Lie Groups

Robert Gilmore
Many years ago I wrote the book *Lie Groups, Lie Algebras, and Some of Their Applications* (NY: Wiley, 1974). That was a big book: long and difficult. Over the course of the years I realized that more than 90% of the most useful material in that book could be presented in less than 10% of the space. This realization was accompanied by a promise that some day I would do just that — rewrite and shrink the book to emphasize the most useful aspects in a way that was easy for students to acquire and to assimilate. The present work is the fruit of this promise.

In carrying out the revision I’ve created a sandwich. Lie group theory has its intellectual underpinnings in Galois theory. In fact, the original purpose of what we now call Lie group theory was to use continuous groups to solve differential (continuous) equations in the spirit that finite groups had been used to solve algebraic (finite) equations. It is rare that a book dedicated to Lie groups begins with Galois groups and includes a chapter dedicated to the applications of Lie group theory to solving differential equations. This book does just that. The first chapter describes Galois theory, and the last chapter shows how to use Lie theory to solve some ordinary differential equations. The fourteen intermediate chapters describe many of the most important aspects of Lie group theory and provide applications of this beautiful subject to several important areas of physics and geometry.

Over the years I have profitted from the interaction with many students through comments, criticism, and suggestions for new material or different approaches to old. Three students who have contributed enormously during the past few years are Dr. Jairzinho Ramos-Medina, who worked with me on Chapter 15 (Maxwell’s Equations), and Daniel J. Cross and Timothy Jones, who aided this computer illiterate with much moral and ether support. Finally, I thank my beautiful wife Claire for her gracious patience and understanding throughout this long creation process.
# Contents

1 Introduction  
1.1 The Program of Lie  
1.2 A Result of Galois  
1.3 Group Theory Background  
1.4 Approach to Solving Polynomial Equations  
1.5 Solution of the Quadratic Equation  
1.6 Solution of the Cubic Equation  
1.7 Solution of the Quartic Equation  
1.8 The Quintic Cannot be Solved  
1.9 Example  
1.10 Conclusion  
1.11 Problems  

2 Lie Groups  
2.1 Algebraic Properties  
2.2 Topological Properties  
2.3 Unification of Algebra and Topology  
2.4 Unexpected Simplification  
2.5 Conclusion  
2.6 Problems  

3 Matrix Groups  
3.1 Preliminaries  
3.2 No Constraints  
3.3 Linear Constraints  
3.4 Bilinear and Quadratic Constraints  
3.5 Multilinear Constraints  
3.6 Intersections of Groups  
3.7 Embedded Groups
3.8 Modular Groups 48
3.9 Conclusion 50
3.10 Problems 50

4 Lie Algebras 61
4.1 Why Bother? 61
4.2 How to Linearize a Lie Group 63
4.3 Inversion of the Linearization Map: EXP 64
4.4 Properties of a Lie Algebra 66
4.5 Structure Constants 68
4.6 Regular Representation 69
4.7 Structure of a Lie Algebra 70
4.8 Inner Product 71
4.9 Invariant Metric and Measure on a Lie Group 74
4.10 Conclusion 76
4.11 Problems 76

5 Matrix Algebras 82
5.1 Preliminaries 82
5.2 No Constraints 83
5.3 Linear Constraints 83
5.4 Bilinear and Quadratic Constraints 86
5.5 Multilinear Constraints 89
5.6 Intersections of Groups 89
5.7 Algebras of Embedded Groups 90
5.8 Modular Groups 91
5.9 Basis Vectors 91
5.10 Conclusion 93
5.11 Problems 93

6 Operator Algebras 98
6.1 Boson Operator Algebras 98
6.2 Fermion Operator Algebras 99
6.3 First Order Differential Operator Algebras 100
6.4 Conclusion 103
6.5 Problems 104

7 EXPonentiation 110
7.1 Preliminaries 110
7.2 The Covering Problem 111
7.3 The Isomorphism Problem and the Covering Group 116
7.4 The Parameterization Problem and BCH Formulas 121
7.5 EXPonentials and Physics 127
7.5.1 Dynamics 127
7.5.2 Equilibrium Thermodynamics 129
7.6 Conclusion 132
7.7 Problems 133

8 Structure Theory for Lie Algebras 145
8.1 Regular Representation 145
8.2 Some Standard Forms for the Regular Representation 146
8.3 What These Forms Mean 149
8.4 How to Make This Decomposition 152
8.5 An Example 153
8.6 Conclusion 154
8.7 Problems 154

9 Structure Theory for Simple Lie Algebras 157
9.1 Objectives of This Program 157
9.2 Eigenoperator Decomposition – Secular Equation 158
9.3 Rank 161
9.4 Invariant Operators 161
9.5 Regular Elements 164
9.6 Semisimple Lie algebras 166
9.6.1 Rank 166
9.6.2 Properties of Roots 166
9.6.3 Structure Constants 168
9.6.4 Root Reflections 169
9.7 Canonical Commutation Relations 169
9.8 Conclusion 171
9.9 Problems 173

