5-1978

Calculation of Galois groups

Daniel Schnackenberg

CALCULATION OF GALOIS GROUPS

by
Daniel Schnackenberg

Bachelor of Science, University of North Dakota, 1976

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota

May
1978

This Thesis submitted by Daniel Schnackenberg in partial fulfillment of the requirements for the Degree of Master of Science from the University of North Dakota is hereby approved by the Faculty Advisory Committee under whom the work has been done.

(Chairman)

Dean of the Graduate School

Permission

Title CALCULATION OF GALOIS GROUPS

Department MATHEMATICS

Degree MASTER OF SCIENCE

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the Library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the Chairman of the Department or the Dean of the Graduate School. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Signature Daniel Schnackenberg

Date April 25, 1978

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to Dr. Jerry M. Metzger for the guidance and encouragement he has given me in writing this thesis and to Dr. Michael B. Gregory and Dr. Edward O. Nelson for serving on my committee.

A special thanks is also extended to Mrs. Phyllis Hellem for the excellent job she did in typing this monstrous paper.

Finally, I would like to express my appreciation to my wife, Janell, for funding this project and for all of the moral support she gave me during the writing of this thesis.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

ABSTRACT

In the 19^{th} Century Galois developed a method for determining whether an equation is solvable. It relied on the close relationship between fields and their automorphism group. This paper is a survey of the techniques of Galois theory. After presenting the main results of elementary Galois theory and some useful facts about factorization, I develop the important methods of calculating the Galois group and give a proof of the Chebotarev density theorem.

NOTATION

ι	is the identity automorphism.
$\|G\|$	is the order of the group G.
$\|\sigma\|$	is the order of the element σ.
[K:F]	is the degree of the field K over the field F.
G(K:F)	is the Galois group of K over F.
G(f,F)	is the Galois group of f(x) over F.
F[x]	is the ring of polynomials with coefficients from F.
[g]	is the degree of the polynomial g(x).
Q	is the field of rational numbers.
Z	is the ring of rational integers.
C	is the field of complex numbers.
$F(\alpha)$	is the finite extension of F formed by adjoining the element α.
$\cong$	means "is isomorphic to."
Z_n	is the group of integers modulo n.
$GF(p^m)$	is the Galois field containing p^m elements.
i(G:H)	is the index of H in G.
S_n	is the symmetric group of degree n.
A_n	is the alternating group of degree n.
I_F	is the ring of integers of the field F.
$N_K(U)$	is the norm in K of the ideal U.
G_P	is the decomposition group of the prime P.
$\alpha_P=(\frac{K/F}{P})$	is the Frobenius automorphism of P.

$(\frac{K/F}{p})$ is the Artin symbol at p.

$C(\sigma)$ is the centralizer of the element σ.

$\zeta_F(s)$ is the Dedikind zeta function.

$I(F)$ is the group of ideals of F whose prime factors are unramified in the finite extension K of F.

χ is a group character.

χ_0 is the trivial character.

G^* is the group of characters of G.

$L(s,\chi;K/F)$ is an abelian L-function.

$d(A)$ is the Dirichlet density of the set A.

$f(P/p)$ is the relative degree of P over p.

$<\sigma>$ is the cyclic group generated by σ.

INTRODUCTION

The theory of Galois groups arose from the problem of trying to calculate the roots of a polynomial equation from the coefficients. If we can write the roots of an equation as a function of its coefficients using addition, subtraction, multiplication, division and extraction of roots, then we say that the equation is solvable by radicals.

Of course, equations of the first degree are always solvable by radicals. If $ax+b=0$, then $x = -\frac{b}{a}$. For quadratic equations, the solution was known several centruies B.C. and is given by the quadratic formula $x = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$ where $ax^2+bx+c = 0$.

Cardan's formulas (Uspensky, 1948, pp. 84-89) give the solution of equations of degree three and four by radicals. For $ax^3+bx^2+cx+d=0$, we let $p = \frac{c}{a} - \frac{b^2}{3a^2}$, $q = \frac{2b^3}{27a^3} - \frac{bc}{3a^2} + \frac{d}{a}$, $A = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$ and $B = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$. Then the solutions are $x = A+B$, $-\frac{A+B}{2} + \frac{A-B}{2}\sqrt{-3}$, $-\frac{A+B}{2} - \frac{A-B}{2}\sqrt{-3}$. When $ax^4 + bx^3 + cx^2 + dx + e = 0$, let $f(x) = x^3 - \frac{c}{a}x^2 + \left(\frac{bd}{a^2} - \frac{4e}{a}\right)x - \frac{b^2e}{a^3} + \frac{4ce}{a^2} - \frac{d^2}{a^2}$ and y be a root of $f(x) = 0$. Put $R = \sqrt{\frac{b^2}{4a^2} - \frac{c}{a} + y}$. If $R \neq 0$, let $D = \sqrt{\frac{3b^2}{4a^2} - R^2 - \frac{2c}{a} + \frac{bc}{a^2R} - \frac{2d}{aR} - \frac{b^3}{4a^3R}}$

and $E = \sqrt{\frac{3b^2}{4a^2} - R^2 - \frac{2c}{a} - \frac{bc}{a^2R} + \frac{2d}{aR} + \frac{b^3}{4a^3R}}$. If R = 0, let

$D = \sqrt{\frac{3b^2}{4a^2} - \frac{2c}{a} + 2\sqrt{y^2 - \frac{4e}{a}}}$ and $E = \sqrt{\frac{3b^2}{4a^2} - \frac{2c}{a} - 2\sqrt{y^2 - \frac{4e}{a}}}$. Then the roots of the quartic equation are $x = \frac{-b}{4a} + \frac{R}{2} \pm \frac{D}{2}$ and $x = \frac{-b}{4a} - \frac{R}{2} \pm \frac{E}{2}$. These formulas were discovered in the 16th Century.

Such a formula for equations of degree greater than four was sought until the 19th Century when it was shown by means of Galois theory that no such formula exists.

Galois theory associates with each polynomial equation a group G called the Galois group. G is said to be solvable provided we can form a finite chain of subgroups $G_0 \subseteq G_1 \subseteq \cdots \subseteq G_n$, with $G_0 = G$, G_n the identity group, G_{i+1} normal in G_i and G_i/G_{i+1} abelian for $i = 0,1,\ldots,n-1$. It can be shown that an equation is solvable by radicals if and only if its Galois group is solvable. Thus, if we can calculate this group, the problem is reduced to determining whether the Galois group is solvable.

This paper is a survey of elementary Galois theory and the techniques used in calculating the Galois group. Chapter 1 deals with the basic concepts of Galois theory. Chapter 2 discusses techniques for factoring polynomials over the rational numbers. Chapter 3 demonstrates some of the methods of calculating the Galois group, while in Chapter 4 I give a proof of the Chebotarev density thoerem and show how it can be used to aid in the calculation of the Galois group.

Some facts concerning the theory of groups and the theory of fields are assumed. This material can be found in any algebra text of the caliber of Herstein (1975). In the discussion of the Zassenhaus Method and the Chebotarev density theorem, I also assume some knowledge of algebraic number theory (Pollard and Diamond, 1975).

CHAPTER I

GALOIS THEORY

A. Basic Concepts

Definition: Let K be a field. A 1-1 function σ from K onto K is an automorphism provided $\sigma(a+b) = \sigma(a) + \sigma(b)$ and $\sigma(ab) = \sigma(a)\sigma(b)$ for all $a,b \in K$.

It is clear that the set of all automorphisms of K forms a group under the operation of composition of functions. We are interested in certain subgroups of this group.

Definition: Let G be a group of automorphisms on K (that is a subgroup of the set of all automorphisms on K). The fixed field of G is the set $F = \{a \in K\colon\ \sigma(a) = a \text{ for all } \sigma \in G\}$.

By the definition of automorphism, if $a,b \in F$ then $a+b$ and ab are in F. Also $0,1 \in F$ since for any automorphism σ, $\sigma(0) = 0$ and $\sigma(1) = 1$. Finally, if $a \in F$ then $a^{-1} = (\sigma(a))^{-1} = \sigma(a^{-1})$ for each $\sigma \in G$. So the fixed field is actually a subfield of K.

Definition: Let K be a field and F a subfield of K. The set of automorphisms of K leaving each element of F fixed is called the Galois group of K over F and is denoted by $G(K:F)$.

To see that $G(K:F)$ is a group, we first note that the identity automorphism is in $G(K:F)$. If $\sigma,\rho \in G(K:F)$, then $\sigma(\rho(a)) = \sigma(a) = a$

for all $a \in F$. Also if $\sigma^{-1}(a) = b$ and $a \in F$, then $\sigma(b) = a = \sigma(a)$, so that $a = b$. Hence $\sigma\rho$ and σ^{-1} are in $G(K:F)$ whenever σ and ρ are.

Lemma 1.1: Any set of distinct automorphism of a field K is linearly independent over K.

Proof: Let $\{\alpha_1,\ldots,\sigma_k\}$ be distinct automorphisms of K. Suppose that there is a set $\{\alpha_1,\ldots,\alpha_k\}$ of elements of K such that at least one of the α_i is nonzero and $\sum_{i=1}^{k} \alpha_i\sigma_i(u) = 0$ for all $u \in K$. Consider all such sets and pick the one with the fewest nonzero elements. Call this set $\{\beta_1,\ldots,\beta_k\}$ and rearrange the β_i so that $\{\beta_1,\ldots,\beta_r\}$ are the nonzero β_i. Then $\sum_{i=1}^{r} \beta_i\sigma_i(u) = 0$ for each $u \in K$. Note that $r \neq 1$, because if $r = 1$, then $\sigma_1(u) = 0$ for all $u \in K$ which cannot happen. Find $c \in K$ such that $\sigma_1(c) \neq \sigma_r(c)$. Such a c must exist since σ_1 and σ_r are distinct. Now $0 = \sum_{i=1}^{r} \beta_i\sigma_i(cu) = \sum_{i=1}^{r} \beta_i\sigma_i(c)\sigma_i(u)$ for all $u \in K$. Also $0 = \sigma_1(c)\sum_{i=1}^{r} \beta_i\sigma_i(u) = \sum_{i=1}^{r} \beta_i\sigma_1(c)\sigma_i(u)$ for all $u \in K$. By subtracting these two sums we get that $\sum_{i=1}^{r} \beta_i(\sigma_i(c)-\sigma_1(c))\sigma_i(u) = 0$ and by setting $\gamma_i = \beta_i(\sigma_i(c)-\sigma_1(c))$ we have that $\sum_{i=2}^{r} \gamma_i\sigma_i(u) = 0$. But the set $\{\gamma_2,\ldots,\gamma_r\}$ is a smaller set than $\{\beta_1,\ldots,\beta_r\}$, and $\gamma_r \neq 0$ since $\beta_r \neq 0$. This is a contradiction of the choice of the β_i and therefore the set $\{\sigma_1,\ldots,\sigma_k\}$ must be linearly independent. //

We now wish to find an upper bound for $|G(K:F)|$.

Theorem 1.2: Let F be a field and K a finite extension of F. Then $|G(K:F)| \leq [K:F]$.

Proof: Suppose $|G(K:F)| > [K:F] = n$, then there are n+1 distinct automorphisms in G(K:F). Let $\omega_1,\ldots,\omega_n$ be a basis for K over F, and $\sigma_1,\ldots,\sigma_{n+1}$ be distinct elements of G(K:F). Now consider the system of n equations in the n+1 unknowns $x_1,\ldots,x_{n+1}$:

$$\begin{array}{c} x_1\sigma_1(\omega_1) + x_2\sigma_2(\omega_1) + \cdots + x_{n+1}\sigma_{n+1}(\omega_1) = 0 \\ x_1\sigma_1(\omega_2) + x_2\sigma_2(\omega_2) + \cdots + x_{n+1}\sigma_{n+1}(\omega_2) = 0 \\ \vdots \\ x_1\sigma_1(\omega_n) + x_2\sigma_2(\omega_n) + \cdots + x_{n+1}\sigma_{x+1}(\omega_n) = 0. \end{array}$$

This system must have a nontrivial solution, say $\alpha_1,\ldots,\alpha_{n+1}$. Then $\sum_{i=1}^{n+1} \alpha_i\sigma_i(\omega_k) = 0$ for $k = 1,\ldots,n$. If $u \in K$, then $u = \sum_{i=1}^{n} \beta_i\omega_i$ with $\beta_i \in F$. Hence $\sum_{i=1}^{n+1} \alpha_i\sigma_i(u) = \sum_{i=1}^{n+1} \alpha_i\sigma_i(\sum_{j=1}^{n} \beta_j\omega_j)$ $= \sum_{i=1}^{n+1} \alpha_i[\sum_{j=1}^{n} \beta_j\sigma_i(\omega_j)] = \sum_{j=1}^{n} \beta_j[\sum_{i=1}^{n+1} \alpha_i\sigma_i(\omega_j)] = 0$. This contradicts Lemma 1.1 so that $|G(K:F)| \leq [K:F]$. //

Under certain conditions we can determine precisely the order of G(K:F).

Definition: Let K be a finite extension of the rational numbers and F a subfield of K. If for every $u \in K-F$ there exists

$\sigma \in G(K:F)$ such that $\sigma(u) \neq u$, then K is said to be a normal extension of F or normal over F. (That is, K is normal over F if F is the fixed field of G(K:F).)

Theorem 1.3: Let G be a group of automorphisms of the field K and let F be the fixed field of G where K is a finite extension of Q. Then $|G| = [K:F]$.

Proof: By the definition of fixed field, we must have that K is normal over F. Theorem 1.2 implies that $|G| \leq [K:F]$ since G must be a subgroup of G(K:F). Suppose that $|G| < [K:F]$. Let $G = \{\sigma_1,\ldots,\sigma_n\}$ and $\omega_1,\ldots,\omega_r$ be a basis for K over F where $[K:F] = r$. The system

$$\begin{matrix} x_1\sigma_1(\omega_1) + \cdots + x_{n+1}\sigma_1(\omega_{n+1}) = 0 \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ x_1\sigma_n(\omega_1) + \cdots + x_{n+1}\sigma_n(\omega_{n+1}) = 0 \end{matrix} \qquad (1)$$

of n equations in n+1 unknowns must have a nontrivial solution. From the set of all solutions pick one $\{\alpha_1,\ldots,\alpha_k,0,\ldots,0\}$ with the fewest number of nonzero elements $\alpha_1,\ldots,\alpha_k$. (We rearrange the ω_i if necessary so that the nonzero elements appear first.) Assume that σ_1 is the identity automorphism.

If $k = 1$, then $\alpha_1\sigma_m(\omega_1) = 0$ for $m = 1,\ldots,n$. This implies that $\alpha_1\sigma_1(\omega_1) = \alpha_1\omega_1 = 0$ so that $\alpha_1 = 0$. But the solution was supposed to be nontrivial, hence $k > 1$. Also we note that not all of the α_i are in F, for if $\alpha_i \in F$ for each i, then $0 = \sum_{i=1}^{k} \alpha_i\sigma_1(\omega_i) = \sum_{i=1}^{k} \alpha_i\omega_i$. This contradicts the linear independence of the ω_i over F.

Without loss of generality, we may assume that $\alpha_k = 1$ and $\alpha_1 \in K-F$. For any fixed $m = 1,2,\ldots,n$ we have,

$$\alpha_1\sigma_m(\omega_1) + \cdots + \alpha_{k-1}\sigma_m(\omega_{k-1}) + \sigma_m(\omega_k) = 0. \qquad (2)$$

Since K is normal over F, there is σ_j such that $\sigma_j(\alpha_1) \neq \alpha_1$. Pick σ_i such that $\sigma_j\sigma_i = \sigma_m$. Now

$$0 = \sigma_j(\alpha_1\sigma_i(\omega_1) + \cdots + \alpha_{k-1}\sigma_i(\omega_{k-1}) + \sigma_i(\omega_k))$$

$$= \sigma_j(\alpha_1)\sigma_j(\sigma_i(\omega_1)) + \cdots + \sigma_j(\alpha_{k-1})\sigma_j(\sigma_i(\omega_{k-1})) + \sigma_j(\sigma_i(\omega_k))$$

$$= \sigma_j(\alpha_1)\sigma_m(\omega_1) + \cdots + \sigma_j(\alpha_{k-1})\sigma_m(\omega_{k-1}) + \sigma_m(\omega_k).$$

Subtracting this from Equation (2) we get that

$$(\alpha_1-\sigma_j(\alpha_1))\sigma_m(\omega_1) + \cdots + (\alpha_{k-1}-\sigma_j(\alpha_{k-1}))\sigma_m(\omega_{k-1}) = 0.$$

This can be done for each m. If we let $\beta_i = \alpha_i-\sigma_j(\alpha_i)$ for $i = 1,\ldots,k-1$, then $\{\beta_1,\ldots,\beta_{k-1},0,\ldots,0\}$ is a nontrivial solution of the system (1) with fewer nonzero elements. This follows from the fact that we chose σ_j such that $\beta_1 = \alpha_1-\sigma_j(\alpha_1) \neq 0$. So we have a contradiction of the choice of the α_i and we must have that $|G| = [K:F]$. //

Note that this theorem implies that if G has fixed field F then $G = G(K:F)$ since $|G| \leq |G(K:F)| \leq [K:F] = |G|$.

<u>Corollary 1.4</u>: Let K and F be finite extensions of Q. K is a normal extension of F if and only if $|G(K:F)| = [K:F]$.

<u>Proof</u>: Theorem 1.3 shows that if K is a normal extension of F, then $|G(K:F)| = [K:F]$. Suppose now that $|G(K:F)| = [K:F]$ and let F_1 be the fixed field of $G(K:F)$. Since F is fixed by $G(K:F)$, we have that $F \subseteq F_1$. Now $[K:F] = [K:F_1][F_1:F]$ and $G(K:F) = G(K:F_1)$. K is normal over F_1, so by Theorem 1.3, $|G(K:F_1)| = [K:F_1]$. Thus we have $|G(K:F)| = [K:F] = [K:F_1][F_1:F] = |G(K:F_1)|[F_1:F] = |G(K:F)|[F_1:F]$, and $[F_1:F] = 1$. Therefore $F = F_1$ and K is normal over F. //

Another characterization of normal extensions is the following.

<u>Theorem 1.5</u>: Let K and F be finite extensions of Q. Then K is normal over F if and only if any polynomial with coefficients from F, which is irreducible over F and has one root in K, has all of its roots in K.

<u>Proof</u>: First let K be a normal extension of F and $f(x) \in F[x]$ be irreducible with root $\alpha \in K$. Let $G(K:F) = \{\sigma_1,\ldots,\sigma_n\}$ and $\alpha_1,\ldots,\alpha_r$ be the distinct values of $\sigma_1(\alpha),\ldots,\sigma_n(\alpha)$. Suppose that $f(x) = \sum_{i=0}^{m} a_i x^i$, then $0 = \sigma(0) = \sigma(f(\alpha)) = \sigma(\sum_{i=0}^{m} a_i\alpha^i) = \sum_{i=0}^{m} a_i\sigma(\alpha)^i$ for each $\sigma \in G(K:F)$. Thus $\sigma(\alpha)$ is a root of $f(x)$ for each $\sigma \in G(K:F)$. Let $g(x) = \prod_{i=1}^{r}(x-\alpha_i)$. Then $g(x) \in F[x]$; indeed if $u \in K$ and $\sigma \in G(K:F)$, then $\sigma(g(u)) = \sigma(\prod_{i=1}^{r}(u-\alpha_i)) = \prod_{i=1}^{r}\sigma(u-\alpha_i) = \prod_{i=1}^{r}(\sigma(u)-\sigma(\alpha_i))$. Since σ is one-to-one, the values $\sigma(\alpha_1),\ldots,\sigma(\alpha_r)$ exhausts the set

$\{\alpha_1,\ldots,\alpha_r\}$. So $\sigma(g(u)) = \prod_{i=1}^{r}(\sigma(u)-\alpha_i) = g(\sigma(u))$ and the coefficients of $g(x)$ must remain fixed by σ. That is $g(x) \in F[x]$. Now $g|f$ because every root of $g(x)$ is a root of $f(x)$. Hence $g(x) = f(x)$ since $f(x)$ is irreducible. Finally $g(x)$ has all of its roots in K and so $f(x)$ has all of its roots in K.

Next suppose that any irreducible polynomial in $F[x]$, which has a root in K, has all of its roots in K. Let $[K:F] = n$ and $K = F(\alpha)$. Suppose that $f(x)$ is the minimal polynomial of α over F and $f(x)$ has roots $\alpha_1,\ldots,\alpha_n$ where $\alpha = \alpha_1$. $f(x)$ must have all of its roots in K by our hypothesis, so that $\alpha_i \in K$ for each i. Any element $\sigma \in G(K:F)$ must have the property that $\sigma(\alpha) = \alpha_i$ for some $i = 1,\ldots,n$. Also $\{1,\alpha,\ldots,\alpha^{n-1}\}$ is a basis for K over F so that the way σ acts on α actually determines its value for all elements of K. Define $\sigma_i(\alpha) = \alpha_i$ for $i = 1,\ldots,n$ and extend σ_i to all of K in a natural way using $\{1,\alpha,\ldots,\alpha^{n-1}\}$ as a basis for K. Each $\sigma_i \in G(K:F)$ so that $n \leq |G(K:F)| \leq [K:F] = n$. Therefore K is normal over F. //

The next theorem illustrates the important relationship between the fields which lie "between" K and F and the normal subgroups of G(K:F). It is called the Fundamental Theorem of Galois Theory.

<u>Theorem 1.6</u>: Let K be a normal extension of F. If $F \subseteq L \subseteq K$ then K is a normal extension of L and $G(K:L) \subseteq G(K:F)$. Furthermore, L is a normal extension of F if and only if $G(K:L)$ is a normal subgroup of $G(K:F)$. In that case $G(L:F) \cong G(K:F)/G(K:L)$.

Proof: First we show that K is a normal extension of L whenever $F \subseteq L \subseteq K$ and K is normal over F. Let $g(x) \in L[x]$ be irreducible over L with root $\alpha \in K$, and $f(x) \in F[x]$ be α's minimal polynomial over F. In L we have $g|f$ and, by Theorem 1.5, f(x) must have all of its roots in K. Hence g(x) must have all of its roots in K and, again by Theorem 1.5, K is normal over L. Clearly $G(K:L) \subseteq G(K:F)$.

Now assume that L is normal over F. L is a finite extension of F, so $L = F(\alpha)$ for some $\alpha \in L$. Let $g(x) \in F[x]$ be α's minimal polynomial over F and suppose $[g] = m$. Now for each $\sigma \in G(K:F)$, $\sigma(\alpha)$ is a root of g(x). Since $\{1,\alpha,\ldots,\alpha^{m-1}\}$ is a basis for L and the roots of g(x) are in L, we have σ mapping L onto L. So $\sigma|L$ is an automorphism of L for each $\sigma \in G(K:F)$. Define a group homomorphism h from G(K:F) to G(L:F) by $h(\sigma) = \sigma|L$. Clearly the kernel of h is G(K:L) since $h(\sigma) = \iota_L$ (the identity in G(L:F)) implies that σ leaves L fixed. Therefore G(K:L) is a normal subgroup of G(K:F). Also $|G(L:F)| = [L:F] = [K:F]/[K:L] = |G(K:F)|/|G(K:L)|$ so that h is onto. Hence $G(L:F) \cong G(K:F)/G(K:L)$.

