FINITE GROUPS OF
FRACTIONAL LINEAR TRANSFORMATIONS

by

VIVIEN BETH KITCHEN
B.Sc., University of British Columbia, 1969

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

in the Department
of
Mathematics

We accept this thesis as conforming to the
required standard

THE UNIVERSITY OF BRITISH COLUMBIA
May, 1972
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Mathematics

The University of British Columbia
Vancouver 8, Canada

Date April 13, 1972
Supervisor: Dr. K. Hoechsmann

ABSTRACT

In this thesis we consider the group of fractional linear transformations of a variable $x$ over an algebraically closed field $k$. The purpose of the thesis is to determine all finite subgroups of this group whose orders are not divisible by the characteristic of $k$. 
ACKNOWLEDGMENTS

I would like to express my gratitude to Dr. Hoechsmann for his guidance and encouragement throughout the preparation of this thesis. Dr. Hoechsmann's enthusiasm for the subject made the task enjoyable. I would also like to thank Dr. Gamst for reading the thesis and the National Research Council and the University of British Columbia for financial assistance.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter I</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter II</td>
<td>Algebra Preliminaries</td>
<td>9</td>
</tr>
<tr>
<td>Chapter III</td>
<td>Differential Algebra</td>
<td>13</td>
</tr>
<tr>
<td>Chapter IV</td>
<td>The Schwarzian Derivative</td>
<td>22</td>
</tr>
<tr>
<td>Chapter V</td>
<td>Ramification Theory</td>
<td>26</td>
</tr>
<tr>
<td>Chapter VI</td>
<td>The Hurwitz Formula</td>
<td>45</td>
</tr>
<tr>
<td>Chapter VII</td>
<td>Consequences of the Hurwitz Formula</td>
<td>50</td>
</tr>
<tr>
<td>Chapter VIII</td>
<td>Proof of the Main Theorem</td>
<td>60</td>
</tr>
<tr>
<td>Bibliography</td>
<td></td>
<td>73</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

The intention of this thesis is to give an up to date version of part of Klein's treatise [3], and to extend Klein's result from the complexes to any algebraically closed field. We consider the group $J$ of fractional linear transformations of a variable $x$

$$x \mapsto \frac{ax + b}{cx + d}$$

where the elements $a, b, c, d$ belong to an algebraically closed field $k$. That is, $\text{PSL}_2(k)$, the invertible $2 \times 2$ matrices over $k$ modulo their centre. Our aim is to determine all finite subgroups of $J$ having order not divisible by the characteristic of $k$. The following observation in the case where $k = \mathbb{C}$ serves to motivate our result in the general case:

For $k = \mathbb{C}$, $J$ is isomorphic to the group of analytic automorphisms of the Riemann sphere. We know that this group has as finite subgroups the cyclic groups, the dihedral groups, the tetrahedral group, the octahedral group and the icosahedral group.

It turns out that these groups also exist as finite subgroups of $J$ for general $k$, if their order is relatively prime to the characteristic of $k$. In fact, we shall show that these are essentially all such finite subgroups of $J$. 
As evidence of the existence of these canonical subgroups we give their generators. If the order of a subgroup is \( N \) and the characteristic of \( k \) is \( p \), assume that \( p \nmid N \).

For the cyclic and dihedral groups it is clear what the generators are.

I The Cyclic Group of order \( N \) is generated by

\[
\begin{pmatrix}
\varepsilon & 0 \\
0 & 1
\end{pmatrix}
\]

where \( \varepsilon \) is a primitive \( N \)th root of unity in \( k \).

II The Dihedral Group of order \( N = 2n \) is generated by

\[
\begin{pmatrix}
\varepsilon & 0 \\
0 & 1
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix}
\]

where \( \varepsilon \) is a primitive \( n \)th root of unity in \( k \).

For the three remaining canonical groups we derive the generators by considering \( k = \mathbb{C} \). However, using well-known properties of the groups it can easily be shown that these are also the generators for \( k \) any algebraically closed field.

Let \( G \) denote one of the three groups. View the figure \( F \) corresponding to \( G \) i.e. the tetrahedron, octahedron or icosahedron, as embedded in the Riemann sphere. Then elements of \( G \) are rotations of the sphere which transform \( F \) into itself. Every element of \( G \) must then be one of
the following: a rotation $R$ of order $i = 3, 4$ or $5$ about a vertex of $F$ ($i$ corresponding to the number of faces meeting at any vertex in the tetrahedron, octahedron or icosahedron), or a rotation $S$ of order 2 about the axis through the midpoint of an edge, or a rotation $T$ of order 3 about the axis through the midpoint of a face. Since elements of each of these three types are conjugate, to generate $G$ it suffices to give one element of each type. (1) We now give the generators $R$, $S$ and $T$ for each group. A rough stereographic projection of each related figure is shown as an aid to the reader interested in checking the generators.

III The Tetrahedral Group

Let $\omega$ be a primitive 3rd root of unity.

\[
\begin{align*}
R &= \begin{pmatrix} \omega & 0 \\ 0 & 1 \end{pmatrix}, \\
T &= \begin{pmatrix} -1 & \omega \\ 2 & \omega \end{pmatrix}, \\
S &= \begin{pmatrix} -1 & 1 \\ 2 & 1 \end{pmatrix}
\end{align*}
\]

(1) $G$ can in fact always be generated by two elements. See, for example §§71-74[5].
IV The Octahedral Group

Let \( \alpha \) be a primitive 8th root of unity.

\[
\begin{align*}
R &= \begin{pmatrix} \alpha^2 & 0 \\ 0 & 1 \end{pmatrix}, \\
T &= \begin{pmatrix} \alpha^3 & 1 \\ 1 & \alpha \end{pmatrix}, \\
S &= \begin{pmatrix} -\alpha & \alpha^2 \\ 1 & \alpha \end{pmatrix}
\end{align*}
\]

V The Icosahedral Group

Let \( \epsilon \) be a primitive 5th root of unity.

Note: \( \epsilon + \epsilon^{-1} - 1 = \frac{-3 + \sqrt{5}}{2} \)
With existence of these subgroups of $\delta$ accepted, we can now state our proposal.

It turns out that on the stereographic projection of the icosahedron, $r = \frac{-3+\sqrt{5}}{2}$ is the distance from the origin of the five outer icosahedron vertices.
MAIN THEOREM

Let \( k \) be an algebraically closed field of characteristic \( p \). Any finite subgroup \( G \) of \( \text{PSL}_2(k) \) for which \( p \nmid |G| \) is isomorphic to one of the canonical groups I-V. In fact, if \( G \) is isomorphic to one of I-IV then \( G \) is actually conjugate to that group. Further, in characteristic 0 all finite subgroups are conjugate to one of I-V.

This result immediately gives the following specialization.

Let \( F \) be a finite subfield of \( k \) containing \( p^n \) elements. Any subgroup of \( \text{PSL}_2(F) \) whose order is not divisible by \( p \) is either conjugate over \( k \) to one of I-IV or is isomorphic to V.

This result is stated in a stronger form in Dickson [1] (§256). Dickson's treatment of this matter deals with all finite subgroups of \( \text{PSL}_2(F) \), not only those with order prime to \( p \). Our reasons for avoiding the most general case will become clear as our technique for proof develops. Note here that our methods will be designed so as to handle any algebraically closed field \( k \), not finite fields.

We now outline our plan for accomplishing the proof. Our first step is to observe that the group of fractional linear transformations of \( x \) over \( k \) is isomorphic to \( \text{Aut}_k k(x) \),
the group of automorphisms of \( k(x) \) over \( k \). We establish
this fact now and then choose to work with \( \text{Aut}_k^*k(x) \)
thereafter.

**Theorem**

\[ k(x) = k(y) \iff y = \frac{ax + b}{cx + d} \]

for some elements \( a, b, c, d \in k \) such that \( ad - bc \neq 0 \).

**Proof**

( \( \iff \) ) \( k(y) \subseteq k(x) \) is clear. To show \( k(x) \subseteq k(y) \) we
must determine \( x \) in terms of \( y \). Rewrite \( y = \frac{ax + b}{cx + d} \) in
matrix form as

\[
\begin{pmatrix}
  y \\
  1
\end{pmatrix} =
\begin{pmatrix}
  a & b \\
  c & d
\end{pmatrix}
\begin{pmatrix}
  x \\
  1
\end{pmatrix}.
\]

From this we have

\[
\begin{pmatrix}
  x \\
  1
\end{pmatrix} = \begin{pmatrix}
  a & b \\
  c & d
\end{pmatrix}^{-1} \begin{pmatrix}
  y \\
  1
\end{pmatrix} = \frac{-1}{ad - bc} \begin{pmatrix}
  -d & b \\
  c & -a
\end{pmatrix} \begin{pmatrix}
  y \\
  1
\end{pmatrix},
\]

where this is possible since \( \det(\begin{pmatrix}
  a & b \\
  c & d
\end{pmatrix}) = ad - bc \neq 0 \). Hence

\[ x = \frac{-dy + b}{cy - a}, \]

and therefore \( k(x) \subseteq k(y) \).

( \( \rightarrow \) ) \( k(x) = k(y) \) implies, in particular, that \( y \in k(x) \).
So \( y = \frac{A(x)}{B(x)} \) for some \( A(x), B(x) \in k[x] \), which we may assume
to be relatively prime.

Let \( I(X) = B(X)y - A(X) \). \( I \) is a non-zero polynomial
satisfied by \( x \) over \( k(y) \). Further, \( I \) is irreducible over
\( k(y) \): otherwise, it would be possible to factor an irreducible
polynomial from each of the relatively prime polynomials
A(X) and B(X). Thus $I(X) = \text{Irr}(x, k(y), X)$. Hence, by field theory, $[k(x) : k(y)] = \deg I(X)$. But $[k(x) : k(y)] = 1$, and $\deg I(X) = \max (\deg A(X), \deg B(X))$. Thus $A(X) = ax + b$ and $B(X) = cx + d$ for some elements $a, b, c, d \in k$ for which $ad - bc \neq 0$. (If $ad - bc = 0$, $(a \ b)$ and $(c \ d)$ would be $k$-multiples of each other and $y$ would therefore be in $k$.)

Hence, we have

$$y = \frac{A(x)}{B(x)} = \frac{ax + b}{cx + d}$$

with $a, b, c, d$ as stated.

That this establishes the desired result is clear. For, any element $\sigma \in \text{Aut}_k k(x)$ is such that $\sigma(k(x)) = k(\sigma(x)) = k(y)$, say; and any element of the group of fractional linear transformations of $x$ over $k$ acts by mapping $x \mapsto \frac{ax + b}{cx + d} = y$, say, for some elements $a, b, c, d \in k$ such that $ad - bc \neq 0$.

In considering finite subgroups of $\text{Aut}_k k(x)$ one naturally thinks of using Galois theory and field theory. With these algebraic techniques and the Hurwitz Formula (a result dependent on differential algebra and ramification theory for its derivation), we deduce the first two conclusions of the Main Theorem. For the final conclusion in characteristic 0, we invoke as well the so-called Schwarzian derivative (3), an operator in the theory of differential algebra.

(3) Our use of the Schwarzian derivative in this context follows Klein [3]. Here he uses the Schwarzian in finding the finite subgroups of $\text{PSL}_2(\mathbb{C})$. 
CHAPTER II  ALGEBRA PRELIMINARIES

Terminology:

By an algebraic function field over $k$ we shall mean a finitely generated extension field of $k$ of transcendence degree 1.

A rational function field is then an algebraic function field having only one generator.

For use in later sections we shall need three special results from the general theory of algebra.

**Theorem 1**

Let $x$ be a transcendental element over a field $k$. Let $G$ be a finite subgroup of $\text{Aut}_k k(x)$. Then there exists an element $y$ transcendental over $k$ such that

$$\text{Fix } G = k(y)$$

Note: This is a special case of Lüroth's Theorem which makes the same claim about any subfield of $k(x)$.

**Proof**

Let $G = \{\sigma_1, \sigma_2, ..., \sigma_n\} \subseteq \text{Aut}_k k(x)$ where $n = |G|$. Define $y = \sigma_1(x) \sigma_2(x)... \sigma_n(x)$. Clearly $y$ is transcendental over $k$. We will show that $y \in \text{Fix } G$ and that, in fact, $k(y) = \text{Fix } G$.

To show that $y \in \text{Fix } G$, we must show that $\sigma_i(y) = y$ for all $i = 1,...,n$. Fix some $i$, $1 \leq i \leq n$. Then
(1) \[ \sigma_i(y) = \sigma_i(\sigma_1(x) \ldots \sigma_n(x)) = (\sigma_i \sigma_1)(x) \ldots (\sigma_i \sigma_n)(x) \]

Clearly, \( \sigma_i \cdot \sigma_j \in G \) for \( j = 1, \ldots, n \). Further, if \( \sigma_j \neq \sigma_k \) then \( \sigma_i \cdot \sigma_j \neq \sigma_i \cdot \sigma_k \). For, \( \sigma_i \cdot \sigma_j = \sigma_i \cdot \sigma_k \) implies that

\[ \sigma_i^{-1} \cdot \sigma_i \cdot \sigma_j = \sigma_i^{-1} \cdot \sigma_i \cdot \sigma_k, \]

so that \( \sigma_j = \sigma_k \).

Hence \( G = \{ \sigma_i \cdot \sigma_1, \ldots, \sigma_i \cdot \sigma_n \} \). Thus from (1) we see that \( \sigma_i(y) = y \) and \( y \) is indeed in \( \text{Fix } G \).

We have now \( k(y) \notin \text{Fix } G \in k(x) \). Hence

(2) \[ [k(x):\text{Fix } G][\text{Fix } G:k(y)] = [k(x):k(y)] \]

By Galois theory we know that \( [k(x):\text{Fix } G] = |G| = n \).

From (2) then we have

(3) \[ [k(x):k(y)] > n \]

We complete the proof by showing that \( [k(x):k(y)] = n \).

This finishes the proof since from (2) it then follows that \( |\text{Fix } G:k(y)| = 1 \), and so, that \( \text{Fix } G = k(y) \).

In the Introduction we established that \( \text{Aut}_k k(x) \) is isomorphic to the group of fractional linear transformations. Thus each \( \sigma_i(x) \) is of the form \( \frac{a_i x + b_i}{c_i x + d_i} \), for some \( a_i, b_i, c_i, d_i \in k \) such that \( a_i d_i - b_i c_i \neq 0 \).

Then \( y = \frac{\sum_{i=1}^{n} \frac{a_i x + b_i}{c_i x + d_i}}{g(x)} = f(x) \), for some relatively prime
polynomials \( f(x), g(x) \in k[x] \) such that \( \deg f, \deg g \leq n \).

Let \( F(X) = g(X)y - f(X) \). \( F(X) \) is irreducible over \( k(y) \)
and \( F(x) = 0 \). i.e. \( F(X) = \text{Irr}(x, k(y), X) \). Hence
\[
[k(x):k(y)] = \deg F(X) = \max(\deg f, \deg g) < n.
\]
Together with (3) this gives the required result.

**Theorem 2** Existence of a Separating Transcendental

Let \( K \) be an algebraic function field over \( k \), where
\( k \) is of characteristic \( p \). Then there exists a transcendental
\( x \in K \) such that \( K/k(x) \) is separable.

**Proof**

Let \( x \in K \) be transcendental over \( k \). Thus we have
\( K \supset k(x) \supset k \) and \( K/k(x) \) algebraic. Suppose
\[
K = k(x, x_2, \ldots, x_n) = k(x_1, x_2, \ldots, x_n),
\]
where the \( x_2, \ldots, x_n \)
are algebraic over \( k \), and \( x = x_1 \). Use induction on \( n \) to show
existence of a separating transcendental.

If \( n = 1 \) then \( K = k(x) \), and this is clearly separable
over \( k(x_1) \).