10 Root Spaces and Dynkin Diagrams 179
10.1 Properties of Roots 179
10.2 Root Space Diagrams 181
10.3 Dynkin Diagrams 185
10.4 Conclusion 189
10.5 Problems 191

11 Real Forms 194
11.1 Preliminaries 194
11.2 Compact and Least Compact Real Forms 197
11.3 Cartan’s Procedure for Constructing Real Forms 199
11.4 Real Forms of Simple Matrix Lie Algebras 200
11.4.1 Block Matrix Decomposition 201
11.4.2 Subfield Restriction 201
11.4.3 Field Embeddings 204
11.5 Results 204
11.6 Conclusion 205
11.7 Problems 206

12 Riemannian Symmetric Spaces 213
12.1 Brief Review 213
12.2 Globally Symmetric Spaces 215
12.3 Rank 216
12.4 Riemannian Symmetric Spaces 217
12.5 Metric and Measure 218
12.6 Applications and Examples 219
12.7 Pseudo Riemannian Symmetric Spaces 222
12.8 Conclusion 223
12.9 Problems 224

13 Contraction 232
13.1 Preliminaries 233
13.2 Inönü–Wigner Contractions 233
13.3 Simple Examples of Inönü–Wigner Contractions 234
13.3.1 The Contraction \( SO(3) \rightarrow ISO(2) \) 234
13.3.2 The Contraction \( SO(4) \rightarrow ISO(3) \) 235
13.3.3 The Contraction \( SO(4,1) \rightarrow ISO(3,1) \) 237
13.4 The Contraction \( U(2) \rightarrow H_4 \) 239
13.4.1 Contraction of the Algebra 239
13.4.2 Contraction of the Casimir Operators 240
13.4.3 Contraction of the Parameter Space 240
13.4.4 Contraction of Representations 241
13.4.5 Contraction of Basis States 241
13.4.6 Contraction of Matrix Elements 242
13.4.7 Contraction of BCH Formulas 242
13.4.8 Contraction of Special Functions 243
13.5 Conclusion 244
13.6 Problems 245

14 Hydrogenic Atoms 250
14.1 Introduction 251
14.2 Two Important Principals of Physics 252
14.3 The Wave Equations 253
14.4 Quantization Conditions 254
14.5 Geometric Symmetry \( SO(3) \) 257
14.6 Dynamical Symmetry \( SO(4) \) 261
14.7 Relation With Dynamics in Four Dimensions 264
14.8 DeSitter Symmetry $SO(4, 1)$ 266
14.9 Conformal Symmetry $SO(4, 2)$ 270
  14.9.1 Schwinger Representation 270
  14.9.2 Dynamical Mappings 271
  14.9.3 Lie Algebra of Physical Operators 274
14.10 Spin Angular Momentum 275
14.11 Spectrum Generating Group 277
  14.11.1 Bound States 278
  14.11.2 Scattering States 279
  14.11.3 Quantum Defect 280
14.12 Conclusion 281
14.13 Problems 282

15 Maxwell’s Equations 293
15.1 Introduction 294
15.2 Review of the Inhomogeneous Lorentz Group 295
  15.2.1 Homogeneous Lorentz Group 295
  15.2.2 Inhomogeneous Lorentz Group 296
15.3 Subgroups and Their Representations 296
  15.3.1 Translations $\{I, a\}$ 297
  15.3.2 Homogeneous Lorentz Transformations 297
  15.3.3 Representations of $SO(3, 1)$ 298
15.4 Representations of the Poincaré Group 299
  15.4.1 Manifestly Covariant Representations 299
  15.4.2 Unitary Irreducible Representations 300
15.5 Transformation Properties 305
15.6 Maxwell’s Equations 308
15.7 Conclusion 309
15.8 Problems 310

16 Lie Groups and Differential Equations 320
16.1 The Simplest Case 322
16.2 First Order Equations 323
  16.2.1 One Parameter Group 323
  16.2.2 First Prolongation 323
  16.2.3 Determining Equation 324
  16.2.4 New Coordinates 325
  16.2.5 Surface and Constraint Equations 326
  16.2.6 Solution in New Coordinates 327
  16.2.7 Solution in Original Coordinates 327
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.3 An Example</td>
<td>327</td>
</tr>
<tr>
<td>16.4 Additional Insights</td>
<td>332</td>
</tr>
<tr>
<td>16.4.1 Other Equations, Same Symmetry</td>
<td>332</td>
</tr>
<tr>
<td>16.4.2 Higher Degree Equations</td>
<td>333</td>
</tr>
<tr>
<td>16.4.3 Other Symmetries</td>
<td>333</td>
</tr>
<tr>
<td>16.4.4 Second Order Equations</td>
<td>333</td>
</tr>
<tr>
<td>16.4.5 Reduction of Order</td>
<td>335</td>
</tr>
<tr>
<td>16.4.6 Higher Order Equations</td>
<td>336</td>
</tr>
<tr>
<td>16.4.7 Partial Differential Equations: Laplace’s Equation</td>
<td>337</td>
</tr>
<tr>
<td>16.4.8 Partial Differential Equations: Heat Equation</td>
<td>338</td>
</tr>
<tr>
<td>16.4.9 Closing Remarks</td>
<td>338</td>
</tr>
<tr>
<td>16.5 Conclusion</td>
<td>339</td>
</tr>
<tr>
<td>16.6 Problems</td>
<td>341</td>
</tr>
</tbody>
</table>