Finally suppose that G(K:L) is a normal subgroup of G(K:F). Now $[K:F] = [K:L][L:F]$ so $[L:F] = [K:F]/[K:L]$. Since K is normal over both L and F, we can apply Corollary 1.4 to get that $[L:F] = [K:F]/[K:L] = |G(K:F)|/|G(K:L)| = |G(L:F)|$. Corollary 1.4 implies that L is normal over F. //

B. The Galois Group of a Polynomial

Definition: Let $f(x) \in F[x]$, where F is a finite extension of the rational numbers. By the fundamental theorem of algebra, we can write $f(x) = \prod_{i=1}^{n} (x-\alpha_i)$ where the α_i are complex numbers. The splitting field of $f(x)$ over F is the field $K = F(\alpha_1,\ldots,\alpha_n)$. If $f(x)$ has all of its roots in some field L, we say $f(x)$ splits in L.

Definition: Let $f(x) \varepsilon F[x]$, where F is a finite extension of the rational numbers. If K is the splitting field of $f(x)$ over F, then the Galois group of $f(x)$ over F is the group $G(K:F)$ and is denoted by $G(f,F)$.

Theorem 1.7: Let $f(x) \varepsilon F[x]$ have distinct roots $\alpha_1,\ldots,\alpha_n$. Then $G(f,F)$ can be embedded in S_n where S_n is the symmetric group of degree n. Therefore $|G(f,F)| \leq n!$.

Proof: We may assume that $f(x)$ has no repeated roots since they can be divided out without changing the Galois group. If

$$f(x) = \prod_{i=1}^{n} (x-\alpha_i)^{e_i} \text{ where } e_i \geq 1, \text{ then } \hat{f}(x) = \prod_{i=1}^{n} (x-\alpha_i) \varepsilon F[x].$$

This is true because the coefficients of $\hat{f}(x)$ are symmetric functions of the roots of $f(x)$ and hence in F.

So $f(x) = \prod_{i=1}^{n} (x-\alpha_i)$. Write $f(x) = \sum_{i=0}^{n} a_i x^i$ and let $\sigma \varepsilon G(f,F)$; then $o = \sigma(0) = \sigma(f(\alpha_j)) = \sigma(\sum_{i=0}^{n} a_i\alpha_j^i) = \sum_{i=0}^{n} a_i\sigma(\alpha_j)^i$. Thus $\sigma(\alpha_j)$ is a root of $f(x)$, say $\sigma(\alpha_j) = \alpha_k$. Then $\sigma|\{\sigma_1,\ldots,\sigma_n\} \varepsilon S_n$ and $G(f,F)$ can be embedded in S_n. Clearly if $\sigma|(\alpha_1,\ldots,\alpha_n) = \tau|(\alpha_1,\ldots,\alpha_n)$, then $\sigma = \tau$ because automorphisms of $G(f,F)$ are determined by how they act on the roots of $f(x)$. So the embedding is 1-1. //

This theorem also tells use that $\sigma \in G(f,F)$ is completely determined by how it behaves on the roots of f. If u is in the splitting field of f(x) over F, then $u = h(\alpha_1,\ldots,\alpha_n)$ where $\alpha_1,\ldots,\alpha_n$ are the roots of f(x) and h is a rational function in n variables with coefficients in F. So $\sigma(u) = h(\sigma(\alpha_1),\ldots,\sigma(\alpha_n))$. Thus we have three ways of describing G(f,F): (1) the automorphism group of the splitting field of f(x) fixing F; (2) a permutation group of the roots of f(x); and (3) a subgroup of the symmetric group of degree n.

Theorem 1.8: Let $f(x) \in F[x]$ be irreducible over F with splitting field K. If $f(x) = \prod_{i=1}^{n}(x-\alpha_i)$ in K[x], then there is an automorphism $\sigma \in G(K:F)$ such that $\sigma(\alpha_1) = \alpha_n$.

Proof: First it is clear that $F(\alpha_1) \cong F(\alpha_n)$ by the isomorphism ψ which holds F fixed and has $\psi(\alpha_1) = \alpha_n$. This is true since $\{1,\alpha_1,\ldots,\alpha_1^{n-1}\}$ is a basis for $F(\alpha_1)$ and $\{1,\alpha_n,\ldots,\alpha_n^{n-1}\}$ is a basis for $F(\alpha_n)$.

We will construct an extension of ψ inductively. Suppose we have extended ψ to ψ_m, an isomorphism of $F(\alpha_1,\ldots,\alpha_m)$ onto $F(\beta_1,\ldots,\beta_m)$ with the following properties:

(1) $\beta_1 = \alpha_n$,

(2) $\{\beta_1,\ldots,\beta_m\} \subseteq \{\alpha_1,\ldots,\alpha_n\}$,

(3) $\psi_m(\alpha_i) = \beta_i$ for $i = 1,\ldots,m$.

We assume that, for $i > m$, $\alpha_i \notin F(\alpha_1,\ldots,\alpha_m)$. For if $\alpha_i \in F(\alpha_1,\ldots,\alpha_m)$, then ψ_m would actually be an isomorphism of $F(\alpha_1,\ldots,\alpha_m,\alpha_i)$ onto $F(\beta_1,\ldots,\beta_m,\psi_m(\alpha_i))$. Factor f(x) over

$F(\alpha_1,\ldots,\alpha_m)[x]$ as $f(x) = g_1(x)\cdots g_k(x)\prod_{i=1}^{m}(x-\alpha_i)$ where the $g_i(x)$ are irreducible and of degree greater than one. Let α_{m+1} be a root of $g_1(x)$, and $h_1(x)$ be the image of $g_1(x)$ under ψ_m. $h_1(x)$ must be irreducible in $F(\beta_1,\ldots,\beta_m)[x]$, since if $h_1(x) = h(x)g(x)$, then the inverse images of $h(x)$ and $g(x)$ would be in $F(\alpha_1,\ldots,\alpha_m)[x]$ and would divide $g_1(x)$. Let β_{m+1} be a root of $h_1(x)$ and define ψ_{m+1} from $F(\alpha_1,\ldots,\alpha_{m+1})$ onto $F(\beta_1,\ldots,\beta_{m+1})$ by $\psi_{m+1}(\alpha_{m+1}) = \beta_{m+1}$ and $\psi_{m+1}(u) = \psi_m(u)$ if $u \in F(\alpha_1,\ldots,\alpha_n)$. ψ_{m+1} is an isomorphism because $\{1,\alpha_{m+1},\ldots,\alpha_{m+1}^{[g_1]-1}\}$ is a basis for $F(\alpha_1,\ldots,\alpha_{m+1})$ over $F(\alpha_1,\ldots,\alpha_m)$, $\{1,\beta_{m+1},\ldots,\beta_{m+1}^{[h_1]-1}\}$ is a basis for $F(\beta_1,\ldots,\beta_{m+1})$ over $F(\beta_1,\ldots,\beta_m)$ and $[g_1] = [h_1]$. ψ_{m+1} also satisfies our 3 conditions. We use this process at most n times to arrive at an automorphism σ from $K = F(\alpha_1,\ldots,\alpha_n)$ onto $F(\beta_1,\ldots,\beta_n) = K$. Finally we have $\sigma(\alpha_1) = \alpha_n$ as required. //

Definition: A subgroup G of S_n is said to be transitive provided for each $i,j \in \{1,\ldots,n\}$ there is $\sigma \in G$ such that $\sigma(i) = j$.

Corollary 1.9: The Galois group of an irreducible polynomial is transitive.

Corollary 1.10: Let $f(x) \in F[x]$ and let $p(x)$ be an irreducible factor of $f(x)$ in $F[x]$. If α_1,α_2 are roots of $p(x)$, then there is $\sigma \in G(f,F)$ such that $\sigma(\alpha_1) = \alpha_2$.

Proof: Let K_1 be the splitting field of $p(x)$ over F. By Theorem 1.8, there is $\psi \in G(K_1:F)$ such that $\psi(\alpha_1) = \alpha_2$. Let $q(x)$ be an irreducible factor of $f(x)$ over K_1. By the method of the proof of Theorem 1.8, we can extend ψ to ϕ, an automorphism of q's splitting field. Proceeding by induction, we can extend ψ to an automorphism $\sigma \in G(K:F)$ such that $\sigma(\alpha_1) = \alpha_2$, where K is the splitting field of $f(x)$. //

Theorem 1.11: Let K and F be finite extensions of Q. K is a normal extension of F if and only if K is the splitting field of some polynomial over F.

Proof: First suppose that K is a normal extension of F, then $K = F(\alpha)$ for some $\alpha \in K$. Let $p(x)$ be α's minimal polynomial over F. Then by Theorem 1.5 $p(x)$ must split in K. Hence $F(\alpha) \subseteq F(\alpha_1,\ldots,\alpha_n) \subseteq K = F(\alpha)$, where $\alpha_1,\ldots,\alpha_n$ are the roots of $p(x)$, and K is p's splitting field.

Next assume that K is the splitting field of $f(x)$ over F. We proceed by induction on $[K:F]$. If $[K:F] = 1$, then $K = F$ and so K is normal over F. Now suppose that $[K:F] = n > 1$, and whenever K_1, F_1 are fields such that $[K_1:F_1] < n$ and K_1 is the splitting field of some polynomial over F_1, then K_1 is normal over F_1.

Since $[K:F] > 1$, $f(x)$ must have an irreducible factor $p(x)$ with degree greater than 1. Let $p(x) = \prod_{i=1}^{m} (x-\alpha_i)$. Now $[K:F(\alpha_1)] < n$ and $f(x) \in F(\alpha_1)[x]$ has splitting field K over $F(\alpha_1)$. Therefore K is normal over $F(\alpha_1)$ be our induction hypothesis.

Let $u \in K$ be such that $\sigma(u) = u$ for all $\sigma \in G(K:F)$. We will show that $u \in F$. Since $G(K:F(\alpha_1)) \subseteq G(K:F)$, u is left fixed by each automorphism of K fixing $F(\alpha_1)$. By the normality of K over $F(\alpha_1)$, $u \in F(\alpha_1)$. $\{1,\alpha_1,\ldots,\alpha_1^{m-1}\}$ is a basis for $F(\alpha_1)$ over F so that $u = \sum_{i=0}^{m-1} a_i\alpha_1^i$ with $a_i \in F$. By Corollary 1.10, there is $\sigma_j \in G(K:F)$ such that $\sigma_j(\alpha_1) = \alpha_j$ for $j = 1,\ldots,m$. We have that $u = \sigma_j(u) =$ $\sigma_j(\sum_{i=0}^{m-1} a_i\alpha_1^i) = \sum_{i=0}^{m-1} a_i\sigma_j(\alpha_1)^i = \sum_{i=0}^{m-1} a_i\alpha_j^i$ and so $(\sum_{i=0}^{m-1} a_i\alpha_j^i) - u = 0$ for $j = 1,\ldots,m$. Let $g(x) = (a_0-u) + \sum_{i=1}^{m-1} a_ix^i$, then g(x) has m roots, namely $\alpha_1,\ldots,\alpha_m$. This can happen only if g(x) is identically zero. In particular $a_0-u = 0$, so $u = a_0 \in F$. Therefore, if $u \in K$ and $\sigma(u) = u$ for each $\sigma \in F(K:F)$, $u \in F$. Thus K is normal over F. //

This theorem is very important in the calculation of the Galois group of a polynomial, because it tells us that any element of K-F must be moved by some $\sigma \in G(f,F)$, where K is the splitting field of f(x) over F. So if we can find $u \in K-F$ such that u is moved by no element of the automorphism group G on K, then G must not be all of G(f,F).

A couple of special polynomials have Galois groups which are relatively easy to calculate.

Theorem 1.12: If F contains a primitive n^{th} root of unity and $f(x) = x^n-a$, where a is a nonzero element of F, then G(f,F) is abelian.

Proof: Let α be a root of f(x) and ξ a primitive n^{th} root of unity. Then $\alpha, \alpha\xi, \ldots, \alpha\xi^{n-1}$ are the distinct roots of $x^n - a$. If $\sigma \in G(f,F)$, then σ is a permutation of the roots of f, so σ is determined by how it acts on α. Suppose $\sigma, \rho \in G(f,F)$ with $\sigma(\alpha) = \alpha\xi^k$ and $\rho(\alpha) = \alpha\xi^m$. Then $\sigma(\rho(\alpha)) = \sigma(\alpha\xi^m) = \sigma(\alpha)\sigma(\xi^m) = \alpha\xi^k\xi^m = \alpha\xi^{k+m}$ and $\rho(\sigma(\alpha)) = \rho(\alpha\xi^k) = \rho(\alpha)\rho(\xi^k) = \alpha\xi^m\xi^k = \alpha\xi^{k+m}$. Hence $\rho\sigma = \sigma\rho$ and G(f,F) is abelian. //

Theorem 1.13: Let F be a subfield of the real nubmers and $f(x) \in F[x]$ be irreducible over F with prime degree p. If f(x) has exactly 2 nonreal roots, then $G(f,F) = S_p$.

Proof: Our goal is to show that every transposition is in G(f,F) and then, since every element of S_p is a product of transpositions, we will have the conclusion of the theorem. Let $f(x) = \prod_{i=1}^{p} (x-\alpha_i)$ and α_1, α_2 be nonreal. Complex conjugation is always an automorphism and can be represented as the transposition (1 2). This is because α_1 is the complex conjugate of α_2 and the rest of the α_i are real. Consider all of the transpositions in G(f,F) involving 1 and arrange the roots of f(x) so that these transpositions are (1 2),(1 3),...,(1 m) for some $m \geq 2$. If $j > m$ and $(j\ i) \in G(f,F)$, then $i > m$. For if $i \leq m$, then $[(j\ i)(i\ 1)]^{-1} = (j\ 1)^{-1} = (1\ j) \in G(f,F)$ which cannot happen. Also G(f,F) contains all transpositions of the form $(i_1\ i_2)$ with $1 \leq i_1, i_2 \leq m$, for $(i_1\ i_2) = (1\ i_1)(1\ i_2)(1\ i_1)$.

Now $m \leq p$ and, if $m < p$, then there is j with $m < j \leq p$. By Corollary 1.9, G(f,F) is transitive and so there is $\sigma \in G(f,F)$ such that $\sigma(1) = j$. Let

$$\sigma = \begin{pmatrix} 1, 2, \cdots, m, \cdots, p \\ \\ j, j_2, \cdots, j_m, \cdots, j_p \end{pmatrix}.$$

Then, for $k = 2, \ldots, m$, $\sigma(1\ k)\sigma^{-1} = (j\ j_k) \in G(f,F)$. By our remarks above, $j_k > m$. We now have $2m$ distinct numbers $1, 2, \ldots, m, j, j_2, \ldots, j_m$ and each is less than or equal to p. So $2m \leq p$ and if $2m < p$ we repeat this process to arrive at $3m \leq p$. We stop when we have exhausted all p numbers. At each step we use exactly m numbers so that $m|p$. Since $m > 1$, $m = p$ and $G(f,F) = S_p$. //

Definition: Let K_1, K_2 be finite extensions of the rational numbers. The compositum of K_1 and K_2 is the smallest field containing both K_1 and K_2. It is denoted by K_1K_2.

Lemma 1.14: If K_1 and K_2 are normal extensions of F, then K_1K_2 is normal over K_1 (and hence over F).

Proof: By Theorem 1.11, we know that K_i is the splitting field of some polynomial $p_i(x) \in F[x]$ for $i = 1,2$. Let K be the splitting field of $p_1(x)p_2(x)$. Then $K = K_1K_2$ because the elements of K are rational functions of the roots of $p_1(x)$ and $p_2(x)$ with coefficients in F as are the elements of K_1K_2. Hence by Theorem 1.11, K_1K_2 is normal over K_1 since it is the splitting field of $p_2(x) \in K_1[x]$. //

Theorem 1.15: If K and L are normal over F, then K is normal over $K \cap L$ and the mapping h from $G(KL:K)$ to $G(L:K \cap L)$ is an isomorphism, where $h(\sigma) = \sigma|L$.

Proof: The mapping h is clearly a homomorphism. Suppose that $h(\sigma)$ is the identity automorphism of $G(L:K \cap L)$ where $\sigma \in G(KL:K)$. Then $\sigma(u) = u$ for each $u \in K$ and $\sigma(v) = v$ for each $v \in L$. Hence $\sigma(\omega) = \omega$ for each $\omega \in KL$ for the members of KL are just rational functions of the elements of K and L. So h is 1-1 and an isomorphism onto its range. Finally we show that $h(G(KL:K)) = G(L:K \cap L)$. Since $h(G(KL:K)) \subseteq G(L:K \cap L)$, it is sufficient to show that $K \cap L$ is the fixed field of the image of G(KL:K) under h. Let $u \in L$ with $(h(\sigma))(u) = u$ for every $\sigma \in G(KL:K)$. Then $u \in K$, for if not then there exists $\sigma \in G(KL:L)$ such that $\sigma(u) \neq u$. But KL is normal over L by Lemma 1.14, so $(h(\sigma))(u) \neq u$ which is a contradiction. Thus $u \in K \cap L$ and by Theorems 1.2 and 1.3, $|h(G(KL:K))| = [L:K \cap L] \geq |G(L:K \cap L)| \geq |h(G(KL:K))|$. So equality must hold. This also shows that L is normal over $K \cap L$ using Corollary 1.4. //

Theorem 1.16: Let K_1 and K_2 be normal over F and define the mapping h from $G(K_1K_2:F)$ to $G(K_1:F) \times G(K_2:F)$ by $h(\sigma) = (\sigma|K_1, \sigma|K_2)$. Then h is a 1-1 homomorphism, and if $K_1 \cap K_2 = F$, then h is an isomorphism.

Proof: It is clear that h is a homomorphism; and if $h(\sigma) = (\iota_1, \iota_2)$, where ι_i is the identity in $G(K_i:F)$, then σ fixes both K_1 and K_2. Hence σ fixes K_1K_2 and must be the identity of $G(K_1K_2:F)$. This implies that h is one-to-one.

Now assume that $K_1 \cap K_2 = F$ and let $(\sigma_1, \sigma_2) \in G(K_1:F) \times G(K_2:F)$. We apply Theorem 1.15, with $K = K_1$ and $L = K_2$, to get a $\sigma \in G(K_1K_2:K_1)$

such that $\sigma|K_2 = \sigma_2$. Again applying Theorem 1.15, only with $K = K_2$ and $L = K_2$, we find $\rho \in G(K_1K_2:K_2)$ such that $\rho|K_1 = \sigma_1$. Then $\rho\sigma|K_1 = \sigma_1$ and $\rho\sigma|K_2 = \sigma_2$. Hence $h(\rho\sigma) = (\sigma_1,\sigma_2)$ which implies that h is onto and an isomorphism. //

An easy induction argument provides the following.

Corollary 1.17: Let $K_1,\ldots,K_n$ be normal extensions of F with Galois groups $G_1,\ldots,G_n$ respectively. Then $G(K_1\cdots K_n:F)$ is isomorphic to a subgroup of $G_1 X \cdots X G_n$. If $K_{i+1} \cap (K_1\cdots K_i) = F$ for $i = 1,\ldots,n-1$, then $G(K_1\cdots K_n:F) \cong G_1 X \cdots X G_n$.

An immediate consequence of Corollary 1.17 is Corollary 1.18, which greatly simplifies the task of calculating the Galois group of an equation.

Corollary 1.18: Let $f(x) = \prod_{i=1}^{n} p_i(x)$ where $f(x),p_1(x),\ldots,p_n(x) \in F[x]$, and suppose that K_i is the splitting field of $p_i(x)$ for each i. Then $G(f,F)$ is isomorphic to a subgroup of $G(p_1,F)X\cdots XG(p_n,F)$. If $K_{i+1} \cap (K_1\cdots K_i) = F$ for $i = 1,\ldots,n-1$, then $G(f,F) \cong G(p_1,F)X\cdots XG(p_n,F)$.

This corollary shows that when trying to calculate the Galois group of a polynomial f(x), we need only search inside the product of the Galois groups of its irreducible factors. Furthermore, if we are fortunate enough to have $K_{i+1} \cap (K_1\cdots K_i) = F$ for $i = 1,\ldots,n-1$, then we can find the Galois group of f(x) directly from the Galois groups of f's irreducible factors. Thus, for most polynomials f(x), our problem is reduced to the problem of factoring and calculating the Galois group of irreducible polynomials.

The Galois group of $f(x) = x^6 - 2x^4 - 2x^2 + 4$ provides an example of when $G(f,Q)$ is not equal to the product of its factors. $f(x) = (x^4-2)(x^2-2)$ which are both irreducible. Let $p_1(x) = x^4-2$, $p_2(x) = x^2-2$, K_1 be the splitting field of $p_1(x)$, K_2 the splitting field of $p_2(x)$ and K the splitting field of $f(x)$. The roots of $p_1(x)$ are $\xi_1 = \sqrt[4]{2}$, $\xi_2 = -\xi_1$, $\xi_3 = \xi_1 i$ and $\xi_4 = -\xi_3$. The roots of $p_2(x)$ are $\xi_5 = \sqrt{2}$ and $\xi_6 = -\sqrt{2}$. Since $\sqrt{2} = \xi_1^2 \in K_1$, $K_2 \subseteq K_1$. Hence $K = K_1$ and $G(f,Q) = G(p_1,Q) \neq G(p_1,Q) \times G(p_2,Q) = G(p_1,Q) \times C_2$ where C_2 is the cyclic group of order 2. To calculate $G(p_1,Q)$, we first observe that $[K_1:Q] = [Q(\sqrt[4]{2},\sqrt[4]{2}i):Q] = [Q(\sqrt[4]{2},\sqrt[4]{2}i):Q(\sqrt[4]{2})][Q(\sqrt[4]{2}):Q]$. $[Q(\sqrt[4]{2}):Q] = 4$ since $p_1(x)$ is irreducible over Q. Also in $Q(\sqrt[4]{2})$, $p_1(x) = (x^2-\sqrt{2})(x^2+\sqrt{2})$ and $x^2 + \sqrt{2}$ is irreducible over $Q(\sqrt[4]{2})$. Thus $\sqrt[4]{2}i$ has degree 2 over $Q(\sqrt[4]{2})$ and $[Q(\sqrt[4]{2},\sqrt[4]{2}i):Q(\sqrt[4]{2})] = 2$. Hence $[K_1:Q] = 8$. Complex conjugation is always an automorphism so that $(3\ 4) \in G(p_1,Q)$. Since $G(p_1,Q)$ is transitive, there must be $\sigma \in G(p_1,Q)$ such that $\sigma(\xi_1) = \xi_2$. Then $\sigma(\xi_2) = \sigma(-\xi_1) = -\sigma(\xi_1) = -\xi_2 = \xi_1$. Thus $(1\ 2)$ or $(1\ 2)(3\ 4) \in G(p_1,Q)$. Because $(3\ 4) \in G(p_1,Q)$, both $(1\ 2),(1\ 2)(3\ 4) \in G(p_1,Q)$. If $\tau \in G(p_1,Q)$ with $\tau(\xi_1) = \xi_3$, then $\tau(\xi_2) = -\tau(\xi_1) = -\xi_3 = \xi_4$. Hence $(1\ 3)(2\ 4)$ or $(1\ 3\ 2\ 4)$ is in $G(p_1,Q)$. But $(1\ 3)(2\ 4)(1\ 2) = (1\ 3\ 2\ 4)$ and $(1\ 3\ 2\ 4)(1\ 2) = (1\ 3)(2\ 4)$, so that if one of $(1\ 3)(2\ 4)$ and $(1\ 3\ 2\ 4)$ is in $G(p_1,Q)$, then both are. Also $(1\ 3\ 2\ 4)^3 = (1\ 4\ 3\ 2) \in G(p_1,Q)$ and $(1\ 4\ 3\ 2)(1\ 2) = (1\ 4)(2\ 3) \in G(p_1,Q)$. Therefore $G(p_1,Q) = \{\iota,(1\ 2),(3\ 4),(1\ 2)(3\ 4),(1\ 3)(2\ 4),(1\ 4)(2\ 3),(1\ 3\ 2\ 4),(1\ 4\ 3\ 2)\}$ which is isomorphic to the dihedral group of order 8.