Assume that \( k(x_2, \ldots, x_{n-1}) \) is separable over \( k(x_1) \).
Since \( x_n \) is algebraic over \( k(x_1) \), there exists an irreducible
polynomial \( f \) such that \( f(x_1, x_n) = 0 \), where the coefficients
of \( f \) are in \( k \). If \( x_n \) is separable over \( k(x_1) \) take the sep-
arating transcendental to be \( x_1 \). Assume that \( x_n \) is not
separable over \( k(x_1) \). Then \( f \) must be a polynomial in \( x_n^p \).
If \( x_1 \) is inseparable over \( k(x_n) \) then \( f \) is also a polynomial
in \( x_1^p \). Thus
\[ f(x_1, x_n) = g(x_1^p, x_n^p) = [g(x_1, x_n)]^p, \]
where \( g \) is some polynomial with coefficients in \( k \). But this implies that \( f \) is reducible. Contradiction. Hence \( x_1 \) must be separable over \( k(x_n) \).

Thus we have \( x_2, \ldots, x_{n-1} \) separable over \( k(x_1) \) and \( x_1 \) separable over \( k(x_n) \). So \( x_1, \ldots, x_{n-1} \) are all separable over \( k(x_n) \). Therefore, take \( x_n \) to be the separating transcendental.

The following is a standard theorem of algebra, a proof of which can be found on p. 171 [4].

**Theorem 3**

Let \( \Lambda \) be an algebraically closed extension of a field \( k \), \( K \) an algebraic extension of \( k \). If \( \mu: k \rightarrow \Lambda \) is a monomorphism then \( \mu \) can be extended to a monomorphism of \( K \) into \( \Lambda \).
CHAPTER III  DIFFERENTIAL ALGEBRA

In the chapter on the Schwarzian derivative we will deal with derivations and will need one general result concerning them. This result is found in the lemma of this section. The theorem following the lemma states that under suitable conditions derivations can be extended. This is used in the proof of the Main Theorem. The latter part of this chapter deals with differentials and related results which are necessary for the development of the Hurwitz Formula.

(1) Derivations

Let $k$ be a commutative ring, $K$ a commutative $k$-algebra and $V$ a $K$-module. Then a $k$-linear map $D:K \rightarrow V$ such that

$$D(fg) = fDg + gDf$$

for any elements $f, g \in K$ is called a $k$-derivation.

Note: This definition will be used in the case where $k$ and $K$ are both fields. Assume in the following then that we have: $k$ a field, $K$ an extension field of $k$, $D:K \rightarrow V$ a $k$-derivation of $K$ into a $K$-module $V$.

We define the field of constants for $D$ to be

$$K_D = \{ f \in K \mid Df = 0 \}.$$  Notice that $k \subseteq K_D$. For, if $a \in k$, $D(a) = D(a \cdot 1) = a \cdot D(1) = a \cdot 0 = 0$. But it is not necessarily true that $K_D \subseteq k$. For, if char $k = p$ and $f \in K$ then $f^p \in K_D$ since $D(f^p) = pf^{p-1}Df = 0$. But $f^p$ need not be in $k$.

Also notice that $D$ is linear over $K_D$. For, if $f \in K_D$ and $g \in K$ then $D(fg) = fDg + gDf = fDg$.

Note: For $f \in K_D(x)$, $Df = D_x f \cdot Dx$ where $D_x$ denotes "formal
differentiation with respect to $x$.

Denote by $D^i(f)$, for $f \in K$, the element
$$D(D(...(D(f))...)).$$
Then we have the following

**Lemma**

Let $K$ be a field extension of $k$, and let $D:K \to K$ be a $k$-derivation with field of constants $K_D$. The elements $f_1, f_2, ..., f_n \in K$ are linearly dependent over $K_D$ if and only if $\det (D^i f_j) = 0$, where $i = 0, 1, ..., n-1$ and $j = 1, 2, ..., n$.

**Proof**

Assume $f_1, ..., f_n$ are linearly dependent over $K_D$. Then, without loss of generality, there exist $\lambda_2, ..., \lambda_n \in K_D$ such that
$$f_1 = \lambda_2 f_2 + ... + \lambda_n f_n.$$  
Since $D$ is linear over $K_D$ it follows that
$$D^i f_1 = \lambda_2 D^i f_2 + ... + \lambda_n D^i f_n$$
for all $i=1, ..., n-1$. The columns of the matrix $(D^i f_j)$ are therefore linearly dependent. Hence
$$\det (D^i f_j) = 0.$$  

Now assume that $\det (D^i f_j) = 0$ and use induction to show that the $f_j$ must be linearly dependent over $K_D$.

For $n = 2$, we have $f_1, f_2 \in K$. Assume both to be nonzero since otherwise they are linearly dependent. Hence we have
$$\begin{vmatrix} f_1 & f_2 \\ Df_1 & Df_2 \end{vmatrix} = 0.$$  
This implies that the columns are linearly dependent over $K$.  

Thus there exists $\lambda \in K$ such that

\begin{align*}
(1) \quad f_1 &= \lambda f_2 \\
(2) \quad Df_1 &= \lambda Df_2
\end{align*}

Clearly we will be done if we can show $\lambda \in K_D$. Applying $D$ to (1) we have $Df_1 = \lambda Df_2 + f_2 D\lambda$. Together with (2) this yields $f_2 D\lambda = 0$. But $f_2 \neq 0$, so that $D\lambda$ must be $0$. Hence $\lambda \in K_D$.

Assume now that $\det (D^j f_j) = 0$ for $j \leq n-1$ implies that the $f_j$ are linearly dependent. We will show this result true for $j = n$.

Let $f_1, \ldots, f_n \in K$ and assume none is $0$. Then assuming $\det (D^j f_j) = 0$, there exist $\lambda_2, \ldots, \lambda_n \in K$ such that

\begin{align*}
(1) \quad f_1 &= \lambda_2 f_2 + \ldots + \lambda_n f_n \\
(2) \quad Df_1 &= \lambda_2 Df_2 + \ldots + \lambda_n Df_n \\
& \quad \vdots \\
(n) \quad D^{n-1} f_1 &= \lambda_2 D^{n-1} f_2 + \ldots + \lambda_n D^{n-1} f_n
\end{align*}

As for the case $n = 2$ we show that the $\lambda_i \in K_D$. Applying $D$ to (1) gives

$$Df_1 = (D\lambda_2 f_2 + \ldots + D\lambda_n f_n) + (\lambda_2 Df_2 + \ldots + \lambda_n Df_n)$$

Subtracting (2) from this gives

$$D\lambda_2 f_2 + \ldots + D\lambda_n f_n = 0$$

Similarly, applying $D$ to (2) we have

$$D^2 f_1 = (D\lambda_2 Df_2 + \ldots + D\lambda_n Df_n) + (\lambda_2 D^2 f_2 + \ldots + \lambda_n D^2 f_n)$$

Subtracting (3) yields

$$D\lambda_2 Df_2 + \ldots + D\lambda_n Df_n = 0$$
Repeating this operation \( n-1 \) times in all we obtain

\[
(1') \quad D\lambda_2 \cdot f_2 + \ldots + D\lambda_n \cdot f_n = 0
\]

\[
(2') \quad D\lambda_2 \cdot Df_2 + \ldots + D\lambda_n \cdot Df_n = 0
\]

\[
\vdots
\]

\[
(n-1') \quad D\lambda_2 \cdot D^{n-2}f_2 + \ldots + D\lambda_n \cdot D^{n-2}f_n = 0
\]

This system can be written as

\[
\begin{pmatrix}
Df_2 & \ldots & Df_n \\
D^2f_2 & \ldots & D^2f_n \\
\vdots & \ddots & \vdots \\
D^{n-2}f_2 & \ldots & D^{n-2}f_n
\end{pmatrix}
\begin{pmatrix}
f_2 \\
f_n \\
\vdots \\
f_n
\end{pmatrix}
= \begin{pmatrix}
0 \\
0 \\
\vdots \\
0
\end{pmatrix}
\]

If \( \det (D^1f_j) = 0 \) for \( j = 2, \ldots, n \) then by the induction assumption we are done. Assuming therefore that \( \det (D^1f_j) \neq 0 \) for \( j = 2, \ldots, n \), we find from the above system that

\[
(D\lambda_2, \ldots, D\lambda_n) = 0.
\]

But then \( D\lambda_i = 0 \) for \( i = 2, \ldots, n \), so that each \( \lambda_i \in K_D \). Thus \( f_1, \ldots, f_n \) are linearly dependent over \( K_D \), as desired.

**Theorem**

Extension of a Derivation

If a finitely generated extension \( K(x) \) over \( K \) is separably algebraic then any derivation \( D \) on \( K \) can be uniquely extended to \( K(x) \).

**Proof**

Since \( D \) is defined on \( K \), to extend \( D \) to \( K(x) \) we need only define \( D(x) \).

Let \( f(x) = \text{Irr}(x, K, X) = a_n X^n + \ldots + a_1 x + a_0 \), for \( a_i \in K \). Then \( f(x) = a_n X^n + \ldots + a_1 x + a_0 = 0 \), and applying any derivation \( d \) to this gives
Thus we must define $Dx$ by

$$Dx = -\frac{f^d(x)}{f'(x)}$$

This defines the only possible extension of $D$ to $K(x)$. Note: division by $f'(x)$ in the definition of $Dx$ makes sense since $f'(x) \neq 0$ for $x$ separable over $K$.

(2) Differentials

Let $K$ be an algebraic function field over $k$. Define $D: K \to \Omega_{K/k}$ to be the universal $k$-derivation in the following sense:

For any $k$-derivation $D: K \to W$, $W$ a $K$-module, there exists a unique $K$-module map $\mu: \Omega_{K/k} \to W$ such that the following diagram commutes

$$
\begin{array}{ccc}
K & \xrightarrow{d} & \Omega_{K/k} \\
\downarrow{D} & & \downarrow{\mu} \\
W & & 
\end{array}
$$

Once we have existence of such a universal pair $(d, \Omega_{K/k})$, being defined by the universal mapping property, it must be unique up to isomorphism.

Existence of $(d, \Omega_{K/k})$:

Take $x$ to be a separating transcendental in $K/k$. (Such $x$ exists by Theorem 2, Chapter II). Thus $K = k(x, \theta)$ for some separable element $\theta \in K$. Now in $k(x)$, formal differentiation with respect to $x$ defines a derivation $d_x$
on $k(x)$. Then, $K$ being separable over $k(x)$, $d_x$ can be extended to give a derivation $d$ on $K$.

Take $\Omega_{K/k} = Kd\theta$.

Given $D: K \rightarrow W$ as above, define $\mu: \Omega_{K/k} = Kd\theta \rightarrow W$ by $\mu(f \cdot d\theta) = f \cdot D\theta$, for $f \in K$. Denoting by $D_\theta$ formal differentiation with respect to $\theta$, we then have for $f \in K$,

$$(\mu \circ d)(f) = \mu(df) = \mu(D_\theta f \cdot d\theta) = D_\theta f \cdot D\theta,$$

while $Df = D_\theta f \cdot D\theta$. Hence $\mu \circ d = D$, as desired.

The $K$-module $\Omega_{K/k}$ is known as the module of differentials for the extension $K/k$.

We now define what we shall mean by a continuous differential. In particular, we shall be interested in continuous differentiation in a field of formal power series.

Let $K = k((t))$ be a field of formal power series. $K$ has a valuation "ord" defined on it by

$$\text{ord}(f) = \text{order of } f \text{ as a function of } t,$$

for any $f \in K$. This valuation ord defines a metric on $K$ in the usual way.

A topological $K$-space is a vector space $W$ with topology in which addition

$$(i) \quad W \times W \rightarrow W$$

and scalar multiplication

$$(ii) \quad K \times W \rightarrow W$$

are continuous.

Define $d: K \rightarrow \Omega_{K/k}^c$ to be the universal continuous
For any continuous $k$-derivation $D: K \to W$, $W$ a topological $K$-space, there exists a unique $K$-module map $\mu$ such that the following diagram is commutative:

\[
\begin{array}{c}
K \\
\downarrow \quad d \\
\Omega^c_{K/k} \\
\downarrow \quad \mu \\
W
\end{array}
\]

For any field of formal power series $K = k((t))$ we will presently show that such a universal pair $(d, \Omega^c_{K/k})$ exists; again, once we have existence, uniqueness is automatic.

Existence of $(d, \Omega^c_{K/k})$:

Set $\Omega^c_{K/k} = K dt$, where $dt$ is the generator of this 1-dimensional module. Define $d: K = k((t)) \to \Omega^c_{K/k}$ by $d: t \mapsto dt$.

Notice that $Df = f' \cdot Dt$ for any $f \in K$ where $f'$ denotes the formal derivative of $f$ with respect to $t$, and $D: K \to W$ is as above.

Proof:

Given $f \in K$, by continuity we can approximate $f$ by functions $f_n \in k[t, t^{-1}]$. Then, since for the $f_n$ we know that $Df_n = f_n' \cdot Dt$,

$$Df = \lim_{n \to \infty} Df_n = \lim_{n \to \infty} (f_n' \cdot Dt)$$

But by (ii) of the topological $K$-space definition

$$\lim_{n \to \infty} (f_n' \cdot Dt) = (\lim_{n \to \infty} f_n') \cdot Dt$$
\textbf{20.}

\begin{align*}
\text{Hence} \quad Df &= (\lim_{n \to \infty} f'_n) \cdot Dt = f' \cdot Dt. \\
\text{Then the universal mapping property follows easily.}
\end{align*}

Define \( \mu : \Omega^c_{K/k} = Kdt \to W \) by \( \mu : f \cdot dt \to f' \cdot Dt \), for \( f \in K \).

Then \( Df = f' \cdot Dt \) and \( (\mu \circ d)(f) = \mu(df) = \mu(f' \cdot dt) = f' \cdot Dt \).

(We have \( df = f' \cdot dt \), as for \( D \), since \( d \) is continuous.)

That is, \( D = \mu \circ d \), as required.

We call \( \Omega^c_{K/k} \) the module of continuous differentials of \( K/k \).

Next we define the order of a continuous differential.

Let \( \omega \in \Omega^c_{K/k} \). Since \( \Omega^c_{K/k} = k((t))dt \), we can write \( \omega \) as \( f \cdot dt \) for some element \( f \in k((t)) \). Then define the order of \( \omega \) to be the order of the function \( f \). Write

\[ \text{ord}(\omega) = \text{ord}(f) \]

Remark: Let \( t' \) be a power series in \( t \) of order 1. Then any \( f \in K = k((t)) \) can be rewritten as a power series in \( t' \). So \( K = k((t')) \). For \( \text{ord}(\omega) \) to be well-defined, it must be independent of the generator chosen for \( K \). We must therefore show that the following holds.

Let \( t, t' \) be generators of \( K \) over \( k \). For \( \omega \in \Omega^c_{K/k} \) we have \( \omega = f \cdot dt \) and \( \omega = g \cdot dt' \) where \( f \in k((t)) \) and \( g \in k((t')) \). Then \( \text{ord}(f) = \text{ord}(g) \)

Proof:

\[ t' \in K = k((t)) \]. Hence we can write \( t' = u(t) \cdot t \) for \( u(t) \) a unit in \( k((t)) \). Then, letting \( D_t \) denote formal
differentiation with respect to \( t \), we have

\[
d t' = d(u(t)) \cdot t + u(t) \cdot dt
\]
\[
= D_t(u(t)) \cdot dt \cdot t + u(t) \cdot dt
\]
\[
= (D_t(u(t)) \cdot t + u(t)) \cdot dt
\]
\[
= u_0(t) \cdot dt, \text{ say.}
\]

It is easy to see that \( u_0(t) \) is a unit in \( k((t)) \).

For, \( u(t) \) a unit implies \( \text{ord} (u(t)) = 0 \) and \( \text{ord} (D_t u(t)) \geq 0 \).