**Bibliography**

**Index**

347

351
1

Introduction

Contents

1.1 The Program of Lie 1
1.2 A Result of Galois 3
1.3 Group Theory Background 4
1.4 Approach to Solving Polynomial Equations 9
1.5 Solution of the Quadratic Equation 10
1.6 Solution of the Cubic Equation 12
1.7 Solution of the Quartic Equation 15
1.8 The Quintic Cannot be Solved 18
1.9 Example 19
1.10 Conclusion 22
1.11 Problems 23

Lie groups were initially introduced as a tool to solve or simplify ordinary and partial differential equations. The model for this application was Galois’ use of finite groups to solve algebraic equations of degree two, three, and four, and to show that the general polynomial equation of degree greater than four could not be solved by radicals. In this chapter we show how the structure of the finite group that leaves a quadratic, cubic, or quartic equation invariant can be used to develop an algorithm to solve that equation.

1.1 The Program of Lie

Marius Sophus Lie (1842 - 1899) embarked on a program that is still not complete, even after a century of active work. This program attempts to use the power of the tool called group theory to solve, or at least simplify, ordinary differential equations.

Earlier in that century, Évariste Galois (1811 - 1832) had used group theory to solve algebraic (polynomial) equations that were quadratic, cubic, and quartic. In fact, he did more. He was able to prove that
Introduction

no closed form solution could be constructed for the general quintic (or any higher degree) equation using only the four standard operations of arithmetic ($+, -, \times, \div$) as well as extraction of the $n$th roots of a complex number.

Lie initiated his program on the basis of analogy. If finite groups were required to decide on the solvability of finite-degree polynomial equations, then ‘infinite groups’ (i.e., groups depending continuously on one or more real or complex variables) would probably be involved in the treatment of ordinary and partial differential equations. Further, Lie knew that the structure of the polynomial’s invariance (Galois) group not only determined whether the equation was solvable in closed form, but also provided the algorithm for constructing the solution in the case that the equation was solvable. He therefore felt that the structure of an ordinary differential equation’s invariance group would determine whether or not the equation could be solved or simplified and, if so, the group’s structure would also provide the algorithm for constructing the solution or simplification.

Lie therefore set about the program of computing the invariance group of ordinary differential equations. He also began studying the structure of the children he begat, which we now call Lie groups.

Lie groups come in two basic varieties: the simple and the solvable. Simple groups have the property that they regenerate themselves under commutation. Solvable groups do not, and contain a chain of subgroups, each of which is an invariant subgroup of its predecessor.

Simple and solvable groups are the building blocks for all other Lie groups. Semisimple Lie groups are direct products of simple Lie groups. Non-semisimple Lie groups are semidirect products of (semi)simple Lie groups with invariant subgroups that are solvable.

Not surprisingly, solvable Lie groups are related to the integrability, or at least simplification, of ordinary differential equations. However, simple Lie groups are more rigidly constrained, and form such a beautiful subject of study in their own right that much of the effort of mathematicians during the last century involved the classification and complete enumeration of all simple Lie groups and the discussion of their properties. Even today, there is no complete classification of solvable Lie groups, and therefore non-semisimple Lie groups.

Both simple and solvable Lie groups play an important role in the study of differential equations. As in Galois’ case of polynomial equations, differential equations can be solved or simplified by quadrature if their invariance group is solvable. On the other hand, most of the classi-
1.2 A Result of Galois

In 1830 Galois developed machinery that allowed mathematicians to definitively resolve questions that had eluded answers for 2000 years or longer. These questions included the three famous challenges to ancient Greek geometers: Whether by ruler and compasses alone it was possible to

- square a circle
- trisect an angle
- double a cube.

His work helped to resolve longstanding questions of an algebraic nature: Whether it was possible, using only the operations of arithmetic together with the operation of constructing radicals, to solve

- cubic equations
- quartic equations
- quintic equations.

This branch of mathematics, now called Galois theory, continues to provide powerful new results, such as supplying answers and solution methods to the following questions:

- Can an algebraic expression be integrated in closed form?
- Under what conditions can errors in a binary code be corrected?

