1}\} \qquad 7.1$$

with

$$\gamma\alpha_i \;=\; \alpha_{i+1} \qquad for \;\; i = 1, 2, \ldots, n \qquad (\alpha_{n+1} = \alpha_1) \qquad 7.2$$

We should note that in our definition of cyclicity we implicitely assume that $E_n(t)$ has distinct roots. This given, the following basic fact is helpful in establishing cyclicity.

Theorem 7.1

$\tilde{E}_n(t)$ is cyclic with respect to $\mathcal{F}[t]$ if and only if

(1) It is irreducible in $\mathcal{F}[t]$,

(2) We have a polynomial $\theta(t) \in \mathcal{F}[t]$ and a labeling $\alpha_1, \alpha_2, \ldots, \alpha_n$ of the roots of $\tilde{E}_n(t)$ such that

$$\alpha_{i+1} \;=\; \theta(\alpha_i) \qquad for \;\; i = 1, 2, \ldots, n \qquad (\alpha_{n+1} = \alpha_1) \qquad 7.3$$

Proof

Suppose $\tilde{E}_n(t)$ is cyclic. Then under the labeling which gives 7.1 and 7.2 set

$$Q(t) \;=\; \sum_{s=1}^{n} \alpha_{s+1} \prod_{j=1}^{(s)} (t - \alpha_j) \;. \qquad (\,\alpha_{n+1} = \alpha_1\,)$$

Note that for $i = 1, 2, \ldots, n$ we have

$$Q(\alpha_i) \;=\; \alpha_{i+1}\,\tilde{E}'_n(\alpha_i) \qquad 7.4$$

Since the roots $\alpha_1, \alpha_2, \ldots, \alpha_n$ are supposed to be distinct $\tilde{E}_n(t)$ and its derivative $\tilde{E}'_n(t)$ have no roots in common. Thus using the Berlekamp algorithm we can construct two polynomials $A(t), B(t) \in \mathcal{F}[t]$

such that $A(t)\tilde{E}_n(t) + B(t)\tilde{E}'_n(t) = 1$. Since setting $t = \alpha_i$ yields $B(\alpha_i)\tilde{E}'_n(\alpha_i) = 1$, we see that we can rewrite 7.4 in the form

$$\theta(\alpha_i) \;=\; \alpha_{i+1}$$

with

$$\theta(t) \;=\; B(t)Q(t) \in \mathcal{F}[t] \;.$$

This proves property (2). To show (1) we need only observe that, in view of the transitivity of the action in 7.2, the irreducibility of $\tilde{E}_n(t)$ follows from Proposition 6.4.

Suppose now that $\tilde{E}_n(t)$ satisfies (1) and (2). Then, using the labeling that gives 7.3, we may define γ as the circular permutation that gives $\gamma\,\alpha_i = \alpha_{i+1}$. Let us also recursively define the polynomials $\theta_i(t) \in \mathcal{F}[t]$ by setting $\theta_o = t$ and

$$\theta_{i+1}(t) \;=\; \theta(\theta_i(t)) \qquad (\; i = 1, 2, \ldots, n-1 \;) \;.$$

Let as also set for any $\Phi(x_1, x_2, \ldots, x_n) \in \mathcal{F}[x_1, x_2, \ldots, x_n]$

$$R_\Phi(t) \;=\; \Phi\big(\theta_o(t), \theta_1(t), \theta_2(t), \ldots, \theta_{n-1}(t)\big) \;. \tag{7.5}$$

Now let $G = \{id, \gamma, \gamma^2, \cdots, \gamma^{n-1}\}$ and suppose that

$$S_\Phi \supseteq G \tag{7.6}$$

Using 7.2 and 7.5 we may rewrite this as

$$R_\Phi(\alpha_1) = R_\Phi(\alpha_2) = \cdots = R_\Phi(\alpha_n) \;. \tag{7.7}$$

In particular, we must have

$$\Phi(\alpha_1, \alpha_2, \ldots, \alpha_n) = R_\Phi(\alpha_1) \;=\; \frac{1}{n}\big(R_\Phi(\alpha_1) + R_\Phi(\alpha_2) + \cdots + R_\Phi(\alpha_n)\big)$$

However since

$$\frac{1}{n}\big(R_\Phi(x_1) + R_\Phi(x_2) + \cdots + R_\Phi(x_n)\big) \in Sym[\mathcal{F}; x_1, x_2, \ldots, x_n]$$

its value at $\alpha_1, \alpha_2, \ldots, \alpha_n$ must be expressible as a polynomial in the coefficients of $\tilde{E}_n(t)$, This gives that

$$\Phi(\alpha_1, \alpha_2, \ldots, \alpha_n) = a \in \mathcal{F} \tag{7.8}$$

and establishes that $G \in \mathcal{A}_{\tilde{E}_n(t)}(\mathcal{F})$. To finish the proof we need only show that we also have

$$G \in \mathcal{B}_{\tilde{E}_n(t)}(\mathcal{F}) \;. \tag{7.9}$$

To this end suppose that 7.8 hold true for some $\Phi \in \mathcal{F}[x_1, x_2, \ldots, x_n]$. Using 7.5 we may translate this property into the statement that the polynomial

$$R_\Phi(t) - a \in \mathcal{F}[t]$$

vanishes for $t = \alpha_1$. Now under (1) $\tilde{E}_n(t)$ is irreducible and this forces it to be a factor of $R_\Phi(t) - a$. In other words we must have

$$R_\Phi(\alpha_i) = a \qquad (\; for \quad i = 1, 2, \ldots, n \;) \;.$$

This shows that $S_\Phi \supseteq G$ and that 7.9 holds true as desired.

Note that although cyclicity is a field dependent property, here and in the following we shall drop the appendage "*in* $\mathcal{F}[t]$" or "*with respect to* $\mathcal{F}[t]$" in all cases in which the identity of the base field $\mathcal{F}$ is clear from the context.

Theorem 7.1 has two immediate important applications:

Corollary 7.1

If a binomial equation $t^p - a = 0$ with $0 \neq a \in \mathcal{F}$ and p a prime is irreducible in $\mathcal{F}[t]$, then it is cyclic if and only if $\mathcal{F}$ contains a primitive p^{th}-root of unity.

Proof

Let $w \in \mathcal{F}$ be a primitive p^{th}-root of unity and let α be a root of $t^p - a$. Then since for no $s < p$ we may have $w^s = 1$ (*) the powers w^i (for $i = 1, .., p-1$) are all distinct. This gives that the roots of $t^p - a$ are simply

$$\alpha_1 = \alpha \;,\; \alpha_2 = w\alpha \;,\; \alpha_3 = w^2\alpha \;,\; \ldots \;,\; \alpha_n = w^{n-1}\alpha$$

and so we have 7.3 with $\theta(t) = wt$. Thus the cyclicity of $t^p - a$ follows from Theorem 7.1.