Thus \( \text{ord} (D(u(t)) \cdot t) \geq 1 \), so that \( \text{ord} (u_0(t)) = \text{ord} (D(u(t)) \cdot t + u(t)) = 0 \).

We now have \( \omega = g \cdot dt' = (g \cdot u_0(t)) \cdot dt \). But also

\( \omega = f \cdot dt \). Hence \( f = g \cdot u_0(t) \). Therefore

\[
\text{ord} (f) = \text{ord} (g \cdot u(t))
\]
\[
= \text{ord} (g) + \text{ord} (u_0(t))
\]
\[
= \text{ord} (g) + 0
\]
\[
= \text{ord} (g)
\]
CHAPTER IV
THE SCHWARZIAN DERIVATIVE

Definition of the Schwarzian Derivative:

Let \( D : E \to E \) be a derivation on a field \( E \), with field of constants \( k \). For \( h \in E - k \) define the Schwarzian Derivative of \( h \) with respect to \( D \) to be

\[
[h]_D = 2\frac{h'''}{h'} - 3\left(\frac{h''}{h'}\right)^2
\]

where \( h' = Dh, h'' = D^2h \) and \( h''' = D^3h \).

Note: Since \( h \in E - k \), \( h' \) is non-zero and therefore \([h]_D\) makes sense.

Theorem

Let \( E \) be a field on which there is defined a derivation \( D : E \to E \). Let \( k \) be the field of constants for \( D \). Then for \( f, g \in E - k \) there exist \( a, b, c, d \in k \) with \( \text{ad-bc} \neq 0 \) such that

\[
f = \frac{ag + b}{cg + d} \iff [f]_D = [g]_D
\]

Proof

Assuming \( f = \frac{ag + b}{cg + d} \) for some \( a, b, c, d \in k \) such that \( \text{ad-bc} \neq 0 \) (1), we see that this is equivalent to

\[
cfg - ag + df - b = 0
\]

(1) Applying \( D \) to this gives the equation

\[
c(f'g + fg') - ag' + df' = 0
\]

(2) This says that \( h_1 = f'g + fg' \), \( h_2 = -g' \) and \( h_3 = f' \) are linearly dependent over \( k \). By the lemma on the Wronskian

(1) \( \text{ad-bc} \neq 0 \) since otherwise \( f \) would be in \( k \), and so \([f]_D\) would not be defined.
(cf. §1, Chapter III) we know that this is equivalent to
\[ \det (D^i h_j) \] being 0. Hence
\[ \det (D^i h_j) = \begin{vmatrix} f'g + fg' & -g' & f' \\ f''g + 2f'g' + fg'' & -g'' & f'' \\ f'''g + 3f''g' + 3f'g'' + fg''' & -g''' & f''' \end{vmatrix} = 0 \]
Adding col 1 plus f(col 2) plus -g(col 3) gives
\[ \begin{vmatrix} 0 & -g' & f' \\ 2f'g' & -g'' & f'' \\ 3f''g' + 3f'g'' & -g''' & f''' \end{vmatrix} = 0 \]
Or, expanding the determinant,
\[ -2f'g'(-g'f''+f'g''')+3(f''g'+3f'g'')(\cdot f''g'+f'g'') = 0 \]
Multiplying this out we have
\[ 2f'f''(g')^2-2(f')^2g'g''-3(f'')^2(g')^2 = 0 \]
Simplifying this gives
\[ (f')^2[-2g'g''+3(g'')^2]+(g')^2[2f'f''-3(f'')^2] = 0 \]
Dividing through by \((f'g')^2\) and transposing one of the two expressions,
\[ \frac{2f'f''}{(f')^2} - 3 \left( \frac{f''}{g'} \right)^2 = \frac{2g'g''}{(g')^2} - 3 \left( \frac{g''}{g'} \right)^2 \]
Or,
\[ \frac{2f''}{f'} - 3 \left( \frac{f''}{f'} \right)^2 = \frac{2g''}{g'} - 3 \left( \frac{g''}{g'} \right)^2 \]
That is, \([f]_D = [g]_D\)
Remark: Our proof is a series of equivalences except at sentences (1) and (2).
Here, assuming \([f]_D = [g]_D\) and following the proof back
to sentence (2) we know that there exist $a, c, d \in k$ such that
\[ c(f'g + fg') - ag' + df' = 0 \]
Recall that this simply means that
\[ D(cfg - ag + df) = 0 \]
So, by definition of $k$, $cfg - ag + df = b$ for some element $b \in k$. Thus
\[ f = \frac{ag + b}{cg + d} \]
for some $a, b, c, d \in k$ such that $ad-bc \neq 0$.

The conclusion of this theorem can be restated in a more useful form as

**Corollary 1**

\[ k(f) = k(g) \iff [f]_D = [g]_D \]

**Proof**

This statement follows directly from the above theorem and the theorem near the end of Chapter I.

Remark: Recall that we are going to need the Schwarzian for proving the Main Theorem in characteristic 0. What enables us to make use of the preceding theorem on the Schwarzian in characteristic 0 is the following fact:

If the field $E$ of the theorem is an algebraic function field over an algebraically closed field $k_0$ of characteristic 0, then the field of $D$-constants is $k^0$.

This holds since in characteristic 0 any transcendental is separating; if $D$ is zero on a separating transcendental then $D$ can only be the trivial derivation.
For non-zero characteristic however, the D-constants may form a non-trivial extension of \( k_0 \). This is the fact which makes the Schwarzian of no use to us in the general case.

The trivial part of the preceding theorem yields the next result.

**Corollary 2**

Let \( H \) be a finite group of automorphisms of \( k(x) \) where \( x \) is an indeterminate over \( k \). Let \( F \) be the fixed field of \( H \) in \( k(x) \). Suppose \( D \) is a derivation on \( F \) with field of constants \( k \). We know that \( D \) extends uniquely to \( k(x) \) (since \( k(x)/F \) is separable algebraic). Then the Schwarzian of \( x \) lies in \( F \). That is,

\[
[x]_D \in F
\]

**Proof**

Let \( \mu \in H = \text{Gal} \ (k(x)/F) \). To show \( [x]_D \in F \) we need only show that \( \mu([x]_D) = [x]_D \). Since \( \mu \) is an automorphism of \( k(x) \), it follows that

\[
\mu([x]_D) = [\mu(x)]_{\mu D}
\]

Now \( \mu D : k(x) \to k(x) \) is a derivation on \( k(x) \), and \( \mu D|_F = D \). Thus, the extension of \( D \) to \( k(x) \) being unique, \( \mu D = D \). Therefore

\[
\mu([x]_D) = [\mu(x)]_D
\]

Now \( \mu \in \text{Gal} \ (k(x)/F) \) implies that \( k(x) = k(\mu(x)) \). But by Corollary 1 this is equivalent to

\[
[x]_D = [\mu(x)]_D
\]

Hence \( \mu([x]_D) = [x]_D \) as required.
Throughout this chapter assume $k$ to be an algebraically closed field, $K$ to be an algebraic function field over $k$. By Theorem 2, Chapter II we have existence of a separating transcendental in $K/k$. Thus we can assume $K$ to be a separable extension of a rational function field $k(x)$, for some transcendental $x \in K$. For our purposes the field $k(x)$ and its extensions are the main source of interest. However, where possible without additional work we state the following theory in the general case.

(1) **Places**

The origin of this theory is in the theory of meromorphic functions on a Riemann surface. Places are introduced here as the concept which supersedes the notion of point on the Riemann surface. It will be convenient in the main text of our argument to be able to view a place in a field as either of two notions. We define each of these here and then note that they are in fact equivalent.

One concept of a place will be as a valuation ring. Let $\mathcal{O}$ be a domain which is not a field. $\mathcal{O}$ is called a valuation ring if there is an irreducible element $\pi \in \mathcal{O}$ such that every $z \in \mathcal{O} - \{0\}$ can be written uniquely in the form $z = u\pi^n$, for some unit $u \in \mathcal{O}$ and some non-negative integer $n$. An element $\pi$ as above is called a local
uniformizing parameter for \( \mathcal{O} \); any other local uniformizing parameter is of the form \( u \pi \), \( u \) a unit in \( \mathcal{O} \). Let \( L \) be the quotient field of \( \mathcal{O} \). Then, for fixed local uniformizing parameter \( \pi \), any \( z \in L - \{ 0 \} \) has a unique expression \( z = u \pi^n \), \( u \) a unit in \( \mathcal{O} \) and \( n \in \mathbb{Z} \). The exponent \( n \) is easily seen to be independent of the particular uniformizing parameter \( \pi \). We call the exponent \( n \) the order of \( z \) and write \( n = \text{ord}(z) \). Define \( \text{ord}(0) = \infty \). Then \( \mathcal{O} = \{ z \in L | \text{ord}(z) \geq 0 \} \) and \( \mathcal{P} = \{ z \in L | \text{ord}(z) > 0 \} \) is the unique maximal ideal of \( \mathcal{O} \).

Our second notion of a place will be as a valuation.

A valuation on a field \( L \) is a function \( v: L \to \mathbb{Z} \cup \{ \infty \} \) satisfying

\[
\begin{align*}
(1) & \quad v(a) = \infty \text{ iff } a = 0 \\
(2) & \quad v(ab) = v(a) + v(b) \\
(3) & \quad v(a+b) \geq \min(v(a), v(b))
\end{align*}
\]

Any such valuation defines a valuation ring in \( L \). For, \( \mathcal{O} = \{ z \in L | v(z) \geq 0 \} \) is a valuation ring with maximal ideal \( \mathcal{P} = \{ z \in L | v(z) > 0 \} \) and quotient field \( L \). Conversely, a valuation ring in \( L \) defines a valuation on \( L \). Namely, if \( \mathcal{O} \) is a valuation ring with quotient field \( L \), then the function \( \text{ord}: L \to \mathbb{Z} \cup \{ \infty \} \) is a valuation on \( L \). Thus giving a valuation ring with quotient field \( L \) is the same as defining a valuation on \( L \). Since these are equivalent notions, we will refer to them both as places. In fact, since specifying the maximal ideal of a valuation ring
determines the ring, we will normally use the following notation:

by a place \( \mathfrak{p} \) of a field \( L \) we denote the maximal ideal of the valuation ring associated with \( \mathfrak{p} \). Denote the valuation ring itself by \( \mathcal{O} \) (or \( \mathcal{O}_x \) if necessary), and the associated valuation by \( \operatorname{ord}_x \). For uniformity call \( \mathcal{O} \) the place ring of \( L \).

Given two fields \( L_1 \) and \( L_2 \), \( L_2 \) a finite extension of \( L_1 \), we will want to be able to extend a given place on \( L_1 \) to a place on \( L_2 \). This process can actually be carried out for \( L_2 \) any extension of \( L_1 \); the reader is referred for proof of this fact to p. 299 [4]. To denote that a place \( \mathfrak{q} \) is an extension of a place \( \mathfrak{p} \), we write \( \mathfrak{p} \mid \mathfrak{q} \), read \( \mathfrak{p} \) lies over \( \mathfrak{q} \).

Consider now the case \( L = k(x) \). To see what the places are in this field we need the following definition. Say a function \( f \in k(x) \) is regular at a point \( a \in k \) if \( f \) can be written as \( \frac{g}{h} \) for some \( g, h \in k[x] \) such that \( h(a) \neq 0 \). For each \( a \in k \), define \( \mathcal{O}_a = \{ f \in k(x) | f \text{ regular at } a \} \). Then it is fairly easy to see that \( \mathcal{O}_a \) is a valuation ring with local uniformizing parameter \( \pi = x - a \) and quotient field \( k(x) \). Also, \( \mathcal{O}_\infty = \{ \frac{f}{g} \in k(x) | \deg g \geq \deg f \} \) is a valuation ring with local uniformizing parameter \( \frac{1}{x} \) and quotient field \( k(x) \). In fact, these rings are the only valuation rings which contain \( k \) and have quotient field \( k(x) \).
So we have the following correspondence.

\[ \{a \in k\} \cup \infty \leftrightarrow \{\mathcal{O}_a | a \in k\} \cup \mathcal{O}_\infty \leftrightarrow \]

{Valuation rings $\mathcal{O}|\mathcal{O} \supset k$ and quotient field of $\mathcal{O}$ is $k(x)$}

Any valuation on $k(x)$ is thus associated to a unique ring $\mathcal{O}_a$ in $k(x)$ ($a$ in $k \cup \infty$). The places in $k(x)$ therefore correspond to local uniformizing parameters $\pi = x-a$, $a \in k$ and $\pi = \frac{1}{x}$. We shall say $\pi$ is a local uniformizing parameter for $\mathcal{O}$ if it is a parameter for the valuation ring $\mathcal{O}_a$ defined by $\mathcal{O}$.

The following result will be useful later in connection with places on $k(x)$.

**Theorem**

Any element of $k(x)$ has an equal number of zeros and poles.

**Proof**

Let $h(x) = \frac{f(x)}{g(x)}$ be an arbitrary element of $k(x)$ where $f$, $g \in k[x]$. Since $k$ is algebraically closed we can factor $f$ and $g$ into linear factors. The zeros of $h$ in $k$ will then be given by the factors of $f$. The number of zeros in $k$, with multiplicities, will therefore be $\deg f = n$, say. Similarly, the number of poles of $h$ in $k$, with multiplicities, will be $\deg g = m$, say. The behaviour of $h$ at $\infty$ balances out the number of zeros and poles. For, if $n = m$, $h$ has no pole or zero at $\infty$ and the proof is complete. If $n < m$, $h$ has a zero of order $m-n$ at $\infty$, so that the number of zeros is $n + (m-n) = m$. Similarly,
if \( n > m \), we find that the number of poles is \( m + (n - m) = n \).

Now we make a short remark about the connection of places with integrality.

Let \( L \) be an extension of \( K \). Suppose \( \mathfrak{q} \) is a place of \( K \) with place ring \( \mathcal{O} \), and \( \mathfrak{p} \) is a place of \( L \) lying over \( \mathfrak{q} \). Let \( \mathcal{O} \) be the place ring of \( \mathfrak{p} \). If \( a \) is an element of \( L \) satisfying an integral equation over \( \mathcal{O} \), i.e. an equation of the form

\[
a^m + a_{m-1}a^{m-1} + \ldots + a_0 = 0, \quad a_i \in \mathcal{O},
\]

then \( a \) must be in \( \mathcal{O} \).

Proof:

Suppose \( \mathcal{O} \) has local uniformizing parameter \( \pi \), and assume \( a \) has a pole at \( \pi \) of order \( n \). Then \( \text{ord}_{\mathfrak{p}}(a) = -n \).

Thus, \( \text{ord}_{\mathfrak{p}}(a^m) = -mn \) while \( \text{ord}_{\mathfrak{p}}(-a_{m-1}a^{m-1} - \ldots - a_0) \geq -(m-1)n = -mn + n \). But since \( a \) satisfies \( a^m = -a_{m-1}a^{m-1} - \ldots - a_0 \), this cannot be true. Hence \( a \) must have \( \text{ord}_{\mathfrak{p}}(a) \geq 0 \).

\( a \) is therefore in \( \mathcal{O} \).

We show now how places give a relation between the extension field \( K \) of \( k \) and a field of formal power series. We will see that each place of \( K \) gives rise to an injection of \( K \) into a suitable power series field.