This beautiful machine, applied to a problem, provides important results. First, it can determine whether a solution is possible or not under the conditions specified. Second, if a solution is possible, it suggests the structure of the algorithm that can be used to construct the solution in a finite number of well-defined steps.

Galois’ approach to the study of algebraic (polynomial) equations involved two areas of mathematics, now called field theory and group theory. One useful statement of Galois’ result is [50, 66]:

**Theorem:** A polynomial equation over the complex field is solvable by radicals if and only if its Galois group $G$ contains a chain of subgroups $G = G_0 ⊃ G_1 ⊃ \cdots ⊃ G_ω = I$ with the properties:
1. $G_{i+1}$ is an invariant subgroup of $G_i$;
2. Each factor group $G_i/G_{i+1}$ is commutative.

In the statement of this theorem the field theory niceties are contained in the term ‘solvable by radicals.’ This means that in addition to the four standard arithmetic operations $+,-,\times,\div$ one is allowed the operation of taking $n$th roots of complex numbers.

The principal result of this theorem is stated in terms of the structure of the group that permutes the roots of the polynomial equation among themselves. Determining the structure of this group is a finite, and in fact very simple, process.

### 1.3 Group Theory Background

A group $G$ is defined as follows: It consists of a set of operations $G = \{g_1, g_2, \ldots\}$, called group operations, together with a combinatorial operation, $\cdot$, called group multiplication, such that the following four axioms are satisfied:

(i) Closure: If $g_i \in G$, $g_j \in G$, then $g_i \cdot g_j \in G$.
(ii) Associativity: for all $g_i \in G$, $g_j \in G$, $g_k \in G$,

\[
(g_i \cdot g_j) \cdot g_k = g_i \cdot (g_j \cdot g_k)
\]

(iii) Identity: There is a group operation, $I$ (identity operator), with the property that

\[
g_i \cdot I = g_i = I \cdot g_i
\]

(iv) Inverse: Every group operation $g_i$ has an inverse (called $g_i^{-1}$):

\[
g_i \cdot g_i^{-1} = I = g_i^{-1} \cdot g_i
\]

The Galois group $G$ of a general polynomial equation

\[
(z - z_1)(z - z_2) \cdots (z - z_n) = 0
\]

\[
z^n - I_1z^{n-1} + I_2z^{n-2} + \cdots + (-1)^nI_n = 0 \quad (1.1)
\]

is the group that permutes the roots $z_1$, $z_2$, $\cdots$, $z_n$ among themselves and leaves the equation invariant:

\[
\begin{bmatrix}
  z_1 \\
  z_2 \\
  \vdots \\
  z_n
\end{bmatrix}
\rightarrow
\begin{bmatrix}
  z_1 \\
  z_2 \\
  \vdots \\
  z_n
\end{bmatrix}
\quad (1.2)
\]
This group, called the permutation group $P_n$ or the symmetric group $S_n$, has $n!$ group operations. Each group operation is some permutation of the roots of the polynomial; the group multiplication is composition of successive permutations.

The permutation group $S_n$ has a particularly convenient representation in terms of $n \times n$ matrices. These matrices have one nonzero element, +1, in each row and each column. For example, the $6=3!$ $3 \times 3$ matrices for the permutation representation of $S_3$ are

\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\rightarrow
\begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0 \\
\end{bmatrix}
\rightarrow
\begin{bmatrix}
0 & 0 & 1 \\
1 & 0 & 0 \\
0 & 1 & 0 \\
\end{bmatrix}
\]

(1.3)

The symbol $(123)$ means that the first root, $z_1$, is replaced by $z_2$, $z_2$ is replaced by $z_3$, and $z_3$ is replaced by $z_1$

\[
\begin{bmatrix}
z_1 \\
z_2 \\
z_3 \\
\end{bmatrix}
\rightarrow
\begin{bmatrix}
z_2 \\
z_3 \\
z_1 \\
\end{bmatrix}
\]

(1.4)

The permutation matrix associated with this group operation carries out the same permutation

\[
\begin{bmatrix}
z_2 \\
z_3 \\
z_1 \\
\end{bmatrix}
=
\begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0 \\
\end{bmatrix}
\begin{bmatrix}
z_1 \\
z_2 \\
z_3 \\
\end{bmatrix}
\]

(1.5)

More generally, a matrix representation of a group is a mapping of each group operation into an $n \times n$ matrix that preserves the group multiplication operation

\[
g_i \cdot g_j = g_i \cdot g_j
\]

\[
\downarrow \quad \downarrow \quad \downarrow
\]

\[
\Gamma(g_i) \times \Gamma(g_j) = \Gamma(g_i \cdot g_j)
\]

(1.6)

Here $\cdot$ represents the multiplication operation in the group (i.e., composition of substitutions in $S_n$) and $\times$ represents the multiplication operation among the matrices (i.e., matrix multiplication). The condition (1.6) that defines a matrix representation of a group, $G \rightarrow \Gamma(G)$, is that the product of matrices representing two group operations $[\Gamma(g_i) \times \Gamma(g_j)]$
is equal to the matrix representing the product of these operations in the group $[\Gamma(g_i \cdot g_j)]$ for all group operations $g_i, g_j \in G$.