Definition: Let F be a finite extension of the rational numbers. The algebraic integers of F (or integers of F) are all of the elements of F which satisfy a monic irreducible polynomial with integer coefficients. This set is denoted by I_F.

We can further simplify our problem by observing that it is necessary to consider only monic polynomials with algebraic integer coefficients. To see this let $f(x) = \sum_{i=o}^{n} a_i x^i \in F[x]$. Then for each i, $a_i = \frac{b_i}{c_i}$, where $b_i, c_i \in I_F$. Let d be the least common multiple in I_F of $c_o, \ldots, c_n$. Then $g(x) = df(x) \in I_F[x]$, and $g(x)$ has the same roots as $f(x)$. Thus $g(x)$ has the same splitting field, and so the same Galois group, as $f(x)$. Suppose $g(x) = \sum_{i=o}^{n} d_i x^i$ with $d_i \in I_F$, and put $h(x) = x^n + \sum_{i=o}^{n-1} d_n^{n-i-1} d_i x^i$. If $g(\alpha) = 0$, then $h(d_n\alpha) =$

$d_n^n \alpha^n + \sum_{i=o}^{n-1} d_n^{n-1} d_i \alpha^i = d_n^{n-1}[\sum_{i=o}^{n} d_i \alpha^i] = d_n^{n-1} g(\alpha) = 0$. Since $d_n \in F$, $h(x)$ has the same splitting field as $g(x)$, and so the same Galois group as $f(x)$.

CHAPTER II

FACTORING

A. Irreducibility Criterion

In factoring polynomials over the rational numbers, it is convenitent to know whether or not the polynomial is reducible in the first place. I present here the three most general irreducibility criterion involving divisibility of coefficients. The most useful is the Theorem of Dumas, which appears first.

Let $f(x) \varepsilon Q[x]$ and p be a prime number. Write $f(x) = \sum_{i=o}^{n} a_i p^{b_i} x^i$, where either $a_i = 0$ or a_i is relatively prime to p for each i. Consider the cartesian coordinates (i,b_i) for each i with $a_i \neq 0$. Let $P_o = (0,b_o)$ and $P_j = (k_j,b_{k_j})$, where k_j is the greatest integer such that no (i,b_i) lies below the line from P_{j-1} to P_j.

Definition: The Newton polygon for f(x) corresponding to p is the set of line segments $P_oP_1,P_1P_2,\ldots,P_{r-1}P_r$, where $P_r = (n,b_n)$. Consider all of the points with integer coordinates which fall on the Newton polygon. The portion of the polygon joining two such points is called an element of the polygon. Note that the number of elements is greater than or equal to r.

Theorem 2.1: Let $f(x),g(x),h(x) \varepsilon Z[x]$ with $f(x) = g(x)h(x)$, and let p be a prime. Then the Newton polygon for g(x) corresponding to p can be formed by joining some of the elements of the Newton polygon for f(x) corresponding to p without changing their lengths or slopes. Furthermore, the Newton polygon for h(x) corresponding to p can be formed by joining the remaining elements.

Proof: Let $f(x) = \sum_{j=o}^{n} a_j p^{b_j} x^j$, $g(x) = \sum_{i=o}^{m} c_i p^{d_i} x^i$ and $h(x) = \sum_{k=o}^{n-m} m_k p^{e_k} x^k$, where the a_j, c_i and m_k are either zero or relatively prime to p. Let $P_i P_{i+1}$ be a segment of the Newton polygon for f(x). Suppose $P_i = (j_q, b_{j_q})$, $P_{i+1} = (j_s, b_{j_s})$, and let d be the greatest common divisor of $j_q - j_s$ and $b_{j_q} - b_{j_s}$. Then $j_q - j_s = Kd$ and $b_{j_q} - b_{j_s} = Bd$ for some B,K; and the slope of $P_i P_{i+1}$ is $\frac{B}{K}$. Also B and K are relatively prime and the equation of the line $P_i P_{i+1}$ is $Ky - Bx = C$, where $C = Kb_{j_q} - Bj_q = Kb_{j_s} - Bj_s$. Now for every (j, b_j) we have that $C \leq Kb_j - Bj$; and if $j < j_q$ or $j > j_s$, then $C < Kb_j - Bj$; and if $j < j_q$ or $j > j_s$, then $C < Kb_j - Bj$.

Notice that these are the defining properties for the endpoints of a segment of a Newton polygon. That is, if j_t and j_r are the least and greatest exponents of x such that $\frac{B_1}{K_1} = \frac{b_{j_t} - b_{j_r}}{j_t - j_r}$, where B_1 and K_1 are relatively prime, and $K_1 b_j - B_1 j > K_1 b_{j_t} - B_1 j_t$ for all j less than j_t or greater than j_r, then (j_t, b_{j_t}), (j_r, b_{j_r}) are endpoints of some segment of the Newton polygon for f(x).

Consider the numbers $Kd_i\text{-}Bi$ for all i where $c_i \neq 0$. Let $D = \min_{\substack{0 \le i \le m \\ c_i \neq 0}}\{Kd_i\text{-}Bi\}$, and let i_q, i_s be the least and greatest exponents of x such that $D = Kd_{i_q}\text{-}Bi_q = Kd_{i_s}\text{-}Bi_s$. Then $D \le Kd_i\text{-}Bi$ for each i.

Also put $E = \min_{\substack{0 \le k \le n-m \\ m_k \neq 0}}\{Ke_k\text{-}Bk\}$, and let k_q and k_s be the least and greatest exponents of x such that $E = Ke_{k_q}\text{-}Bk_q = Ke_{k_s}\text{-}Bk_s$. Then $E \le Ke_k\text{-}Bk$ for each k.

We have that $a_{[i_q+k_q]}p^{b_{[i_q+k_q]}}x^{[i_q+k_q]} = \sum_{i+k=i_q+k_q}(c_i p^{d_i}x^i)(m_k p^{e_k}x^k)$. Also for $i < i_q$, $d_i > \frac{D+Bi}{K}$ and for $k < k_q$, $e_k > \frac{E+Bk}{K}$.

So if $i \neq i_q$, but $i+k = i_q+k_q$, then $d_i+e_k > \frac{D+E+B(i+k)}{K} = \frac{D+E+B(i_q+k_q)}{K}$ $= d_{i_q}+e_{k_q}$. Thus $\sum_{i+k=i_q+k_q}(c_i p^{d_i}x^i)(m_k p^{e_k}x^k) = p^{[d_{i_q}+e_{k_q}]}$ $(c_{i_q}m_{k_q}+p\sum_{i+k=i_q+k_q}c_i m_k p^{[d_i+e_k-d_{i_q}-e_{k_q}-1]})x^{i_q+k_q}$, and the part in the parentheses is relatively prime to p. So $b_{i_q+k_q} = d_{i_q}+e_{k_q}$ and $Kb_{i_q+k_q}\text{-}B(i_q+k_q) = D+E$. Also for $j < i_q+k_q$, $Kb_j\text{-}Bj > D+E$; while if $j > i_q+k_q$ then $Kb_j\text{+}Bj \ge D+E$. Therefore $D+E = C$ and $i_q+k_q = j_q$. In a similar manner we get that $i_s+k_s = j_s$. Hence $0 < j_s-j_q = (i_s-i_q)+(k_s-k_q)$

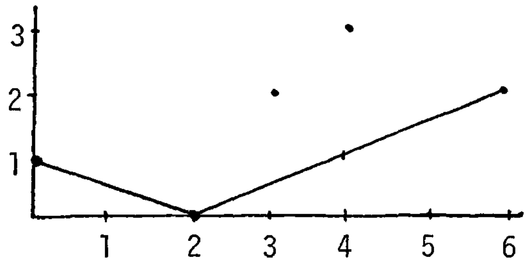
and either $i_s - i_q > 0$ or $k_s - k_q > 0$.

If $i_s = i_q$, then $k_s - k_q = j_s - j_q$ and (k_q, e_{k_q}), (k_s, e_{k_s}) are endpoints of a segment of the Newton polygon for h(x) with slope B/K.

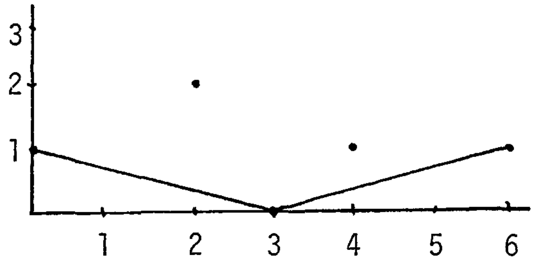
If $k_s = k_q$, then $i_s - i_q = j_s - j_q$ and $(i_q, d_{i_q}), (i_s, d_{i_s})$ are endpoints for a segment of the Newton polygon for g(x) with slope B/K.

If both are greater than zero, then $(i_q, d_{i_q}), (i_s, d_{i_s})$ are endpoints of a segment of the Newton polygon for g(x) and (k_q, e_{k_q}), (k_s, e_{k_s}) are endpoints of a segment of the Newton polygon for h(x). Both segments have slope B/K. In all three cases, the conclusions of the theorem hold. //

The Theorem of Dumas can be used to test for irreducibility as the following example illustrates. Let $f(x) = 63x^6 + 189x^4 + 18x^3 + 49x^2 + 42$. With p = 3 we have $f(x) = 7 \cdot 3^2 x^6 + 7 \cdot 3^3 x^4 + 2 \cdot 3^2 x^3 + 49 \cdot 3^0 x^2 + 14 \cdot 3^1$. The Newton polygon of f(x) corresponding to 3 has 3 elements each of length 2. So if f(x) has a factor, it must have one of degree 2. With p = 7, $f(x) = 9 \cdot 7^1 x^6 + 27 \cdot 7^1 x^4 + 18 \cdot 7^0 x^3 + 1 \cdot 7^2 x^2 + 6 \cdot 7^1$. The Newton polygon of f(x) corresponding to 7 has 2 elements both of length 3. Thus if f(x) has a factor, it must have degree 3. Therefore f(x) is irreducible.



Newton polygon for p=3



Newton polygon for p=7

As an immediate corollary of the Theorem of Dumas we get Eisenstein's irreducibility criterion:

Corollary 2.2: Let $f(x) = \sum_{i=o}^{n} a_i x^i$ be a polynomial in Z[x] and p a prime. If $p|a_i$ for $i = 0,1,\ldots,n-1$, but $p \nmid a_n$ and $p^2 \nmid a_0$, then f(x) is irreducible over the rational numbers.

The final irreducibility criterion follows from the next theorem. For this theorem, we will use the following notation. Let $t(x) \in Z[x]$ with $t(x) = \sum_{i=o}^{k} a_i x^i$, and let p be an odd prime. By $\hat{a}_i$ we mean that unique integer such that $-\frac{p-1}{2} \leq \hat{a}_i \leq \frac{p-1}{2}$ and $\hat{a}_i = a_i + c_i p$, where $c_i \in Z$. If p = 2, then $\hat{a}_i = 0$ if a_i is even and $\hat{a}_i = 1$ if a_i is odd. Then we let $\hat{t}(x) = \sum_{i=o}^{k} \hat{a}_i x^i$.

Definition: Let $f(x), g(x) \in Z[x]$ and n be an integer. Then $f(x) \equiv g(x) \pmod{nZ[x]}$ provided n divides all of the coefficients of $f(x)-g(x)$.

Theorem 2.3: Let $f(x), g(x), h(x) \in Z[x]$ with $f(x) = g(x)h(x)$. Then for any prime p, $f(x) \equiv \hat{g}(x)\hat{h}(x) \pmod{pZ[x]}$.

Proof: Let $g(x) = \sum_{i=o}^{n} a_i x^i$, $h(x) = \sum_{j=o}^{m} b_j x^j$ and $\hat{a}_i = a_i + c_i p$, $\hat{b}_j = b_j + d_j p$. Also for $i > n$ let $c_i = a_i = \hat{a}_i = 0$, and for $j > m$ let $d_j = b_j = \hat{b}_j = 0$. Then

$$\hat{g}(x)\hat{h}(x) \equiv \sum_{i=o}^{n+m} x^i \sum_{j=o}^{i} \hat{a}_j \hat{b}_{i-j}$$

$$\equiv \sum_{i=o}^{n+m} x^i \sum_{j=o}^{i} [(a_j + c_j p)(b_{i-j} + d_{i-j} p)]$$

$$\equiv \sum_{i=o}^{n+m} x^i \sum_{j=o}^{i} [a_j b_{i-j} + p(a_j d_{i-j} + b_{i-j} c_j) + p^2 c_j d_{i-j}]$$

$$\equiv \sum_{i=o}^{n+m} x^i \sum_{j=o}^{i} a_j b_{i-j} \equiv f(x) \pmod{pZ[x]}. \quad //$$

Corollary 2.4: If $f(x) \in Z[x]$ is irreducible modulo $pZ[x]$ for some prime p, then $f(x)$ is irreducible over Z (and hence over Q).

B. Factorization Over Q[x]

Definition: A valuation on a field K is a function ϕ from K to the real numbers such that for all $a,b \in K$:

(1) $\phi(a) > 0$ if $a \neq 0$,

(2) $\phi(0) = 0$,

(3) $\phi(ab) = \phi(a)\phi(b)$,

(4) $\phi(a+b) \leq \phi(a) + \phi(b)$.

An example of a valuation is the absolute value function on the real numbers.

<u>Definition</u>: A valuation ϕ on K is said to be <u>non-Archimedean</u> provided $\phi(a+b) \leq \max\{\phi(a),\phi(b)\}$ for $a,b \in K$.

Let p be a prime integer and let a be a rational number. Write $a = \frac{s}{t} p^n$ where $p \nmid s$, $p \nmid t$ and $s,t \in Z$. If we let

$$\phi_p(a) = \begin{cases} p^{-n} & \text{if } a \neq 0 \\ 0 & \text{if } a = 0 \end{cases}$$

then ϕ_p satisfies (1), (2) in the definition of valuation. If $a,b \in Q$, say $a = \frac{s_1}{t_1} p^{n_1}$, $b = \frac{s_2}{t_2} p^{n_2}$, where p does not divide s_1, s_2, t_1, t_2, then $ab = \frac{s_1 s_2}{t_1 t_2} p^{n_1+n_2}$. So $\phi_p(ab) = p^{-(n_1+n_2)} = p^{-n_1} p^{-n_2} = \phi_p(a)\phi_p(b)$.

Also if $n_1 \leq n_2$, then $a+b = \frac{s_1 t_2 + s_2 t_1 p^{n_2-n_1}}{t_1 t_2} p^{n_1}$ and $\phi_p(a+b) \leq p^{-n_1} = \max\{\phi_p(a),\phi_p(b)\} \leq \phi_p(a) + \phi_p(b)$. Hence ϕ_p is a non-Archimedean valuation of the rational numbers.

<u>Definition</u>: The valuation ϕ_p constructed above is called the <u>p-adic valuation</u> of Q.

Any field with a valuation has a completion. This completion can be constructed in the usual way by using Cauchy sequences, that is, sequences $\{x_n\}$ from the field for which $\lim_{m,n\to\infty} \phi(x_n - x_m) = 0$.

Definition: The completion of the rational numbers using the p-adic valuation is called the p-adic numbers and is denoted by Q_p.

Definition: For a non-Archimedean valuation ϕ on the field K, the set $\{a \in K: \phi(a) \leq 1\}$ is called the integral elements of K or the integers of K. Note that the set of integral elements of K is a ring, and the set $\{a \in K: \phi(a) < 1\}$ is an ideal in that ring. It is the ideal of non-units of the ring.

Definition: Let D be a ring and P an ideal of D. Then for $f(x),g(x) \in D[x]$, $f(x) \equiv g(x) \pmod{PD[x]}$ means that the coefficients of $f(x)-g(x)$ are in P.

Definition: Let D be a ring and P an ideal of D. $f(x),g(x) \in D[x]$ are relatively prime modulo P provided there exist $s(x),t(x) \in D[x]$ such that $s(x)f(x) + t(x)g(x) \equiv 1 \pmod{PD[x]}$. $f(x) \in D[x]$ is said to be primitive if the only elements of D which divide all of the coefficients of $f(x)$ are units.

In factoring polynomials over the rational numbers, a reducibility criterion called Hensel's lemma is useful. It is presented here in a general setting and a bit later in a manner more applicable to the rational numbers.

Theorem 2.5: Let K be a complete field under the non-Archimedean valuation ϕ, D the set of integral elements of K and $P = \{a \in K: \phi(x) < 1\}$. Suppose $f(x),g_0(x),h_0(x) \in D[x]$ such that $f(x)$ is primitive, $g_0(x)$ and $h_0(x)$ are relatively prime modulo P, and $f(x) \equiv g_0(x)h_0(x) \pmod{PD[x]}$. The there are polynomials $g(x),h(x) \in D[x]$ such that

(1) $f(x) = g(x)h(x)$,

(2) $g(x) \equiv g_0(x) \pmod{PD[x]}$,

(3) $h(x) \equiv h_0(x) \pmod{PD[x]}$,

(4) $[g] = [g_0]$.

Proof: If an element of P divides one of the coefficients of $g_0(x)$ or $h_0(x)$, we may set that coefficient equal to 0; so we assume that the leading coefficient of both $g_0(x)$ and $h_0(x)$ is a unit. Also assume that the leading coefficient of $g_0(x)$ is 1. If not, divide $g_0(x)$ by its leading coefficient and multiply $h_0(x)$ by the same number. Let $[g_0] = m$ and $[f] = n$; then $[h_0] \leq n-m$.

Since the coefficients of $f-g_0h_0$ are elements of P they have ϕ-value greater than or equal to 0, but strictly less than 1. Let δ_1 be the greatest of these values. If $\delta_1 = 0$, then $f(x) = g_0(x)h_0(x)$ and we are done; if not, then $0 < \delta_1 < 1$.

Since $g_0(x)$ and $h_0(x)$ are relatively prime modulo P, there are $s(x), t(x) \in D[x]$ such that $s(x)g_0(x) + t(x)h_0(x) \equiv 1 \pmod{PD[x]}$. As before, the coefficients of $s(x)g_0(x) + t(x)h_0(x) - 1$ have ϕ-values between 0 and 1. Let δ_2 be the greatest of these coefficients; then $0 \leq \delta_2 < 1$. Set $\varepsilon = \max\{\delta_1, \delta_2\}$ and let $\pi \in K$ with $\phi(\pi) = \varepsilon$. Such a π must exist since one of the coefficients of $f-g_0h_0$ or $sg_0 + th_0-1$ has ϕ-value ε. So we have

$$f(x) \equiv g_0(x)h_0(x) \pmod{\pi D[x]},$$

$$s(x)g_0(x) + t(x)h_0(x) \equiv 1 \pmod{\pi D[x]}.$$

We will construct two sequences $\{g_k(x)\}$, $\{h_k(x)\}$ of polynomials in $D[x]$ with the following properties:

(a) $f(x) \equiv g_k(x)h_k(x) \pmod{\pi^{k+1}D[x]}$,

(b) $g_k(x) \equiv g_0(x) \pmod{\pi D[x]}$,

(c) $h_k(x) \equiv h_0(x) \pmod{\pi D[x]}$,

(d) $g_k(x)$ is monic, $[g_k] = [g_0]$ and $[h_k] \leq n-m$.

We proceed by induction. Suppose we have constructed such polynomials g_i, h_i for $i = 1,2,\ldots,k-1$. We will now construct g_k and h_k.

By (a) we see that $f(x)-g_{k-1}(x)h_{k-1}(x) = \pi^k p(x)$ for some $p(x) \in D[x]$. Then $p(x)s(x)g_0(x) + p(x)t(x)h_0(x) \equiv p(x) \pmod{\pi D[x]}$. If we divide $p(x)t(x)$ by $g_0(x)$, we get a quotient $q(x)$ and a remainder $r(x)$ with $[r] < m$. So $p(x)t(x) = g_0(x)q(x) + r(x)$, and $p(x)s(x)g_0(x) + [g_0(x)q(x) + r(x)]h_0(x) \equiv p(x) \pmod{\pi D[x]}$, or $[p(x)s(x) + q(x)h_0(x)]g_0(x) + r(x)h_0(x) \equiv p(x) \pmod{\pi D[x]}$. Let $u(x) \equiv p(x)s(x) + q(x)h_0(x) \pmod{\pi D[x]}$, where the coefficients of $u(x)$ are units or zero. Then $u(x)g_0(x) + r(x)h_0(x) \equiv p(x) \pmod{\pi D[x]}$.

Put $g_k(x) = g_{k-1}(x) + \pi^k r(x)$ and $h_k(x) = h_{k-1}(x) + \pi^k u(x)$, then $[g_k] = [g_{k-1}] = [g_0]$. Also $[h_k] \leq n-m$, for if not, then $[u] > n-m$ and $[ug_0] > n$. Now $[rh_0] < m + [h_0] \leq m+n-m = n$, so that $[ug_0+rh_0] > n$ and $[p] > n$. But by the selection of p, $[p] \leq n$. So (d) has been verified.

To see that (b) and (c) hold, we note that $g_k(x) \equiv g_{k-1}(x) \equiv g_0(x) \pmod{\pi D[x]}$ and $h_k(x) \equiv h_{k-1}(x) \equiv h_0(x) \pmod{\pi D[x]}$.