Let \( \mathfrak{p} \) be a place in \( K \). As usual, denote by \( \mathcal{O} \) the place ring defined by \( \mathfrak{p} \), and by \( \pi \) a local uniformizing parameter of \( \mathfrak{p} \). We want to show that to any element \( z \in \mathcal{O} \) there is associated a unique power series in the local uniformizing parameter \( \pi \).
Let $z \in \mathcal{O}$. Then $z$ is congruent modulo $\mathfrak{p}$ to a unique element $\lambda_0 \in k$. (Obtain $\lambda_0$ through the isomorphism $k \to \mathcal{O} \to \mathcal{O}/\mathfrak{p}$.) Thus $z - \lambda_0 = z_1 \pi$ for some $z_1 \in \mathcal{O}$. Repeating this argument for $z_1$ and using induction on $n$, we obtain unique $\lambda_0$, $\lambda_1$, $\lambda_2$, ..., $\in k$ such that

$$z = \lambda_0 + \lambda_1 \pi + \lambda_2 \pi^2 + \ldots$$

Thus we have a map $z \mapsto \lambda_0 + \lambda_1 \pi + \lambda_2 \pi^2 + \ldots$ which defines an injection of $\mathcal{O}$ into $k[[\pi]]$, the ring of formal power series over $k$. This homomorphism extends in the obvious way to an injection of $K$ into $k((\pi))$, as desired.

Note: If $\mathfrak{p}$ is a place of $K$, then $\mathfrak{p}$ defines a metric on $K$ and we can then form the completion of $K$ with respect to this metric. Denote this completion by $K_{\mathfrak{p}}$. If $\pi$ is a local uniformizing parameter for $\mathfrak{p}$ then it is not difficult to see that $K_{\mathfrak{p}} = k((\pi))$. Hence the above map gives an injection of $K$ into its completion with respect to $\mathfrak{p}$.

(2) **Ramification**

In this section we discuss the interconnection between places and their extensions or restrictions.

Assume $L$ and $K$ are both algebraic function fields over $k$, and $L$ is a finite extension of $K$. Take a place $\mathfrak{p}$ in $L$. Denote by $\text{ord}_{\mathfrak{p}}|_K$ the restriction of $\text{ord}_{\mathfrak{p}}$ to $K$. The values of $\text{ord}_{\mathfrak{p}}|_K$ form some subgroup of $\mathbb{Z}$—not necessarily all of $\mathbb{Z}$, of course. In any case, there is a smallest positive integer $e$ occurring in this subgroup, and all values of
ord\(p\) on \(K\) will be multiples of \(e\). Call \(e\) the ramification index of \(p\). If \(e \neq 1\), say \(p\) is ramified. In future in referring to the restriction of \(\text{ord}_p\) to \(K\), we will mean the normalized restriction i.e. the valuation \(\frac{\text{ord}_p|_K}{e}\).

This normalization ensures that the restriction has smallest positive value 1. By the restriction of \(p\) to \(K\), we will mean the place of \(K\) defined by \(\frac{\text{ord}_p|_K}{e}\).

From this definition we deduce easily the following two results.

(1) Suppose \(p\) is a place in \(L\) with ramification index \(e\) and \(\overline{p}\) is the restriction of \(p\) to \(K\). Let \(\Pi, \pi\) be local uniformizing parameters for \(p, \overline{p}\) respectively. Then

\[
\pi = U\Pi^e,
\]

where \(U\) is some unit in \(\mathcal{O}_p\).

For, \(\Pi\) a uniformizing parameter for \(p\) implies that \(\text{ord}_p(\Pi) = 1\), and hence that \(\text{ord}_p(\Pi^e) = e \cdot 1 = e\). So \(\text{ord}_p(\Pi^e) = \frac{\text{ord}_p(\Pi^e)}{e} = 1\). Thus \(\Pi^e\) is an element of \(L\) having the same order as \(\pi\) with respect to \(\overline{p}\). This implies that \(\frac{\pi}{\Pi^e}\) must be a unit in \(\mathcal{O}_p\).

(2) Ramification indices multiply in a tower of fields.

Suppose \(L_2/L_1/K\) is a tower of fields and \(p\) is a place of \(L_2\). Let \(\overline{q}, \overline{r}\) be the restrictions of \(p\) to \(L_1\) and \(K\) respectively. Suppose \(q\) has ramification index \(e_1\) over \(\overline{q}\), \(p\) ramification index \(e_2\) over \(q\) and ramification index \(e\) over \(\overline{q}\). Then \(e = e_1 \cdot e_2\). The proof of this is obvious.
We now state some of the basic facts about ramification needed. Result I gives the major fact. The results following this are corollaries of I in the special case where the field extension is Galois.

In most texts on ramification theory, the result corresponding to our result I takes on a more complicated form. The simplicity in our case is due to the fact that \( k \) is algebraically closed.

I Let \( L/K \) be a finite separable field extension. Suppose \( \mathfrak{p}_0 \) is a place on \( L \), and \( \mathfrak{q} \) is the restriction of \( \mathfrak{p}_0 \) to \( K \). Consider the set of all places \( \mathfrak{p} \) on \( L \) lying over \( \mathfrak{q} \). Let \( e_\mathfrak{p} \) denote the ramification index of \( \mathfrak{p} \). Then

\[
\prod_{\mathfrak{p} | \mathfrak{q}} e_\mathfrak{p} = [L:K]
\]

Instead of giving a complete proof of this result, we shall merely indicate how it comes about that the ramification indices are connected to the degree of the extension. For a complete proof the reader is referred to Corollary 2, p. 308 [4].

Consider the place \( \mathfrak{q} \) on \( K \). We are interested in its extensions to \( L \). We can view this as a homomorphism-extension problem in the following manner.

Consider the homomorphism \( \psi: \mathcal{O} \to k \), where \( \mathcal{O} \) is the place ring of \( \mathfrak{q} \). (Actually, \( \psi: \mathcal{O} \to \mathcal{O}/\mathfrak{q} \to k \)) If we extend \( \psi \) to a subring of \( L \) which is as large as possible, this subring will be a place ring for an extension \( \mathfrak{p} \) of \( \mathfrak{q} \) in \( L \).
In fact, there is a 1-1 correspondence between \( \mathfrak{P} / \mathfrak{g} \) and such maximal extensions of \( \psi \). We view our problem therefore in terms of these extensions of \( \psi \).

Since \( L \) is separable over \( K \), we can choose a generator \( \theta \) for \( L \) over \( K \) such that \( \theta \) is integral over \( \mathfrak{o} \). Say \( \theta \) satisfies the equation

\[ F(X) = X^N + a_{N-1}X^{N-1} + \ldots + a_0 = 0 \]

where the \( a_i \in \mathfrak{o} \).

Consider \( \mathcal{O}_0 = \mathfrak{o}[\theta] = \mathfrak{o}(X)/F(X) \). The quotient field of \( \mathcal{O}_0 \) is \( L \). Thus, if we have extended \( \psi \) to \( \mathcal{O}_0 \), it is fairly easy to see that this determines a unique maximal extension of \( \psi \). Our problem therefore is to extend \( \psi \) to \( \mathcal{O}_0 \).

\( \psi \) is already defined on \( \mathfrak{o} \). To extend \( \psi \) to \( \mathfrak{o}[\theta] \), we must find the possible values for \( \psi(\theta) \). \( \theta \) satisfies \( F(X) = 0 \); hence \( \psi(\theta) \) must satisfy the equation

\[ \overline{F}(X) = X^N + \psi(a_{N-1})X^{N-1} + \ldots + \psi(a_0) = 0 \]

where the coefficients \( \psi(a_i) \in k \). Since \( k \) is algebraically closed, \( \overline{F}(X) \) can be factored into linear polynomials, say as

\[ \overline{F}(X) = (X-\lambda_1)^{e_1} \ldots (X-\lambda_r)^{e_r} \]

where the \( \lambda_i \in k \). But then we see that the possible values for \( \psi(\theta) \) are precisely the roots of \( \overline{F} \), namely, \( \lambda_1, \ldots, \lambda_r \).

This gives us the following information:

1. there always will be extensions of \( \psi \) (there are always roots of \( \overline{F}(X) = 0 \)), and
2. there are at most \( N \) extensions of \( \psi \) (\( r < \deg \overline{F} = N \)).
Clearly there are ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ of $\mathcal{O}_0$ lying over $\mathfrak{p}$, namely $\mathfrak{p}_i = (\theta - \lambda_i)\mathcal{O}_0$. The ramification index of $\mathfrak{p}_i$ is the largest power of $(\theta - \lambda_i)\mathcal{O}_0$ containing $\mathfrak{p}$. This turns out to be the same as the largest power of $(x - \lambda_i)$ dividing $F(X)$, namely $e_i$. All in all then, we have $\sum e_i = \deg F(X) = [L:K]$.

Note: As remarked in the proof, we obtain 2 important facts about the extensions of $\mathfrak{p}$. Namely, that $\mathfrak{p}$ always has extensions in $L$ and that these extensions are finite in number where this number is less than or equal to $N$.

If there is precisely one extension of $\mathfrak{p}$ in $L$, we say that $\mathfrak{p}$ is **totally ramified**. In this case, $I$ tells us that the ramification index of the extension of $\mathfrak{p}$ is $N$, the degree of $L$ over $K$. Also, we find that the local uniformizing parameter of the extension in $L$ satisfies a particularly nice equation.

**Lemma (Eisenstein)**

Let $L/K$ be a field extension of degree $N$. Suppose the place $\mathfrak{p}$ on $K$ is totally ramified. Let $\mathfrak{p}$ be the place on $L$ lying over $\mathfrak{p}$ and take $\Pi$ to be a local uniformizing parameter for $\mathfrak{p}$. Then there exist $\alpha_i \in K$ such that

$$\Pi^N = \alpha_{N-1} \Pi^{N-1} + \ldots + \alpha_1 \Pi + \alpha_0$$

where the $\alpha_i \in \mathfrak{p}$ for $1 \leq i \leq N-1$ and $\text{ord} \mathfrak{p}(\alpha_0) = 1$. 
Proof

\( \Pi \) is an element of \( L \) and hence satisfies an equation of degree \( n \) over \( K \), for some \( n \mid N \). Thus, without loss of generality, there exist \( \alpha_i \in K, i = 0, 1, \ldots, n-1 \), such that

\[
(*) \quad \Pi^n = \alpha_{n-1} \Pi^{n-1} + \cdots + \alpha_1 \Pi + \alpha_0
\]

From this equation we see that \( \text{ord}(\Pi^n) = n \), and hence that \( \text{ord}(\alpha_{n-1} \Pi^{n-1} + \cdots + \alpha_0) \) must also be \( n \).

Now consider each term on the right of the equation. For each element \( \alpha_i \in K \), \( \text{ord}(\alpha_i) = N \text{ ord}(\alpha_i) \), since \( \varphi \) is totally ramified. Thus \( \text{ord}(\alpha_i) \equiv 0 \pmod{N} \), and therefore \( \text{ord}(\alpha_i) \equiv 0 \pmod{n} \). But then \( \text{ord}(\alpha_i \Pi^i) = \text{ord}(\alpha_i) + \text{ord}(\Pi^i) \equiv i \pmod{n} \).

Hence

1. For each \( i, j \) from 0 to \( n-1 \), if \( i \neq j \) then \( \text{ord}(\alpha_i \Pi^i) \neq \text{ord}(\alpha_j \Pi^j) \), and
2. \( \text{ord}(\alpha_i \Pi^i) \neq 0 \pmod{n} \) for \( 1 \leq i \leq n-1 \).

From (1) we see that the terms on the right of (*) are distinct. Therefore

\[
\text{ord}(\alpha_{n-1} \Pi^{n-1} + \cdots + \alpha_0) = \min_{0 \leq i \leq n-1} (\text{ord}(\alpha_i \Pi^i))
\]

We know this order must be \( n \). By (2) the only term whose order can possibly be \( n \) is \( \alpha_0 \). Thus \( \text{ord}(\alpha_0) \) must be \( n \), and \( \text{ord}(\alpha_i \Pi^i) \) must be \( >n \). In fact, since \( \text{ord}(\alpha_0) \equiv 0 \pmod{N} \) and \( n \mid N \), we find \( n = N \). Summing up, we have \( \text{ord}(\alpha_0) = N \) and \( \text{ord}(\alpha_i) = m_i N \), for some integer \( m_i > 1 \) and \( 1 \leq i \leq n-1 \). This is the desired conclusion.
Our second basic result gives the relationship among all extensions of a given place in $K$, assuming that $L/K$ is Galois.

II Let $G$ be the Galois group of $L/K$ and let $\mathfrak{q}$ be a place on $K$. Then $G$ acts transitively on $
$ \{ $\mathfrak{p} \mid \mathfrak{p}$ a place on $L$ and $\mathfrak{p} \mid \mathfrak{q}$ \}. That is, given places $\mathfrak{p}_1, \mathfrak{p}_2$ lying over $\mathfrak{q}$, there exists $\sigma \in G$ such that $\mathfrak{p}_1 = \sigma \mathfrak{p}_2$

For the proof of this standard result the reader is referred to Proposition 11, p. 244 [4].

Assume $L/K$ to be Galois in the following. We develop a connection between ramification indices and certain subgroups of the Galois group of $L/K$.

Let $G$ be the Galois group of $L/K$. Let $\mathfrak{p}$ be a place on $L$ with place ring $\mathcal{O}$. Define the inertia group of $\mathfrak{p}$ to be

$$I_\mathfrak{p} = \{ \sigma \in G \mid \sigma(a) \equiv a \pmod{\mathfrak{p}} \text{ for all } a \in \mathcal{O} \}$$

Each $\sigma \in I_\mathfrak{p}$ leaves all elements of the place ring of $\mathfrak{p}$ fixed modulo $\mathfrak{p}$. Define the decomposition group of $\mathfrak{p}$, $D_\mathfrak{p}$, to be composed of all elements of $G$ leaving $\mathfrak{p}$ fixed. That is, take

$$D_\mathfrak{p} = \{ \sigma \in G \mid \sigma \mathfrak{p} = \mathfrak{p} \}$$

Then, it turns out that in our set-up, since $k$ is algebraically closed, these two groups are identical.

**Lemma**

For any valuation $\mathfrak{p}$ on $L$,
38.

\[ I_\mathfrak{p} = D_\mathfrak{p} \]

**Proof**

It is clear that \( I_\mathfrak{p} \subseteq D_\mathfrak{p} \). Assume \( \sigma \in D_\mathfrak{p} \), and let \( a \) be an arbitrary element of \( \mathcal{O} \). Then there is a unique \( \lambda \in k \), the field of constants, such that \( a \equiv \lambda \pmod{\mathfrak{p}} \). (Obtain \( \lambda \) through the isomorphism \( k \to \mathcal{O} \to \mathcal{O}/\mathfrak{p} \).) Applying \( \sigma \) to this we have \( \sigma(a) \equiv \sigma(\lambda) \pmod{\mathfrak{p}} \). But \( \sigma \lambda = \lambda \) since \( \lambda \in k \), and \( \sigma \mathfrak{p} = \mathfrak{p} \) since \( \sigma \in D_\mathfrak{p} \). Hence \( \sigma(a) \equiv \lambda \pmod{\mathfrak{p}} \), so that \( \sigma \) does belong to \( I_\mathfrak{p} \).

We can now easily derive from I and II the following important relation: namely, that the inertia group of a place has order precisely equal to the ramification index of the place.

**III**

Let \( \mathfrak{p} \) be a place on \( L \), and let \( e_\mathfrak{p} \) be its ramification index. Then

\[ |I_\mathfrak{p}| = e_\mathfrak{p} \]

**Proof**

Let \( \mathfrak{q} \) be a place on \( K \) such that \( \mathfrak{p} \mid \mathfrak{q} \). Take \( S = \{ \mathfrak{p} \mid \mathfrak{p} \) a place on \( L \) and \( \mathfrak{p} \mid \mathfrak{q} \} \). We know \( S \) is finite by result I; let \( |S| \) denote the number of elements in \( S \). Then our situation is as follows: we have the Galois group \( G \) of \( L/K \) acting transitively on \( S \) (by II), and the subgroup \( D_\mathfrak{p} \) of \( G \) is \( \text{Stab}_G(\mathfrak{p}) \), the stabilizer of \( \mathfrak{p} \) in \( G \). Hence by the fundamental theorem on transformation groups, \( |G:D_\mathfrak{p}| = |S| \),
or,
\[ |D_{\mathcal{P}}| = \frac{|G|}{|S|} \]

Also from II, we deduce that all the \( e_{\mathcal{P}} \) must be the same; say they are all \( e \). Then \( \sum_{\mathcal{P} \mid \mathcal{Q}} e_{\mathcal{Q}} = \sum e = e|S| \).
From I we find that \( e|S| = [L:k] \); by Galois theory this is just \( |G| \). Thus \( e = \frac{|G|}{|S|} \), so that \( |D_{\mathcal{P}}| = |I_{\mathcal{P}}| = e \), as desired.