This permutation representation of $S_3$ is 1:1, or a faithful representation of $S_3$, since knowledge of the $3 \times 3$ matrix uniquely identifies the original group operation in $S_3$.

A subgroup $H$ of the group $G$ is a subset of group operations in $G$ that is closed under the group multiplication in $G$.

**Example:** The subset of operations $I, (123), (321)$ forms a subgroup of $S_3$. This particular subgroup is denoted $A_3$ (alternating group). It consists of those operations in $S_3$ whose determinants, in the permutation representation, are $+1$. The group $S_3$ has three two-element subgroups:

$$S_2(12) = \{I,(12)\}$$
$$S_2(23) = \{I,(23)\}$$
$$S_2(13) = \{I,(13)\}$$

as well as the subgroup consisting of the identity alone. The alternating subgroup $A_3 \subset S_3$ and the three two-element subgroups $S_2(ij)$ of $S_3$ are illustrated in Fig. 1.1. The set of operations $I, (123), (12)$ does not constitute a subgroup because products of operations in this subset do not lie in this subset: $(123) \cdot (123) = (321), (123) \cdot (12) = (23)$, etc. In fact, the two operations $(123), (12)$ generate $S_3$ by taking products of various lengths in various order.

![Fig. 1.1. Subgroups of $S_3$](image-url)

A group $G$ is commutative, or abelian, if

$$g_i \cdot g_j = g_j \cdot g_i$$

(1.7)
for all group operations \(g_i, g_j \in G\).

**Example:** \(S_3\) is not commutative, while \(A_3\) is. For \(S_3\) we have

\[(12)(23) = (321) \quad \text{and} \quad (123) \neq (321) \quad (1.8)\]

\[(23)(12) = (123) \quad (1.9)\]

Two subgroups of \(G\), \(H_1 \subset G\) and \(H_2 \subset G\) are *conjugate* if there is a group element \(g \in G\) with the property

\[gH_1g^{-1} = H_2 \quad (1.9)\]

**Example:** The subgroups \(S_2(12)\) and \(S_2(13)\) are conjugate in \(S_3\) since

\[(23)S_2(12)(23)^{-1} = (23) \{I, (12)\} (23)^{-1} = \{I, (13)\} = S_2(13) \quad (1.10)\]

On the other hand, the alternating group \(A_3 \subset S_3\) is *self-conjugate*, since any operation in \(G = S_3\) serves merely to permute the group operations in \(A_3\) among themselves:

\[(23)A_3(23)^{-1} = (23) \{I, (123), (321)\} (23)^{-1} = \{I, (321), (123)\} = A_3 \quad (1.11)\]

A subgroup \(H \subset G\) which is self-conjugate under all operations in \(G\) is called an *invariant subgroup* of \(G\), or *normal subgroup* of \(G\).

In constructing group-subgroup diagrams, it is customary to show only one of the mutually conjugate subgroups. This simplifies Fig. 1.1 to Fig. 1.2.

![Fig. 1.2. Subgroups of \(S_3\), combining conjugate subgroups](image)

A mapping \(f\) from a group \(G\) with group operations \(g_1, g_2, \ldots\) and group multiplication \(\cdot\) to a group \(H\) with group operations \(h_1, h_2, \ldots\)
and group multiplication $\times$ is called a homomorphism if it preserves group multiplication:

$$
g_i \cdot g_j = g_i \cdot g_j
$$

(1.12)

The group $H$ is called a homomorphic image of $G$. Several different group elements in $G$ may map to a single group element in $H$. Every element $h_i \in H$ has the same number of inverse images $g_j \in G$. If each group element $h \in H$ has a unique inverse image $g \in G$ ($h_1 = f(g_1)$ and $h_2 = f(g_2)$, $h_1 = h_2 \Rightarrow g_1 = g_2$) the mapping $f$ is an isomorphism.

**Example:** The 3:1 mapping $f$ of $S_3$ onto $S_2$ given by

$$
\begin{align*}
S_3 & \xrightarrow{f} S_2 \\
I, (123), (321) & \longrightarrow I \\
(12), (23), (31) & \longrightarrow (12)
\end{align*}
$$

is a homomorphism.

**Example:** The 1:1 mapping of $S_3$ onto the six $3 \times 3$ matrices given in (1.3) is an isomorphism.

**Remark:** Homomorphisms of groups to matrix groups, such as that in (1.3), are called matrix representations. The representation in (1.3) is 1:1 or faithful, since the mapping is an isomorphism.