Conversely suppose that $t^p - a$ is cyclic. Let $\alpha_1, \alpha_2, \ldots, \alpha_p$ be the labeling of its roots that gives 7.2 so that its Galois group is $G = \{id, \gamma, \ldots, \gamma^{p-1}\}$. Set $w = \alpha_2/\alpha_1 = \alpha_2\alpha_1^{p-1}/a$. Clearly we must have $w^p = 1$ and we can't have $w = 1$ since that would contradict the irreducibility of $t^p - a$. This given, since the the elements $w^{i-1}\alpha_1$ are distinct and all satisfy $t^p - a = 0$ they must be a permutation of $\alpha_1, \alpha_2, \ldots, \alpha_p$. Thus

$$t^p - a \;=\; \prod_{i=1}^{p} (t - w^{i-1}\alpha_1) \;.$$

We thus deduce that

$$\alpha_i = w^{h_i}\alpha \qquad (\; i = 1, \ldots p \;)$$

where $h_1, h_2, \ldots, h_p$ is a permutation of $0, 1, \ldots, p-1$. From this we derive that $\gamma w = \alpha_3/\alpha_2 = w^{h_3 - 1}$. Since $\alpha_3 \neq \alpha_2$ we see that $h_3 \neq 1 \;\; mod \;\; p$. Setting $h_3 - 1 = e$ we get $\gamma^k w = w^{e^k}$. But then a theorem of Euler gives that $s^{p-1} \cong 1 \bmod p$. In particular we get that $\gamma^{p-1}w = w$. Now, $\gamma^{p-1} = \gamma^{-1}$ and γ^{-1} generates G as well as γ. Thus w is invariant under G and must necessarily belong to $\mathcal{F}$ as we wanted to show.

(*) the smallest such s would have to divide p

Corollary 7.2

The cyclotomic polynomial

$$\Phi_p(t) = 1 + t + t^2 + \cdots + t^{p-1} \qquad (for\ p\ a\ prime\ ) \tag{7.10}$$

is cyclic with respect to the field $\mathcal{Q}$ of rational numbers

Proof

We start by showing that Φ_p is irreducible in $\mathcal{Q}[t]$. Suppose not. Then by Gauss theorem we will have a non trivial factorization

$$\Phi_p(t) \;=\; A(t)B(t) \tag{7.11}$$

where $A(t)$ and $B(t)$ are both monic polynomials with integer coefficients. Setting $t = 1$ gives

$$p = \Phi_p(1) \;=\; A(1)B(1)\ .$$

So one of $A(1), B(1)$ must be ± 1. Say it is $A(1)$. Now let α be a root of Φ_p. Since $\alpha^p = 1$ we can't have $\alpha^s = 1$ for any $1 \le s \le p-1$. In particular, for such an s, the elements

$$\alpha^s\ ,\ \alpha^{2s}\ ,\ \ldots\ ,\ \alpha^{(p-1)s}$$

are just a rearrangement of the roots of Φ_p. Since at least one of them is a root of $A(t)$, we are forced to conclude that

$$A(\alpha^s)A(\alpha^{2s})\cdots A(\alpha^{(p-1)s}) \;=\; 0\ . \qquad (\ \forall\ 1 \le s \le p-1\ )$$

This yields that the polynomial

$$R(t) = A(t)A(t^2)\cdots A(t^{(p-1)})$$

must be divisible by Φ_p. Since $R(t)$ is monic with integer coefficients, again by Gauss theorem, we shall have the factorization

$$R(t) = (1 + t + t^2 + \cdots + t^{p-1})R_1(t)$$

with $R_1(t)$ also monic with integer coefficients. Now setting $t = 1$ we are forced to the impossible conclusion that

$$\pm 1 \;=\; p\,R_1(1)\ .$$

Thus Φ_p must be irreducible as asserted.

We know from number theory that for any prime p we can find a primitive exponent $e \in [1, p-1]$ which has the propery that the integers

$$e\ ,\ e^2\ ,\ \ldots e^{p-2}$$

are distinct and, in fact, give (modulo p) a permutation of

$$1\ ,\ 2\ ,\ \ldots, p-1$$

Choosing one such exponent, we derive that if α is any of the roots of Φ_p then the powers

$$\alpha\ ,\ \alpha^e\ ,\ \ldots \alpha^{e^{p-2}}$$

give back again all the roots of Φ_p. But this means that if we label the roots of Φ_p by setting $\alpha_i = \alpha^{e^{i-1}}$ we shall have 7.3 with

$$\theta(t) \;=\; t^e \in \mathcal{Q}[t]\ .$$

This completes our proof.

There is a further property of the binomial equation with prime exponent we need to know here.

Proposition 7.1

The binomial $t^p - a$ ($a \in \mathcal{F}$) is reducible in $\mathcal{F}[t]$ if and only if a is the p^{th} power of an element of $\mathcal{F}$.

proof

Suppose that $t^p - a$ is reducible in $\mathcal{F}[t]$ and let

$$t^p - a \;=\; A(t)B(t) \qquad (\ A(t), B(t) \in \mathcal{F}[t]\ )$$

be a non trivial factorization. We may then write

$$A(t) = \prod_{i=1}^{s}(t - w^{h_i}\alpha)$$

where $s < p$, α is any root of $t^p - a$, and w can be chosen to be the ratio of any two roots. However, $A(t) \in \mathcal{F}[t]$ implies in particular that its constant term c is in $\mathcal{F}$. In other words

$$\alpha^s w^{(h_1+h_2+\cdots+h_s)} = c\ \in \mathcal{F} \tag{7.12}$$

Since $1 \le s < p$ we can find integers h, k such that $hs = 1 + kp$. Raising both sides of 7.12 to the power h yields

$$\alpha\, a^k w^{h'} = c^h\ ,$$

for a suitable integer $0 \le h' \le p-1$. Thus one of the roots of $t^p - a$, namely $b = \alpha w^{h'}$ lies in $\mathcal{F}$ and a must necessarily the p^{th} power of an element of $\mathcal{F}$ as we asserted.

The converse is entirely trivial since when $a = b^p$ with $b \in \mathcal{F}$ we have the factorization

$$t^p - a = (t - b)(b^{p-1} + b^{p-2}t + \cdots + bt^{p-2} + t^{p-1})\ .$$

To proceed any further in this section we need to update the meaning of ***solving by radicals*** in the Galois setting. It is natural to assume that in this setting ***root extraction*** should simply mean extending a given field $\mathcal{P}$, by the ***adjunction*** of a root of a binomial equation

$$t^n - a = 0\ . \qquad (\ a \in \mathcal{P}\ )$$

This given, solving by radicals the equation $\tilde{E}_n(t)=0$, in the Galois setting means constructing the roots of $\tilde{E}_n(t)\in\mathcal{F}[t]$ by a sequence of extensions

$$\mathcal{F}_{k-1}\ \rightarrow\ \mathcal{F}_k\ =\ \mathcal{F}_{k-1}[\xi_k] \qquad (\ k=0,1,2,..,k_o\ ) \tag{7.13}$$

with

$$\begin{aligned}&(1)\quad \xi_k^{p_k}\in\mathcal{F}_{k-1}\\ &(2)\quad \mathcal{F}_o=\mathcal{F}\ ,\\ &(3)\quad \alpha_1,\alpha_2,\ldots,\alpha_n\in\mathcal{F}_{k_o}\ .\end{aligned} \tag{7.14}$$