This result gives us two more facts which will prove very useful later on.

**Corollary 1**

Suppose \( L_2/L_1/K \) is a Galois tower of fields. Let \( \mathcal{P} \) be a place on \( L_2 \). If \( I_{\mathcal{P}} \) is contained in \( \text{Gal} \left( L_2/L_1 \right) \), then \( \mathcal{P} \) is not ramified in \( L_1/K \) - all ramification occurs in \( L_2/L_1 \).

This is clear using the fact that ramification indices are multiplicative in a tower of fields.

**Corollary 2**

The number of ramified places in any separable extension \( L/K \) is finite.
Notice that here \( L/K \) need not be Galois.

**Proof**

Suppose \( k \) is the ground field for \( L \) and \( K \), as usual.
Take \( x \) to be a separating transcendental in \( K/k \), and let \( \theta \) be the generator of \( L \) over \( K \). Thus we have \( L = K(\theta) \) and \( K(\theta) \) separable over \( k(x) \).
Before getting into the proof we make two simplifications of the situation. To show that there are finitely many ramified places on \( L/K \), it is clearly sufficient to show that only finitely many places of \( L/k(x) \) are ramified. Similarly, if we can show only finitely many places ramified in \( \bar{L}/k(x) \) for some extension \( \bar{L} \) of \( L \), then we are also done. In particular, we take \( \bar{L} \) to be the field obtained by adjoining to \( L \) all conjugates of \( \theta \). For simplicity, denote \( \bar{L} \) again by \( L \). We now are considering the Galois extension \( L/k(x) \). Denote its Galois group by \( G \).

By modifying the generator \( \theta \) of \( L \) slightly, if necessary, we can assume that \( \theta \) satisfies an equation of the form

\[
F(T) = T^n + f_{n-1}T^{n-1} + \ldots + f_0
\]

where the \( f_i \in k[x] \). Then, \( L \) being normal over \( k[x] \), we can factor \( F(T) \) in \( L \) as

\[
F(T) = \prod_{\sigma \in G} (T - \sigma \theta)
\]

Taking the formal derivative of \( F \) with respect to \( T \), and evaluating at \( \theta \) we find

\[
F'(\theta) = \prod_{\sigma \in G} (\theta - \sigma \theta)
\]

Now we come to counting the number of ramified places of \( L/k(x) \). Suppose \( \mathfrak{p} \) is a ramified place on \( L \), not lying over the place corresponding to \( \infty \) in \( k(x) \). (This excludes only finitely many places by I.) Then \( \theta \) lies in the place ring of \( \mathfrak{p} \) in \( L \). For, if \( \mathfrak{q} \) is the place on \( k(x) \) induced by \( \mathfrak{p} \) and \( \sigma \) is its place ring, then \( \sigma \not\in k[x] \).
Thus $\theta$ satisfies an equation with coefficients in $\mathcal{O}$. By the remark on integrality of section 1, $\theta$ must therefore lie in $\mathcal{O}_p$.

By III, $\mathfrak{p}$ ramified implies that its inertia group $I_{\mathfrak{p}}$ is non-trivial. Hence there is some automorphism $\sigma \neq \text{id}$ in $G$ such that $\theta - \sigma \theta \equiv 0 \pmod{\mathfrak{p}}$. But then $F'(\theta) = \prod (\theta - \sigma \theta) \equiv 0 \pmod{\mathfrak{p}}$. This means in particular that $F'(\theta)$ is divisible by the uniformizing parameter for $\mathfrak{p}$.

This statement holds for any ramified place not over $\infty$ in $L/k(x)$. Since $F'(\theta)$ has only finitely many zeros, the number of ramified places must be finite.

As an example of our ramification theory we derive the ramification indices of the field extensions corresponding to the canonical groups.\(^{(1)}\)

Assume that $k = \mathbb{C}$ and let $G$ denote one of the canonical groups. Say $|G| = N$. As in the Introduction, we view the figure $F$ corresponding to $G$ i.e. the plane $N$-gon, the dihedron, the tetrahedron, the octahedron or the icosahedron, as embedded in the Riemann sphere. Recall that elements of $G$ are then rotations of the sphere which leave $F$ fixed.

Now suppose that $G$ is one of the tetrahedral, octahedral or icosahedral groups. Each element of $G$ moves

\(^{(1)}\) Refer to §§1-9 Chp. 1[3] for further details on this example. Klein looks at this situation from a geometric point of view.
a point $P_0$ on the sphere to another point on the sphere. Call the set of points to which $P_0$ is moved by $G$ its orbit. Most points have orbit containing $N$ points. However, there are three types of points which have smaller orbits, namely: midpoints of edges of the figure $F$, centrepoints of faces of $F$ and vertices of $F$. It can easily be seen that the midpoint $P$ of an edge is left fixed by precisely two elements of $G$ (the half turn about the axis through $P$ and the midpoint of the opposite edge, and the identity); the centrepoint of any face is left fixed by three elements of $G$; any vertex is fixed by $i$ elements of $G$ where $i = 3, 4$ or $5$, the number of faces adjoining at a vertex of $F$. Also, it is clear that the points in any one of these three sets form a conjugacy class under $G$. Each such class of points corresponds to the places in $k(x)$ lying over a given place in $\text{Fix } G$. (Refer to pp. 28-29 for the correspondence between points in $k$ and places in $k(x)$. ) The elements of $G$ leaving a point in any class fixed form the inertia group at that point. Since we know that the order of the inertia group is equal to the ramification index at a point, 2, 3 and $i$ are the only non-unit ramification indices for $G$. This triple, $(2, 3, i)$, of ramification indices for the three non-trivial conjugacy classes will be called the ramification pattern of $G$.

Suppose now that $G$ is a cyclic group of order $N$. Viewed as a group of automorphisms of the sphere, the
elements of G are obtained by repetition of a single rotation. As such we see immediately that there are only two points on the sphere which have orbits containing less than N points—the two points at either end of the axis of rotation. These points are fixed by all N elements and are therefore non-conjugate. The ramification pattern of G must then be \((1, N, N)\).

For the dihedral group of order \(N = 2n\) the points on the dihedron having an orbit of less than \(2n\) points are: the two points at the end of the axis of the cyclic order \(n\) rotation (these are conjugate and fixed for \(n\) rotations); the \(n\) vertices (these are conjugate and fixed by a half-turn about the axis through the vertex and the identity); the \(n\) mid-edge points (these are also conjugate and fixed by 2 rotations). Thus there are three distinct conjugacy classes with orbits smaller than normal. We see by the above considerations that the ramification pattern of G must therefore be \((2, 2, \frac{N}{2})\).

For future reference, a list of the ramification patterns is given here.

<table>
<thead>
<tr>
<th></th>
<th>Cyclic group, order N</th>
<th>(1, N, N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>Dihedral group, order N</td>
<td>(2, 2, (\frac{N}{2}))</td>
</tr>
<tr>
<td>III</td>
<td>Tetrahedral group</td>
<td>(2, 3, 3)</td>
</tr>
<tr>
<td>IV</td>
<td>Octahedral group</td>
<td>(2, 3, 4)</td>
</tr>
<tr>
<td>V</td>
<td>Icosahedral group</td>
<td>(2, 3, 5)</td>
</tr>
</tbody>
</table>
For the general situation, k any algebraically closed field, in determining the ramification patterns of the canonical groups, the geometry of our situation gives way to combinatorics of G-action on places of k(x). We will see later that this does result in the same patterns. In fact, we will determine that any finite subgroup of Aut\_k k(x) has to have one of the above ramification patterns.
(1) **Order of a Differential**

Let $K$ be an algebraic function field over $k$, with $k$ algebraically closed. Let $d: K \to \Omega_{K/k}$ be the universal $k$-derivation. (Cf. §2, Chapter III) Fix a place $\mathfrak{p}$ on $K$, and denote by $K_{\mathfrak{p}}$ the completion of $K$ with respect to $\mathfrak{p}$. Then, for any local uniformizing parameter $\pi$ for $\mathfrak{p}$, $K_{\mathfrak{p}} = k((\pi))$.

Let $d_{\mathfrak{p}}: K_{\mathfrak{p}} \to \Omega^c_{K_{\mathfrak{p}}/k}$ be the universal continuous $k$-derivation of $K_{\mathfrak{p}}/k$. Take $i: K \to K_{\mathfrak{p}}$ to be the injection of $K$ into $K_{\mathfrak{p}}$. Then universality of $d$ guarantees existence of a unique $K$-module map $i_\Omega: \Omega_{K/k} \to \Omega^c_{K_{\mathfrak{p}}/k}$ completing the following diagram.

![Diagram](chart.png)

Let $\omega \in \Omega_{K/k}$. Define $\text{ord}_\mathfrak{p}(\omega)$, the order of $\omega$ at $\mathfrak{p}$, as follows:

Map $\omega$ to $i_\Omega(\omega) = \omega_{\mathfrak{p}}$, say.

Define $\text{ord}_\mathfrak{p}(\omega) = \text{ord}(\omega_{\mathfrak{p}})$

(recall that the order of a continuous differential was defined in §2, Chapter III)

(2) For any differential $\omega \in \Omega_{K/k}$, $\text{ord}_\mathfrak{p}(\omega) = 0$ for almost
all places \( \mathfrak{p} \) on \( K \).

Let \( x \) be a separating transcendental in \( K/k \). Then
\[
\Omega_{K/k} = Kdx.
\]
Thus for any \( \omega \in \Omega_{K/k} \) we have \( \omega = fdx \), for some \( f \in K \).

Claim: \( \text{ord}_{\mathfrak{p}}(\omega) = 0 \) for almost all places \( \mathfrak{p} \) on \( K \).

Proof:

\[
\text{ord}_{\mathfrak{p}}(\omega) = \text{ord}_{\mathfrak{p}}(f) + \text{ord}_{\mathfrak{p}}(dx)
\]

We know that \( \text{ord}_{\mathfrak{p}}(f) = 0 \) almost everywhere since \( f \in K \). (\( f \) has order 0 at all points except at its zeros and poles, and these are finite in number.) So we must show that \( \text{ord}_{\mathfrak{p}}(dx) = 0 \) almost everywhere.

Let \( \mathfrak{p} \) be any place on \( K \) which is unramified and which does not lie over the place \( \mathfrak{q}_0 : x \mapsto \infty \) in \( k(x) \). (This excludes \( \mathfrak{p} \) from being only a finite number of places)

Suppose \( \mathfrak{p} \) lies over the place \( \mathfrak{q} \) of \( k(x) \). \( \mathfrak{q} \), not being \( \mathfrak{q}_0 : x \mapsto \infty \), must be of the form \( \mathfrak{q} : x \mapsto a \), for some \( a \in k - \infty \). Thus \( \mathfrak{q} \) has local uniformizing parameter \( \pi = x - a \).

But then \( \mathfrak{p} \) has local uniformizing parameter \( \Pi = \pi = x - a \).

Then \( x = \Pi + a \), and \( dx = 1 \cdot d\Pi \). Therefore, \( \text{ord}_{\mathfrak{p}}(dx) = \text{ord}(1) = 0 \).

(3) Definition of \( \delta(K/k) \)

Let \( \omega \in \Omega_{K/k} \). Define \( \text{div}(\omega) = \prod_{\mathfrak{p}} \text{ord}_{\mathfrak{p}}(\omega) \) where the product is taken over all places \( \mathfrak{p} \) on \( K \). This is a finite product by section (2); the following definition therefore makes sense. We define

\[
\text{deg}(\text{div} \ \omega) = \sum_{\mathfrak{p}} \text{ord}_{\mathfrak{p}}(\omega)
\]
Claim: \( \deg (\text{div } \omega) \) is independent of \( \omega \).

Proof:

Let \( x \) be a separating transcendental in \( K \). Then \( \Omega_{K/k} = Kdx \), so that any two differentials \( \omega, \omega' \) differ at most by an element of \( K \). (For, \( \omega = f dx \) and \( \omega' = g dx \) for some \( f, g \in K \mapsto \omega = (fg^{-1})\omega' \), where \( fg^{-1} \in K \).) Say \( \omega = h\omega' \) then. We can now compare \( \deg (\text{div } \omega) \) and \( \deg (\text{div } \omega') \).

\[
\deg (\text{div } \omega) = \sum_p \text{ord}_p(\omega) = \sum_p \text{ord}_p(h\omega')
\]

\[
= \sum_p (\text{ord}_p h + \text{ord}_p \omega')
\]

\[
= (\sum_p \text{ord}_p h) + (\sum_p \text{ord}_p \omega') \quad (1)
\]

\[
= \sum_p \text{ord}_p \omega' = \deg (\text{div } \omega').
\]

Since \( \deg (\text{div } \omega) \) is an invariant depending only on the field extension \( K \) over \( k \), we denote it with reference to \( K/k \) alone, by \( \delta(K/k) \).

Note: If \( K \) is a rational function field, say \( K = k(x) \) where \( x \in K \) is some indeterminate over \( k \), then \( \delta(K/k) = -2 \).

Proof:

Assume \( K = k(x) \), where \( x \in K \) is some indeterminate over \( k \). Let \( \omega = dx \in \Omega_{K/k} \). Then \( \delta(K/k) = \sum_p \text{ord}_p(dx) \).

Now, all places \( p \) on \( K \) are of the form

(i) \( p : x \mapsto a, a \in k \) or

(ii) \( p : x \mapsto \infty \)

Case (i) Take \( \pi = x - a \) to be the local uniformizing parameter for \( p \). Then \( dx = l \cdot d\pi \), so that \( \text{ord}_p(dx) = 0 \).

\[
(1) \sum_p \text{ord}_p h = 0 \text{ by the theorem of p. 29.}
\]
Case (ii) Take \( \pi = \frac{1}{x} \) to be the local uniformizing parameter for \( \mathfrak{p} \). Then we have \( x = \frac{1}{\pi} \), so that \( dx = -\frac{1}{\pi^2} d\pi \). Thus \( \text{ord}_\mathfrak{p}(dx) = -2 \).

Hence \( \deg(\text{div } dx) = \sum \text{ord}_\mathfrak{p}(dx) \)

\[ = 0 + (-2) \]

\[ = -2 \]

(4) The Hurwitz Formula

Let \( L \) and \( K \) be field extensions of \( k \) of transcendence degree 1, \( L \) algebraic of degree \( N \) over \( K \), and \( k \) of characteristic \( p \). Let \( e_\mathfrak{p} \) denote the ramification index of the place \( \mathfrak{p} \) on \( L \). Write \( \mathfrak{q} \in L \) to denote that \( \mathfrak{q} \) is a place on \( L \). Then

\[ \delta(L/k) = \delta(K/k) \cdot N + \sum_{\mathfrak{q} \in L} \varepsilon_{\mathfrak{q}} \]

where \( \varepsilon_{\mathfrak{q}} = e_\mathfrak{q} - 1 \) if \( p \nmid e_\mathfrak{q} \) and \( \varepsilon_{\mathfrak{q}} > e_\mathfrak{q} - 1 \) if \( p | e_\mathfrak{q} \). \( (2) \)

Proof:

Let \( \omega \in \Omega_{K/k} \) (and so \( \omega \in \Omega_{L/k} \)). Idea of proof:

Consider a single place \( \mathfrak{q} \) on \( K \). Compare \( \delta_{\mathfrak{q}}(K) = \text{ord}_{\mathfrak{q}}(\omega) \) to \( \delta_{\mathfrak{q}}(L) = \sum_{\mathfrak{q} \in L} \text{ord}_{\mathfrak{q}}(\omega) \). Then to relate \( \delta(K/k) \) and \( \delta(L/k) \), notice that \( \delta(K/k) = \sum_{\mathfrak{q} \in K} \delta_{\mathfrak{q}}(K) \) and \( \delta(L/k) = \sum_{\mathfrak{q} \in K} \delta_{\mathfrak{q}}(L) \).