**Remark:** Isomorphic groups are indistinguishable at the algebraic level. Thus, when an isomorphism exists between a group and a matrix group, it is often preferable to study the matrix representation of the group since the properties of matrices are so well known and familiar. This is the approach we pursue in Chapter 3 when discussing Lie groups.

If $H$ is a subgroup of $G$, it is possible to write every group element in $G$ as a product of an element $h$ in the subgroup $H$ with a group element in a ‘quotient,’ or coset (denoted $G/H$). A coset is a subset of $G$. If the order of $G$ is $|G|$ ($S_3$ has $3! = 6$ group elements, so the order of $S_3$ is 6), then the order of $G/H$ is $|G|/|H|$. For example, for subgroups $H = A_3 = \{I, (123), (321)\}$ and $H = S_2(23) = \{I, (23)\}$ we have

$$
\begin{align*}
G/H \cdot H &= G \\
\{I, (12)\} \cdot \{I, (123), (321)\} &= \{I, (123), (321), (12), (13), (23)\} \\
\{I, (12), (321)\} \cdot \{I, (23)\} &= \{I, (23), (12), (123), (321), (13)\}
\end{align*}
$$

(1.14)

The choice of the $|G|/|H|$ group elements in the quotient space is not unique. For the subgroup $A_3$ we could equally well have chosen $G/H =$
1.4 Approach to Solving Polynomial Equations

The general $n$th degree polynomial equation over the complex field can be expressed in terms of the $k$th order symmetric functions $I_k$ of the roots $z_i$ as follows:

$$(z - z_1)(z - z_2) \ldots (z - z_n) = z^n - I_1z^{n-1} + I_2z^{n-2} - \cdots + (-)^nI_n = 0$$

The $n$ functions $I_k \ (k = 1, 2, \ldots, n)$ of the $n$ roots $(z_1, z_2, \ldots, z_n)$ are symmetric: this means that they are invariant under the Galois group $S_n$ of this equation. Further, any function $f(z_1, z_2, \ldots, z_n)$ that is invariant under $S_n$ can be written as a function of the invariants $I_1, I_2, \ldots, I_n$. The invariants are easily expressed in terms of the roots [cf., Eq. (1.15) above]. The inverse step, that of expressing the roots in terms of the invariants, or coefficients of the polynomial equation, is the problem of solving the polynomial equation.

Galois' theorem states that a polynomial equation over the complex field can be solved if and only if its Galois group $G$ contains a chain of
subgroups [50, 66]

$$G = G_0 \supset G_1 \supset \cdots \supset G_\omega = I$$  \hspace{1cm} (1.16)

with the properties

(i) $G_{i+1}$ is an invariant subgroup of $G_i$

(ii) $G_i/G_{i+1}$ is commutative

The procedure for solving polynomial equations is constructive. First, the last group-subgroup pair in this chain is isolated: $G_{\omega-1} \supset G_\omega = I$. The character table for the commutative group $G_{\omega-1}/G_\omega = G_{\omega-1}$ is constructed. This lists the $|G_{\omega-1}|/|G_\omega|$ inequivalent one-dimensional representations of $G_{\omega-1}$. Linear combinations of the roots $z_i$ are identified that transform under (i.e., are basis functions for) the one-dimensional irreducible representations of $G_{\omega-1}$. These functions are

(i) symmetric under $G_\omega = I$

(ii) not all symmetric under $G_{\omega-1}$.

Next, the next pair of groups $G_{\omega-2} \supset G_{\omega-1}$ is isolated. Starting from the set of functions in the previous step, one constructs from them functions that are

(i) symmetric under $G_{\omega-1}$

(ii) not all symmetric under $G_{\omega-2}$.

This bootstrap procedure continues until the last group-subgroup pair $G = G_0 \supset G_1$ is treated. At this stage the last set of functions can be solved by radicals. These solutions are then fed down the group-subgroup chain until the last pair $G_{\omega-1} \supset G_\omega = I$ is reached. When this occurs, we obtain a linear relation between the roots $z_1, z_2, \ldots, z_n$ and functions of the invariants $I_1, I_2, \ldots, I_n$.

This brief description will now be illustrated by using Galois theory to solve quadratic, cubic, and quartic equations by radicals.

### 1.5 Solution of the Quadratic Equation

The general quadratic equation has the form

$$ (z - r_1)(z - r_2) = z^2 - I_1 z + I_2 = 0 \hspace{1cm} (1.17)$$

$$I_1 = r_1 + r_2$$

$$I_2 = r_1 r_2$$
The Galois group is $S_2$ with subgroup chain shown in Fig. 1.3.

$$S_2 = \{ I, (12) \}$$

![Fig. 1.3. Group chain for the Galois group $S_2$ of the general quadratic equation.](image)

The character table for the commutative group $S_2$ is

<table>
<thead>
<tr>
<th></th>
<th>$I$</th>
<th>(12)</th>
<th>Basis Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Gamma^1$</td>
<td>1</td>
<td>1</td>
<td>$u_1 = r_1 + r_2$</td>
</tr>
<tr>
<td>$\Gamma^2$</td>
<td>1</td>
<td>-1</td>
<td>$u_2 = r_1 - r_2$</td>
</tr>
</tbody>
</table>