Finally for (a), we have that

$$g_k(x)h_k(x)-f(x) = g_{k-1}(x)h_{k-1}(x)-f(x)$$

$$+ \pi^k[r(x)h_{k-1}(x) + u(x)g_{k-1}(x)] + \pi^{2k}r(x)u(x)$$

$$= \pi^k[r(x)h_{k-1}(x) + u(x)g_{k-1}(x)-p(x)] + \pi^{2k}r(x)u(x)$$

so that

$$g_k(x)h_k(x)-f(x)$$

$$\equiv \pi^k[r(x)h_{k-1}(x) + u(x)g_{k-1}(x)-p(x)] \pmod{\pi^{k+1}D[x]}.$$

Also

$$r(x)h_{k-1}(x) + u(x)g_{k-1}(x) \equiv p(x) \pmod{\pi D[x]},$$

so

$$g_k(x)h_k(x)-f(x) \equiv 0 \pmod{\pi^{k+1}D[x]}.$$

Since

$$g_{k+1}(x) \equiv g_k(x) \pmod{\pi^{k+1}D[x]}$$

and

$$h_{k+1}(x) \equiv h_k(x) \pmod{\pi^{k+1}D[x]},$$

we must have the coefficients of $g_k(x)$ and $h_k(x)$ converging. For if $g_k(x) = \sum_{i=o}^{m} a_{i,k}x^i$ then $\pi^{k+1} | (a_{i,k+1}-a_{i,k})$. So $\phi(\frac{a_{i,k+1}-a_{i,k}}{\pi^{k+1}}) \leq 1$ or $\phi(a_{i,k+1}-a_{i,k}) \leq \varepsilon^{k+1}$, which tends to 0. Hence, since K is complete, $\{a_{i,k}\}_{k=o}^{\infty}$ converges. In a similar manner the coefficients of $h_k(x)$ converge. Let $g(x) = \lim_{k\to\infty} g_k(x)$ and $h(x) = \lim_{k\to\infty} h_k(x)$.

Each $g_k(x)$ is congruent to $g_o(x)$ modulo $\pi D[x]$ and each $h_k(x)$ is congruent to $h_o(x)$ modulo $\pi D[x]$. Hence the limits, $g(x)$ and $h(x)$, must be congruent to $g_o(x)$ and $h_o(x)$, respectively, modulo $\pi D[x]$.

Also $[g] = [g_0]$, and since $f(x) \equiv g_k(x)h_k(x) \pmod{\pi^{k+1}D[x]}$ for each k, it follows that $f(x) = g(x)h(x)$. //

Lemma 2.6: Let p be a rational prime, k an integer and $f(x),g(x),h(x) \in Z[x]$, where p does not divide the leading coefficient of $g(x)$. If $[f] < [g]$ and $f(x) \equiv p^k g(x)h(x) \pmod{p^{2k}Z[x]}$, then $h(x) \equiv 0 \pmod{p^k Z[x]}$, and consequently $f(x) \equiv 0 \pmod{p^{2k}Z[x]}$.

Proof: Write $h(x) = \sum_{i=0}^{n} a_i x^i$ and let j be such that $p^k \nmid a_j$, but for $i > j$, $p^k | a_i$. Suppose, for $i > j$, $a_i = p^k b_i$ with $b_i \in Z$, and let b be the leading coefficient of $g(x)$. Then

$$f(x) \equiv p^k g(x)h(x) \equiv (\sum_{i=0}^{n} p^k a_i x^i)g(x) \equiv (\sum_{i=j+1}^{n} p^{2k} b_i x^i + \sum_{i=0}^{j} p^k a_i x^i)g(x)$$

$$\equiv (\sum_{i=0}^{j} p^k a_i x^i)g(x) \pmod{p^{2k}Z[x]}.$$ Because $[g] > [f]$, we must have the leading term on the right congruent to 0 modulo p^{2k}. So $p^k b a_j \equiv 0 \pmod{p^{2k}}$ and $b a_j \equiv 0 \pmod{p^k}$. Since $p \nmid b$, it must be that $p | a_j$. Thus no such j exists, and $h(x) \equiv 0 \pmod{p^k Z[x]}$. //

A more useful form of Theorem 2.5 is the following. It gives an algorithm for factoring polynomials with integer coefficients modulo $p^k Z[x]$ for arbitrarily large k.

Definition: Let n be an integer and $f(x) \in Z[x]$. $f(x)$ is said to be reduced modulo n provided the coefficients of $f(x)$ are in the interval $(-\frac{n}{2}, \frac{n}{2}]$.

Theorem 2.7: Let $f(x), g_0(x), h_0(x), s_0(x), t_0(x), r_0(x)$, $u_0(x) \in Z[x]$, p be a prime and k a positive integer. If

(a) f, g_0, h_0 are monic and non-constant,

(b) g_0, h_0, s_0, t_0 are reduced modulo p^k,

(c) $[s_0] < [h_0]$, $[t_0] < [g_0]$,

(d) $f = g_0 h_0 + p^k r_0$,

(e) $s_0 g_0 + t_0 h_0 = 1 + p^k u_0$,

(f) r_0 is not identically zero,

then there are polynomials $g_1, h_1, s_1, t_1, r_1, u_1 \in Z[x]$ such that

(1) g_1, h_1 are monic and non-constant,

(2) g_1, h_1, s_1, t_1 are reduced modulo p^{2k},

(3) $[s_1] < [h_1]$, $[t_1] < [g_1]$,

(4) $f = g_1 h_1 + p^{2k} r_1$,

(5) $s_1 g_1 + t_1 h_1 = 1 + p^{2k} u_1$,

(6) $g_1 \equiv g_0 \pmod{p^k Z[x]}$, $h_1 \equiv h_0$ (mod $p^k Z[x]$.

Proof: The following is the algorithm for obtaining $g_1, h_1, s_1, t_1, r_1, u_1$.

Divide $t_0 r_0$ by g_0 and $s_0 r_0$ by h_0 modulo $p^k Z[x]$ to get remainders d_0 and d_0^*. Then $t_0 r_0 \equiv d_0 \pmod{(p^k, g_0) Z[x]}$ and $s_0 r_0 \equiv d_0^* \pmod{(p^k, h_0) Z[x]}$. Let

$$\phi_0 = g_0 + p^k d_0, \quad \phi_0^* = h_0 + p^k d_0^*. \tag{7}$$

Reduce ϕ_0 and ϕ_0^* modulo p^{2k} to obtain g_1 and h_1.

Then

$$g_1 = \phi_0 + p^{2k} \beta_0, \quad h_1 = \phi_0^* + p^{2k} \beta_0^* \text{ for some } \beta_0, \beta_0^* \in Z[x]. \tag{8}$$

Set

$$\sigma_o = d_o + p^k\beta_o \text{ and } \sigma_o^* = d_o^* + p^k\beta_o^*. \tag{9}$$

Then

$$g_1 = g_o + p^k\sigma_o \text{ and } h_1 = h_o + p^k\sigma_o^*. \tag{10}$$

Now let

$$L_o = -(u_o + s_o\sigma_o + t_o\sigma_o^*), \tag{11}$$

and divide L_os_o by h_o, L_ot_o by g_o modulo $p^kZ[x]$ to obtain remainders P_o and P_o^*. Then

$$L_os_o \equiv P_o \ (\text{mod}(p^k,h_o)Z[x]),\ L_ot_o \equiv P_o^* \ (\text{mod}(p^k,g_o)Z[x]). \tag{12}$$

$$\text{Next put } \alpha_o = s_o + p^kP_o \text{ and } \alpha_o^* = t_o + p^kP_o^*. \tag{13}$$

Reduce these modulo p^{2k} to get s_1 and t_1. Then

$$s_1 = \alpha_o + p^{2k}\psi_o \text{ and } t_1 = \alpha_o^* + p^{2k}\psi_o^*, \text{ where } \psi_o,\psi_o^* \in Z[x]. \tag{14}$$

If we let

$$\pi_o = P_o + p^k\psi_o \text{ and } \pi_o^* = P_o^* + p^k\psi_o^*, \tag{15}$$

then

$$s_1 = s_o + p^k\pi_o \text{ and } t_1 = t_o + p^k\pi_o^*. \tag{16}$$

Finally we let

$$r_1 = (r_o + \sigma_o^*g_o - \sigma_oh_o)/p^k - \sigma_o\sigma_o^* \tag{17}$$

and

$$u_1 = (-L + \pi_og_o + \pi_o^*h_o)/p^k + \pi_o\sigma_o + \pi_o^*\sigma_o^*. \tag{18}$$

Now we will show that conditions (1)-(6) hold for g_1,h_1,s_1,t_1,r_1,u_1. Clearly (6) is satisfied because of (10).

To prove (1), notice that $[d_0] < [g_0]$ and $[d_0^*] < [h_0]$. Also $[\beta_0] < [\phi_0]$ and $[\beta_0^*] < [\phi_0^*]$ so that, by (7), $[\phi_0] = [g_0]$, $[\phi_0^*] = [h_0]$ and by (8), $[g_1] = [\phi_0]$, $[h_1] = [\phi_0^*]$. ϕ_0 and ϕ_0^* must then be monic; hence g_1 and h_1 are monic. Neither g_1 nor h_1 could be constant since $[g_1] = [g_0]$ and $[h_1] = [h_0]$.

By the construction of g_1, h_1, s_1, t_1, they are all reduced modulo $p^{2k}Z[x]$, and so (2) is satisfied.

By the definition of P_0 and P_0^*, we have that $[P_0] < [h_0]$ and $[P_0^*] < [g_0]$. So, by (13), $[\alpha_0] < [h_0]$ and $[\alpha_0^*] < [g_0]$. Now $[s_1] \leq [\alpha_0]$ and $[t_1] \leq [\alpha_0^*]$, hence $[s_1] < [h_0] = [h_1]$ and $[t_1] < [g_0] = [g_1]$. This proves (3).

For (4) we see that

$$
\begin{aligned}
f-g_1h_1 &= f-(g_0+p^k\sigma_0)(h_0+p^k\sigma_0^*) \text{ by (10)} \\
&= f-g_0h_0-p^k(\sigma_0^*g_0 + \sigma_0h_0)-p^{2k}\sigma_0\sigma_0^* \\
&= p^k(r_0-\sigma_0^*g_0-\sigma_0h_0)-p^{2k}\sigma_0\sigma_0^* \text{ by (d)} \\
&= p^{2k}r_1 \text{ by (17).}
\end{aligned}
$$

We still must show that $r_1 \in Z[x]$. Since $t_0r_0 \equiv d_0 \pmod{(p^k,g_0)Z[x]}$, and $s_0r_0 \equiv d_0^* \pmod{p^k,h_0)Z[x]}$, from (9) it follows $t_0r_0 = q_0g_0 + p^kb_0 + \sigma_0$ and $s_0r_0 = q_0^*h_0 + p^kb_0^* + \sigma_0^*$, where $q_0, q_0^*, b_0, b_0^* \in Z[x]$. Then

$$
\begin{aligned}
f-g_1h_1 &\equiv p^k(r_0-\sigma_0^*g_0-\sigma_0h_0) \\
&\equiv p^k[r_0-(s_0r_0-q_0^*h_0)g_0-(t_0r_0-q_0g_0)h_0]
\end{aligned}
$$

$$\equiv p^k[r_o(1-s_og_o-t_oh_o) + (q_o+q_o^*)h_og_o]$$

$$\equiv p^k[r_o(-p^ku_o) + (q_o+q_o^*)(f-p^kr_o)]$$

$$\equiv p^k(q_o+q_o^*)f \pmod{p^{2k}Z[x]}.$$

Since $q_o + q_o^* \in Z[x]$ and $[f-g_1h_1] < [f]$, we can apply Lemma 2.6 to get that $f-g_1h_1 \equiv 0 \pmod{p^{2k}Z[x]}$. Hence $r_1 \in Z[x]$.

Finally we prove (5). By (10) and (16),

$$s_1g_1 + t_1h_1-1 = (s_o+p^k\pi_o)(g_o+p^k\sigma_o) + (t_o+p^k\pi_o^*)(h_o+p^k\sigma_o^*)-1$$

$$= s_og_o + t_oh_o-1 + p^k(s_o\sigma_o+t_o\sigma_o^*+\pi_og_o+\pi_o^*h_o)$$

$$+ p^{2k}(\pi_o\sigma_o+\pi_o^*\sigma_o^*)$$

$$= p^k(u_o+s_o\sigma_o+t_o\sigma_o^*+\pi_og_o+\pi_o^*h_o)$$

$$+ p^{2k}(\pi_o\sigma_o+\pi_o^*\sigma_o^*) \text{ by (e)}$$

$$= p^k(-L_o+\pi_og_o+\pi_o^*h_o) + p^{2k}(\pi_o\sigma_o+\pi_o^*\sigma_o^*) \text{ by (11)}$$

$$= p^{2k}u_1 \text{ by (18)}.$$

Now we need only show that $u_1 \in Z[x]$. By (12) and (15), there are polynomials $G_o, G_o^*, H_o, H_o^* \in Z[x]$ such that $L_os_o = G_oh_o + p^kH_o + \pi_o$ and $L_ot_o = G_o^*g_o + p^kH_o^* + \pi_o^*$. Then

$$s_1g_1 + t_1h_1-1 \equiv p^k(-L_o+\pi_og_o+\pi_o^*h_o)$$

$$\equiv p^k[-L_o + (L_os_o-G_oh_o)g_o + (L_ot_o-G_o^*g_o)h_o]$$

$$\equiv p^k[-L_o(1-s_og_o-t_oh_o)-(G_o+G_o^*)g_oh_o]$$

$$\equiv p^k[-L_o(-p^k u_o)-(G_o+G_o^*)(f-p^k r_o)]$$

$$\equiv -p^k(G_o+G_o^*)f \pmod{p^{2k}Z[x]}.$$

Again we can apply Lemma 2.6 since $[s_1g_1+t_1h_1-1] < [f]$ and $(G_o+G_o^*) \in Z[x]$. Thus $s_1g_1 + t_1h_1-1 \equiv 0 \pmod{p^{2k}Z[x]}$ and $u_1 \in Z[x]$. //

A look at this theorem shows that it is just a constructive form of Hensel's lemma with $K = Q_p$ and $\phi = \phi_p$. We are guaranteed that the sequences $\{g_k\}$ and $\{h_k\}$ converge in $Q_p[x]$. At times we are lucky and this form of Hensel's lemma leads to a solution in the integers, but the method need not converge in the integers. We can always start the algorithm modulo p unless f(x) is irreducible. If f(x) is reducible, Theorem 2.3 says that f(x) can be factored modulo p. Now these factors can be chosen so that they are relatively prime modulo p. One problem is that there may be more than one way to choose $g_o(x)$ and $h_o(x)$ so that they are relatively prime. For the method to have a chance to converge in the rational integers, we must pick $g_o(x)$ and $h_o(x)$ so that they have the same degree as factors of f(x) in Z[x].

As an example, consider $f(x) = x^6+3x^5+x^4+7x^3-3x^2+5x-5$. f(x) can be factored modulo 2Z[x] into $g_o(x) = x^3+x+1$ and $h_o(x) = x^3+x^2+1$, which are relatively prime modulo 2. We can find $s_o(x)$ and $t_o(x)$ by solving the congruence $g_o(x)(a_1x^2+b_1x+c_1) + h_o(x)(a_2x^2+b_2x+c_2) \equiv 1$ (mod 2Z[x]) for a_1,b_1,c_1,a_2,b_2,c_2. We find $s_o(x) = x$ and $t_o(x) = x+1$. A simple calculation gives $r_o = x^5+2x^3-2x^2+2x-3$ and $u_o = x^4+x^3+x^2+x$. The computation proceeds as follows:

$$t_or_o \equiv x^6+x^5+2x^4-x-3 \equiv 1 \pmod{(2,x^3+x+1)Z[x]}$$

$$s_or_o \equiv x^6+2x^4-2x^3+2x^2-3x \equiv x^2 \pmod{(2,x^3+x^2+1)Z[x]}$$

So

$$d_o = 1,\ d_o^* = x^2,\ \phi_o = x^3+x+3 \text{ and } \phi_o^* = x^3+3x^2+1.$$

Now

$$x^3+x+3 \equiv x^3+x-1 \pmod{2^2Z[x]},\ x^3+3x^2+1 \equiv x^3-x^2+1 \pmod{2^2Z[x]}$$

so

$g_1 = x^3+x-1$, $h_1 = x^3-x^2+1$, $\beta_o = -1$, $\beta_o^* = -x^2$, $\sigma_o = -1$, $\sigma_o^* = -x^2$ and $L_o = -x^4$.

$$L_os_o \equiv -x^5 \equiv x+1 \pmod{(2,x^3+x^2+1)Z[x]}$$

$$L_ot_o \equiv 1 \pmod{(2,x^3+x+1)Z[x]}.$$

Thus

$$P_o = x+1,\ P_o^* = 1,\ \alpha_o = 3x+2 \text{ and } \alpha_o^* = x+3.$$

$$\alpha_o \equiv 3x+2 \equiv -x+2 \pmod{2^2Z[x]} \text{ and } \alpha_o^* \equiv x+3 \equiv x-1 \pmod{2^2Z[x]}$$

so

$s_1 = -x+2$, $t_1 = x-1$, $\psi_o = -x$, $\psi_o^* = -1$, $\pi_o = -x+1$, $\pi_o^* = -1$, $r_1 = x^5+2x^3-x^2+x-1$ and $u_1 = x-1$.

Since $r_1 \neq 0$ we use the algorithm again.

$$t_1r_1 \equiv x^6-x^5+2x^4-3x^3+2x^2-2x+1 \equiv 0 \pmod{(2^2,x^3+x-1)Z[x]}$$

$$s_1r_1 \equiv -x^6+2x^5-2x^4+5x^3-3x^2+3x-2 \equiv x^2+1 \pmod{(2^2,x^3-x^2+1)Z[x]}.$$

So

$$d_o = 0,\ d_1^* = x^2+1,\ \phi_1 = x^3+x-1 \text{ and } \phi_1^* = x^3+3x+5.$$

$$x^3+x-1 \equiv x^3+x-1 \pmod{2^4 Z[x]} \text{ and } x^3+3x+5 \equiv x^3+3x+5 \pmod{2^4 Z[x]}$$

so

$$g_2 = x^3+x-1,\ h_2 = x^3+3x+5,\ \beta_1 = 0,\ \beta_1^* = 0,\ \sigma_1 = 0,$$

$\sigma_1^* = x^2+1$ and $L_1 = -x^3+x^2-2x+2$.

$$L_1 s_1 \equiv x^4-3x^3+4x^2-6x+4 \equiv 2x^2+x+2 \pmod{(2^4, x^3-x^2+1)Z[x]}$$

$$L_1 t_1 \equiv -x^4+2x^3-3x^2+4x-2 \equiv 2x^2+x \pmod{(2^4, x^3+x-1)Z[x]}.$$

Thus

$$P_1 = 2x^2+x+2,\ P_1^* = 2x^2+x,\ \alpha_1 = 8x^2+3x+10 \text{ and } \alpha_1^* = 8x^2+5x-1.$$

$$8x^2+3x+10 \equiv 8x^2+3x-6 \pmod{2^4 Z[x]}$$

and

$$8x^2+5x-1 \equiv 8x^2+5x-1 \pmod{2^4 Z[x]},$$

so

$$s_2 = 8x^2+3x-6,\ t_2 = 8x^2+5x-1,\ \psi_1 = -1,\ \psi_1^* = 0,\ \pi_1 = 2x^2+x-2,$$

$\pi_1^* = 2x^2+2$ and $r_2 = 0$. Therefore $f(x) = (x^3+x-1)(x^3+3x^2+5)$.

We can modify this method so that we will always get a solution over the integers provided $f(x)$ is reducible over the integers. The goal is to find a constant M such that, if $g(x) = x^m + \sum_{i=0}^{m-1} b_i x^i$ is a factor of the monic polynomial $f(x)$, then $|b_i| \leq M$ for each i. If we can find such an M, then for any prime p we find r such that $p^r \geq 2M$. Using Hensel's lemma, we factor $f(x)$ modulo p^r, say $f(x) = \prod_{i=1}^{k} g_i(x) \pmod{p^r Z[x]}$. Then $g(x) \equiv \prod_{i=n_1}^{n_t} g_i(x) \pmod{p^r Z[x]}$ for some subset $\{n_1,\ldots,n_t\}$ of $\{1,\ldots,k\}$. Since $g(x)$ is reduced modulo p^r, the problem of factorization is reduced to seeing which products of the irreducible factors of f(x) modulo $p^r Z[x]$, when reduced modulo p^r, actually divide f(x) in Z[x]. Theorem 2.13 provides us with an appropriate M, but first we need a few lemmas. For convenience, let

$$||f|| = \left(\sum_{i=0}^{n} |a_i|^2\right)^{1/2}, \text{ where } f(x) = \sum_{i=0}^{n} a_i x^i \in C[x].$$

Lemma 2.8: Let $f(x) = \sum_{i=0}^{n} a_i x^i \in C[x]$ and $\alpha \in C$ with $\alpha \neq 0$. If $g(x) = (x + \alpha)f(x)$ and $h(x) = (x + \bar{\alpha}^{-1})f(x)$, where $\bar{\alpha}$ denotes the complex conjugate of α, then $||g|| = |\alpha|\,||h||$.

Proof: $g(x) = \sum_{i=0}^{n+1} (a_{i-1} + \alpha a_i)x^i$ and $h(x) = \sum_{i=0}^{n+1} (a_{i-1} + \bar{\alpha}^{-1} a_i)x^i$, where $a_{-1} = a_{n+1} = 0$.

So

$$||g||^2 = \sum_{i=0}^{n+1} |a_{i-1} + \alpha a_i|^2$$

$$= \sum_{i=0}^{n+1} (a_{i-1} + \alpha a_i)\overline{(a_{i-1} + \alpha a_i)}$$

$$= \sum_{i=0}^{n+1} (|a_{i-1}|^2 + \alpha \bar{a}_{i-1} a_i + \bar{\alpha} a_{i-1} \bar{a}_i + |\alpha|^2 |a_i|^2).$$

Also $$|\alpha|^2 ||h||^2 = |\alpha|^2 \sum_{i=0}^{n+1} |a_{i-1} + \bar{\alpha}^{-1} a_i|^2 = \sum_{i=0}^{n+1} |\alpha a_{i-1} + \alpha\bar{\alpha}^{-1} a_i|^2$$

$$= \sum_{i=0}^{n+1} (\alpha a_{i-1} + \alpha\bar{\alpha}^{-1} a_i)\overline{(\alpha a_{i-1} + \alpha\bar{\alpha}^{-1} a_i)}$$

$$= \sum_{i=0}^{n+1} (\alpha a_{i-1} + \alpha\bar{\alpha}^{-1} a_i)(\bar{\alpha}\bar{a}_{i-1} + \bar{\alpha}\alpha^{-1} \bar{a}_i)$$

$$= \sum_{i=0}^{n+1} (|\alpha|^2|a_{i-1}|^2 + \alpha\bar{a}_{i-1}a_i + \bar{\alpha}a_{i-1}\bar{a}_i + |a_i|^2).$$

Each term in this sum is also in the sum for $||g||^2$. So we have $||g||^2 = |\alpha|^2||h||^2$ and $||g|| = |\alpha|\,||h||$. //

Lemma 2.9: Let $\xi_1, \xi_2, \ldots, \xi_n$ be complex numbers such that $0 < |\xi_1| \leq \cdots \leq |\xi_q| < 1 \leq |\xi_{q+1}| \leq \cdots \leq |\xi_n|$ for some $q \geq 0$. Put $f(x) = \prod_{i=1}^{n} (x-\xi_i)$ and $g(x) = [\prod_{i=1}^{q} (x-\bar{\xi}_i^{-1})][\prod_{i=q+1}^{n} (x-\xi_i)]$ then

$$||f|| = (\prod_{i=1}^{q} |\xi_i|)||g||.$$

Proof: We proceed by induction on q. For $q = 0$ we have $f(x) = g(x)$, so the conclusion holds. Now assume $q > 0$, and set

$f^*(x) = \dfrac{f(x)}{x-\xi_1}$ and $g^*(x) = \dfrac{g(x)}{x-\bar{\xi}_1^{-1}}$. Then $||f|| = ||(x-\xi_1)f^*(x)||$

$= |\xi_1|\,||(x-\bar{\xi}_1^{-1})f^*(x)|| = |\xi_1||\xi_2 \cdots \xi_q|\,||(x-\bar{\xi}_1^{-1})g^*(x)||$ using Lemma 2.8, our induction hypothesis and the fact that $|\bar{\xi}^{-1}| = |\xi_1|^{-1} \geq 1$.