In particular, this will enable us to construct a formula, say for α_1, which will be of the form

$$\alpha_1\ =\ \cdots\ +\ \sqrt[p]{\cdots\sqrt[q]{\cdots+\sqrt[r]{\cdots}+\sqrt[s]{\cdots}\cdots}} \tag{7.15}$$

Where the radicals appearing in it will be appropriately chosen solutions of the equations

$$t_k^p-a_k\ =\ 0 \qquad \text{with}\quad a_k\in\mathcal{F}_{k-1} \tag{7.16}$$

If we chose to represent the element ξ_k appearing in 7.14 (1) by the symbol $\sqrt[p_k]{a_k}$, we are confronted with the ambiguity resulting from the multivalued nature of the symbol "$\sqrt[p]{\ldots}$". For instance if, at the k^{th} step of the extension process, ξ_k is to be a primitive 6^{th}-root of unity. It would be better to represent ξ_k by $-\frac{1}{2}+\frac{\sqrt[2]{-3}}{2}$ rather by $\sqrt[6]{1}$. This is because the expression $-\frac{1}{2}+\frac{\sqrt[2]{-3}}{2}$ yields only these two primitive roots as we specialize $\sqrt[2]{-3}$ to the two conjugate roots of the equation $x^2+3=0$. Clearly, it is in the nature of the problem that any formula we may construct for α_1 should have multivalued symbols appearing in it. For α_1 itself is in essence an n-valued function. Indeed, labeling the roots of $\tilde{E}_n$ $\alpha_1,\alpha_2,\ldots,\alpha_n$ is an artificial device, and "α_1" should really represent only a ***generic*** root of $\tilde{E}_n$. Now, we have seen already in the Lagrange setting, that there is no loss in requiring that in the successive radicals $\sqrt[p_k]{a_k}$ the exponent p_k should be a prime number. Now if the equation $t^{p_k}-a_k=0$ is ***cyclic*** in $\mathcal{F}_{k-1}$ the adjonction of any one of its roots to $\mathcal{F}_{k-1}$ will result in the same field $\mathcal{F}_k$. So writing $\mathcal{F}_k=\mathcal{F}_{k-1}[\sqrt[p_k]{a_k}]$ does not produce any ambiguity as far $\mathcal{F}_k$ is concerned. Moreover, we see that if we require that each symbol $\xi_k=\sqrt[p_k]{a_k}$ appearing in 7.14 represents the same root of $t^{p_k}-a_k=0$ and if we let each of these symbols in turn and independently describe each of the other solutions of the corresponding equation $t^{p_k}-a_k=0$ then α_1 as expressed by 7.14 will represent a $p_1p_2\cdots p_{k_o}$-valued function. So that if $p_1p_2\cdots p_{k_o}=n$ then 7.14 by this process will deliver each of the roots of $\tilde{E}_n$. Here and after a formula 7.15 satisfying these requirements will be called a ***tight*** formula and the symbols $\sqrt[p_k]{a_k}$) appearing in it will likewise be called ***tight radicals***. More generally we shall refer to an adjunction $\mathcal{P}\rightarrow\mathcal{P}[\xi=\sqrt[p]{a}]$ as a ***tight radical extraction*** if and only if t^p-a is cyclic in $\mathcal{P}$. All other formulas and root extractions will be referred to as ***loose***. It is not difficult to verify that the formulas for the roots of the cubic and quartic which can be obtained by the process given in section 5 are, in fact, tight. However as we pointed out, the process in section 5 has the additional property that at each step the element adjoined (roots of unity apart), is a root of a polynomial function of the roots $x_1,x_2,\ldots,x_n$. This brings us to define 7.15 as a ***natural*** formula if every ξ_k is also a polynomial in $\alpha_1,\alpha_2,\ldots,\alpha_n$ with coefficients in $\mathcal{F}$. In this terminology, we can

say that Lagrange showed that (as long as $\mathcal{F}$ contains all the needed roots of unity) the roots of the general cubic and quartic may be given by a tight natural formula, Ruffini showed that the roots of the general quintic have no loose natural formulas and Abel showed that they may not even be solved by loose radical extractions.

It appears that we are now faced with the additional problems of finding out which equations have roots with tight and/or loose and/or natural ...-etc formulas. However, we shall soon see that things are not that complicated. To begin with we show that the roots of a cyclic equation of prime degree p, after the adjunction of a primitive p^{th}-root of unity, may be given a tight natural formula. In fact, to do so we need only add a Galois twist to Lagrange's identity 5.32.

Proposition 7.2

Let p be a prime and let $B(t)=\prod_{i=0}^{p-1}(t-\beta_i)\in\mathcal{F}[t]$ be cyclic with Galois group $G=G_B(\mathcal{F})=\{id,\gamma,\gamma^2,\ldots,\gamma^{p-1}\}$ where

$$\gamma\beta_i=\beta_{i+1}\ . \tag{7.17}$$

Let u be a primitive p^{th}-root of unity. Then we have

$$\beta_s=\frac{1}{p}\sum_{i=0}^{p-1}u^{-is}\,c_i\,\theta^i \qquad (\ for\quad s=0,..,p-1) \tag{7.18}$$

where each $c_i\in\mathcal{F}[u]$ and θ is a root of a binomial equation

$$t^p-\Xi=0 \qquad (\ \Xi\in\mathcal{F}[u]\ ) \tag{7.19}$$

which is irreducible in $\mathcal{F}[t]$. Moreover, if u is properly chosen, we have also the expansion

$$\theta=\beta_o+u\beta_1+u^2\beta_2+\cdots+u^{p-1}\beta_{p-1}\ . \tag{7.20}$$

Proof

Set

$$\theta_s=\sum_{i=0}^{p-1}u^{is}\,\beta_i \qquad (\ for\quad s=0,..,p-1) \tag{7.21}$$

Note that since the matrices $\|u^{rs}\|_{r,s=0,..p-1}$ and $\frac{1}{p}\|u^{-rs}\|_{r,s=0,..p-1}$ are inverses of each other, the relation in 7.21 may be inverted to

$$\beta_s\ =\ \frac{1}{p}\sum_{i=0}^{p-1}u^{-is}\,\theta_i \qquad (\ for\quad s=0,..,p-1) \tag{7.22}$$

Note next that we may write

$$\begin{aligned}\theta_s^p\ &=\ \sum_{i_1,i_2,\ldots,i_p}u^{i_1+i_2+\cdots+i_p}\,\beta_{i_1}\beta_{i_2}\cdots\beta_{i_p}\\ &=\ \sum_{r=0}^{p-1}u^r\sum_{i_1+i_2+\cdots+i_p\cong r\ \ mod\ p}\beta_{i_1}\beta_{i_2}\cdots\beta_{i_p}\end{aligned}\ . \tag{7.23}$$