Linear combinations of the roots that transform under the one-dimensional irreducible representations $\Gamma^1, \Gamma^2$ are

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} r_1 + r_2 \\ r_1 - r_2 \end{bmatrix}$$

(1.18)

That is, the function $r_1 - r_2$ is mapped into itself by the identity, and into its negative by (12)

$$r_1 - r_2 \xrightarrow{(12)} -(r_1 - r_2)$$

(1.20)

As a result, $(r_1 - r_2)$ is not symmetric under the action of the group $S_2$. It transforms under the irreducible representation $\Gamma^2$, not the identity representation $\Gamma^1$.

Since the square $(r_1 - r_2)^2$ is symmetric (transforms under the identity representation of $S_2$), it can be expressed in terms of the two invariants $I_1, I_2$ as follows

$$(r_1 - r_2)^2 = r_1^2 - 2r_1r_2 + r_2^2$$

$$= r_1^2 + 2r_1r_2 + r_2^2 - 4r_1r_2 = I_1^2 - 4I_2 = D$$

(1.21)

where $D$ is the discriminant of the quadratic equation. Since $(r_1 - r_2) = \ldots$
\(\pm \sqrt{D}\), we have the following linear relation between roots and symmetric functions:

\[
\begin{bmatrix}
  1 & 1 \\
  1 & -1
\end{bmatrix}
\begin{bmatrix}
  r_1 \\
  r_2
\end{bmatrix} = \begin{bmatrix}
  I_1 \\
  \pm [I_1^2 - 4I_2]^{1/2}
\end{bmatrix}
\]  
(1.22)

Inversion of a square matrix involves a sequence of linear operations. We find

\[
\begin{bmatrix}
  r_1 \\
  r_2
\end{bmatrix} = \frac{1}{2} \begin{bmatrix}
  1 & 1 \\
  1 & -1
\end{bmatrix}
\begin{bmatrix}
  I_1 \\
  \pm \sqrt{D}
\end{bmatrix}
\]  
(1.23)

The roots are

\[r_1, r_2 = \frac{1}{2}(I_1 \pm \sqrt{D})\]  
(1.24)

We solve the quadratic equation by another procedure, which we use in the following two sections to simplify the cubic and quartic equations. This method is to move the origin to the mean value of the roots by defining a new variable, \(x\), in terms of \(z\) \([c.f., \text{Equ. (1.15)}]\) by a Tschirnhaus transformation

\[z = x + \frac{1}{2}I_1\]  
(1.25)

The quadratic equation for the new coordinate is

\[x^2 - I'_1 x + I'_2 = x^2 + I'_2 = 0\]

\[I'_1 = 0\]  
(1.26)

\[I'_2 = I_2 - \left(\frac{1}{2}I_1\right)^2\]

The solutions for this auxiliary equation are constructed by radicals

\[x = \pm \sqrt{-I'_2}\]  
(1.27)

from which we easily construct the roots of the original equation

\[r_{1,2} = \frac{1}{2} \left( I_1 \pm \sqrt{I_1^2 - 4I_2} \right)\]  
(1.28)

### 1.6 Solution of the Cubic Equation

The general cubic equation has the form

\[(z - s_1)(z - s_2)(z - s_3) = z^3 - I_1 z^2 + I_2 z - I_3 = 0\]
1.6 Solution of the Cubic Equation

\[ I_1 = s_1 + s_2 + s_3 \]
\[ I_2 = s_1 s_2 + s_1 s_3 + s_2 s_3 \]
\[ I_3 = s_1 s_2 s_3 \]

(1.29)

The Galois group is \( S_3 \) with subgroup chain shown in Fig. 1.4.

![Fig. 1.4. Group chain for the Galois group \( S_3 \) of the general cubic equation.](image)

Since \( A_3 \) is an invariant subgroup of \( S_3 \) and \( I \) is an invariant subgroup of \( A_3 \), the first of the two conditions of the Galois theorem (there exists a chain of invariant subgroups) is satisfied. Since \( S_3/A_3 = S_2 \) is commutative and \( A_3/I = A_3 \) is commutative, the second condition is also satisfied. This means that the general cubic equation can be solved.