Now $(x-\bar{\xi}_1^{-1})g^*(x) = g(x)$ so we have $||f|| = (\prod_{i=1}^{q}|\xi_i|)||g||$. //

Lemma 2.10: Let $f(x) = \sum_{i=0}^{n} a_i x^i = a_n \prod_{i=1}^{n} (x-\xi_i) \in C[x]$ and $|\xi_1| \leq \cdots \leq |\xi_q| < 1 \leq |\xi_{q+1}| \leq \cdots \leq |\xi_n|$ for some $q \geq 0$. Then

$$||f||^2 \geq |a_n|^2|\xi_{q+1} \cdots \xi_n|^2 + |a_o|^2|\xi_{q+1} \cdots \xi_n|^{-2}.$$

<u>Proof</u>: Let $g(x) = a_n \prod_{k=1}^{q} (x-\overline{\xi}_i^{-1}) \prod_{i=q+1}^{n} (x-\xi_i) = \sum_{i=o}^{n} b_i x^i$.

First assume that $\xi_1 \neq 0$, then by Lemma 2.9 $||f|| = |\xi_1 \cdots \xi_q|||g||$.

Hence $||f||^2 = |\xi_1 \cdots \xi_q|(\sum_{i=o}^{n} |b_i|^2) \geq |\xi_1 \cdots \xi_q|^2|b_o|^2$

$$+ |\xi_1 \cdots \xi_q|^2|b_n|^2.$$

Now

$$|b_o| = |a_n(\prod_{i=1}^{q} \overline{\xi}_i^{-1})(\prod_{i=q+1}^{n} \xi_i)|$$

so

$$|\xi_1 \cdots \xi_q|^2|b_o|^2 = |a_n|^2|\xi_{q+1} \cdots \xi_n|^2.$$

Also

$$|a_o| = |a_n \prod_{i=1}^{n} \xi_i| \text{ and } b_n = a_n.$$

Thus

$$|\xi_1 \cdots \xi_q|^2|b_n|^2 = |\xi_1 \cdots \xi_q|^2|a_n|^2$$

$$= |\xi_1 \cdots \xi_q|^2|a_o|^2|\xi_1 \cdots \xi_n|^{-2} = |a_o|^2|\xi_{q+1} \cdots \xi_n|^{-2}.$$

Hence

$$||f||^2 + |a_n|^2|\xi_{q+1} \cdots \xi_n|^2 + |a_n|^2|\xi_{q+1} \cdots \xi_n|^{-2}.$$

Now suppose that $\xi_1 = \xi_2 = \cdots = \xi_m = 0$, while $0 < |\xi_{m+1}|$, with $m \leq q$. Then $a_o = 0$ and we need only show that $||f||^2 \geq |a_n|^2|\xi_{q+1} \cdots \xi_n|^2$. But $||f|| = ||f(x)/x^m||$ and $||f(x)/x^m|| \geq |a_n|^2|\xi_{q+1} \cdots \xi_n|^2 + |a_m|^2|\xi_{q+1} \cdots \xi_n|^{-2}$ by the first part of our proof. Hence $||f|| \geq |a_n|^2|\xi_{q+1} \cdots \xi_n|^2$. //

Corollary 2.11: Let $f(x) = \sum_{i=o}^{n} a_i x^i = a_n \prod_{i=1}^{n} (x-\xi_i) \in C[x]$ with $|\xi_1| \leq \cdots \leq |\xi_q| < 1 \leq |\xi_{q+1}| \leq \cdots \leq |\xi_n|$ for some $q \geq 0$. Then $|a_n| \prod_{i=q+1}^{n} |\xi_i| \leq ||f||$.

Lemma 2.12: Let $f(x) = \sum_{i=o}^{n} a_i x^i = a_n \prod_{i=1}^{n} (x-\xi_i) \in C[x]$ with $|\xi_1| \leq \cdots \leq |\xi_q| < 1 \leq |\xi_{q+1}| \leq \cdots \leq |\xi_n|$ for some $q \geq 0$. Then

$$|a_i| \leq \binom{n}{i} |\xi_{q+1} \cdots \xi_n| |a_n| \text{ for } i = 0,1,\ldots,n$$

and

$$\sum_{i=o}^{n} |a_i| \leq 2^n |\xi_{q+1} \cdots \xi_n| |a_n|.$$

Proof: Let $\sigma,\tau \in S_n$ and say $\sigma \equiv_j \tau$ provided $\{\sigma(i): i=1,\ldots,j\} = \{\tau(i): i=1,\ldots,j\}$. This defines an equivalence relation on S_n. Put $S_{n,j}$ equal to the set of equivalence classes with respect to this equivalence relation. Note that $S_{n,j}$ has $\binom{n}{j}$ elements and $a_j = (-1)^{n-j} a_n \sum_{\sigma \in S_{n,j}} \prod_{i=1}^{j} \xi_{\sigma(i)}$. Then $|a_j| = |a_n| |\sum_{\sigma \in S_{n,j}} \prod_{i=1}^{j} \xi_{\sigma(i)}|$

$$\leq |a_n| | \sum_{\sigma \in S_{n,j}} \prod_{i=q+1}^{n} \xi_i| = |a_n| \binom{n}{j} |\xi_{q+1} \cdots \xi_n|.$$

Also

$$\sum_{i=o}^{n} |a_i| \leq \sum_{i=o}^{n} \binom{n}{i} |\xi_{q+1} \cdots \xi_n| |a_n| = |a_n| |\xi_{q+1} \cdots \xi_n| \sum_{i=o}^{n} \binom{n}{i}$$

$$= 2^n |\xi_{q+1} \cdots \xi_n| |a_n|. \quad //$$

Theorem 2.13: Let $f(x) \in Z[x]$ with $f(x) = g_1(x)\cdots g_k(x)$, where each $g_i(x) \in Z[x]$. If $[f] = n$ and $g_i(x) = \sum_{j=o}^{m_i} b_{i,j}x^i$, then

$$\prod_{i=1}^{k} \Big(\sum_{j=o}^{m_i} |b_{i,j}| \Big) \leq 2^n ||f||$$

and

$$|b_{i,j}| \leq \binom{m_i}{j} ||f|| \text{ for each } i \text{ and } j.$$

Proof: If $\xi_{i,1} \ldots, \xi_{i,m_i}$ are the roots of $g_i(x)$ for $i = 1,\ldots,k$, then $g_i(x) = b_{i,m_i} \prod_{j=1}^{m_i} (x-\xi_{i,j})$. Suppose $|\xi_{i,1}| \leq \cdots \leq |\xi_{i,q_i}| < 1 \leq |\xi_{i,q_i+1}| \leq \cdots \leq |\xi_{i,m_i}|$ with $q_i \geq 0$, then by Lemma 2.22

$$\sum_{j=o}^{m_i} |b_{i,j}| \leq 2^{m_i} |\xi_{i,q_i+1} \cdots \xi_{i,m_i}| |b_{i,m_i}|.$$

Now if

$$a_n = \prod_{i=1}^{k} b_{i,m_i}, \text{ then } f(x) = a_n \prod_{i=1}^{k} \prod_{j=1}^{m_i} (x-\xi_{i,j}).$$

So

$$\prod_{i=1}^{k} \sum_{j=o}^{m_i} |b_{i,j}| \leq \prod_{i=1}^{k} 2^{m_i} |\xi_{i,q_i+1} \cdots \xi_{i,m_i}| |b_{i,m_i}| = 2^{\sum_{i=1}^{k} m_i} \prod_{i=1}^{k} |b_{i,m_i}| \prod_{j=q_i+1}^{m_i} |\xi_{i,j}|$$

$$= 2^n |a_n| \prod_{i=1}^{k} \prod_{j=q_i+1}^{m_i} |\xi_{i,j}| \leq 2^n ||f|| \text{ by Corollary 2.11.}$$

To prove the second inequality we use Lemma 2.12 and Corollary 2.11 to get

$$|b_{i,j}| \leq \binom{m_i}{j} |\xi_{i,q_i+1} \cdots \xi_{i,m_i}| |b_{i,m_i}| \leq \binom{m_i}{j} |a_n| \prod_{i=1}^{k} \prod_{j=q_i+1}^{m_i}$$

$|\xi_{i,j}| \leq \binom{m_i}{j}||f||$. //

If we let k be the greatest integer in n/2, then we have $|b_{i,j}| \leq \binom{n}{k}||f||$ for each i,j. Therefore a suitable constant would be $M = \binom{n}{k}||f||$. Zassenhaus (1969) uses as a bound on the roots of f(x) the number $\phi f = \{\max_{1 \leq i \leq n} [|a_i|/\binom{n}{i}]^{1/i}\}/(\sqrt[n]{2} - 1)$, and notes that $|b_{i,j}| \leq \binom{m_i}{j}(\phi f)^j$. Another bound for the roots of f(x) is $A = \max_{1 \leq i \leq n} |a_i| + 1$ (Mignotte, 1974). Mignotte, however, claims that the bound provided by Theorem 2.13 is the best in general.

For completeness, I now include Kronecker's method of factorization, as it was the first finite method. Let $f(x) \varepsilon Z[x]$ and suppose [f] = n. If f(x) has a factor g(x) in Z[x], say f(x) = g(x)h(x), then either $[g] \leq \frac{n}{2}$ or $[h] \leq \frac{n}{2}$. So to test for divisors of f(x), we need only check for polynomials of degree less than or equal to $\frac{n}{2}$. Let m be the greatest integer in $\frac{n}{2}$, and pick $\alpha_o, \alpha_1, \ldots, \alpha_m$, distinct elements of Z. Calculate $f(\alpha_i)$ for each i. If $g(x)|f(x)$, then $g(\alpha_i)|f(\alpha_i)$ for each i. Pick one set of integers $b_o, b_1, \ldots, b_m$ such that $b_i|f(\alpha_i)$ for each i. Let

$$g(x) = \sum_{i=o}^{m} \frac{b_i(x-\alpha_o)(x-\alpha_1)\cdots(x-\alpha_{i-1})(x-\alpha_{i+1})\cdots(x-\alpha_m)}{(\alpha_i-\alpha_o)\cdots(\alpha_i-\alpha_{i-1})(\alpha_i-\alpha_{i+1})\cdots(\alpha_i-\alpha_m)},$$

then $g(\alpha_i)|f(\alpha_i)$ for each i since $g(\alpha_i) = b_i$. Also g(x) is the only polynomial of degree less than or equal to m such that $g(\alpha_i) = b_i$ for each i. (If there was another, then their difference would have m+1 distinct roots which cannot be.) For each set $b_o, b_1, \ldots, b_m$, there is a unique polynomial g(x). Hence a divisor of f(x) must be selected from one of these. Since there are only a finite number of choices for $b_o,\ldots,b_m$, this is a finite method of factorization.

CHAPTER III

CALCULATION OF THE GALOIS GROUP

A. Early Methods

The purpose of this section is to present two of the first methods of calculating the Galois group of a polynomial over the rational numbers. The first method involves the calculation of a polynomial called the Galois resolvent. We will let ι denote the identity permutation.

Let $f(x) \in Z[x]$ with $f(x) = a \prod_{i=1}^{n} (x-\xi_i)$, where each $\xi_i \in C$ and $\xi_i \neq \xi_j$ if $i \neq j$. Let $G(x_1,\ldots,x_n) = \sum_{i=1}^{n} \xi_i x_i$. For each $\sigma \in S_n$, put $G_\sigma(x_1,\ldots,x_n) = \sum_{i=1}^{n} \xi_{\sigma(i)} x_i$. We pick $c_1,\ldots,c_n \in Z$ such that $G_\sigma(c_1,\ldots,c_n) \neq G_\rho(c_1,\ldots,c_n)$ if $\sigma \neq \rho$. Put $t_\sigma = G_\sigma(c_1,\ldots,c_n)$ for each $\sigma \in S_n$, and $F(x) = \prod_{\sigma \in S_n} (x-t_\sigma)$. By the theory of symmetric functions, $F(x) \in Z[x]$. Factor $F(x)$ in $Z[x]$ so that $F(x) = F_1(x) \cdots F_r(x)$, where $F_1(t_\iota) = 0$. Note that if $F(x)$ is irreducible, then $F_1(x) = F(x)$. $F_1(x)$ is called the Galois resolvent of $f(x)$. Write $F_1(x) = \prod_{\sigma \in G} (x-t_\sigma)$, where G is the set of permutations from S_n from which $F_1(x)$ is derived. We will see that G is actually the Galois group of $f(x)$.

Lemma 3.1: Each root of $f(x)$ can be written as a polynomial in t_ι.

Proof: For this discussion recall that we may think of elements of S_n as either a permutation of the numbers 1,2,...,n or as a permutation of the roots of f(x). So, with this convention, $\sigma(\xi_i)$ and $\xi_{\sigma(i)}$ will have the same meaning. This notion can be extended to all rational functions of the ξ_i. We will identify $\sigma(h(\xi_1,\ldots,\xi_n))$ with $h(\xi_{\sigma(1)},\ldots,\xi_{\sigma(n)})$, where h is a rational function over the rational numbers.

Let ξ_1 be any root of f(x), and let $\sigma_1,\ldots,\sigma_m$ be the permutations in S_n such that $\sigma_i(1) = 1$ for each i. Put $H(t) = \prod_{i=1}^{m}(t-t_{\sigma_i}) = \sum_{i=0}^{m}\alpha_i t^i$. Each $\alpha_i \in Q(\xi_1)$ since they are the elementary symmetric functions of the roots of $\frac{f(x)}{x-\xi_1} \in (Q(\xi_1))[x]$.

Thus each α_i can be expressed as a polynomial in ξ_1 with coefficients in Q, say $\alpha_i = p_i(\xi_1)$ where $p_i(x) \in Q[x]$. Now $H(t) = \sum_{i=0}^{m} p_i(\xi_1)t^i$. Let $S(x) = \sum_{i=0}^{m} p_i(x)t_\iota^i$; then $S(\xi_1) = H(t_\iota) = 0$.

If $j > 1$, we have $S(\xi_j) \neq 0$. To see this, let $\rho_i = (1\ j)\sigma_i$ and compute $\sigma(\alpha_i) = \sigma(p_i(\xi_1)) = p_i(\xi_{\sigma(1)}) = p_i(\xi_j)$ where $\sigma = (1\ j)$. Also $\alpha_i = \sum_{k=0}^{i} t_{\sigma_k}t_{\sigma_{i-k}}$ and so $\sigma(\alpha_i) = \sum_{k=0}^{i} t_{\sigma\sigma_k}t_{\sigma\sigma_{i-k}} = \sum_{k=0}^{i} t_{\rho_k}t_{\rho_{i-k}}$.

Let $H_1(t) = \prod_{i=1}^{m}(t-t_{\rho_i}) = \sum_{i=0}^{m}\beta_i t^i$, where $\beta_i = \sum_{k=0}^{i} t_{\rho_k}t_{\rho_{i-k}} = \sigma(\alpha_i)$ $= p_i(\xi_j)$. Finally $S(\xi_j) = H_1(t_\iota) \neq 0$ since $\iota \notin \{\rho_1,\ldots,\rho_m\}$. For if $\iota = \rho_i$, then $\iota = (1\ j)\sigma_i$ and $\sigma_i = (1\ j)$ which does not hold 1 fixed.

Both $f(x)$ and $S(x)$ have their coefficients in $Q(t_\iota)$, so we can find the greatest common divisor of $f(x)$ and $S(x)$ in $(Q(t_\iota))[x]$. If they are relatively prime, then there are polynomials $h(x), g(x) \in (Q(t_\iota))[x]$ such that $1 = h(x)f(x) + g(x)S(x)$. But then $1 = h(\xi_1) f(\xi_1) + g(\xi_1)S(\xi_1) = 0$. Thus $f(x)$ and $S(x)$ cannot be relatively prime. Because they share only one root, the greatest common divisor must be $x - \xi_1$. This implies that $\xi_1 \in Q(t_\iota)$, and hence a polynomial in t_ι. //

<u>Corollary 3.2</u>: If t_σ is a root of $F(x)$, then t_σ is a rational function of t_ι.

<u>Proof</u>: Each t_σ is a rational function of the ξ_i, and each ξ_i is a polynomial in t_ι. Therefore each t_σ can be expressed as a rational function of t_ι. //

<u>Theorem 3.3</u>: $G = G(f,Q)$.

Let K be the splitting field of $f(x)$ over Q and suppose $\sigma \in G$. If $u \in Q$, then let $\sigma(u) = u$. If $u \in K-Q$, then $u = g(\xi_1, \ldots, \xi_n)$, where $g(x_1, \ldots, x_n) \in Q(x_1, \ldots, x_n)$, and we will put $\sigma(u) = g(\xi_{\sigma(1)}, \ldots, \xi_{\sigma(n)})$.

First we must show that G is a group. Let $\sigma, \rho \in G$. By Corollary 3.2 $t_\sigma = g(t_\iota)$, where $g(x) \in Q(x)$. Now $\rho(t_\sigma) = t_{\rho\sigma} = g(t_\rho)$. Let $H(x) = F_1(g(x))$; then $H(t_\iota) = 0$. But $F_1(x)$ is irreducible, so $F_1 | H$ and $H(t_\rho) = 0$. $H(t_\rho) = F_1(g(t_\rho)) = F_1(t_{\rho\sigma})$, thus $\rho\sigma \in G$ and G is a group.

We now show that the fixed field of G is Q. Let $u \in K$ such that $\sigma(u) = u$ for all $\sigma \in G$. Since $u \in K$, there is $h(x_1,\ldots,x_n) \in Q(x_1,\ldots,x_n)$ with $u = h(\xi_1,\ldots,\xi_n)$. We use Lemma 3.1 to find $g_i(x) \in Q(x)$ such that $\xi_i = g_i(t_\iota)$ for $i = 1,2,\ldots,n$. Then $u = h(g_1(t_\iota),\ldots,g_n(t_\iota)) = T(t_\iota)$ for some $T(x) \in Q(x)$. Now $u = \sigma(u) = T(t_{\sigma\iota}) = T(t_\sigma)$ for each $\sigma \in G$. Let $G = \{\sigma_1,\ldots,\sigma_m\}$. Then $u = \frac{1}{m}[T(t_{\sigma_1}) + T(t_{\sigma_2}) + \cdots + T(t_{\sigma_m})]$ which is a symmetric function of the roots of $F_1(x)$ and hence in Q. Therefore $u \in Q$ and Q is the fixed field of G. By Theorem 1.3 $|G| = [K:Q] = |G(f,Q)|$. Also $G \subseteq G(f,Q)$ so that $G = G(f,Q)$. //

Using the Galois resolvent is a method of finding the Galois group of an equation over the rational numbers in a finite number of steps, but the calculations required are formidable if n is large. First we need to find the roots of the equation. Next, to find $c_1,\ldots,c_n$, we must check $\frac{n!(n!-1)}{2}$ equations of the form $G_\sigma(x_1,\ldots,x_n) = G_\rho(x_1,\ldots,x_n)$. Finally we must factor $F(x)$, where the degree of $F(x)$ is $n!$. In general this is unreasonable for $n > 4$.

Theorem 1.11 gives us an alternative to using the Galois resolvent. Let $g(\xi_1,\ldots,\xi_n)$ be a rational function of the roots of $f(x)$ with coefficients in Q, where $f(x) \in Q[x]$. Let $u = g(\xi_1,\ldots,\xi_n)$. If $u \in Q$, then $\sigma(u) = u$ for each $\sigma \in G(f,Q)$. So to see that a permutation σ is not in $G(f,Q)$, we only need to find a $g(\xi_1,\ldots,\xi_n) \in Q$ such that $g(\xi_{\sigma(1)},\ldots,\xi_{\sigma(n)}) \notin Q$. Also if $g(\xi_1,\ldots,\xi_n) \notin Q$, then at least one of the permutations σ such that $g(\xi_{\sigma(1)},\ldots,\xi_{\sigma(n)}) \neq$

$g(\xi_1,\ldots,\xi_n)$ is in $G(f,Q)$.

As an example, consider the polynomial $f(x) = x^4 + x^2 - 6$. $f(x)$ has roots $\xi_1 = \sqrt{2}$, $\xi_2 = -\sqrt{2}$, $\xi_3 = \sqrt{3}$ and $\xi_4 = -\sqrt{3}$. $\xi_1 + \xi_2 = 0$ and $\xi_3 + \xi_4 = 0$, but $\xi_1 + \xi_3 \neq 0$, $\xi_1 + \xi_4 \neq 0$, $\xi_2 + \xi_3 \neq 0$ and $\xi_2 + \xi_4 \neq 0$. Since $\xi_1 + \xi_2 = 0$, while $\xi_1 + \xi_3 \neq 0$, the permutation (2 3) cannot be in $G(f,Q)$. Similar observations with the remaining relations eliminate all of the permutations of S_4 except ι, (1 2), (3 4), (1 2)(3 4), (1 3 2 4) and (1 4 3 2). Also if K is the splitting field of $f(x)$ over Q, then $K = Q(\sqrt{2},\sqrt{3}) = Q(\sqrt{2} + \sqrt{3})$ and $\sqrt{2} + \sqrt{3}$ has x^4-10x^2-35 for a minimal polynomial over Q. Hence $[K:Q] = 4$ and $G(f,Q) = \{\iota,(1\ 2),(3\ 4),(1\ 2)(3\ 4)\}$ or $G(f,Q) = \{\iota,(1\ 3\ 2\ 4),(1\ 4\ 2\ 3),(1\ 2)(3\ 4)\}$. These are isomorphic, and so we have calculated $G(f,Q)$.

B. Method of Zassenhaus

Another method, which is a bit more practical, involves calculating $G(f,Q)$ by finding the subgroups of S_n which contain $G(f,Q)$.

<u>Definition</u>: Let $F(x_1,\ldots,x_n) \in Z[x_1,\ldots,x_n]$ and $G \subseteq S_n$. F <u>belongs to G</u> provided $F(x_1,\ldots,x_n) = F(x_{\sigma(1)},\ldots,x_{\sigma(1)})$ if and only if $\sigma \in G$.