Now it is easy to see that the polynomial

$$P_r(\beta_0,\beta_1,\dots,\beta_{p-1}) \;=\; \sum_{i_1+i_2+\cdots+i_p\cong r \mod p} \beta_{i_1}\beta_{i_2}\cdots\beta_{i_p}$$

is invariant under the cyclic shift $\beta_i \to \beta_{i+1}$ thus its value must necessarily lie in $\mathcal{F}$. Formula 7.23 then yields that

$$\theta_s^p = \Xi_s \in \mathcal{F}[u] \qquad 7.24$$

Similarly, we can write

$$\theta_1^{p-s}\theta_s \;=\; \sum_{r=0}^{p-1} u^r \sum_{i_1+i_2+\cdots+i_{p-s}+sj\cong r \mod p} \beta_{i_1}\beta_{i_2}\cdots\beta_{i_{p-s}}\beta_j$$

and deduce from it that

$$\theta_1^{p-s}\theta_s = d_s \in \mathcal{F}[u]\ . \qquad 7.25$$

Now observe that we cannot have $\theta_1=\theta_2=\cdots=\theta_{p-1}=0$ for otherwise formula 7.22 would yield $\beta_1=\beta_2=\cdots=\beta_{p-1}$ which contradicts the irreducibility of $B(t)$. But this means that we can assume $\theta_1\neq 0$. In fact, the case when θ_s is the first one that is $\neq 0$ can be reduced to the case $\theta_1\neq 0$ by replacing u with $v=u^s$. This given, we can invert Ξ_1 in $\mathcal{F}[u]$ and, combining 7.24 with 7.25, derive that

$$\theta_s = c_s\,\theta_1^s \qquad 7.26$$

with

$$c_s \;=\; d_s/\Xi_1 \in \mathcal{F}[u]\ .$$

Using 7.26 in 7.22 gives 7.18 with $\theta=\theta_1$. Since θ_1 satisfies the equation 7.19 with $\Xi=\Xi_1$ and 7.20 is simply the definition 7.21 of θ_1, to complete the proof we need only show that the binomial $t^p-\Xi_1$ is irreducible in $\mathcal{F}[u]$. Now suppose it is reducible. Then Proposition 7.1 gives that it factors in the form

$$t^p-\Xi_1 \;=\; (t-b)(t-ub)\cdots(t-u^{p-1})$$

with $b\in\mathcal{F}[u]$. Thus we must have $\theta_1=u^hb\in\mathcal{F}[u]$ and in particular 7.18 gives that $\beta_1=\theta(u)$ with $\theta(t)\in\mathcal{F}[t]$. Now let H be the Galois group of the cyclotomic polynomial $\Phi_p=1+t+\cdots+t^{p-1}$ in $\mathcal{F}[t]$ and let H_{β_1} denote the Galois stabilizer of $\theta(u)$ in H. From the left coset decomposition

$$H=\tau_1H_{\beta_1}+\tau_2H_{\beta_1}+\cdots+\tau_kH_{\beta_1} \qquad (\ \tau_1=id\ )$$

construct the polynomial

$$B_1(t) \;=\; \prod_{i=1}^{k}(t-\tau_i\beta_1)\in\mathcal{F}[t]\ .$$

Note that $B_1(t)$ must have $B(t)$ as a factor since it shares the root β_1 with $B(t)$ and $B(t)$ is irreducible in $\mathcal{F}[t]$. Now all this leads to an impossibility since $B_1(t)$ is of degree $k\le p-1$. In fact, k must be a divisor of the order of H and H is (by Corollary 7.1) a subgroup of the cyclic group of order $p-1$.

Thirty years before Galois and ninety years before the concept of field was introduced and developped by Dedekind [], Gauss, essentially showed [] that all roots of unity can be obtained by a sequence of *normal field extensions*, in fact by *tight root extractions* in our terminology. His proof his algorithmic and is therefore very much in style with present preferences for explicit constructions. This given, we would be amiss not presenting it here. Gauss' argument relies on a family of remarkably beautiful identities. They may be stated as follows

Lemma 7.1 (Gauss)

Let p be a prime, let $w=e^{2\pi i/p}$ and e be a primitive exponent modulo p. For a given factorization

$$p-1 \;=\; a\,b \qquad 7.27$$

set

$$\Pi_a(t) \;=\; t+t^{e^a}+t^{e^{2a}}+\cdots+t^{e^{(b-1)a}} \;=\; \sum_{i=0}^{b-1} t^{e^{ia}}\ . \qquad 7.28$$

Then for any pair $0\le r,s\le a-1$ we have

$$\Pi_a(w^{e^r})\Pi_a(w^{e^s}) \;=\; \sum_{i=0}^{b-1}\Pi_a(w^{e^r+e^{s+ia}})\ . \qquad 7.29$$

Proof

Note first that, for the purpose of establishing 7.28, the polynomial $\Pi_a(t)$ is essentially invariant under the substitution $t\to t^{e^a}$. More precisely, since $e^{p-1}\cong 1$ modulo p we have

$$\Pi_a(t^{e^a}) \;=\; t^{e^a}+t^{e^{2a}}+\cdots+t^{e^{ba}} \;=\; \Pi_a(t) \qquad (\ mod\ t^p-1\ ) \qquad 7.30$$

This given we have

$$\begin{aligned}\Pi_a(w^{e^r})\Pi_a(w^{e^s}) &= \sum_{j=0}^{b-1}\left(w^{e^r}\right)^{e^{ja}}\Pi_a(w^{e^s}) = \sum_{j=0}^{b-1}\left(w^{e^r}\right)^{e^{ja}}\Pi_a\big((w^{e^s})^{e^{ja}}\big)\\ &= \sum_{j=0}^{b-1}\sum_{i=0}^{b-1}\left(w^{e^r}\right)^{e^{ja}}\left((w^{e^s})^{e^{ja}}\right)^{e^{ia}} = \sum_{i=0}^{b-1}\sum_{j=0}^{b-1}\left(w^{e^r}\right)^{e^{ja}}\left((w^{e^{s+ia}}\right)^{e^{ja}}\\ &= \sum_{i=0}^{b-1}\Pi_a(w^{e^r+e^{s+ia}})\ .\end{aligned}$$

Q.E.D.