Suppose \( \mathfrak{q} \) is a place on \( L \) with ramification index \( e \).

Let \( \mathfrak{f} \) be the place in \( K \) lying under \( \mathfrak{q} \). Let \( \pi, \Pi \) be local uniformizing parameters for \( \mathfrak{q} \) in \( K \) and \( \mathfrak{q} \) in \( L \) respectively.

Then \( \pi = U\pi^e \) for some unit \( U \) in \( L \).

(2) The original deduction of this formula can be found in § 2 Chapter XXI - V [2]
Thus
\[ d_\pi = e \mu \pi e^{-1} d_\pi + (D_\pi U) \pi e d_\pi \]
\[ = [e \mu \pi e^{-1} + (D_\pi U) \pi e ] d_\pi \]

Now we can relate \( d_{\mathfrak{q}}(K) \) and \( d_{\mathfrak{q}}(L) \).

\[ \omega \in \Omega_{K/k}, \text{ Thus } i_{\Omega}(\omega) = \omega = f_\pi d_\pi \text{ for some } f_\pi \in K((\pi)), \]
and \( \text{ord}_{\mathfrak{q}}(\omega) = \text{ord}_{\mathfrak{q}}(f_\pi) = \delta_{\mathfrak{q}}(K) \).

What is \( \delta_{\mathfrak{q}}(L) ? \)

\[ \omega_{\mathfrak{q}} = f_\pi d_\pi = f_\pi [e \mu \pi e^{-1} + (D_\pi U) \pi e ] d_\pi \]

Therefore \( \text{ord}_{\mathfrak{q}}(\omega) = \text{ord}_{\mathfrak{q}}(f_\pi) + \text{ord}_{\mathfrak{q}}(e \mu \pi e^{-1} + (D_\pi U) \pi e ) \)
\[ = \text{ord}_{\mathfrak{q}}(f_\pi) + e, \text{ say.} \]

Notice: \( e = e - 1 \) if \( p \nmid e \), and \( e > e - 1 \) if \( p | e \).

With \( \mathfrak{p} \) now any place above \( \mathfrak{q} \) and \( e_\mathfrak{p} \) its ramification index, we have
\[ \delta_{\mathfrak{q}}(L) = \sum_{\mathfrak{p} | \mathfrak{q}} \text{ord}_{\mathfrak{p}}(\omega) \]
\[ = \sum_{\mathfrak{p} | \mathfrak{q}} \text{ord}_{\mathfrak{p}}(f_\pi) + \sum_{\mathfrak{p} | \mathfrak{q}} e_\mathfrak{p} \]
\[ = \sum_{\mathfrak{p} | \mathfrak{q}} e \cdot \text{ord}(f_\pi) + \sum_{\mathfrak{p} | \mathfrak{q}} e_\mathfrak{p} \]
\[ = \delta_{\mathfrak{q}}(K) \cdot \sum_{\mathfrak{p} | \mathfrak{q}} e + \sum_{\mathfrak{p} | \mathfrak{q}} e_\mathfrak{p} \]
\[ = \delta_{\mathfrak{q}}(K) \cdot (N) + \sum_{\mathfrak{p} | \mathfrak{q}} e_\mathfrak{p} \]
\[ = \delta_{\mathfrak{q}}(K) \cdot N + \sum_{\mathfrak{p} | \mathfrak{q}} e_\mathfrak{p} \]

We therefore have the following relation between \( \delta(L/k) \) and \( \delta(K/k) \).

\[ \delta(L/k) = \sum_{\mathfrak{q} \in K} \delta_{\mathfrak{q}}(L) \]
\[ = \sum_{\mathfrak{q} \in K} \delta_{\mathfrak{q}}(K) \cdot N + \sum_{\mathfrak{q} \in K} \sum_{\mathfrak{p} | \mathfrak{q}} e_\mathfrak{p} \]
\[ = \delta(K/k) \cdot N + \sum_{\mathfrak{p} \in L} e_\mathfrak{p} \]
Assume throughout this chapter that $K$ is a fixed field of rational functions, purely transcendental of degree 1 over an algebraically closed field $k$. We study Galois extensions of $K$ which are also rational function fields. Although this is not the natural point of view arising from our problem, (the natural one being to study subfields of a fixed extension) it will be seen later that the problem does reduce to this situation.

(1) In the theorem on the Hurwitz Formula, if both $L$ and $K$ are purely transcendental of degree 1 over $k$, then the formula takes on a particularly simple form. Namely, it yields a diophantine equation, the solutions of which give enough information to determine $L$ to isomorphism over $K$ in most cases.

We now determine the diophantine equation and its solutions. With $L$ and $K$ as above, from §3, Chapter VI we know that $\delta(L/k) = \delta(K/k) = -2$. If furthermore, the characteristic $p$ of $k$ does not divide $e$ for any ramification index $e$ in $L/K$, then the Hurwitz Formula becomes

$$2N - 2 = \sum_{\hat{\phi} \in L} \varepsilon_{\hat{\phi}} = \sum_{\hat{\phi} \in L} (e_{\hat{\phi}} - 1)$$

$$= \sum_{\hat{\phi} \in K} \frac{N_{\hat{\phi}/k}(e_{r_{\hat{\phi}}})}{r_{\hat{\phi}}} - 1$$

$$= \sum_{i=1}^{\infty} \frac{N_{e_{i}/k}(e_{i} - 1)}{e_{i}},$$
indexing the \( \phi_j \) in \( K \) from 1 to \( r \). Or, dividing by \( N \), we have

\[
(*) \quad 2 - \frac{2}{N} = \sum_{i=1}^{r} \left( 1 - \frac{1}{e_i} \right)
\]

Regarding this as a diophantine equation in \( N \) and \( e_i \), it is fairly easy to see that it has relatively few solutions. Assuming \( N > 1 \), then \( r = 2 \) or 3.

For, \( r = 1 \) implies R.S. \((*) < 1\), while L.S. \((*) \geq 1\); and \( r \geq 4 \) implies R.S. \((*) \geq 2\), while L.S. \((*) < 2\).

(i) Letting \( r = 2 \) \((*)\) becomes

\[
\frac{2}{N} = \frac{1}{e_1} + \frac{1}{e_2}
\]

Since \( e_1, e_2 \leq N \), we have \( \frac{1}{e_1}, \frac{1}{e_2} \geq \frac{1}{N} \). It follows from our equation that \( \frac{1}{e_1}, \frac{1}{e_2} \) must be \( \frac{1}{N} \). Hence \( e_1 = e_2 = N \).

(ii) Letting \( r = 3 \) \((*)\) becomes

\[
1 + \frac{2}{N} = \frac{1}{e_1} + \frac{1}{e_2} + \frac{1}{e_3}
\]

Since not all \( e_i \) can be \( \geq 3 \), one, say \( e_1 \), must be 2. So

\[
\frac{1}{2} + \frac{2}{N} = \frac{1}{e_2} + \frac{1}{e_3}
\]

If another \( e_1 \), say \( e_2 \), is 2 then

\( e_1 = e_2 = 2, \quad e_3 = n \) and \( N = 2n \) for any integer \( n \geq 1 \).

If neither \( e_2 \) nor \( e_3 \) is 2 then not both can be \( \geq 4 \).

Assume \( e_2 = 3 \). Then

\[
\frac{1}{6} + \frac{2}{N} = \frac{1}{e_3}
\]

Thus \( e_3 < 6 \), and \( e_3 \) must therefore be 3, 4 or 5. Here we
have
\[ e_1 = 2, \quad e_2 = 3, \quad e_3 = 3, \quad N = 12 \]
\[ e_1 = 2, \quad e_2 = 3, \quad e_3 = 4, \quad N = 24 \]
\[ e_1 = 2, \quad e_2 = 3, \quad e_3 = 5, \quad N = 60 \]

So the Hurwitz Formula in this case has only five systems of solutions. Notice that these systems (N and ramification indices) correspond to the group orders and ramification indices for the five canonical groups: cyclic dihedral, tetrahedral, octahedral and icosahedral. This is the major fact donated by the Hurwitz Formula towards the proof of the Main Theorem.

(2) The Hurwitz Formula shows that with \( L, K \) both as described in (1), \([L:K]\) and the ramification indices of \( L/K \) can be one of only five types. Since the set of ramification indices is different for each type, it is enough to specify this to determine the type. We set up a notation for easy reference to each type.

Let \( \mathcal{g}_1, \mathcal{g}_2, \mathcal{g}_3 \) be the three places in \( K \) at which ramification occurs. (It may occur at only two places, in which case it does not matter what the third place is.) Let \( e_1, e_2, e_3 \) be the respective ramification indices at these places. Then we use the ordered triple \((e_1, e_2, e_3)\) to show the ramification at \( \mathcal{g}_1, \mathcal{g}_2, \mathcal{g}_3 \) in \( L/K \). Call this the ramification pattern \( R(L/K) \) of \( L/K \). Hurwitz says that \( R(L/K) \) is one of types
\[
\begin{array}{c|c}
I & (1, N, N) \\
II & (2, 2, \frac{N}{2}) \\
III & (2, 3, 3) \\
IV & (2, 3, 4) \\
V & (2, 3, 5) \\
\end{array}
\]

In what follows, we shall show that knowing the ramification pattern of an extension of \( K \) is enough to determine the extension in cases I-IV, and in case V it determines the group of the extension (assuming always that the extension is Galois).

For the lemma and theorem following, we fix an extension \( L \) of \( K \). Say \( L = k(x) \), for some \( x \in L \) transcendental over \( k \). Also assume \([L:K] = N\).

**Lemma**

In \( L/K \) we have

\[
(*) \quad \text{Irr} \ (x, K, T) = T^N + a_{N-1}T^{N-1} + \ldots + a_0,
\]

where either the \( a_i \in k \) or \( K = k(\alpha_i) \), and all nonconstant \( a_i \) have the same pole.

**Proof**

Choose some element \( y \in K \) such that \( K = k(y) \).

Let \( \text{Irr} \ (x, k(y), T) = f_N(y)T^N + f_{N-1}(y)T^{N-1} + \ldots + f_0(y) \).

We know \( y = \frac{h(x)}{g(x)} \) for some \( h, g \in k[x] \), so that

\( \text{Irr} \ (x, k(y), T) = h(T) - yg(T) \). Thus \( y \) occurs to the first power only in the minimal polynomial for \( x \). Since we have unique factorization in \( k(y)[T] \), we see that all
f(y) must be of the form \( a_i^y + b_i \) where \( a_i, b_i \in k \). Then we have

\[
\text{Irr} (x, K, T) = T^N + \frac{a_{N-1}^y + b_{N-1}}{a_N^y + b_N} T^{N-1} + \ldots + \frac{a_0^y + b_0}{a_N^y + b_N}
\]

\[
= T^N + a_{N-1}^y T^{N-1} + \ldots + a_0^y, \text{ say.}
\]

By definition then, all \( a_i \) which are nonconstant have the same pole (given by \( a_N^y + b_N \)); and all are of the form

\[
\frac{a_i^y + b_i}{a_N^y + b_N}. \quad \text{This is either in } k \text{ (if } a_i b_N - a_N b_i = 0 \text{) or generates } K \text{ (by the theorem of Chapter I).}
\]

Remark: If \( R(L/K) \) is type I with ramification at \((x)\) and \((\frac{1}{x})\), then in the above lemma we can conclude that \( K = k(a_i) \) for all \( i \). For, by the lemma following I of Chapter V, we know \( x \) satisfies an equation of form (*) where each \( a_i \) has order \( \geq 1 \) at 0. This together with the above lemma implies that each \( a_i \) must have order precisely 1, as desired.

**Theorem 1**

Suppose \( R(L/K) \) is type I with ramification at the places \((x)\) and \((\frac{1}{x})\). Then \( K = k(x_N) \).

**Proof**

By the preceding remark we know that

\[
\text{Irr} (x, K, T) = T^N + a_{N-1}^y T^{N-1} + \ldots + a_0^y
\]

where each \( a_i \) has order 1. Hence we can fix \( y = a_0 \), in particular, as generator of \( K \). Then all other \( a_i \) are of
55.

the form \( \frac{a_i y + b_i}{c_i y + d_i} \) for some \( a_i, b_i, c_i, d_i \in k \) such that \( a_i d_i - b_i c_i \neq 0 \). Now the preceding lemma states that all \( \alpha_i \) have the same pole. Since \( \alpha_0 \) has a pole at \( \infty \) (meaning \( \frac{1}{y} \)) all other \( \alpha_i \) must have a pole at \( \infty \). That is, they must be of the form \( \alpha_i = \bar{a}_i y \) for some \( \bar{a}_i \in k \). Thus

\[
(**) \quad \text{Irr} (x, K, T) = T^N + \bar{a}_N^{-1} y T^{N-1} + \ldots + y.
\]

Now look at the equation satisfied by \( \frac{1}{x} \). Take \( T = \frac{1}{U} \) and divide by \( y \) in \((**')\) to find that \( \frac{1}{x} \) satisfies

\[
\frac{1}{y} + \bar{a}_N^{-1} U + \bar{a}_N^{-2} U^2 + \ldots + \bar{a}_1 U^{N-1} + U^N = 0
\]

Again applying the preceding lemma we find that all \( \bar{a}_i \) must be multiples of \( \frac{1}{y} \). That is, they must all be 0. From \((**')\) then we have \( y = -x^N \), or \( K = k(y) = k(x^N) \).

Remark: In Theorem 1 we assumed that ramification occurred at \( (x) \) and \( (\frac{1}{x}) \) merely for convenience in computation. The theorem could have been stated as: if \( R(L/K) \) is type I and the places at which ramification occurs are known, then the extension is determined as cyclic of degree \( N \). Thus, if \( R(L/K) \) and \( R(L'/K) \) are both type I with ramification at the same places, then the extensions \( L \) and \( L' \) are isomorphic over \( K \). This is the basic fact upon which the proof of the following theorem relies.

We also make use of the following facts from Chapter V.

(1) ramification index = order of inertia group

for a fixed place (result III)
(2) ramification indices multiply in a tower of fields.

(3) degree of field extension = sum of ramification indices (result I)

**Theorem 2**

Let \( L/K \) be a Galois extension and \( L \) a rational function field. Suppose the group of \( L/K \) is solvable. Then \( R(L/K) \) determines \( L \) up to an isomorphism over \( K \).

**Proof**

We consider the three cases where \( R(L/K) \) is type II, III or IV, I having been dealt with in Theorem 1.

II \((2, 2, \frac{N}{2})\)

Assume first that \( \frac{N}{2} \neq 2 \). Let \( H \) be the inertia group of \( \mathfrak{q}_3 \), the place with ramification index \( \frac{N}{2} \). Let \( L_0 = \text{Fix} \ H \). Then \( \text{Gal} (L_0/K) \) is a finite subgroup of \( \text{Aut}_k(x) \) and so must have ramification pattern type I-V. Since \( |\text{Gal} (L_0/K)| = 2 \), \( R(L_0/K) \) can only be type I with \( N = 2 \). That is, \( R(L_0/K) = (2, 2, 1) \). By Theorem 1 this means that \( L_0 \) is determined to isomorphism over \( K \). Let \( \mathfrak{q}_1, \mathfrak{q}_2 \) be the two places in \( L_0 \) above \( \mathfrak{q}_3 \). These places must ramify of index \( \frac{N}{2} \) in \( L/L_0 \). Hence \( R(L/L_0) = (\frac{N}{2}, \frac{N}{2}, 1) \).