<u>Theorem 3.4</u>: If $G \subseteq S_n$, then there is $F(x_1,\ldots,x_n) \in Z[x_1,\ldots,x_n]$ such that F belongs to G.

<u>Proof</u>: Let $H(x_1,\ldots,x_n) = x_1x_2^2 \cdots x_n^n$ and $F(x_1,\ldots,x_n) = \sum_{\sigma\in G} H(x_{\sigma(1)},\ldots,x_{\sigma(n)})$. If $\rho \in G$, then $F(x_{\rho(1)},\ldots,x_{\rho(n)}) = \sum_{\sigma\in G} H(x_{\rho\sigma(1)},\ldots,x_{\rho\sigma(n)}) = \sum_{\sigma\in G} H(x_{\sigma(1)},\ldots,x_{\sigma(n)}) = F(x_1,\ldots,x_n)$. If

$\rho \notin G$, then the sum $\sum_{\sigma\varepsilon G} H(x_{\rho\sigma(1)},\ldots,x_{\rho\sigma(n)})$ contains the term $H(x_{\rho(1)},\ldots,x_{\rho(n)})$ since the identity is in G. But this term is not in the original sum because $\rho \notin G$. Hence $F(x_{\rho(1)},\ldots,x_{\rho(n)}) \neq F(x_1,\ldots,x_n)$. //

Definition: Let G,H be subgroups of S_n. $F(x_1,\ldots,x_n) \varepsilon Z[x_1,\ldots,x_n]$ belongs to G in H provided for every $\sigma \varepsilon H$, $F(x_1,\ldots,x_n) = F(x_{\sigma(1)},\ldots,x_{\sigma(n)})$ if and only if $\sigma \varepsilon G$.

Definition: Let G,H be subgroups of S_n, and suppose $F(x_1,\ldots,x_n)$ belongs to G in H. If $G \subseteq H$ and $\sigma_1,\ldots,\sigma_m$ is a representative set for the right cosets of G in H, then

$R(x) = \prod_{i=1}^{m} [x-F(x_{\sigma_i(1)},\ldots,x_{\sigma_i(n)})]$ is the resolvent polynomial of G in H corresponding to F. If $f(x) \varepsilon Z[x]$ and $f(x) = a \prod_{i=1}^{n} (x-\xi_i)$, then the resolvent polynomials of G in H corresponding to F for f(x) is $R(x) = \prod_{i=1}^{m} [x-F(\xi_{\sigma_i(1)},\ldots,\xi_{\sigma_i(n)})]$.

Theorem 3.5: Let $f(x) = a \prod_{i=1}^{n} (x-\xi_i) \varepsilon Z[x]$ be irreducible over Z and H a transitive subgroup of S_n. Suppose also that G is a subgroup of H and $F(x_1,\ldots,x_n)$ is a polynomial in n variables which belongs to G in H, with $F(\xi_1,\ldots,\xi_n)$ not a repeated root of the resolvent polynomial of G in H corresponding to F for f(x). Then $G(f,Q) \subseteq G$ if and only if $F(\xi_1,\ldots,\xi_n) \varepsilon Z$.

Proof: First we note that $R(x) = \prod_{i=1}^{m} [x-F(\xi_{\sigma_i(1)},\ldots,\xi_{\sigma_i(n)})]$ ε Z[x]. The coefficients are products and sums of the $F(\xi_{\sigma_i(1)},\ldots,\xi_{\sigma_i(n)})$, which are products and sums of $\xi_1,\ldots,\xi_n$, which are algebraic integers. To see that the coefficients are in Q, let $\sigma \varepsilon G(f,Q)$. Then $\sigma\sigma_1,\ldots,\sigma\sigma_m$ forms a representative set for the right cosets of G in H. Hence the coefficients of R(x) are left fixed by the elements of G(f,Q) and $R(x) \varepsilon Q[x]$. Thus the coefficients of R(x) are both algebraic integers and rational numbers, and hence they are rational integers.

Now suppose that $G(f,Q) \subseteq G$. Then for each $\sigma \varepsilon G(f,Q)$, $\sigma(F(\xi_1,\ldots,\xi_n)) = F(\xi_{\sigma(1)},\ldots,\xi_{\sigma(n)}) = F(\xi_1,\ldots,\xi_n)$ since $\sigma \varepsilon G$. So $F(\xi_1,\ldots,\xi_n) \varepsilon Q$ because Q is the fixed field of G(f,Q). But $F(\xi_1,\ldots,\xi_n)$ is an algebraic integer and so a rational integer.

Finally, let $F(\xi_1,\ldots,\xi_n) \varepsilon Q$. Then $F(\xi_{\sigma(1)},\ldots,\xi_{\sigma(n)}) = F(\xi_1,\ldots,\xi_n)$ for each $\sigma \varepsilon G(f,Q)$. This implies that $\sigma \varepsilon G$ since $F(\xi_1,\ldots,\xi_n)$ is not a repeated root of R(x). Hence $G(f,Q) \subseteq G$. //

Corollary 3.6: Let $f(x) = a \prod_{i=1}^{n} (x-\xi_i) \varepsilon Z[x]$ be irreducible over Z and H a transitive subgroup of S_n. Suppose also that G is a subgroup of H and $F(x_1,\ldots,x_n)$ is a polynomial in n variables over the integers belonging to G in H. If $R(x) = \prod_{i=1}^{m} [x-F(\xi_{\sigma_i(1)},\ldots,\xi_{\sigma_i(n)})]$ is the resolvent polynomial of G in H corresponding to F for f(x), then $G(f,Q) \subseteq G$ (for some arrangement of the roots of f(x)) if and only if $F(\xi_{\sigma_i(1)},\ldots,\xi_{\sigma_i(n)}) \varepsilon Z$ for some i, provided

$F(\xi_{\sigma_i(1)},\ldots,\xi_{\sigma_i(n)})$ is not a repeated root of $R(x)$.

Proof: If $G(f,Q) \subseteq G$, then Theorem 3.5 says that $F(\xi_1,\ldots,\xi_n) \in Z$, provided $F(\xi_1,\ldots,\xi_n)$ is not a repeated root of $R(x)$. Now suppose that $F(\xi_{\sigma_i(1)},\ldots,\xi_{\sigma_i(n)}) \in Z$ is not a repeated root of $R(x)$. Then for each $\sigma \in G(f,Q)$, $F(\xi_{\sigma\sigma_i(1)},\ldots,\xi_{\sigma\sigma_i(n)}) = F(\xi_{\sigma_i(1)},\ldots,\xi_{\sigma_i(n)})$. So $\sigma_i F$ belongs to $\sigma_i G \sigma_i^{-1}$ in H and $G(f,Q) \subseteq \sigma_i G \sigma_i^{-1}$. If we reorder the roots of $f(x)$ so that $\alpha_j = \xi_{\sigma_i(j)}$ for $j = 1,\ldots,n$, then

$f(x) = a \prod_{i=1}^{n} (x-\alpha_i)$ and $G(f,Q) \subseteq G$. //

Definition: Let $f(x) = \prod_{i=1}^{n} (x-\xi_i)$, then the number

$D(f) = \prod_{i<j} (\xi_i-\xi_j)^2$ is called the discriminant of f(x).

An important consequence of Theorem 3.5 is the following:

Theorem 3.7: Let $f(x) \in Z[x]$. Then $G(f,Q) \subseteq A_n$ if and only if $\sqrt{D(f)} \in Z$, where A_n is the alternating group of degree n.

Proof: Let $F(x_1,\ldots,x_n) = \prod_{i<j} (x_i-x_j)$; then F belongs to A_n in S_n. For if σ is a transposition in S_n, say $\sigma = (k\ m)$ where $k < m$, then $F(x_{\sigma(1)},\ldots,x_{\sigma(n)}) = \prod_{i<j} (x_{\sigma(i)}-x_{\sigma(j)}) = (x_{\sigma(k)}-x_{\sigma(m)})$

$$\prod_{\substack{i<j \\ (i,j)\neq(k,m)}} (x_i-x_j) = (x_m-x_k) \prod_{\substack{i<j \\ (i,j)\neq(k,m)}} (x_i-x_j) = -\prod_{i<j} (x_i-x_j).$$

Thus if $\rho \in S_n$, then $F(x_{\rho(1)},\ldots,x_{\rho(n)}) = F(x_1,\ldots,x_n)$ if and only if ρ can be written as an even number of permutations.

Now let $R(x) = [x-F(\xi_1,\ldots,\xi_n)][x-F(\xi_{\sigma(1)},\ldots,\xi_{\sigma(n)})]$ where $\sigma \notin A_n$. By Theorem 3.5, $G(f,Q) \subseteq A_n$ if and only if $F(\xi_1,\ldots,\xi_n) \in Z$, that is $\sqrt{D(f)} \in Z$. //

Corollary 3.6 and Theorem 3.7 give us an important method of calculating the Galois group of an irreducible polynomial, and this method is a definite improvement on the use of the Galois resolvent. Here is a summary of the method. First we calculate the roots and discriminant of f(x). Next we find a maximal transitive subgroup H of S_n, where n = [f]. Theorem 3.4 guarantees that we can find a function F which belongs to H in S_n. Actually, F can be constructed so that the resolvent R(x) has no repeated roots. We test R(x) for integer roots. If there are none, then we find a new maximal subgroup to work with. If no maximal transitive subgroup has a resolvent with an integer root, then $G(f,Q) = S_n$. Now suppose that the resolvent computed for H has an integer root. If $\sigma(F)$ is that root, then we rearrange the roots of f(x) by letting $\sigma_i = \xi_{\sigma(i)}$. According to Corollary 3.6, with this root arrangement, we must have $G(f,Q) \subseteq H$. Next we find a maximal transitive subgroup H_1 of H and a function F_1 belonging to H_1 in H. We test to see if $G(f,Q) \subseteq H_1$. This process is terminated when either we reach a minimal transitive subgroup of S_n (which then must be G(f,Q)), or we have $G(f,Q) \subseteq H_k$ and there is no maximal transitive subgroup H_{k+1} of H_k such that $G(f,Q) \subseteq H_{k+1}$. In this case $G(f,Q) = H_k$. Of course the method is accelerated by knowledge of the discriminant of f(x). If D(f) is a perfect square, then by Theorem 3.7 we need only search in A_n for G(f,Q). If not, then we may omit from our technique all subgroups of A_n. The main difficulties of this method come from the need to know, with a great deal of accuracy, what the roots of f(x) are; the fact that we must somehow come up with all of the transitive subgroups in S_n; and the calculation of suitable functions F. The latter two problems have been solved in part by Stauduhar (1973) who has produced tables for this purpose. (See appendix.) Figure 1 indicates the order in which we select our maximal transitive subgroups, while Table 1 describes the groups listed in Figure 1 and exhibits an appropriate function F. For that function F, we use the right coset representatives listed in Table 2. If the function given in Table 1 gives rise to repeated roots in the integers, then we can use Table 2 to construct our own resolvent. We must find, on our own, a function belonging to G in H, then Table 2 gives us the right coset representatives which we use to calculate the resolvent. Zassenhaus (1971) also suggests a particular function that we may use. If $G \subseteq H$ and K is the splitting field of f(x), then set $tr_G(\alpha) = \sum_{\sigma\varepsilon G} (\alpha)$. If $\alpha = h(\xi_1,\ldots,\xi_n)$, where the ξ_i are the roots of f(x), then by $\sigma(\alpha)$ we mean $h(\xi_{\sigma(1)},\ldots,\xi_{\sigma(n)})$. By the selection of a suitable α, $tr_G(\alpha)$ belongs to G in H. Observe that if $\alpha = \xi_1\xi_2^2 \cdots \xi_n^n$, then $tr_G(\alpha)$ yields the same function as given in Theorem 3.4.

Notice that in Corollary 3.6 we must have that f(x) is irreducible. However, we may still apply this method to any polynomial over the rationals by using Corollary 1.18. We factor the given polynomial over the integers and apply the method to each factor.

Then the Galois group must be a subgroup of the product of the groups of the factors. The following example of the method of Corollary 3.6 is due to Stauduhar (1973).

Let $f(x) = x^6-42x^4+80x^3+441x^2-1680x+4516$. The roots of f(x) are $\xi_1 = 4.392-1.570i$, $\xi_2 = \overline{\xi}_1$, $\xi_3 = -5.490-.780i$, $\xi_4 = \overline{\xi}_3$, $\xi_5 = 1.098-2.355i$ and $\xi_6 = \overline{\xi}_5$ and f(x) is irreducible over the integers. Also a routine calculation shows that $D(f) < 0$ and hence not a perfect square.

We now refer to Figure 1 to see that a maximal transitive subgroup of S_6 is G_{72}. (The subscript denotes the order of the group.) Table 2 gives as right coset representatives ι, (2 5 4 3),(2 3 6)(4 5), (2 5 4 3 6),(2 5)(3 4),(2 4 5 3),(2 5),(2 3 4 5),(2 4 5 3 6) and (3 6 4 5); and Table 1 suggests the use of $F_1(x_1,\ldots,x_6) = x_1x_2x_3 + x_4x_5x_6$. We use this information to calculate $R_1(x) = x^{10} + 80x^9 - 59166x^8 - 4390320x^7 + 1200615393x^6 + 88076918880x^5 - 7198940057856x^4 - 388801984512000x^3 + 20193311991398400x^2 + 595967000182784000x - 4689149328097280000$. $R_1(x)$ has a root -80 corresponding to the coset representative (2 3 6)(4 5). By letting $\alpha_1 = \xi_1, \alpha_2 = \xi_3$, $\alpha_3 = \xi_6$, $\alpha_4 = \xi_5$, $\alpha_5 = \xi_4$ and $\alpha_6 = \xi_2$ we have, according to Corollary 3.6, $G(f,Q) \subseteq G_{72}$.

Figure 1 now implies that we should use either G^2_{36} or G^1_{36}. But $G^1_{36} \subseteq A_6$, so by Theorem 3.7 $G(f,Q) \not\subseteq G^1_{36}$. Table 2 yields right coset representatives ι and (5 6) for G^2_{36} in G_{72}. Table 1 gives us the function $F_2(x_1,\ldots,x_6) = (x_1-x_2)(x_2-x_3)(x_3-x_1)(x_4-x_5)(x_5-x_6)(x_6-x_4)$, and the resolvent is $R_2(x) = (x+137376)(x-137376)$. Thus $G(f,Q) \subseteq G^2_{36}$.

Now G_{36}^2 has two maximal transitive subgroups. To see if $G(f,Q) \subseteq G_{18}$, we note that G_{36}^2 contains two isomorphic copies of G_{18} which are conjugate in G_{72}, but not in G_{36}^2. So either we test both of these, or we test one as a subgroup of G_{72}. For the latter choice we use the coset representatives ι, (1 2)(4 5),(5 6),(1 2)(4 6 5) given in Table 2 and compute $R_3(x) = (x + 360i)(x - 360i)(x + 648)(x - 648)$, so that $R_3(x)$ has a root corresponding to the coset representative (5 6). If we let $\beta_1 = \alpha_1$, $\beta_2 = \alpha_2$, $\beta_3 = \alpha_3$, $\beta_4 = \alpha_4$, $\beta_5 = \alpha_6$ and $\beta_6 = \alpha_5$ then $G(f,Q) \subseteq G_{18}$.

G_{18} has two maximal transitive subgroups G_6^1 and G_6^2. For G_6^1, we use the right coset representatives ι, (1 2 3) and (1 3 2), and the function $F_4(x_1,\ldots,x_6) = x_1x_4 + x_2x_6 + x_3x_5$. Then $R_4(x) = x^3 - 1323x + 7722 = (x - 33)(x - 6)(x + 39)$. Hence $G(f,Q) \subseteq G_6^1$, and since G_6^1 is a minimal transitive subgroup, we have that $G(f,Q) = G_6^1 = \{\iota,(1\ 2\ 3)(4\ 6\ 5),(1\ 3\ 2)(4\ 5\ 6),(1\ 4)(2\ 5)(3\ 6),(1\ 5)(2\ 6)(3\ 4),(1\ 6)(2\ 4)(3\ 5)\} \cong S_3$.

CHAPTER IV

CHEBOTAREV-VAN DER WAERDEN METHOD

A. The Chebotarev Density Theorem

Definition: A finite field containing p^m elements, where p is a rational prime and m is a positive integer, is called a Galois field. It is denoted by $GF(p^m)$.

It is a well known fact from the theory of fields that every finite field is a Galois field. We will use Z_p to denote the field of integers modulo p.

Theorem 4.1: The Galois group of $GF(p^{mn})$ over $GF(p^n)$ is a cyclic group. The automorphism σ defined by $\sigma(a) = a^{p^n}$ generates this group.

Proof: $GF(p^{mn})$ has characteristic p so that $(a+b)^{p^n}=a^{p^n}+b^{p^n}$ and $(ab)^{p^n}=a^{p^n}b^{p^n}$. Also σ is 1-1 since if $a^{p^n}=b^{p^n}$, then $0 = a^{p^n}-b^{p^n}=(a-b)^{p^n}$. So a-b = 0 and a=b. Due to the fact that $GF(p^{mn})$ is finite, it must be that σ is onto. If $a \in GF(p^n)$, then $a^{p^n}=a$ so σ fixes $GF(p^n)$. Hence σ is in the Galois group of $GF(p^{mn})$ over $GF(p^n)$. Now $\sigma,\sigma^2,\ldots,\sigma^m$ are all distinct since if $0 \leq j < i \leq m$ and $\sigma^i(a) = \sigma^j(a)$, then $a^{p^{ni}}=a^{p^{nj}}$. So $a^{p^{nj}}(a^{p^{ni}-p^{nj}}-1)=0$ and either a=0 or a has degree $p^{ni}-p^{nj}$. For each i and j we can find a nonzero element b in $GF(p^{mn})$ whose degree is not $p^{ni}-p^{nj}$. Then $\sigma^i(b)\neq\sigma^j(b)$ and hence $\sigma,\sigma^2,\ldots,\sigma^m$ are distinct elements of $G(GF(p^{mn}):GF(p^n))$.

Now $[GF(p^{mn}):GF(p^n)] = m$, so that, by Theorem 1.2, the Galois group can have at most m elements. Therefore $G(GF(p^{mn}):GF(p^n)) = \{\sigma,\sigma^2,\ldots,\sigma^m\}$. //

Definition: Let p be a prime in a finite extension F of Q, and suppose $p = P_1 \cdots P_k$ is the factorization of p into primes in the finite extension K of F. If the P_i are distinct, then p is unramified in K.

For the remainder of this section, p, P and B will represent unramified primes and K,F will be finite extensions of Q. Observe that if $p \in F$ and $P \in K$ with K a finite extension of F, then I_F/p can be considered as a subfield of I_K/P when $P|p$. Define h from I_F/p to I_K/P by $h(a+p) = a + P$, where $a \in I_F$; then h is a 1-1 mapping. Indeed if $h(a+p) = h(b+p)$, then $a + P = b + P$. Hence $P|(a-b)$ and $p|(a-b)$ because each prime of K divides only one prime of F.

Definition: Let K be a finite extension of F,$p \in F$ and $P \in K$ with $P|p$. The relative degree of P over p is the number $f(P/p) = [I_K/P:I_F/p]$. If F = Q, then we say that P has relative degree $f(P/p)$ over Q.

Definition: If K is a field and U is an ideal in I_K, then the norm of U, $N_K(U)$, is the number of elements in I_K/U.

It can be shown that the following properties of N_K hold (Pollard and Diamond, 1975):

(1) $N_K(U_1U_2) = N_K(U_1)N_K(U_2)$,

(2) If K is a finite extension of F and $a \varepsilon F$, then $N_K(a) = N_F(a)^{[K:F]}$,

(3) If K is a finite extension of F and $a \varepsilon K$, then

$N_K(a) = \prod_{\sigma \varepsilon G(K:F)} \sigma(a)$,

(4) If K is a finite extension of F, $p \varepsilon F$ and $p = P_1 \cdots P_k$ in K, then $N_K(P_i) = N_K(P_j)$ and $N_K(p) = N_K(P_i)^k$.

Definition: Let K be a finite extension of F and $P \varepsilon K$. The decomposition group of P is $G_P = \{\sigma \varepsilon G(K:F): \sigma(P) = P\}$.

Theorem 4.2: Let K be a normal extension of F, $p \varepsilon F$ and $P \varepsilon K$ with $P|p$. Suppose also that L is the fixed field of G_P and $B \varepsilon L$ with $B|P$. Define the mapping h from I_F/p to I_L/B by $h(a+p)=a+B$ for $a \varepsilon I_F$. Then h is an isomorphism.

Proof: h is clearly a homomorphism, and if $h(a+p) = h(b+p)$, then $B|(a-b)$. But $a-b \varepsilon F$, so $p|(a-b)$ since $B|p$. Hence h is 1-1.

To see that h is onto let $b \varepsilon I_L$. For each $\sigma \varepsilon G(K:F) - G_P$ we have $\sigma(P) \neq P$ and $\sigma^{-1}(P) \neq P$. Let $B_\sigma \varepsilon L$ be such that $\sigma^{-1}(P)|B_\sigma$. We use the Chinese remainder theorem to find $a \varepsilon I_L$ such that

$a \equiv b \pmod{B}$

$a \equiv 1 \pmod{B\sigma}$

for each $\sigma \varepsilon G(K:F)-G_P$. Then $a \equiv b \pmod{P}$ and $a \equiv 1 \pmod{\sigma^{-1}P}$; thus $\sigma(a) \equiv 1 \pmod{P}$ for each $\sigma \varepsilon G(K:F)-G_P$. Now $G(L:F) \cong G(K:F)/G_P$ so that $N_L(a) = \prod_{\sigma \varepsilon G(L:F)} \sigma(a)$, and σ runs through a set of right coset representatives of G_P in $G(K:F)$. Thus $N_L(a) \equiv b \pmod{P}$. Also

$N_L(a) \in Z$ and hence is in I_F. Finally $N_L(a) \equiv b \pmod{B}$ since $P|(N_L(a)-b)$ and $N_L(a)-b \in I_L$. Therefore $h(N_L(a)+p) = b+B$ and h is onto. //

Lemma 4.3: Let K be a normal extension of F and $p \in F$ an unramified prime in K. If $p = P_1 \cdots P_k$, then G(K:F) is transitive on the P_i.

Proof: Suppose $\sigma(P_1) \neq P_2$ for all $\sigma \in G(K:F)$. P_1 and $P_2 \cdots P_k$ are relatively prime so that there exist $s,t \in I_K$ such that $sP_1 + tP_2 \cdots P_k = 1$. Now $P_2 \nmid N_K(sP_1)$ since $N_K(sP_1) = \prod_{\sigma \in G(K:F)} \sigma(sP_1)$. But $p|N_K(sP_1)$ since $p|N_K(P_1)$, and $P_2|p$. Hence $P_2|N_K(sP)$, a contradiction. Therefore G(K:F) is transitive on the P_i. //

Theorem 4.4: Let K be a normal extension of F and $p \in F$ an unramified prime in K. If $P \in K$ with $P|p$, then $G_P \simeq G(I_K/P:I_F/p)$.