Gauss' construction of the roots of unity is best understood if illustrated in special cases. Nevertheless it will be helpful if we start with some general remarks. Suppose that we want all the primitive p^{th} roots of unity for a certain prime p. Since they are the roots of the cyclotomic polynomial

$$\Phi_p \;=\; 1+t+t^2+\cdots+t^{p-1}\ , \qquad 7.31$$

our starting point should be the cyclicity of this polynomial with respect to the rationals. Following the proof of Corollary 7.2, we let e be a primitive exponent modulo p, let α be one of the roots of

7.31 and label all the roots by setting

$$\alpha_i \;=\; \alpha^{e^i} \qquad (\; for \quad i=0,1,\dots,p-2\;) \tag{7.32}$$

We have seen that with this labeling the Galois group $G=G_{\Phi_p}(\mathcal{Q})$ is the cyclic group of order $p-1$ generated by the cyclic permutation

$$\gamma\,\alpha_i \;=\; \alpha_{i+1}\;. \tag{7.33}$$

Now for a given factorization $p-1=ab$ we let G_a denote the cyclic subgroup of G generated by the cycle γ^a. That is

$$G_a = 1+\gamma^a+\gamma^{2a}+\cdots+\gamma^{(b-1)a}\;. \tag{7.34}$$

We then have the left coset decomposition

$$G = G_a+\gamma G_a+\gamma^2 G_a+\cdots+\gamma^{a-1}G_a\;. \tag{7.35}$$

Now we know from Theorem 6.1 that we can find a polynomial $\Psi_a(x_1,x_2,\dots,x_n)\in\mathcal{Q}[x_1,x_2,\dots,x_n]$ such that $\tilde{G}_\Psi = G_a$. Actually in this case we have a very simple choice for $\tilde{\Psi}_a$, namely the linear expression

$$\tilde{\Psi}_a=\beta_o=G_a\alpha_o=\alpha_o+\alpha_a+\alpha_{2a}+\cdots+\alpha_{(b-1)a} \tag{7.36}$$

Since the stability of $\tilde{\Psi}_a$ under G_a is clear, to assure that $\tilde{G}_{\Psi_a}=G_a$ we need only verify that no other element of G leaves $\tilde{\Psi}_a$ unchanged. But, because of 7.35, this will be so if and only if the conjugates

$$\beta_i=\gamma^i\,\beta_o \;=\; \alpha_i+\alpha_{i+a}+\alpha_{i+2a}+\cdots+\alpha_{i+(b-1)a} \qquad (\; for \quad i=1,2,\dots,a-1\;) \tag{7.37}$$

are all distinct. However, since

$$\beta_o=\sum_{r=0}^{b-1}\alpha^{e^{ra}} \tag{7.38}$$

the equality

$$\beta_i \;=\; \beta_j$$

holds true if and only if α is a root of the polynomial

$$P_{ij}(t) \;=\; \sum_{r=0}^{b-1}t^{e^{i+ra}} \;-\; \sum_{r=0}^{b-1}t^{e^{j+ra}}\;.$$

Adding $\pm\Phi_p(t)$ to $P_{ij}(t)$, if necessary to cancel the the term in t^{p-1}, we would then obtain a polynomial in $\mathcal{Q}[t]$ of degree less than $p-1$ which shares a root with Φ_p contradicting the irreducibility of Φ_p. Thus $\tilde{G}_{\beta_o}=G_a$ and we can use Theorem 6.5 to conclude that

(i)

The polynomial

$$B_a(t) \;=\; \prod_{i=0}^{a-1}(t-\beta_i) \;\in\; \mathcal{Q}[t]$$

is irreducible in $\mathcal{Q}[t]$

(ii)

By adjoining β_o to $\mathcal{Q}$ the Galois group of Φ_p reduces to

$$G_a = 1+\gamma^a+\gamma^{2a}+\cdots+\gamma^{(b-1)a}\;.$$

(iii)

The Galois group of $B_a(t)$ is the cyclic group

$$G/G_a \;=\; 1+\gamma+\gamma^2+\cdots+\gamma^{a-1}\;. \tag{7.39}$$

Here γ can keep the same meaning as before since we may let it act on the β_i's as they are given by the defining identities 7.37.

This establishes that $B_a(t)$ itself is cyclic and therefore, if a is a prime, we can construct its roots according to formula 7.18 of Proposition 7.2. All this may be very nice but yet not very explicit! The beauty of Gauss identities is that they enable us to compute all that we need with a minimum of effort. Indeed, (using the notation of Lemma 7.1), we see that we have

$$\beta_i \;=\; \Pi_a(\alpha^{e^i})\;, \tag{7.39}$$

thus we may use 7.29 to construct a multiplication table for the roots of $B_a(t)$ and obtain $B_a(t)$ itself as well as the ingredients entering in formula 7.18 quite explicitely.

This is but the first step in the construction of the roots of Φ_p. It reduces us to work with Φ_p in the extended field $\mathcal{Q}[\beta_o]$ which now (by the cyclicity of $B_a(t)$) contains $\beta_1,\beta_2,\dots,\beta_{a-1}$ as well. The next step is to factorize in turn the new Galois group of Φ_p, which as we have seen reduces to

$$G_a = 1+\gamma^a+\gamma^{2a}+\cdots+\gamma^{(b-1)a}\;,$$

and then proceed to split each of the β_i into sums λ_j of powers of α which are invariant under a normal subgroup of G_a. Then, after the adjonction of the λ_j to $\mathcal{Q}[\beta_o]$, we further reduce the Galois group of Φ_p to this normal subgroup. We proceed in this manner until the β_i's are split all the way down to their individual summands, which are of course the roots of Φ_p. To describe the process explicitely at this point and in full generality would require more notation that would only blur the beauty of the argument. Imitating established tradition, we will avoid this difficulty by just saying that after we split the number $p-1$ into its prime factors.

$$p-1=p_1p_2\cdots p_{k_o} \qquad \text{with} \qquad p_1\ge p_2\ge\cdots\ge p_{k_o}\;, \tag{7.40}$$

we proceed to construct the *composition series* of $G=1+\gamma+\cdots+\gamma^{p-1}$

$$G\,\triangleright_{p_1}G_1\,\triangleright_{p_2}G_2\,\triangleright_{p_3}\cdots\triangleright_{p_{k_o}}G_{k_o}=\{id\}\;.$$

Then by a sequence of adjunctions $\xi_k \in \mathcal{Q}[\alpha_o, \alpha_1, \ldots, \alpha_{p-1}]$, with $\tilde{G}_{\xi_k} = G_k$ we arrive at a final tight and natural expression for the roots of Φ_p. We must also point out that this argument can only be completed by an induction process. Since at each step, as we have seen we need to adjoin primitive roots of unity of lower order which inductively must be assumed to have already been given tight formulas. Before we indulge into this type of mental gymnastics it will be good to work out a few revealing examples.