Again using Theorem 1 we find that \( L \) is determined to isomorphism over \( L_0 \), and so over \( K \).

If \( \frac{N}{2} = 2 \), then \( R(L/K) = (2, 2, 2) \). We know that the only groups of order 4 are the cyclic group or the Klein 4-group; and the ramification pattern of the cyclic
group is type I. Hence Gal (L/K) must be the Klein 4-group.

Let H be the inertia group of any point $\mathfrak{P}_1$ with ramification index 2. Take $L_0 = \text{Fix } H$. Again we find that $R(L_0/K)$ is type I with $N = 2$, but now the place unramified in $L_0/K$ is not uniquely determined unless the above $\mathfrak{P}_1$ is specified.

We find that $R(L_0/K)$ is either $(2, 2, 1)$, $(2, 1, 2)$ or $(1, 2, 2)$. Theorem 1 then yields that to isomorphism there are precisely three extension fields $L_0$ of $K$.

Choose one of these, say $L_0$ such that $R(L_0/K) = (2, 2, 1)$. As in the argument for $N \neq 4$ we obtain $R(L/L_0) = (2, 2, 1)$, and $L$ is determined.

**III (2, 3, 3)**

Since $[L:K] = 12$, Gal (L/K) has a subgroup $H$ of index 3. Take $L_0 = \text{Fix } H$. Then $|\text{Gal } (L_0/K)| = 3$, and $R(L_0/K)$ is therefore $(1, 3, 3)$. $L_0$ is then determined over $K$. Let $\mathfrak{P}_1, \mathfrak{P}_2, \mathfrak{P}_3$ be the three places in $L_0$ lying over $\mathfrak{P}_1$. Each of these must ramify of index 2 in $L/L_0$.

Hence $R(L/L_0) = (2, 2, 2)$. By the argument for II we know that $L$ is determined to isomorphism for such a ramification pattern. That is, $L$ is determined to isomorphism over $K$, as desired.

**IV (2, 3, 4)**

In a group of order 24 there is always a subgroup of index 2. Let $H$ be such a subgroup in Gal (L/K), and take $L_0 = \text{Fix } H$, as usual. We find that $R(L_0/K)$ must be $(2, 1, 2)$, so that $L_0$ is determined. Let $\mathfrak{P}_1, \mathfrak{P}_2$ be the
two places in \( L_0 \) above \( \wp_2 \); let \( \wp_3 \) be the place in \( L_0 \) above \( \wp_3 \). Then clearly \( R(L/L_0) = (3, 3, 2) \). From the previous case, \( L \) is determined to isomorphism over \( L_0 \), and so over \( K \), as stated.

**Theorem 3**

Let \( L/K \) be a Galois extension and let \( L \) be a rational function field. Let \( G \) be \( \text{Gal} (L/K) \). If \( R(L/K) \) is type \( V \) then \( G \) is simple.

**Proof**

Suppose \( H \) is a normal subgroup in \( G \) of order \( m > 1 \). Let \( L_0 = \text{Fix} H \). Then \( |\text{Gal} (L_0/K)| = m \). We show that such an extension \( L_0/K \) cannot exist.

Recall that \( R(L/K) = (2, 3, 5) \). If \( \wp_3 \) is ramified in \( L_0/K \) it must be totally ramified here, so that \( R(L_0/K) \) must be either \( (1, 5, 5) \) or \( (2, 2, 5) \). But this is impossible since two places in \( L/K \) would then have to ramify, both with indices a multiple of 5 or 2. Thus \( \wp_3 \) is unramified in \( L_0/K \).

Hence there exist at least two places \( \wp_1, \wp_2 \) in \( L_0 \) above \( \wp_3 \).

These ramify totally in \( L/L_0 \) so that \( R(L/L_0) \) must be \( (5, 5, 1) \). But this implies \( [L_0:K] = 12 \). \( \wp_3 \) being unramified in \( L_0/K \) determines \( R(L_0/K) \) as \( (12, 12, 1) \). Contradiction.

**Remark:**

Theorems 1-3 actually establish the first claim of the Main Theorem, namely, that a finite subgroup \( G \) of \( \text{Aut}_k(x) \) for which \( p \nmid |G| \) must be isomorphic to a canonical
group. For, $G$ determines $K = \text{Fix } G$, a rational function field $(1) \subseteq L = k(x)$, and by (1) we know that $R(L/K)$ is type $I-V$. Theorems 1 and 2 then show that if $R(L/K)$ is type $I-IV$, the field extension $L/K$ is determined to isomorphism— but this clearly implies that the group of the extension, $G$, must be determined to isomorphism. Since we know that the canonical groups $I-IV$ have extensions determined by $R(L/K)$, $G$ must be isomorphic to one of these canonical groups. Theorem 3 shows that if $R(L/K)$ is type $V$ then $G = \text{Gal } (L/K)$ is simple of order 60. But the icosahedral group is the only simple group of order 60. Hence $G$, in this case too, is isomorphic to a canonical group.

(1) Theorem 1, Chapter II guarantees that $\text{Fix } G$ is a rational function field.
CHAPTER VIII PROOF OF THE MAIN THEOREM

Recall the hypotheses of the theorem; we have a finite subgroup $G$ of $\text{Aut}_k k(x)$, where $k$ is an algebraically closed field of characteristic $p$ with $p \nmid |G|$. The aim is to show that $G$ must be conjugate to one of the five canonical groups if the ramification pattern of $G$ is type I-IV or if $p = 0$.

The first connection we obtain between $G$ and a canonical group is derived from the Hurwitz Formula, as in Chapter VII. We found there that $G$ must be isomorphic to a canonical group, say $G_0$. Our aim is to show existence of $\sigma \in \text{Aut}_k k(x)$ such that $G^\sigma \equiv \sigma G \sigma^{-1} = G_0$. We can reduce this to a field isomorphism problem by the following argument.

Let $F = \text{Fix } G$ and $F_0 = \text{Fix } G_0$. Then for $\sigma \in \text{Aut}_k k(x)$,

$$\sigma F = F_0 \longrightarrow G^\sigma = G_0$$

Proof:

$G^\sigma \subseteq G_0$: Let $\sigma g \sigma^{-1} \in G^\sigma$. Then, for $f_0 \in F_0$,

$$(\sigma g \sigma^{-1})(f_0) = (\sigma g \sigma^{-1})(\sigma f)$$

for some $f \in F$ such that $\sigma f = f_0$.

But $(\sigma g \sigma^{-1})(\sigma f) = (\sigma g)(f) = \sigma f = f_0$. Thus $\sigma g \sigma^{-1}$ is in $G_0$.

$G_0 \subseteq G^\sigma$: Let $g_0 \in G_0$ and $f_0 \in F_0$. Then, writing $f_0$ as $\sigma f$ for some $f \in F$, we have $g_0(\sigma f) = \sigma f$. But then $(\sigma^{-1} g_0 \sigma)(f) = \sigma^{-1}(\sigma f) = f$. This implies $\sigma^{-1} g_0 \sigma = g$, for some $g \in G$, so that $g_0 = \sigma g \sigma^{-1} \in G^\sigma$.

Thus, if $\text{Fix } G = k(y)$ and $\text{Fix } G_0 = k(y_0)$ for some
elements $y, y_0 \in k(x)$ transcendental over $k$, then the problem is to find $\sigma \in \text{Aut}_{k} k(x)$ such that

$$\sigma(k(y)) = k(y_0) \quad (1)$$

Define $\sigma$ to be the natural isomorphism from $k(y)$ to $k(y_0)$. That is, define $\sigma: k(y) \rightarrow k(y_0)$ by $\sigma: y \mapsto y_0$ and $\sigma: a \mapsto a$ for all $a \in k$. Then $\sigma$ will be the required map if it can be extended to an automorphism of $k(x)$. Extend $\sigma$ as follows:

Let $K$ be an algebraic closure of $k(x)$. Then $\sigma$ can be extended to give an automorphism of $K$ (by Theorem 3, Chapter II). Thus we have the set-up

$$
\begin{array}{ccc}
K & \overset{\sim}{\longrightarrow} & K \\
\downarrow{\sigma} & & \downarrow{\sigma} \\
k(x) & \overset{\sim}{\longrightarrow} & k(\sigma(x)) \\
\downarrow{\sigma} & & \downarrow{\sigma} \\
k(y) & \overset{\sim}{\longrightarrow} & k(y_0)
\end{array}
$$

where now our job is to show that $k(x) = k(\sigma(x))$.

If the ramification pattern of $G$ and $G_0$ is type I-IV, then Theorems 1 and 2 of Chapter VII show that the extensions $k(x)$ and $k(\sigma(x))$ are uniquely determined in the algebraic closure. i.e. $k(x) = k(\sigma(x))$ for these cases. Hence $G$ is conjugate to $G_0$ here.

The remaining case we have to consider is $p = 0$ and

(1) Assume, without loss of generality, $y, y_0$ chosen such that ramification occurs at 1, 0 and $\infty$. 

the ramification pattern of $G$ being type $V$. For this we need several technical results concerning the Schwarzian of $x$ with respect to certain derivations. These results are listed in the following lemmas.

Through the lemmas let $k(z) = \text{Fix} \, G$ for some $z \in k(x)$, $z$ transcendental over $k$.\(^{(2)}\) If $D$ is a derivation on $k(z)$ we know that it can be extended to $k(x)$ since $k(x)/k(z)$ is separable. Thus differentiation with respect to some function of $z$, $f(z)$, on $k(z)$ can be extended to $k(x)$. Denote the Schwarzian of $x$ with respect to this derivation by $[x]_{f(z)}$.

In the first lemma we relate the Schwarzian of $x$ with respect to $z$ \(^{(3)}\) to the Schwarzian of $x$ with respect to $\frac{1}{z}$ and with respect to $z - c$, for some $c \in k$.

**Lemma 1**

(i) $[x]_{z} = \frac{1}{z^4} [x]_{\frac{1}{z}}$

(ii) $[x]_{z} = [x]_{z-c}$ for $c \in k$.

**Proof**

Let $D$ denote differentiation with respect to $z$ on $k(x)$ and let $D_0$ denote differentiation with respect to $\frac{1}{z}$ on $k(x)$.

(i) Recall that $[x]_{z} = 2 \frac{D^3 x}{Dx} - 3 \left( \frac{D^2 x}{Dx} \right)^2$

\(^{(2)}\) Possible by Theorem 1, Chapter II,

\(^{(3)}\) Actually we should write "The Schwarzian of $x$ with respect to the derivation on $k(x)$ induced by differentiation with respect to $z$ on $k(z)$".
and \( \frac{1}{z} \frac{D_0^3 x}{D_0 x} = 2 \left( \frac{D_0^2 x}{D_0 x} \right)^2 - 3 \left( \frac{D_0^2 x}{D_0 x} \right)^2 \). To compare these Schwarzians we first of all compare "like" terms. Notice that extensive use of the Chain Rule is made.

\[
D_0 x = D_x \cdot D_0 z
\]

and \( D_0^2 x = D_0 (D_0 x) = D(D_0 x) \cdot D_0 z \)

Thus \( \frac{D_0^2 x}{D_0 x} = \frac{D(D_0 x) \cdot D_0 z}{D_x \cdot D_0 z} = \frac{D(D_0 x)}{D_x} = \frac{D(D(x \cdot D_0 z)}{D_x} = \frac{D^2 x \cdot D_0 z + D_x \cdot D(D_0 z)}{D_x} \)

But \( D_0 z = -z^2 \). For, let \( y = \frac{1}{z} \). Then \( D_0 \) is differentiation with respect to \( y \), and \( D_0 z = D_0 (\frac{1}{y}) = -\frac{1}{y^2} = -z^2 \). Therefore

(a) \( \frac{D_0^2 x}{D_0 x} = \frac{D^2 x}{D_x} \cdot (-z^2) + (-2z) \)

similarly we have

\[
\frac{D_0^3 x}{D_0 x} = \frac{D(D_0^2 x) \cdot D_0 z}{D_x \cdot D_0 z} = \frac{D(D_0^2 x)}{D_x} = \frac{D[D^2 x \cdot z^4 + D_x \cdot 2z^3]}{D_x} = \frac{D^3 x \cdot z^4 + D^2 x \cdot 4z^3 + D^2 x \cdot 2z^3 + D_x \cdot 6z^2}{D_x}
\]

Thus

(b) \( \frac{D_0^3 x}{D_0 x} = \frac{D^3 x}{D_x} \cdot z^4 + \frac{D^2 x}{D_x} \cdot 6z^3 + 6z^2 \)

From (a) and (b) we have
\[
\frac{[x]}{z} = 2\left[ \frac{D^3}{Dx} \cdot z^4 + \frac{D^2}{Dx} \cdot 6z^3 + 6z^2 \right] - 3\left( \frac{D^2}{Dx} \cdot z^4 + \frac{D^2}{Dx} \cdot 4z^3 + 4z^2 \right) \\
= 2 \frac{D^3}{Dx} \cdot z^4 - 3\left( \frac{D^2}{Dx} \cdot z^4 \right) \\
= [x] \cdot z^4
\]

(ii) This is clear since differentiation with respect to \( z \) is the same as differentiation with respect to \( z - c \) for \( c \in k \).

Define the principal term of \( [x] \) at \( z - c \) for \( c \in k \), to be that part of the Laurent series for \( [x] \) at \( z - c \) which contains negative powers of \( z - c \).

Lemma 2

Let \( e \) be the ramification index of \( k(x)/k(z) \) at \( z - c \), for \( c \in k \). Then the principal term of \( [x] \) at \( z - c \) is

\[(i) \quad \frac{e^2 - 1}{e^2} (z - c)^{-2} + A_0 (z - c)^{-1} \quad \text{if} \quad e > 1, \]

where \( A_0 \) is some constant element.

(ii) 0 if \( e = 1 \).

Proof

(i) Since the ramification index at \( z - c \) is \( e \), we can write \( U(x - C)^e = z - c \) for \( U \) a unit in \( k(x) \) and \( C \) a constant in \( k(x) \).

Notation: Notice that \( [x]_z = [x - C]_{z-c} \). Thus we want the principal term of \( [x - C]_{z-c} \) where \( U(x - C)^e = z - c \).

For simplicity of notation, write \( x \) for \( x - C \), \( z \) for \( z - c \).
The problem then is to find the principal term of \([x]_z\) where \(Ux^e = z\).

Denote differentiation with respect to \(z\) by \(D\), with respect to \(x\) by \(d\).

From \(Ux^e = z\) we have

\[
DUx^e + UDx^e = 1
\]

or,

\[
dUx^e + Uex^{e-1}Dx = 1.
\]

Thus

\[
Dx(dUx^e + eUx^{e-1}) = 1
\]

\[
Dx(1 + \frac{dU}{eU}x) = \frac{1}{eU}x^{1-e}
\]

Let \(W^{-1} = 1 + \frac{dU}{eU}x\)

Then

(a) \(Dx = \frac{W}{eU}x^{1-e}\)

Notice that \(W\) is a 1-unit i.e. a unit whose initial term is 1. For, \(U\) a unit implies \(dU\) has order \(\geq 0\), and \(\frac{1}{eU}\) is also a unit. Hence \(\text{ord} \left(\frac{dU}{eU}x\right) > 0\), so that the constant term of \(W^{-1}\) must be 1.

\[
D^2x = D(Dx) = d(Dx)\cdot Dx
\]

From (a) then we have

(b) \(D^2x = (\frac{1-e}{e} \frac{W}{U} x^{-e} + \frac{1}{e} d(\frac{W}{U}) x^{1-e})Dx\)

Therefore

(c) \(\frac{D^2x}{Dx} = \frac{1-e}{e} \frac{W}{U} x^{-e} + \frac{1}{e} d(\frac{W}{U}) x^{1-e}\)

\(\frac{W}{U}\) is a unit, so \(\text{ord} \left(\frac{d(W)}{U}\right) \geq 0\). Hence

\(\text{ord} \left(\frac{1}{e} d(\frac{W}{U}) x^{1-e}\right) \geq 1-e > -e\). The initial term of \(\frac{D^2x}{Dx}\)

is therefore at worst (4) that of

(4) One initial term is worse than another if it has a larger negative exponent. Notice that we cannot say above that the initial term is... because 1-e may be 0.
\[
\frac{1-e}{e} W x^{-e} = \frac{1-e}{e} (Ux^e)^{-1} W = \frac{1-e}{e} z^{-1} W.
\]

This has initial term \(\frac{1-e}{e} z^{-1}\) since \(W\) is a 1-unit.