Proof: If $\sigma \in G_P$, define σ' on I_K/P by $\sigma'(a+P) = \sigma(a) + P$ for $a \in I_K$. It is easy to see that σ' is a homomorhpism. Also if $\sigma'(a+P) = \sigma'(b+P)$, then $\sigma(a) + P = \sigma(b) + P$ and $P|\sigma(a-b)$. Hence $\sigma(a-b) = a-b$ since $\sigma \in G_P$ and $P|(a-b)$. Thus σ' is an isomorphism. To see that σ' fixes I_F/p, let $a + P \in I_F/p$. Then $\sigma'(a+P) = \sigma(a)+P = a + P$. Therefore $\sigma' \in G(I_K/P: I_F/p)$.

Define the mapping h on G_P by $h(\sigma) = \sigma'$. h is clearly a homomorphism. If $h(\sigma)(a+P) = a + P$ for each $a \in I_K$, then $\sigma(a) = a$ for each $a \in I_K$ and σ is the identity automorphism. This shows that h is 1-1.

To see that h is onto we show that $|G(I_K/P: I_F/p)| = |G_P|$. Let L be the fixed field of G_P and $B \in L$ with $P|B$. Lemma 4.3 implies that P is the only prime in K such that $P|B$. For if $P_1|B$, then there is $\sigma \in G(K:L) = G_P$ such that $\sigma(P) = P_1$. Thus $P_1 = P$. Now $N_L(B)^{[K:L]} = N_K(B) = N_K(P)$. Suppose that $N_L(B) = q^m$ and $N_K(P) = q^n$ for some rational prime q. Then $q^{m[K:L]} = q^n$ so that $\frac{n}{m} = [K:L]$. Also $[I_K/P: I_L/B] = \frac{n}{m}$, and by Theorem 4.2 $[I_K/P: I_L/B] = [I_K/P:I_F/p]$. Thus $|G(I_K/P: I_F/p)| = [I_K/P: I_F/p] = \frac{n}{m} = [K:L] = |G_P|$.
//

Theorem 4.1 implies that $G(I_K/P: I_F/p)$ is cyclic and generated by σ_P' where $\sigma_P'(a + P) = a^{N_F(p)} + P$. Use the isomorphism of Theorem 4.4 to find $\sigma_P \in G_P$. Then $\sigma_P(a) \equiv a^{N_F(p)} \pmod{P}$ for each $a \in I_K$.

Definition: σ_P is called the Frobenius automorphism of P. We will use both σ_P and $(\frac{K/F}{P})$ to represent the Frobenius automorphism.

Suppose K is a normal extension of F, $p \in F$, $P \in K$ with $P|p$ and σ_P is the Frobenius automorphism of P. If P_1 is another prime divisor of p in K, then there is $\tau \in G(K:F)$ such that $\tau(P) = P_1$. If $n = N_F(p)$, then $\sigma_P(a) \equiv a^n \pmod{P}$ for each $a \in I_K$. So $\sigma_P(\tau^{-1}(a)) \equiv (\tau^{-1}(a))^n \equiv \tau^{-1}(a^n) \pmod{P}$. Thus $\tau\sigma_P\tau^{-1}(a) \equiv a^n \pmod{\tau(P)}$. Hence the Frobenius automorphism of P_1 is $\sigma_{P_1} = \tau\sigma_P\tau^{-1}$ and σ_{P_1} and σ_P are conjugate.

Definition: Let K be a normal extension of F and p a prime in F, unramified in K. The Artin symbol at p, $(\frac{K/F}{p})$, is the conjugary class of the Frobenius automorphisms of the primes in K which divide p.

Observe that if G(K:F) is abelian, then the Artin symbol consists of a single element. So for a prime p ε F and $P|p$, we have $(\frac{K/F}{P}) = (\frac{K/F}{p})$.

Definition: Let K be a finite extension of F and p a prime in F. If p has just one prime divisor in K, then p is said to be undecomposed in K.

Definition: The centralizer of an element σ in a group G is the subgroup $C(\sigma) = \{\tau \varepsilon G: \tau\sigma = \sigma\tau\}$.

Lemma 4.5: Let K be normal over Q, p ε Q and P ε K with $P|p$. Suppose that F is a field with $Q \subseteq F \subseteq K$ and every prime divisors B of p in F is undecomposed in K. Let $C(\sigma_P)$ be the centralizer of the Frobenius automorphism σ_P of P over Q. Then there are $i(C(\sigma_P): <\sigma_P>)$ prime divisors B of p in F such that $(\frac{F/Q}{B}) = (\frac{K/Q}{P})$.

Proof: Let B_1 be a prime divisor of p and P_1 the unique prime divisor of B_1 in K. If $\sigma_{P_1} = (\frac{K/Q}{P_1}) = (\frac{K/Q}{P})$, then $P_1|(\sigma_{P_1}(a)-a^p)$ for all a ε I_F. Hence $B_1|(\sigma_{P_1}(a)-a^p)$ and $(\frac{F/Q}{B_1}) = (\frac{K/Q}{P_1})$. Thus it is sufficient to show that there are $i(C(\sigma_P): <\sigma_P>)$ prime divisors P_i of p in K such that $(\frac{K/Q}{P_i}) = (\frac{K/Q}{P})$.

Any prime divisor of p in K is of the form $\tau(P)$ for some $\tau \varepsilon G(K:Q)$. We know that $(\frac{K/Q}{\tau(P)}) = \tau(\frac{K/Q}{P})\tau^{-1}$, so there are $|C(\sigma_P)|$

elements of G(K:Q) which are conjugate to $(\frac{K/Q}{P})$. To see how many of these yield distinct prime divisors of p, we note that $\tau_1(P) = \tau_2(P)$ if and only if $\tau_2^{-1}\tau_1 \in G_P$. Since $G_P = \langle\sigma_P\rangle$ we have the result. //

<u>Definition</u>: Let K be a finite extension of F and I(F) the group of ideals of F whose prime factors are unramified in K. The <u>Dedikind zeta function of F</u> is the complex valued function

$$\zeta_F(s) = \sum_{U \in I(F)} N_F(U)^{-s}.$$

<u>Definition</u>: A <u>character</u> of a group G is a homomorphism of G into the complex unit circle. The <u>trivial character</u> χ_0 has the property that $\chi_0(\sigma) = 1$ for all $\sigma \in G$.

The set of characters can be made into a group G* by defining $\chi_1\chi_2(\sigma) = \chi_1(\sigma)\chi_2(\sigma)$. The trivial character is the identity of G*.

Suppose that K is a normal extension of F and G(K:F) is abelian. Then we can define a group of characters on the group I(F) by letting $\chi(p) = \chi(\frac{K/F}{p})$, where p is a prime of F unramified in K, and $\chi \in G^*(K:F)$. We extend χ to all of I(F) by letting $\chi(U) = \prod_{i=1}^{k} \chi(p_i)$ where $U = p_1 \cdots p_k$.

<u>Definition</u>: Let $L(s,\chi;K/F) = \sum_{U \in I(F)} \chi(U)N_F(U)^{-s}$, $(\mathrm{Re}(s)>1)$, then $L(s,\chi;K/F)$ is called a <u>abelian L-function</u>.

Notice that $L(s,\chi_0;K/F) = \zeta_F(s)$.

<u>Lemma 4.6</u>: $L(s,\chi;K/F) = \prod_{p \in F} (1-\chi(p)N_F(p)^{-s})^{-1}$

Proof: Let $T(x) = \prod_{N_F(p)\leq x} \{1+\chi(p)N_F(p)^{-s}+\chi(p)^2N_F(p)^{-2s}+\cdots\}$.

Observe that since $Re(s) > 1$, $L(s,\chi;K/F)$ converges uniformly so that we can rearrange the terms without altering the sum. If $U \in I(F)$ and $U = p_1^{e_1}\cdots p_k^{e_k}$ with $N_F(p_i) \leq x$ for each i, then $\chi(U)N_F(U)^{-s}$ is in the product $T(x)$. Let $A = \{U \in I(F)$: U has a prime factor p with $N_F(p) > x\}$. Then $|L(s,\chi;K/F)-T(x)| = \sum_{U\in A} \chi(U)N_F(U)^{-s}$ which tends to zero. Hence $\lim_{x\to\infty} T(x) = L(s,\chi;K/F)$ and $L(s,\chi;K/F) =$

$\prod_{p\in F} \{1+\chi(p)N_F(p)^{-s}+\chi(p)^2N_F(p)^{-2s}+\cdots\} = \prod_{p\in F} (1-\chi(p)N_F(p)^{-s})^{-1}$. //

Corollary 4.7: $\zeta_F(s) = \prod_{p\in F} (1-N_F(p)^{-s})^{-1}$.

Lemma 4.8: If G is cyclic, then $\sum_{\chi\in G^*} \chi(\iota)=|G|$ and $\sum_{\chi\in G^*} \chi(\sigma)=0$ for $\sigma \neq \iota$.

Proof: Let G be cyclic of order n, say $G = \langle\sigma\rangle$. Then $\chi(\tau)^n = \chi(\tau^n) = \chi(\iota) = 1$ for all $\tau \in G$, $\chi \in G^*$. So $\chi(\tau)$ must be an n^{th} root of unity. Also $\chi(\sigma) = e^{2k\pi i/n}$ is clearly a character for $k = 0,1,\ldots,n-1$. Hence the group of characters of G has order n since there are n n^{th} roots of unity.

Now $\chi(\iota) = 1$ for each $\chi \in G^*$. So $\sum_{\chi\in G^*} \chi(\iota) = |G^*| = |G|$.

Also if $\tau \neq \iota$, then there is $\chi_1 \in G^*$ such that $\chi_1(\tau) \neq 1$. $\chi_1\chi$ runs through G^* as χ runs through G^* so that $\sum_{\chi\in G^*} \chi(\tau) = \sum_{\chi\in G^*} \chi_1(\tau)\chi(\tau) = \chi_1(\tau)\sum_{\chi\in G^*} \chi(\tau)$. Thus $(1-\chi_1(\tau))\sum_{\chi\in G^*} \chi(\tau) = 0$ and $\sum_{\chi\in G^*} \chi(\tau) = 0$. //

Definition: Let A be a set of primes of K. Then the Dirichlet density of A is

$$d(A) = \lim_{s\to 1^+} \frac{\log \prod_{P\varepsilon A} (1-N_K P^{-s})^{-1}}{\log \zeta_K(s)}$$ whenever the limit exists.

It can be shown that $\zeta_K(s)$ has a simple pole at $s = 1$, (Janusz, 1973, p. 125). Thus $\log \zeta_K(s) = -\log(s-1) + O(1)$, where $f(s) = O(g(s))$ means $\frac{f(s)}{g(s)}$ remains bounded as $s \to 1^+$. By $f(s) = o(g(s))$ we mean that $\lim_{s\to 1^+} \frac{f(s)}{g(s)} = 0$.

Lemma 4.9: Let A be a set of primes in K. Then $d(A) = a$ if and only if $\sum_{P\varepsilon A} N_K(P)^{-s} = -a \log(s-1) + o(\log(s-1))$.

Proof: Suppose $d(A) = a$. $\log \prod_{P\varepsilon A} (1-N_K(P)^{-s})^{-1}$

$$= -\sum_{P\varepsilon A} \log(1-N_K P^{-s}) = \sum_{P\varepsilon A} \sum_{m=1}^{\infty} m^{-1} N_K(P)^{-ms}$$ since

$$\log(1-Z) = -\sum_{m=1}^{\infty} \frac{Z^m}{m}.$$ So

$$0 = \lim_{s\to 1^+} \frac{\log \prod_{P\varepsilon A} (1-N_K(P)^{-s})^{-1}}{\log\zeta_K(s)} - a$$

$$= \lim_{s\to 1^+} \frac{\sum_{P\varepsilon A} \sum_{m=1}^{\infty} m^{-1} N_K(P)^{-ms} - a \log \zeta_K(s)}{\log \zeta_K(s)}$$

$$= \lim_{s\to 1^+} \frac{\sum_{P\varepsilon A} N_K(P)^{-s} + a(\log(s-1)+O(1)) + \sum_{P\varepsilon A} \sum_{m=2}^{\infty} m^{-1} N_K(P)^{-ms}}{-\log(s-1) + O(1)}$$

$$= \lim_{s\to 1^+} \frac{\sum_{P\varepsilon A} N_K(P)^{-s} + a \log(s-1)}{-\log(s-1) + O(1)}$$

$$+ \lim_{s\to 1^+} \frac{O(1) + \sum_{P\varepsilon A} \sum_{m=2}^{\infty} m^{-1} N_K(P)^{-ms}}{-\log(s-1) + O(1)}$$

and the second limit clearly goes to zero. Hence

$$\lim_{s\to 1^+} \frac{\sum_{P\varepsilon A} N_K(P)^{-s} + a\log(s-1)}{-\log(s-1) + O(1)} = 0 \text{ and}$$

$\sum_{P\varepsilon A} N_K(P)^{-s} = -a\log(s-1) + o(\log(s-1))$. These steps are clearly reversible, so that the result is obtained. //

<u>Lemma 4.10</u>: Let K be normal over F and G(K:F) be abelian with [K:F] = n. If $B \varepsilon F$, then $\prod_{\chi\varepsilon G^*(K:F)} (1-\chi(B)N_F(B)^{-s}) =$ $(1-N_F(B)^{-sm})^{n/m}$ where $m = |G_P|$, $P \varepsilon K$ and $P|B$.

<u>Proof</u>: Consider the mapping h from G*(K:F) to G_P^* defined by $h(\chi) = \chi|G_P$. h is a homomorphism with kernel $H = \{\chi\varepsilon G^*(K:F): \chi(B)=1\}$. $|H| = |G^*(K:F)|/|G_P^*| = \frac{n}{m}$ so that $\prod_{\chi\varepsilon G^*(K:F)} (1-\chi(B)N_F(B)^{-s}) =$

$\prod_{\chi\varepsilon G_P^*} (1-\chi(B)N_F(B)^{-s})^{n/m}$. As in the proof of Lemma 4.8, the m elements of G_P^* are the characters of G_P which send $(\frac{K/F}{B})$ to the m^{th} roots of unity. Let ξ be a primitive m^{th} root of unity. Then

$$\prod_{\chi\varepsilon G^*(K:F)} (1-\chi(B)N_F(B)^{-s}) = \prod_{\chi\varepsilon G_P^*} (1-\chi(\tfrac{K/F}{B})N_F(B)^{-s})^{n/m} = \prod_{i=0}^{m-1} (1-\xi^i N_F(B)^{-s})^{n/m}$$

$$= (1-N_F(B)^{-ms})^{n/m}. \quad //$$

<u>Theorem 4.11</u>: Let K be a normal over F and G(K:F) be abelian. Then $\prod_{\chi \varepsilon G^*(K:F)} L(s,\chi;K/F) = \zeta_K(s)$, $(\mathrm{Re}(s) > 1)$.

<u>Proof</u>: Since $L(s,\chi;K/F) = \prod_{B\varepsilon F} (1-\chi(B)N_F(B)^{-s})^{-1}$ and $\zeta_K(s) = \prod_{P\varepsilon K} (1-N_K(P)^{-s})^{-1}$, it suffices to show that

$\prod_{\chi \varepsilon G^*(K:F)} (1-\chi(B)N_F(B)^{-s}) = \prod_{P|B} (1-N_K(P)^{-s})$. First we note that if

$P_1|B$, then $N_K(P_1) = N_F(B)^m$ where $m = f(P_1/B)$. This is because $f(P_1/p) = |G_{P_1}| = \frac{n}{k}$, where $n = [K:F]$ and k is the number of prime divisors of B in K. By Lemma 4.10, $\prod_{\chi \varepsilon G^*(K:F)} (1-\chi(B)N_F(B)^{-s}) =$

$(1-N_F(B)^{-sm})^{n/m} = (1-N_K(P_1)^{-s})^k = \prod_{P|B} (1-N_K(P)^{-s})$. //

Because $\zeta_K(s)$ and $\zeta_F(s)$ have simple poles at s = 1, we have that $\frac{\zeta_K(s)}{\zeta_F(s)}$ is analytic at s = 1. Hence $\frac{\zeta_K(1)}{\zeta_F(1)} = \prod_{\chi \neq \chi_0} L(1,\chi;K/F) \neq 0,\infty$.

We use this fact to get Dirichlet's theorem.

<u>Theorem 4.12</u>: Let K be normal over F and G(K:F) be abelian of order n. If $\sigma \varepsilon G(K:F)$ and $A = \{B\varepsilon F: (\frac{K/F}{B}) = \sigma\}$, then A has Dirichlet density $\frac{1}{n}$.

<u>Proof</u>: We must show that $\lim_{s\to 1^+} \dfrac{\log \prod_{B\varepsilon A} (1-N_F(B)^{-s})^{-1}}{\log \zeta_F(s)} = \frac{1}{n}$.

As in the proof of Lemma 4.9, this limit is equal to

$$\lim_{s\to 1^+} \frac{\sum_{B\varepsilon A} \sum_{m=1}^{\infty} m^{-1} N_F(B)^{-sm}}{\log \zeta_F(s)} .$$

Let $T(s) = n^{-1} \sum_{\chi\varepsilon G^*(K:F)} \chi(\sigma^{-1}) \log L(s,\chi;K/F)$. Then

$$T(s) = n^{-1}\chi_0(\sigma^{-1})\log L(s,\chi_0;K/F) + n^{-1}\sum_{\chi\neq\chi_0} \chi(\sigma^{-1})\log L(s,\chi;K/F)$$

$$= n^{-1}\log \zeta_F(s) + n^{-1}\sum_{\chi\neq\chi_0} \chi(\sigma^{-1})\log L(s,\chi;K/F). \text{ So}$$

$$\lim_{s\to 1^+} \frac{T(s)}{\log \chi_F(s)} = \frac{1}{n} + \lim_{s\to 1^+} \frac{\sum_{\chi\neq\chi_0} \chi(\sigma^{-1})\log L(s,\chi;K/F)}{\log \zeta_F(s)}$$

By the remarks following Theorem 4.11, the above limit tends to $\frac{1}{n}$.

Hence $\lim_{s\to 1^+} \frac{T(s)}{\log \zeta_F(s)} = \frac{1}{n}$.

We also have that

$$T(s) = n^{-1} \sum_{\chi\varepsilon G^*(K:F)} \chi(\sigma^{-1})\log L(s,\chi;K/F)$$

$$= n^{-1} \sum_{\chi\varepsilon G^*(K:F)} \chi(\sigma^{-1})\log \prod_{B\varepsilon F} (1-\chi(B)N_F(B)^{-s})^{-1}$$

$$= n^{-1} \sum_{\chi\varepsilon G^*(K:F)} \chi(\sigma^{-1}) \sum_{B\varepsilon F} \sum_{m=1}^{\infty} m^{-1}\chi(B)N_F(B)^{-sm}$$

$$= n^{-1} \sum_{B\varepsilon F} \sum_{m=1}^{\infty} m^{-1}N_F(B)^{-sm} \sum_{\chi\varepsilon G^*(K:F)} \chi(\sigma^{-1})\chi\left(\frac{K/F}{B}\right)$$

$$= n^{-1} \sum_{B\varepsilon F} \sum_{m=1}^{\infty} m^{-1} N_F(B)^{-sm} \sum_{\chi\varepsilon G^*(K:F)} \chi(\sigma^{-1}(\tfrac{K/F}{B})).$$

By Lemma 4.7, $\sum_{\chi\varepsilon G^*(K:F)} \chi(\sigma^{-1}(\frac{K/F}{B}))$ is zero if $(\frac{K/F}{B}) \neq \sigma$ and n if $(\frac{K/F}{B}) = \sigma$. Hence

$$T(s) = n^{-1} \sum_{B\varepsilon A} \sum_{m=1}^{\infty} m^{-1} N_F(B)^{-ms} n = \sum_{B\varepsilon A} \sum_{m=1}^{\infty} m^{-1} N_F(B)^{-ms}$$

$$= \log \prod_{B\varepsilon A} (1-N_F(B)^{-s})^{-1}.$$ That is

$$\lim_{s\to 1^+} \frac{\log \prod_{B\varepsilon A} (1-N_F(B)^{-s})^{-1}}{\log \zeta_F(s)} = \lim_{s\to 1^+} \frac{T(s)}{\log \zeta_F(s)} = \frac{1}{n}. \quad //$$

<u>Lemma 4.13</u>: Let A be a set of primes in K and A_1 the set of primes of A with relative degree one over the rationals. Then the Dirichlet density of $A-A_1$ is zero so that $d(A) = d(A_1)$.

<u>Proof</u>: Let $P \varepsilon A-A_1$; then $N_K(P) = p^k$ for some $p \varepsilon Q$ and $k \geq 2$. Let S be the set of rational primes p such that $P|p$ for some $P \varepsilon A-A_1$. There are at most [K:Q] primes in $A-A_1$ which divide p for any $p \varepsilon S$. Now

$$\sum_{P\varepsilon A-A_1} N_K(P)^{-s} \leq [K:F] \sum_{p\varepsilon S} p^{-2s} = O(1).$$ Thus $$\lim_{s\to 1^+} \frac{\sum_{P\varepsilon A-A_1} N_K(P)^{-s}}{\log(s-1)} = 0$$

and $d(A-A_1) = 0$ by Lemma 4.9. //

We now come to the main theorem of this section, the Chebotarev density theorem.

Theorem 4.14: Let K be normal over Q, C a conjugary class of G(K:Q) with c elements and $A = \{p \varepsilon Q: (\frac{K/Q}{p}) = C\}$. Then A has Dirichlet density $\frac{c}{n}$, where $n = |G(K:Q)|$.

Proof: Let $\sigma \varepsilon C$ and F the fixed field of $<\sigma>$. G(K:F) is cyclic so that the set $A_1 = \{B \varepsilon F:(\frac{K/F}{B}) = \sigma\}$ has Dirichlet density $\frac{1}{|\sigma|}$ by Dirichlet's theorem. Using Lemma 4.9 we have

$\sum_{B\varepsilon A_1} N_F(B)^{-s} = \frac{-1}{|\sigma|} \log(s-1) + o(\log(s-1))$. If $A_2 = \{B\varepsilon A_1 : B$ has relative degree one over Q$\}$, then $\sum_{B\varepsilon A_2} N_F(B)^{-s} = \sum_{\substack{B\varepsilon A_2 \\ B|p}} N_F(p)^{-s}$

$= \frac{-1}{|\sigma|} \log(s-1) + o(\log(s-1))$ by Lemma 4.13.