$\underline{p = 13}$:

Here, $p - 1 = 12$ and we may take $e = 2$ as a primitive exponent *mod* 13. This given, from Corollary 7.2 we get that the Galois group of the equation

$$\Phi_{13}(t) = 1 + t + t^2 + \cdots + t^{12}$$

with respect to the field $\mathcal{Q}$ of rational numbers, is the cyclic group generated by the operation $t \to t^2$. More precisely, if α is our desired 13^{th} root of unity and the roots of $\Phi_{13}(t)$ are labeled by setting

$$\alpha_o = \alpha\ ,\ \alpha_1 = \alpha^2\ ,\ \alpha_2 = \alpha^{2^2}\ ,\ \ldots, \alpha_{11} = \alpha^{2^{11}}\ ,$$

then $G = G_{\Phi_{13}}(\mathcal{Q})$ is generated by the cyclic permutation

$$\gamma\alpha_i \ = \ \alpha_i^2 \ = \ \alpha_{i+1}\ .$$

This gives

$$\begin{array}{ccccc}
\alpha_o & = & \alpha^{2^0} & = & w \\
\alpha_1 & = & \alpha^{2^1} & = & \alpha^2 \\
\alpha_2 & = & \alpha^{2^2} & = & \alpha^4 \\
\alpha_3 & = & \alpha^{2^3} & = & \alpha^8 \\
\alpha_4 & = & \alpha^{2^4} & = & \alpha^3 \\
\alpha_5 & = & \alpha^{2^5} & = & \alpha^6 \\
\alpha_6 & = & \alpha^{2^6} & = & \alpha^{12} \\
\alpha_7 & = & \alpha^{2^7} & = & \alpha^{11} \\
\alpha_8 & = & \alpha^{2^8} & = & \alpha^9 \\
\alpha_9 & = & \alpha^{2^9} & = & \alpha^5 \\
\alpha_{10} & = & \alpha^{2^10} & = & \alpha^{10} \\
\alpha_{11} & = & \alpha^{2^11} & = & \alpha^7
\end{array} \qquad 7.41$$

We start by factoring the Galois group of Φ_{13}

$$1 + \gamma + \cdots + \gamma^{11} \ = \ 1 + \gamma^3 + \gamma^6 + \gamma^9 + \gamma(1 + \gamma^3 + \gamma^6 + \gamma^9) + \gamma^2(1 + \gamma^3 + \gamma^6 + \gamma^9)\ . \qquad 7.42$$

Then we seek for an element $\beta_o \in \mathcal{Q}[\alpha_o\alpha_1, \ldots, \alpha_{11}]$ whose Galois stabilizer is

$$\tilde{G}_{\beta_o} \ = \ 1 + \gamma^3 + \gamma^6 + \gamma^9$$

We may take

$$\beta_o \ = \ \alpha_o + \alpha_3 + \alpha_6 + \alpha_9 = \alpha + \alpha^8 + \alpha^{12} + \alpha^5 \ = \ \Pi_3(\alpha)\ . \qquad 7.43$$

Its conjugates are

$$\beta_1 \ = \ \alpha_1 + \alpha_4 + \alpha_7 + \alpha_{10} = \alpha^2 + \alpha^3 + \alpha^{11} + \alpha^{10} \ = \ \Pi_3(\alpha_1)\ . \qquad 7.44$$

and

$$\beta_2 \ = \ \alpha_2 + \alpha_5 + \alpha_8 + \alpha_{11} = \alpha^4 + \alpha^6 + \alpha^9 + \alpha^7 \ = \ \Pi_3(\alpha_2)\ . \qquad 7.45$$

Now, using the Gauss relations with parameters $p = 13$, $a = 3$, $b = 4$ and $e = 2$ we get the multiplication table

$$\begin{array}{lll}
\beta_o\beta_o = 4 + \beta_1 + 2\beta_2 & \beta_o\beta_1 = \beta_o + 2\beta_1 + \beta_2 & \beta_o\beta_2 = 2\beta_o + \beta_1 + \beta_2 \\
\beta_1\beta_o = \beta_o + 2\beta_1 + \beta_2 & \beta_1\beta_1 = 4 + 2\beta_o + \beta_2 & \beta_1\beta_2 = \beta_o + \beta_1 + 2\beta_2 \\
\beta_2\beta_0 = 2\beta_o + \beta_1 + \beta_2 & \beta_2\beta_1 = \beta_o + \beta_1 + 2\beta_2 & \beta_2\beta_2 = 4 + \beta_o + 2\beta_1
\end{array} \qquad 7.46$$

From which we derive that

$$(t - \beta_o)(t - \beta_1)(t - \beta_2) \ = \ 1 - 4t + t^2 + t^3\ .$$

Theorem 6.5 then gives that this equation is cyclic with Galois group

$$1 + \gamma + \gamma^2$$

So if we let u denote a primitive cube root of unity, that is

$$u = \frac{-1 + \sqrt[3]{-3}}{2}$$

and set

$$\begin{aligned}
\theta_o &= \beta_o + \beta_1 + \beta_2\ , \\
\theta_1 &= \beta_o + u\beta_1 + u^2\beta_2\ , \\
\theta_2 &= \beta_o + u^2\beta_1 + u^4\beta_2\ ,
\end{aligned}$$

then formula 7.22 gives

$$\beta_o = (\theta_o + \theta_1 + \theta_2)/3\ .$$

Note that we must have $-\theta_o = 1$ since it must equal the coefficient of t^{11} in Φ_{13}. To get β_o into the form given in 7.18 we must compute the coefficient

$$c_2 \ = \ \theta_2/\theta_1^2\ .$$

Now we can write

$$c_2 \ = \ \frac{\theta_2\theta_1}{\theta_1^3}\ .$$

and the table in 7.46 gives

$$\theta_2\theta_1 \ = \ 13 \quad , \quad \theta_1^3 \ = \ -13(4 + 3u) \ = \ \frac{13^2}{-1 + 3u}$$

Thus

$$c_2 \;=\; (-1+3u)/13$$

and

$$\beta_o \;=\; \frac{-1+\theta_1+(-1+3u)\theta_1^2/13}{3}\ . \qquad 7.47$$

where we may write

$$\theta_1 \;=\; \sqrt[3]{-13(4+3u)}\ ,$$

which is assured by Proposition 7.2 to be a tight radical in $\mathcal{Q}[u]$.

By Theorem 6.5 the Galois group of Φ_{13} in $\mathcal{Q}[\beta_o]$ is

$$1+\gamma^3+\gamma^6+\gamma^9$$

So our next step is to construct a polynomial in $\alpha_o,\alpha_1,\ldots,\alpha_{11}$ whose stabilizer in $1+\gamma^6$. We may take

$$\lambda_o \;=\; \alpha_o+\alpha_6 \;=\; \alpha+\alpha^{2^6} \;=\; \alpha+\alpha^{-1}$$

Now we need to use the Gauss relations with $a=6$ $e=2$. In fact, if we set

$$\lambda_r \;=\; \Pi_6(\alpha^{2^r}) \qquad (\ r=0,1,\ldots,5\ )$$

then

$$\lambda_o \;=\; \alpha_o+\alpha_6 \quad \text{and} \quad \lambda_3 \;=\; \alpha_3+\alpha_9$$

thus

$$\beta_o \;=\; \lambda_o+\lambda_3\ .$$

From Gauss formula (or even by direct computation in this case) we get that

$$\lambda_o\times\lambda_3 \;=\; \lambda_2+\lambda_5 \;=\; \beta_2$$

and λ_o can be obtained by solving the equation

$$(t-\lambda_o)(t-\lambda_3) \;=\; t^2-\beta_o t+\beta_2\ .$$

This gives that

$$\alpha+\alpha^{-1}=\lambda_o \;=\; \frac{-\beta_o+\sqrt[2]{\beta_o^2-4\beta_2}}{2} \qquad 7.48$$

If we prefer to write λ_o only in terms of β_o the we resort again to the table in 7.46 which gives

$$\beta_1=-\beta_o^2-2\beta_o+2 \quad \text{and} \quad \beta_2=\beta_o^2+\beta_o-3$$

this agrees with the fact (implied by the cyclicity of the equation satisfied by β_o) that both β_1 and β_2 must be in $\mathcal{Q}[\beta_o]$. We should note again that, as assured by Proposition 7.2, the square root in 7.48 must necessarily be tight in $\mathcal{Q}[\beta_o]$.