From (a) and (b) we have

\[
\frac{D^2 x}{D^3 x} = \frac{1-e}{e^2} (\frac{W}{U}) x^{1-2e} + \frac{1}{e^2} \frac{W}{U} d(W) x^{2(1-e)}
\]

Then, since \(D^3 x = d(D^2 x) \cdot Dx\), we have

\[
\frac{D^3 x}{Dx} = d(D^2 x) = \frac{(1-e)(1-2e)}{e^2} (\frac{W}{U}) x^{2-2e} + \text{(terms of higher degree)}
\]

\[
= \frac{(1-e)(1-2e)}{e^2} (Ux^e)^{-2} W^2 + \ldots
\]

\[
= \frac{(1-e)(1-2e)}{e^2} z^{-2} + \ldots
\]

where, again, we have used the fact that \(W\) is a 1-unit.

Thus the term with largest possible negative exponent in \([x]_z\) is

\[
[2\frac{(1-e)(1-2e)}{e^2} - 3\left(\frac{1-e}{e}\right)^2]z^{-2}
\]

\[
= \frac{2 - 6e + 4e^2 - 3 + 6e - 3e^2}{e^2} z^{-2}
\]

\[
= \frac{e^2 - 1}{e^2} z^{-2}
\]

Hence the principal term of \([x]_z\) is \(\frac{e^2 - 1}{e^2} z^{-2} + A_0 z^{-1}\) for some constant \(A_0\). Replacing \(z\) by \(z-c\) gives the result as stated.

(ii) In the proof of (i) we did not ever use the fact that \(e > 1\). So for \(e = 1\), proof (i) is valid. But now we want to show that for \(e = 1\), the entire principal term is 0.
For $e = 1$, line (c) of (1) reads

$$\frac{D^2 x}{Dx} = d\left(\frac{W}{U}\right).$$

This we know has principal term 0. Line (d), for $e = 1$, becomes

$$\frac{D^3 x}{Dx} = \frac{W}{U} d\left(\frac{W}{U}\right)$$

Thus

$$\frac{D^3 x}{Dx} = d(D^2 x)$$

$$= \left[d\left(\frac{W}{U}\right)\right]^2 + \frac{W}{U} d\left(\frac{W}{U}\right)$$

and this too clearly has principal term 0. Hence the principal term of $2\frac{D^3 x}{Dx} - 3\left(\frac{D^2 x}{Dx}\right)^2$ must be 0, as desired.

**Lemma 3**

Let $e$ be the ramification index of $k(x)/k(z)$ at $\infty$, and assume $e > 1$. Then the expansion of $[x]_z$ at $\infty$ is of the form

$$\frac{e^2 - 1}{e^2} \frac{1}{z^2} + A_0 \frac{1}{z^3} + A_1 \frac{1}{z^4} + A_2 \frac{1}{z^5} + \ldots$$

for some constants $A_0, A_1, A_2, \ldots$

**Proof**

From Lemma 1(i), $[x]_z = \frac{1}{z^4} [x]_z^1$. Using this fact together with Lemma 2(i) we find that the terms of least possible degree in the expansion of $[x]_z$ at $\infty$ are

$$\frac{1}{z^4} \left(\frac{e^2 - 1}{e^2} \left(\frac{1}{z}\right)^{-2} + A_0 \left(\frac{1}{z}\right)^{-1}\right)$$

$$= \frac{e^2 - 1}{e^2} \frac{1}{z^2} + A_0 \frac{1}{z^3}, \text{ for some constant } A_0.$$
Note: Lemma 3 tells us that if there is ramification at \( \infty \) then \([x]_z\) has no pole at \( \infty \): it has a zero at \( \infty \) of order at least 2.

The next lemma, deduced using the previous two, gives the basic result needed to conclude the Main Theorem. Namely, that the Schwarzian of \( x \) with respect to \( z \) is a rational function whose coefficients depend only on the ramification in \( k(x)/k(z) \).

**Main Lemma**

There is a rational function \( r(z) \in k(z) \) such that

\[
[x]_z = r(z)
\]

where the coefficients of \( r(z) \) depend only on the ramified places and ramification indices of \( k(x)/k(z) \).

**Proof**

Rationality of \([x]_z\) follows from Corollary 2 of Chapter IV. Suppose \([x]_z = r(z) \in k(z)\).

Assume without loss of generality the generator \( z \) of \( \text{Fix} \ G \) chosen such that ramification occurs at 1, 0 and \( \infty \) (that is, at the places \((z-1)\), \((z)\) and \((\frac{1}{z})\)) with ramification indices \( e_1 \), \( e_2 \) and \( e_3 \) respectively. From Lemmas 2 and 3 we find that \( r(z) \) has poles at most at 0 and 1: not at \( \infty \) or at any unramified place. Thus \( r(z) \) has the form

\[
(*) \quad r(z) = \frac{e_1^2 - 1}{e_1^2(z-1)^2} + \frac{A}{z-1} + \frac{e_2^2 - 1}{e_2^2 z^2} + \frac{B}{z} + C
\]

for some constants \( A, B, C \). (\( C \) is a constant, since if
it were a nonconstant polynomial, \( r(z) \) would have a pole at \( \infty \). Contradiction.)

From (*) we can calculate the initial terms of the expansion of \( r(z) \) at \( \infty \) in terms of the constants \( A, B \) and \( C \). But Lemma 3 already tells us what these initial terms are. Equating like coefficients in the expansions will allow us to solve for \( A, B \) and \( C \), and so to write down an explicit formula for \( r(z) \).

In (*) the initial terms of \( r(z) \) at \( \infty \) are found by rewriting \( r(z) \) in terms of \( u = \frac{1}{z} \). Then \( z = \frac{1}{u} \) and (*) yields

\[
r(z) = r\left(\frac{1}{u}\right) = \frac{e_1^2 - 1}{e_1^2 (1-u)^2} + \frac{A}{1-u} + \frac{e_2^2 - 1}{e_2^2 (1-u)^2} + \frac{B}{1-u} + C
\]

\[
= \frac{e_1^2 - 1}{e_1^2} \frac{u^2}{1-u^2} + A \frac{u}{1-u} + \frac{e_2^2 - 1}{e_2^2} u^2 + Bu + C
\]

\[
= \frac{e_1^2 - 1}{e_1^2} (u^2 + 2u^3 + 3u^4 + ...) + A(u + u^2 + u^3 + ...)
\]

\[
+ \frac{e_2^2 - 1}{e_2^2} u^2 + Bu + C
\]

\[
= C + (A + B)u + \left( \frac{e_1^2 - 1}{e_1^2} + A + \frac{e_2^2 - 1}{e_2^2} \right) u^2 + ...
\]

Lemma 3 states that the initial terms of this expansion must be

\[
\frac{e_3^2 - 1}{e_3^2} \frac{1}{(1-u)^2} + A_0 \frac{1}{(1-u)^3} + A_1 \frac{1}{(1-u)^4}
\]
\[ e_1^2 - 1 \frac{u^2}{e_3^2} + A_0 u^3 + A_1 u^4 \]

Equating coefficients of \( u^0, u, u^2 \) gives

\[ C = 0 \]
\[ A + B = 0 \]
\[ \frac{e_1^2 - 1}{e_3^2} + A + \frac{e_2^2 - 1}{e_2} = \frac{e_3^2 - 1}{e_3} \]

Therefore \( A = \frac{1}{e_1^2} + \frac{1}{e_2} - \frac{1}{e_3} - 1 \) and \( B = -A \)

Substituting these values into equation (*) gives

\[ r(z) = \frac{e_1^2 - 1}{e_1}(z-1)^{-2} + \frac{e_2^2 - 1}{e_2}z^{-2} + \left(\frac{1}{e_1} + \frac{1}{e_2} - \frac{1}{e_3} - 1\right)[(z)(z-1)]^{-1} \]

The coefficients of \( r(z) \) do, therefore, depend only on \( e_1, e_2, e_3 \).

Now we are in shape to prove the remaining case in the Main Theorem.

Let \( D, D_0 \) denote the extensions to \( k(x) \) of differentiation with respect to \( y, y_0 \) in \( k(y), k(y_0) \) respectively. Assuming, without loss of generality, that the generators \( y, y_0 \) have been chosen such that ramification occurs at \( 1, 0 \) and \( \infty \), then \( k(x)/k(y) \) and \( k(x)/k(y_0) \) have the same ramified places and ramification indices.

By the Main Lemma then, there is a rational function \( r \) with coefficients in \( k \) depending only on the ramification
indices of 1, 0 and \( \infty \) such that
\[
[x]_D = r(y) \quad \text{and} \quad [x]_{D_0} = r(y_0).
\]

But then it is easy to see that
\[
\sigma([x]_D) = [x]_{D_0}
\]

For, \( \sigma([x]_D) = \sigma(r(y)) = r(\sigma(y)) = r(y_0) = [x]_{D_0} \). Also,

since \( \sigma \) is an isomorphism we find that
\[
\sigma([x]_D) = [\sigma(x)]_{D_0}
\]

Then, from these last two equations, we have
\[
[x]_D = [\sigma(x)]_{D_0}
\]

By Corollaries 1 and 2 of Chapter IV it follows that \( k(x) = k(\sigma(x)) \), as desired.

Remarks:

(1) In the hypotheses of the Main Theorem we state \( p \nmid |G| \). In the proof (derivation of the Hurwitz Formula in particular), we used only the fact that \( p \nmid e \) for any ramification index \( e \). So our original hypothesis might seem too strong, in that it is plausible that the case \( p \mid |G| \) but \( p \nmid e \) for any ramification index \( e \) arise. But the result of the theorem guarantees that this cannot happen. For, the theorem says that if \( p \nmid e \) for all \( e \), then \( |G| \) and the ramification indices \( e \) must be one of the five canonical types. We know for these that if \( p \mid |G| \) (\( p \) a prime) then \( p = e \), for some ramification index \( e \).
(2) It turns out to be true that for any algebraically closed field \( k \), a finite subgroup of \( \text{Aut}_k k(x) \) must be conjugate to a canonical group. That is, the major claim of our Main Theorem does in fact hold for the icosahedral case in characteristic \( p > 0 \), although it does not yield to proof using the techniques here developed for the case of characteristic 0.

That the theorem is true for all cases in characteristic \( p > 0 \) follows from Dickson [1]. From §259 we have: if \( k_0 \) is the algebraic closure of the prime field \( \mathbb{Z}/p\mathbb{Z} \) and if a subgroup of \( \text{Aut}_{k_0} k_0(x) \) is isomorphic to the icosahedral group then that subgroup is conjugate to the icosahedral group. By standard techniques of group representations and group extensions this result holds for any algebraically closed field \( k \).

Dickson's techniques, however, do not work for both characteristic 0 and characteristic \( p \). Dickson uses finite permutation groups to deduce his result for characteristic \( p \)-- in characteristic 0 the corresponding devices are simply not finite. Thus, in characteristic 0, some other method such as the one used here is necessary.
Bibliography


Postscript to Beth Kitchen's Thesis

by Klaus Hoechsmann

After the completion of this thesis and shortly before it was due to be handed in, we stumbled on a simple proof of the icosahedral case, which works in any characteristic 2, 3, 5. It hinges on the easy arithmetic of what we shall call affine extensions $L/K : L = k(x), K = k(y)$ and, $k[y] \subseteq k[x]$. In other words, a pair $L \supseteq K$ of rational function fields is affine, if generators $x, y$ can be found such that $y$ is a polynomial in $x$. As above, $k$ will be algebraically closed. The following lemma is a synthesis of the lemmas on pages 35 and 53.

**Lemma:** $L/K$ affine $\iff$ some place of $K$ is totally ramified in $L$.

**Proof:** ($\Leftarrow$) Let $\mathfrak{p}, \mathfrak{q}$ be the places in question. As in the lemmas on pages 35 and 53, we have

\[(*) \quad w^N + a_{N-1}u w^{N-1} + a_{N-2}u w^{N-2} + \ldots + u = 0\]

where $w \in L, u \in K$ are local uniformizing parameters for $\mathfrak{p}, \mathfrak{q}$, respectively. Put $x = \frac{1}{w}, y = \frac{1}{u}$. ($\Rightarrow$) From an equation $F(x) = y$ with a polynomial $F$ of degree $N$, we get an Eisenstein equation $(*)$ for $w = \frac{1}{x}, u = \frac{1}{y}$; hence total ramification.
Remark: In an affine extension, the splitting of any place of $K$ (other than the $y$ singled out above) is reflected completely in the factorization of its uniformizing parameter $y - c = F(x) - c$ in $k[x]$. For instance, if there is another totally ramified place, we obtain a pure equation

$$y - c = (x - b)^N.$$ 

Our derivation of the icosahedral equation will follow the same lines.

To prove the conjugacy of two icosahedral groups, we choose a subgroup of index 5 in each of them. These are conjugate; making them equal by a suitable inner automorphism, we now have two icosahedral groups which intersect in a tetrahedral group hence a field $L$ containing fields $K_1, K_2$ such that $[L : K_i] = 5$, $(i = 1,2)$. It is easy to see that ramification in each $L/K_i$ must necessarily be as follows:

We choose the generator $x$ of $L$ so that the uniformizing parameters of $\mathcal{D}, \mathcal{D}_1, \mathcal{R}_1, \mathcal{R}_2$ are $\frac{1}{x}, x, x - \alpha, x - \beta$, respectively, where $\alpha + \beta = -4$.

By our lemma and its proof, each $K_1$ is generated by a $y = F(x)$. Let us insist that $F(x)$ be monic and that $y$ be a local parameter for the
place \( \mathcal{C} \). We must show that \( F(x) \) is uniquely determined by these data.

Ramification of \( \mathcal{C} \) and \( \mathcal{Y} \) gives us

\[
(1) \quad y = x^3(x^2 + ax + b) \quad \text{and}
\]

\[
(2) \quad y - \delta = (x-a)^2(x-\beta)^2(x-\gamma) \quad \text{respectively.}
\]

Accordingly, \( F(x) = x^5 + ax^4 + 6x^3 \). We shall see that \( a = 5 \), \( b = 40 \).

Obviously, \( x^2(x-a)(x-\beta) \) must divide the derivative \( F'(x) = (5x^2 + 4ax + 3b)x^2 \), whence \( (x-a)(x-\beta) \) is determined. Thus

\[
(2') \quad y - c = (x^2 + \frac{4}{3}ax + \frac{3}{5}b)^2(x-\gamma).
\]

Comparing the linear and quadratic terms of \( (2') \) to those of \( (1) \) yields

\[
(i) \quad 3b - 8\gamma a = 0,
\]

\[
(ii) \quad 12ab - 15\gamma b - 8\gamma a^2 = 0, \quad \text{whence} \quad 3a = 5\gamma.
\]

Therefore \( (i) \) becomes: \( b = \frac{8}{5}a \). Since we had normalized \(-4 = a + \beta = -\frac{4}{5}a\), we have \( a = 5 \) and \( b = 40 \). Quod erat demonstrandum.