Now if $(\frac{K/F}{B}) = \sigma$, then $\sigma \varepsilon G_B \subseteq G(K:F) = <\sigma>$. Hence $G_B = <\sigma>$ and $f(P/B) = |G_B| = [K:F]$ for $P|B$. If $B = P_1 \cdots P_k$, then $f(P_1/B)k = [K:F]$ and $k = 1$. Therefore each $B \varepsilon A_2$ is undecomposed in K. Also if $B|p$, then $(\frac{K/F}{B}) = \sigma$ if and only if $(\frac{K/Q}{p}) = C$. Thus the hypothesis of Lemma 4.5 are satisfied and we have $i(C(\sigma): <\sigma>) \sum_{p\varepsilon A} N_Q(p)^{-s} = \frac{-1}{|\sigma|}$ $\log(s-1) + o(\log(s-1))$. Hence

$$\sum_{p\varepsilon A} N_Q(p)^{-s} = \frac{-\log(s-1)}{i(C(\sigma):<\sigma>)|\sigma|} + o(\log(s-1)).$$

Finally $\frac{1}{i(C(\sigma): <\sigma>)|\sigma|} = \frac{1}{|C(\sigma)|} = \frac{i(G(K:Q): C(\sigma))}{|G(K:Q)|} = \frac{c}{n}$. //

B. A Theorem of Van der Waerden

Theorem 4.15: Let $f(x) \in Z[x]$ and p be a rational prime with p not dividing the leading coefficient of f(x). Then $G(f,Z_p)$ in a subgroup of G(f,F).

Proof: As at the beginning of Chapter 3 we form the polynomial $F(x) = F_1(x)\cdots F_k(x)$, where $F_1(x)$ is the Galois resolvent of f(x). If $\sigma \in G(f,Z_p)$, then σ holds the coefficients of $F_i(x)$ fixed for each i because $F_i(x) \in Z_p[x]$. Hence $\sigma F_1 = F_1$ which, according to Theorem 3.3, is precisely the condition necessary for $\sigma \in G(f,Q)$. //

Theorem 4.16: Let $f(x) \in Z[x]$ and p be a rational prime with p not dividing the leading coefficient of f(x). If $f(x) \equiv f_1(x)^{e_1} f_2(x)^{e_2} \cdots f_k(x)^{e_k}$ (mod pZ[x]), where the $f_i(x)$ are distinct irreducible polynomials of $Z_p[x]$, then G(f,Q) contains a permutation consisting of k cycles and the i^{th} cycle has length $[f_i]$.

Proof: By Theorem 4.1, $G(f,Z_p)$ is cyclic. Let σ be an automorphism generating $G(f,Z_p)$. Now σ is transitive on the roots of $f_i(x)$ for each i, while σ does not send a root of $f_i(x)$ to a root of $f_j(x)$ if $i \neq j$. Since σ, acting on the roots of $f_i(x)$, must be a cycle of length $[f_i]$, σ has the desired form. Finally $\sigma \in G(f,Q)$ because $G(f,Z_p) \subseteq G(f,Q)$ by Theorem 4.15. //

Theorem 4.17: Let $f(x) \in Z[x]$ and p be a rational prime with p not dividing the leading coefficient of f(x). Suppose also that K is the splitting field of f(x) and P is a prime in I_K with $P|p$. If the Frobenius automorphism of P is σ and $f(x) \equiv \prod_{i=1}^{k} f_i(x)^{e_i}$ (mod pZ[x]), where the $f_i(x)$ are distinct irreducible polynomials of $Z_p[x]$, then σ has k cycles and the i^{th} cycle has length $[f_i]$.

Proof: By the definition of σ, it is the automorphism of G(K:Q) such that $\sigma(a) \equiv a^p$ (mod P) for all $a \in I_K$. By Theorem 4.1, $\sigma'(a) = a^p$ for all $a \in I_K/P$ generates $G(I_K/P: Z_p)$. Since $G_P \cong G(I_K/P: Z_p)$, and using the proof of Theorem 4.17, σ has the appropriate cycle structure. //

We can use Theorems 4.14, 4.16 and 4.17 to aid us in calculating the Galois group of f(x) over the rational numbers. Factoring f(x) modulo p for a sufficient number of primer p will yield the cycle structure of each permutation in G(f,Q). If we want an approximation as to what proportion of the elements of G(f,Q) have the same cycle structure as a particular element σ, we use Theorems 4.14 and 4.17. We let A_x be the set of rational primes p for which $p \leq x$ and factoring f(x) modulo p yields the same cycle structure as σ. Let S_x be the set of all rational primes p for which $p \leq x$. If τ and σ have the same cycle structure, then they are conjugates. So the Chebotarev density theorem says that $\lim_{x\to\infty} \frac{|A_x|}{|S_x|} = \frac{c}{|G(f,Q)|}$ where c is the number of elements of G(K:Q) conjugate to σ. For any x, we can use the approximation $\frac{|A_x|}{|S_x|}$ for the proportion of elements of G(f,Q) which have the same cycle structure as σ. One problem here is that we need to know how large to pick x. Lagaries and Odlyzko (1977) indicate how one might calculate such a bound, but the bound is quite difficult to compute.

Fortunately we seldom need to know these bounds. Generally if we know the cyclic structure of the elements of G(f,Q), we can determine G(f,Q). It is advantageous to have a listing of the permutation groups of degree [f] with entries describing the degree, order, transitivity and cycle structure of these groups. Such a listing can be found for degrees up to seven in Zassenhaus (1971), but there are a number of errors in the tables (Neuman, 1975).

To use this method, you must be able to factor polynomials modulo p for a prime p. This factorization is done by trial and error. It amounts to solving, for each possible degree of a factor, a system of n+1 congruences modulo p, where n = [f]. If $f(x) = \sum_{i=o}^{n} a_i x^i$ and we want to determine whether f(x) is congruent to the product of an m^{th} degree polynomial with an $n-m^{th}$ degree polynomial, then we set up the n+1 congruences $\sum_{j=o}^{i} b_j c_{i-j} \equiv a_i \pmod{p}$ for $i = 0,1,\ldots,n$, where the b_i, c_i are unknowns, $b_i = 0$ if $i > m$ and $c_j = 0$ if $j > n-m$.

As an example of Van der Waerden's method of determining the Galois group, consider the polynomial $f(x) = x^5+2x^4+8x^3+3x^2+5x+1$.

By trial and error it can be shown that f(x) is irreducible modulo 2; has factors of degree 1,2 and 2 modulo 3; and has factors of degree 2 and 3 modulo 5. So G(f,Q) contains a 5 cycle σ_1, a permutation σ_2 with 2 cycles of length 2 and a permutation σ_3 with a 2 cycle and a 3 cycle. Note that σ_3^2 is a 3 cycle and σ_3^3 is a 2 cycle. Hence the order of G(f,Q) must be at least $2 \cdot 3 \cdot 5 = 30$. Thus G(f,Q) is S_5 or A_5, but σ_3^3 is an odd permutation. Therefore G(f,Q) = S_5.

Another example is $f(x) = x^4+2x^3+2x+2$. f(x) factors into 2 quadratics modulo 3, and so G(f,Q) is either the cyclic group of order 4 or the Klein 4 group. f(x) is irreducible modulo 5, hence G(f,Q) is the cyclic group of order 4.

An alternative to using the Chebotarev density theorem in cases where the use of Van der Waerden's theorem is inconclusive, is to use the method of Zassenhaus. We use Van der Waerden's method for a few "small" primes to narrow down the choices for G(f,Q), and then apply the Zassenhaus method to determine which of these choices is actually G(f,Q). This is actually the most efficient procedure in general because it usually avoids calculating resolvent equations for subgroups in groups where the index is large. Also it avoids the most difficult part, as far as the computation goes, of the Van der Waerden method—factoring modulo large primes.

Zassenhaus (1971) suggest two other methods, a p-adic method and a ring theoretic approach. The numerous errors and misprints, along with the sketchy proofs and explanations, make these methods difficult to understand. Both involve deep ring theoretic results.

which are beyond the scope of this paper. Some of the details for the second method are filled in by the papers Zassenhaus (1967) and Zassenhaus (1974), although there are still numerous gaps even in these articles.

Van der Waerden's method generally needs no help. Gallagher (1973) shows that "almost all" monic polynomials of a given degree are irreducible and have Galois group equal to the symmetric group. Zassenhaus (1971) claims that if the Galois group of an equation is the symmetric group, Van der Waerden's method will usually quickly realize this by showing that the Galois group contains a transposition and a p-cycle for some $p > n/2$, where n is the degree of the equation.

It is worth noting that computers can be used in some of the techniques described in this paper to do the tedious calculations. For instance, Hensel's lemma applied to the p-adic numbers provides an algorithm that can easily be used on a computer. Many of the computations of the Zassenhaus method can be done by computers (Stauduhar, 1973). Also factoring modulo p can be done by computers as it is just a matter of testing a finite number possibilities.

APPENDIX

TABLE 1

TRANSITIVE SUBGROUPS OF S_n FOR n=4,...,7

Degree	Group	Contained in	Function	Generators, Description
4	G_8	S_4	$x_1x_3+x_2x_4$	(1234), (13) group of the square
4	G_4^1	G_8	$x_1x_2^2+x_2x_3^2+x_3x_4^2+x_4x_1^2$	(1234) cyclic four group
4	G_4^2			(12)(34), (13)(24) Klein 4-group
5	G_{20}	S_5	$[x_1x_2+x_2x_3+x_3x_4+x_4x_5+x_5x_1 -x_1x_3-x_2x_5-x_5x_2-x_2x_4-x_4x_1]^2$	(12345), (2354) metacyclic five group
5	G_{10}			(12345), (25)(34)
5	G_5	G_{10}	$x_1x_2^2+x_2x_3^2+x_3x_4^2+x_4x_5^2+x_5x_1^2$	(12345) cyclic five group
6	G_{72}	S_6	$x_1x_2x_3+x_4x_5x_6$	(123), (456), (12), (45), (14)(25)(36) maximal group imprimitive on two sets of three letters

TABLE 1--Continued

Degree	Group	Contained in	Function	Generators, Description
6	G_{36}^{1}			(123), (456), (12)(45), (1425)(36) G_{72} A_6
6	G_{36}^{2}	G_{72}	$(x_1-x_2)(x_2-x_3)(x_3-x_1)(x_4-x_5)$ $\cdot (x_5-x_6)(x_6-x_4)$	(123), (456), (12)(45), (14)(25)(36)
6	G_{18}	G_{36}^{2}	$(x_1-x_2)(x_2-x_3)(x_3-x_1)$ $+ (x_4-x_5)(x_5-x_6)(x_6-x_4)$	(123), (456), (14)(25)(36)
6	G_{12}^{1}	G_{36}^{2}	$x_1x_4+x_2x_5+x_3x_6$	(123)(456), (12)(45), (14)(25)(36) metacyclic six group
6	G_{6}^{1}	G_{18}	$x_1x_4+x_2x_6+x_3x_5$	(123)(465), (14)(25)(36) isomorphic to S_3
6	G_{6}^{2}	G_{18}	$x_1x_6^2+x_2x_4^2+x_3x_5^2+x_4x_2^2+x_5x_1^2$ $+ x_6x_2^2$	(123)(456), (14)(25)(36) cyclic six group

TABLE 1--Continued

Degree	Group	Contained in	Function	Generators, Description
6	G_{48}	S_6	$x_1x_2+x_3x_4+x_5x_6$	(12), (34), (56), (135)(246), (13)(24) maximal group imprimitive on three sets of two letters
6	G_{24}^1	G_{48}	$(x_1+x_2-x_3-x_4)(x_3+x_4-x_5-x_6)$ $\cdot (x_5-x_6-x_1-x_6)(x_1-x_2)$ $\cdot (x_3-x_4)(x_5-x_6)$	(12)(34), (34)(56), (12)(56), (135)(246), (14)(23)(56)
6	G_{24}^2	G_{48}	$(x_1+x_2-x_3-x_4)(x_3+x_4-x_5-x_6)$ $\cdot (x_5+x_6-x_1-x_2)$	(12)(34)(56), (34)(56), (56), (135)(246)
6	G_{24}^3			(135)(246), (13)(24), (12)(34), (34)(56) G_{48} A_6 isomorphic to S_4
6	G_{12}^2	G_{24}^3	see G_{24}^2	(12)(34), (34)(56), (12)(56), (135)(246) isomorphic to A_4

TABLE 1--Continued

Degree	Group	Contained in	Function	Generators, Description
6	G_{120}	S_6	$[x_1x_2+x_3x_5+x_4x_6]\cdot[x_1x_3+x_4x_5+x_2x_6]$ $\cdot\ [x_3x_4+x_1x_6+x_2x_5]\cdot[x_1x_5+x_2x_4+x_3x_6]$ $\cdot\ [x_1x_4+x_2x_3+x_5x_6]$	(126)(354), (12345), (2354) isomorphic to S_5
6	G_{60}			(126)(354), (12345), (25)(34) G_{120} A_6 isomorphic to A_5
7	G_{168}	S_7	$x_1x_2x_4+x_1x_3x_7+x_1x_5x_6+x_2x_3x_5$ $+\ x_2x_6x_7+x_3x_4x_6+x_4x_5x_7$	(1234567), (235)(476), (2743)(56)
7	G_{42}	S_7	$x_1x_2x_4+x_1x_2x_6+x_1x_3x_4+x_1x_3x_7$ $+x_1x_5x_6+x_1x_5x_7+x_2x_3x_5+x_2x_3x_7$ $+x_2x_4x_5+x_2x_6x_7+x_3x_4x_6+x_3x_5x_6$ $+x_4x_5x_7+x_4x_6x_7$	(1234567), (243756) metacyclic seven group
7	G_{21}	G_{168}	See G_{42} S_7	(1234567), (235)(476)

TABLE 1--Continued

Degree	Group	Contained in	Function	Generators, Description
7	G_{14}	G_{42}	$x_1x_2+x_2x_3+\cdots+x_6x_7+x_7x_1$	(1234567), (27)(45)(36)
7	G_7	G_{21}	See G_{14} G_{42}	(1234567) cyclic 7 group

TABLE 2

RIGHT COSET REPRESENTATIVES

<u>Degree 4</u>	
$G_8 \subset S_4$	ι, (23), (34)
$G_4^1 \subset G_8$	ι, (12)(34)
<u>Degree 5</u>	
$G_{20} \subset S_5$	ι, (12)(34), (12435), (15243), (12453), (12543)
$G_5 \subset G_{10}$	ι, (12)(35)
<u>Degree 6</u>	
$G_{72} \subset S_6$	ι, (2543), (236)(45), (25436), (25)(34), (2453), (25), (2345), (24536), (3645)
$G_{36}^2 \subset G_{72}$	ι, (56)
$G_{18} \subset G_{72}$	ι, (12)(45), (56), (12)(465)
$G_6^1 \subset G_{18}$	ι, (123), (132)
$G_6^2 \subset G_{18}$	ι, (123), (132)
$G_{12} \subset G_{72}$	ι, (123), (132), (56), (123)(56), (132)(56)
$G_{48} \subset S_6$	ι, (24635), (26)(35), (354), (2345), (253), (345), (256)(34), (26435), (2346), (234), (25)(36), (2435), (24)(35), (26543)
$G_{24}^1 \subset G_{48}$	ι, (12)
$G_{24}^2 \subset G_{48}$	ι, (13)(24)
$G_{12}^2 \subset G_{24}^3$	ι, (13)(24)
$G_{120} \subset S_6$	ι, (13), (23), (123), (132), (12)

TABLE 2--Continued

Degree 7

$G_{168} \subset S_7$	ι, (356), (365), (34)(56), (354), (364), (456), (345), (36)(45), (465), (35)(46), (346), (47)(56), (35)(47), (36)(47), (243756), (243675), (243)(57), (2475), (247536), (247563), (246375), (246)(57), (246753), (24)(375), (24)(36)(57), (24)(567), (245)(37), (245736), (245673)
$G_{42} \subsetneq S_7$	Let A be the set consisting of the even coset representatives for G_{168} in S_7. Let B be the set of all coset representatives for G_{21} in G_{168}. Then the required 120 coset representatives here are given by A • B.
$G_{21} \subset G_{168}$	ι, (37)(56), (23)(74), (2347)(56), (24)(56), (24)(37), (2743)(56), (27)(34)
$G_{14} \subset G_{42}$	ι, (235)(476), (253)(467)
$G_7 \subset G_{21}$	ι, (235)(476), (253)(467)

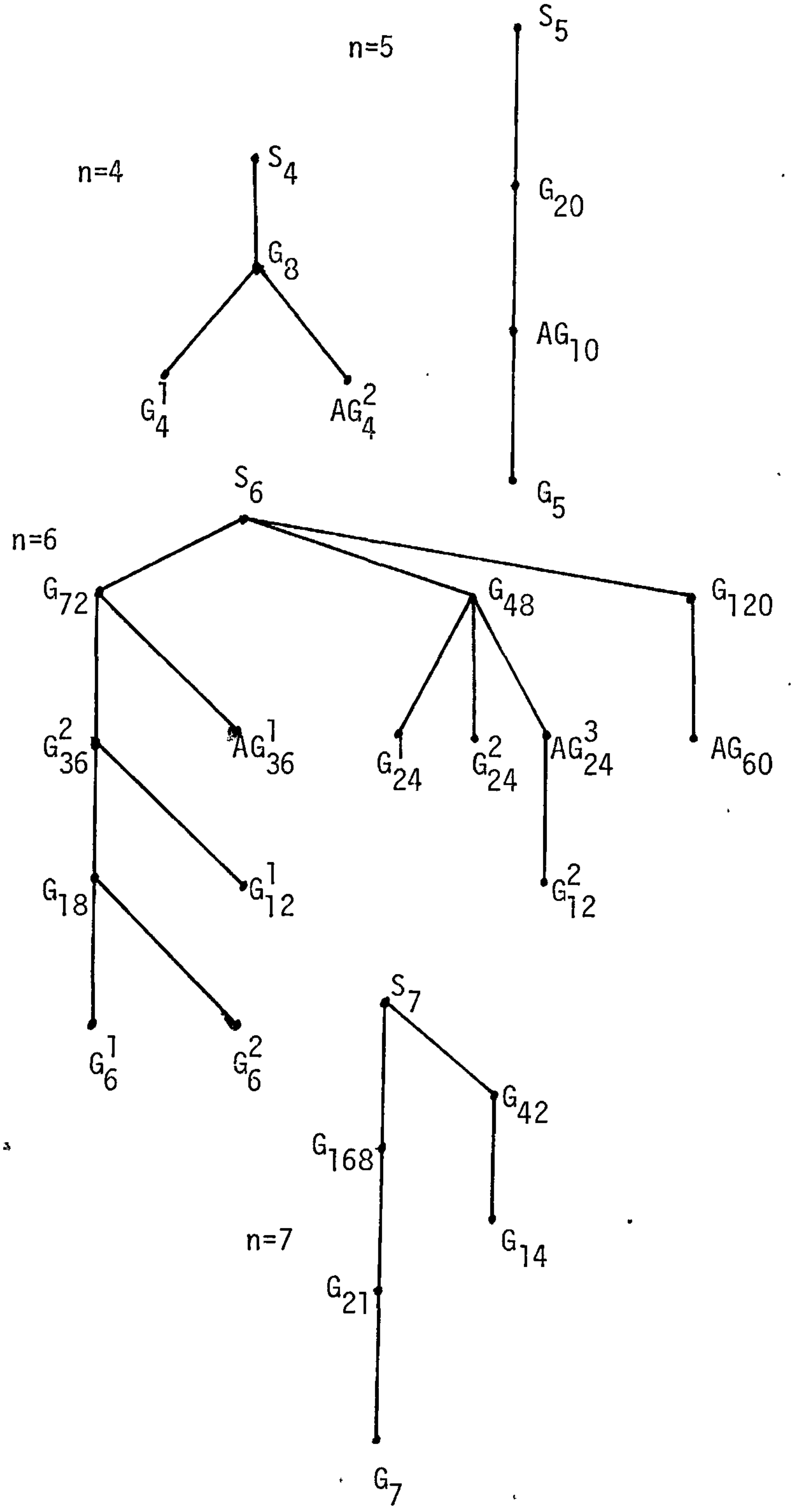


Figure 1. The Order of Subgroup Choices
for the Zassenhaus Method

BIBLIOGRAPHY

Clark, Allan. Elements of Abstract Algebra. Belmont, Calif.: Wadsworth Publishing Company, 1971.

Gallagher, P. X. "The Large Sieve and Probabalistic Galois Theory." Analytic Number Theory. Vol. XXIX: Proceedings of the Symposium in Pure Mathematics. Edited by Harold G. Diamond. Providence, R.I.: American Mathematical Society, 1973.

Goldstein, Larry Joel. Analytic Number Theory. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1971.

Herstein, I. N. Topics in Algebra. 2nd ed. Lexington, Mass.: Xerox College Publishing, 1975.

Janusz, Gerald J. Algebraic Number Fields. New York: Academic Press, 1973.

Lagarias, J. C., and Odlyzko, A. M. "Effective Versions of the Chebotarev Density Theorem." Algebraic Number Fields. Edited by A. Fröhlich. New York: Academic Press, 1977.

Lang, Serge. Algebraic Numbers. Reading, Mass.: Addison-Wesley Publishing Company, 1964.

________. Algebra. New York: Addison-Wesley Publishing Company, 1965.

Lieber, Lillian R. Galois and the Theory of the Groups. Lancaster, Penn.: The Science Press Printing Company, 1932.

Mignotte, M. "An Inequality About Factors of Polynomials." Mathematics of Computation, XXVIII (October, 1974), pp. 1153-1157.

Neumann, Peter M. Review of "On the Group of an Equation," by Hans Zassenhaus. Mathematical Reviews, Vol. 49: March, 1975, p. 917.

Pollard, Harry, and Diamond, Harold G. The Theory of Algebraic Numbers. 2nd ed. The Carus Mathematical Monographs, Vol. IX. New York: The Mathematical Association of America, 1975.

Postnikov, M. M. Foundations of Galois Theory. Translated by Ann Swinfen. Vol. XXIX of the International Series of Monographs on Pure and Applied Mathematics. Edited by I. N. Sneddon, M. Stark and S. Ulam. New York: The MacMillan Company, 1962.

Stauduhar, Richard P. "The Determination of Galois Groups." Mathematics of Computation, XXVII (October, 1973), pp. 981-996.

Uspensky, J. V. Theory of Equations. New York: McGraw-Hill Book Company, Inc., 1948.

Van der Waerden, B. L. Modern Algebra. Vol. I. Translated by Fred Blum. New York: Frederick Ungar Publishing Co., 1949.

Wahab, J. H. "New Cases of Irreducibility for Legendre Polynomials." Duke Mathematical Journal, XIX (1952), pp. 167-169.

Zassenhaus, Hans. "The Group of an Equation." Nachrichten der Akademie der Wissenschaften in Göttingen, Mathematisch-Physikalische Klasse, II (1967), pp. 147-166.

________. "On Hensel Factorization, I." Journal of Number Theory, I (1969), pp. 291-311.

________. "On the Group of an Equation." Computers in Algebra and Number Theory. Vol. IV: SIAM-AMS Proceedings. Edited by Garret Birkhoff and Marshall Hall, Jr. Providence, R.I.: American Mathematical Society, 1971.

________. "How to Find the Group of an Equation." Classroom Notes, Ohio State University, October, 1974. (Typewritten.)

Zimmer, Horst G. "Factorization of Polynomials According to a Method of Zassenhaus." University of California, Los Angeles, 1969. (Typewritten.)