The last step is to reduce the Galois group of Φ_{13} to the identity by adjoining α. From 7.48 we get that α is obtained by solving the equation

$$(t-\alpha)(t-\alpha^{-1}) \;=\; t^2-\lambda_o t+1=0$$

which gives that

$$\alpha \;=\; \frac{\lambda_o+\sqrt[2]{\lambda_o^2-4}}{2} \qquad 7.49$$

In summary, by combining 7.47, 7.48 and 7.49, we can construct a tight natural formula for all the primitive 13^{th} roots of unity. In fact the successive adjunctions may be taken to be

$$\beta_o=\alpha_o+\alpha_3+\alpha_6+\alpha_9 \quad , \quad \lambda_o \;=\; \alpha_o+\alpha_6 \quad , \quad \alpha_o=\alpha$$

The corresponding reductions of the Galois group of Φ_{13} being

$$1+\gamma+\cdots+\gamma^{11} \ \triangleright_3\ 1+\gamma^3+\gamma^6+\gamma^9 \ \triangleright_2\ 1+\gamma^6 \ \triangleright_2\ \{id\}$$

Remark 7.1

The fact that in the previous calculation we found that $\theta_1\theta_2=13\in\mathcal{Q}$ is not an accident. In general, for any cyclic equation, formulas 7.21 give that

$$f \;=\; \theta_1\theta_2\cdots\theta_{p-1} \;=\; \sum_{r=0}^{p-1}u^r \sum_{i_1+2i_2+\cdots+(p-1)i_{p-1}\cong r\ mod\ p} \beta_{i_1}\beta_{i_2}\cdots\beta_{i_{p-1}}\ .$$

Since the coefficient of u^r is clearly invariant under the cyclic shift $\beta_i\to\beta_{i+1}$, we should not be surprised if f comes out to be in $\mathcal{Q}[u]$. However, it is easy to see that f also remains unchanged by the replacement of u in f by any other primitive p^{th}-root of unity. This forces $f\in\mathcal{Q}$ as well.

<u>$p=11$</u>:

Here we may take $e=2$ again and set $\alpha_i=\alpha^{2^i}$ for $i=0,1,\ldots,9$ with α our desired primitive 11^{th}-root of unity. Letting $\gamma\alpha_i=\alpha_{i+1}$ again Corollary 7.2 gives

$$G_{\Phi_{11}}(\mathcal{Q}) \;=\; 1+\gamma+\gamma^2+\cdots+\gamma^9$$

Since $p-1=10=5\times 2$ we can at once reduce this Galois group to the subgroup

$$1+\gamma^5$$

by the adjonction of

$$\beta_o \;=\; \alpha_o+\alpha_5\ . \qquad 7.50$$

To use Gauss machinery with $p=11$ $e=2$ and $a=5$ we set

$$\beta_r \;=\; \Pi_5(\alpha_r)\ . \qquad (\ r=0,1,\ldots,4\ )$$

In this case a repetitive application of 7.29 gives

$$(t-\beta_o)(t-\beta_1)\cdots(t-\beta_4) \;=\; 1+3t-3t^2-4t^3+t^4+t^5 \;,$$

which by Theorem 6.4 must be cyclic with Galois group

$$1+\gamma+\gamma^2+\gamma^3+\gamma^4$$

We can then solve it with the formulas of Proposition 7.2. So we pick a primitive 5^{th}-root of unity u and set

$$\theta_r \;=\; \sum_{i=0}^{4} u^{ri}\,\beta_i \qquad (\ for\ r=0,1,\dots,4\ )\ .$$

Using the Gauss identities, (this time MAPLE comes in handy) we get that

$$\theta_1^5 \;=\; -11u(26+20u-15u^2+10u^3)\ .$$

So θ_1 is obtained by the tight root extraction

$$\theta_1 \;=\; \sqrt[5]{-11u(26+20u-15u^2+10u^3)} \qquad 7.51$$

and β_o is then given by

$$\beta_o \;=\; (\theta_o+\theta_1+\cdots+\theta_4)/5$$

Now again we have $\theta_o=-1$. To express β_o entirely in terms of θ_1 we need to compute the ratios $c_i=\theta_i/\theta_1^i$. We shall compute c_2 here and leave it to the reader to compute the other c_i's.

In complete agreement with our Remark 7.1 we find that

$$\theta_1\theta_2\theta_3\theta_4 \;=\; 121=11^2\in\mathcal{Q}$$

So we may write

$$\theta_2=c_2\theta_1^2 \;=\; \frac{121}{\theta_1^3\theta_3\theta_4}\,\theta_1^2$$

Again the Gauss formulas (and MAPLE) give that

$$\theta_1^3\theta_3\theta_4 \;=\; 121u(2-2u-u^2)\ ,$$

and we can then easily get that

$$\theta_2 \;=\; \frac{4+2u+2u^2+u^3}{11u}\ \theta_1^2\ .$$

Since $2^5\cong -1 \bmod 11$ we see that

$$\beta_o \;=\; \alpha+\alpha^{-1}$$

so in one more step we can get our desired α by solving the equation

$$(t-\alpha_o)(t-\alpha_5) \;=\; t^2-\beta_o t+1 \;=\; 0\ .$$

This gives again

$$\alpha \;=\; \frac{\beta_o+\sqrt[2]{\beta_o-4}}{2}$$

In accordance with the fact that $p-1=5\times 2$ here we have only needed two natural adjunctions to reduce the Galois group to the identity. Namely, $\alpha_o+\alpha_5$ and α_o.

$\underline{p=7}$:

We will be brief here. In this case we must take $e=3$ as a primitive exponent and set $\alpha_i=\alpha^{3^i}$ for $i=0,1,\dots,5$, where α is our desired 7^{th}-root of unity. The Galois group in this case is

$$G_{\Phi_7}(\mathcal{Q}) \;=\; 1+\gamma+\cdots+\gamma^5 \qquad (\ with\ \gamma\alpha_i=\alpha_{i+1}\ )$$

Since $p-1=3\times 2$ we start with

$$\beta_r \;=\; \Pi_3(\alpha_r) \qquad (\ for\ r=0,1,2\ )$$

Then Gauss' formulas yield us the table

$$\begin{array}{lll}
\beta_o\beta_o=2+\beta_2 & \beta_o\beta_1=\beta_1+\beta_2 & \beta_o\beta_2=\beta_o+\beta_1\\
\beta_1\beta_o=\beta_1+\beta_2 & \beta_1\beta_1=2+\beta_o & \beta_1\beta_2=\beta_o+\beta_2\\
\beta_2\beta_o=\beta_o+\beta_1 & \beta_2\beta_1=\beta_o+\beta_2 & \beta_2\beta_2=2+\beta_1
\end{array}$$

From which we get that

$$(t-\beta_o)(t-\beta_1)(t-\beta_2) \;=\; t^3+t^2-2t-1\ .$$

Theorem 6.5 gives that this polynomial is cyclic with Galois group $1+\gamma+\gamma^2$. So we may use Proposition 7.2 with $p=3$ and set again

$$\begin{aligned}
\theta_o &= \beta_o+\beta_1+\beta_2\ ,\\
\theta_1 &= \beta_o+u\beta_1+u^2\beta_2\ ,\\
\theta_2 &= \beta_o+u^2\beta_1+u^4\beta_2\ ,
\end{aligned}$$

This gives us

$$\beta_o=(-1+\theta_1+\theta_2)/2$$

Using the table we easily derive that

$$\theta_1\theta_2 \;=\; 7 \quad and \quad \theta_1^3 \;=\; -7-21u$$

so we may write

$$\theta_2 \;=\; \frac{7}{\theta_1} \;=\; \frac{7}{\theta_1^3}\ \theta_1^2 \;=\; \frac{2+3u}{7}\ \theta_1^2\ .$$

So we get

$$\theta_1 \;=\; \sqrt[3]{-7-21u}$$

and

$$\beta_o = \frac{-1+\theta_1+(2+3u)\theta_1^2/7}{2} \ .$$

Now here

$$\beta_o = \alpha_o + \alpha_3 \ , \ \beta_1 = \alpha_1 + \alpha_4 \ , \ \beta_2 = \alpha_2 + \alpha_5$$

and since $\alpha_3 = \alpha^{-1}$ we can find α by solving the equation

$$(t-\alpha_o)(t-\alpha_3) \ = \ t^2 - \beta_o t + 1$$

which gives

$$\alpha = \frac{\beta_o + \sqrt[2]{\beta_o^2 - 4}}{2}$$

$\underline{p=5}$:

Here we may take the exponent $e = 2$. So if α denotes our desired primitive 5^{th}-root, we need to set $\alpha_i = \alpha^{2^i}$ for $i = 0, 1, \ldots, 3$. Since $p-1 = 2\times 2$ we start by constructing the two elements

$$\beta_o = \alpha_o + \alpha_2 \quad , \quad \beta_1 = \alpha_1 + \alpha_3$$

by solving the quadratic

$$(t-\beta_o)(t-\beta_1) \ = \ t^2 + t - 1 \ .$$

Thus we may take

$$\beta_o \ = \ \frac{-1+\sqrt[2]{5}}{2} \quad , \quad \beta_1 \ = \ \frac{-1-\sqrt[2]{5}}{2} \ ,$$

and α can be obtained by solving

$$(t-\alpha_o)(t-\alpha_2) \ = \ t^2 - \beta_o t + 1 \ = \ 0 \ .$$

Since $\beta_o^2 = \beta_1 + 2$ we finally obtain

$$\alpha = \frac{\beta_o + \sqrt[2]{\beta_1 - 4}}{2} \ = \ \frac{\frac{-1+\sqrt[2]{5}}{2} + i\ \sqrt[2]{\frac{5+\sqrt[2]{5}}{2}}}{2}$$

We are now in a position to establish the basic result of this section

Theorem 7.2

Every cyclic equation can be solved by a sequence of tight and natural radical extractions

Proof

Suppose that $\tilde{E}_n(t)$ is cyclic in $\mathcal{F}$, and let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be the labeling of its roots under which

$$G_{\tilde{E}_n}(\mathcal{F}) \ = \ 1 + \gamma + \gamma^2 + \cdots + \gamma^{n-1} \ . \qquad (\ with\ \ \gamma\alpha_i = \alpha_{i+1}\ )$$

If $p_1 \geq p_2 \geq \cdots \geq p_m$ are the primes factoring n we proceed by constructing a sequence of subgroups G_k yielding the composition series

$$G_{\tilde{E}_n}(\mathcal{F}) \triangleright_{p_1} G_1 \triangleright_{p_2} G_2 \triangleright_{p_2} \cdots \triangleright_{p_m} G_m = \{id\}$$

This given we construct (via Theorem 6.1) a sequence of polynomials $\Psi_k(x_1, x_2, \ldots, x_n) \in \mathcal{F}[x_1, x_2, \ldots, x_n]$ with $\tilde{G}_{\Psi_k} = G_k$. Now Theorem 6.4 and $\tilde{G}_{\Psi_k} \triangleleft_{p_k} \tilde{G}_{\Psi_{k-1}}$ give that

$$\mathcal{F}[\tilde{\Psi}_k] \supseteq \mathcal{F}[\tilde{\Psi}_{k-1}] \ .$$

Moreover from Theorem 6.5 we also deduce that $\tilde{\Psi}_k$ is a root of a polynomial $B_{\tilde{\Psi}_k}(t) \in \mathcal{F}[\tilde{\Psi}_{k-1}][t]$ of degree p_k which is cyclic in $\mathcal{F}[\tilde{\Psi}_{k-1}]$. Thus its roots may be constructed by means of the formulas given by Proposition 7.2. Let $\sqrt[p_k]{\Xi_k}$ with $\Xi_k \in \mathcal{F}_{k-1}$ denote the radical we must extract to obtain $\tilde{\Psi}_k$. Now Proposition 7.2 assures that $\sqrt[p_k]{\Xi_k}$ is tight, and formula 7.20 gives that this a natural radical extraction as long as we are in possession of a primitive p_k^{th} root of unity. Finally, the identity in 7.18 shows that $\mathcal{F}[\tilde{\Psi}_k]$ may also be obtained from $\mathcal{F}[\tilde{\Psi}_{k-1}]$ by the adjonction of $\sqrt[p_k]{\Xi_k}$ itself. This given, since the Galois group of $\tilde{E}_n$ reduces to the identity after the extraction of the last radical $\sqrt[p_m]{\Xi_m}$, the process will yield a tight natural formula for each of the roots of $\tilde{E}_n$.

We should note that the tight extraction of prime p^{th} roots of unity can also included in this process. This is because, as we have seen, the formulas giving a primitive p_k^{th}-root depend on the solution of cyclic equations whose degrees are prime factors of $p_k - 1$. But since each $p_k \leq n$, we see that we are appropriately setup for an induction argument. We can in fact assume from the start that the Theorem is true for all cyclic equations of prime degree less or equal than a certain prime p then carry out the constrution outlined above for all cylic equations of any degree n whose prime factors are all less or equal to the next prime. The induction can of course start with $p = 2$ where the Theorem is easily verified to be true.