We begin the solution with the last group-subgroup pair in this chain: \( A_3 \supset I \). The character table for the commutative group \( A_3 \) is

<table>
<thead>
<tr>
<th>( \Gamma )</th>
<th>( I )</th>
<th>( (123) )</th>
<th>( (321) )</th>
<th>Basis Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Gamma^1 )</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>( v_1 = s_1 + s_2 + s_3 )</td>
</tr>
<tr>
<td>( \Gamma^2 )</td>
<td>1</td>
<td>( \omega )</td>
<td>( \omega^2 )</td>
<td>( v_2 = s_1 + \omega s_2 + \omega^2 s_3 )</td>
</tr>
<tr>
<td>( \Gamma^3 )</td>
<td>1</td>
<td>( \omega^2 )</td>
<td>( \omega )</td>
<td>( v_3 = s_1 + \omega^2 s_2 + \omega s_3 )</td>
</tr>
</tbody>
</table>

(1.30)

where

\[ \omega^3 = +1, \quad \omega = e^{2\pi i/3} = \frac{-1 + i\sqrt{3}}{2} \]

(1.31)

Linear combinations of the roots that transform under each of the three one-dimensional irreducible representations are easily constructed

\[
\begin{bmatrix}
  v_1 \\
  v_2 \\
  v_3 
\end{bmatrix} =
\begin{bmatrix}
  1 & 1 & 1 \\
  1 & \omega & \omega^2 \\
  1 & \omega^2 & \omega 
\end{bmatrix}
\begin{bmatrix}
  s_1 \\
  s_2 \\
  s_3 
\end{bmatrix} =
\begin{bmatrix}
  s_1 + s_2 + s_3 \\
  s_1 + \omega s_2 + \omega^2 s_3 \\
  s_1 + \omega^2 s_2 + \omega s_3 
\end{bmatrix}
\]

(1.32)
For example, the action of $(123)^{-1}$ on $v_2$ is
\[(123)^{-1}v_2 = (321)v_2 = (321)(s_1 + \omega s_2 + \omega^2 s_3) = s_3 + \omega s_1 + \omega^2 s_2 = \omega(s_1 + \omega s_2 + \omega^2 s_3) = \omega v_2 \tag{1.33}\]

Since $v_1$ is symmetric under both $A_3$ and $S_3$, it can be expressed in terms of the invariants $I_k$:
\[v_1 = I_1 \tag{1.34}\]

The remaining functions, $v_2$ and $v_3$, are symmetric under $I$ but not under $A_3$.

We now proceed to the next group-subgroup pair: $S_3 \supset A_3$. To construct functions symmetric under $A_3$ but not under $S_3$ we observe that the cubes of $v_2$ and $v_3$ are symmetric under $A_3$ but not under $S_3$:
\[(12)(v_2)^3 = (12)(s_1 + \omega s_2 + \omega^2 s_3)^3 = (s_2 + \omega s_1 + \omega^2 s_3)^3 = \omega^3(s_1 + \omega^2 s_2 + \omega s_3)^3 = (v_4)^3 \tag{1.35}\]
\[(12)(v_3)^3 = (12)(s_1 + \omega s_2 + \omega^2 s_3)^3 = (s_2 + \omega^2 s_1 + \omega s_3)^3 = \omega^6(s_1 + \omega s_2 + \omega^2 s_3)^3 = (v_2)^3 \tag{1.35}\]

Since $S_2 = S_3/A_3$ permutes the functions $v_2^3$ and $v_3^3$, it is the Galois group of the resolvent quadratic equation whose two roots are $v_2^3$ and $v_3^3$. This equation has the form
\[(x - v_2^3)(x - v_3^3) = x^2 - J_1 x + J_2 = 0 \tag{1.36}\]
\[J_1 = v_2^3 + v_3^3 \quad J_2 = v_2^3 v_3^3 \tag{1.36}\]

Since $J_1, J_2$ are symmetric under $S_3$, they can be expressed in terms of the invariants $I_1, I_2, I_3$ of the original cubic. Since $J_1$ has order 3 and $J_2$ has order 6, we can write the invariants of the quadratic equation (1.36) in terms of the invariants $I_1, I_2, I_3$ (of orders 1, 2, 3) of the original cubic equation (1.29) as follows:
\[J_1 = \sum_{i+2j+3k=3} A_{ijk} I_1^i I_2^j I_3^k \tag{1.37}\]
\[J_2 = \sum_{i+2j+3k=6} B_{ijk} I_1^i I_2^j I_3^k \tag{1.37}\]

These relations can be computed, but they simplify considerably if $I_1 = s_1 + s_2 + s_3 = 0$. This can be accomplished by shifting the origin using a Tschirnhaus transformation as before, with
\[z = y + \frac{1}{3} I_1 \tag{1.38}\]
1.7 Solution of the Quartic Equation

The auxiliary cubic equation has the structure

\[ y^3 - 0y^2 + I'_2y - I'_3 = 0 \]

\[
I'_1 = s'_1 + s'_2 + s'_3 = 0 \\
I'_2 = s'_1s'_2 + s'_1s'_3 + s'_2s'_3 = I_2 - (1/3)I'_1^2 \\
I'_3 = s'_1s'_2s'_3 = I_3 - (1/3)I_2I_1 + (2/27)I'_1^3
\] (1.39)

The invariants \( J_1 = v^3_2 + v^3_3 \) and \( J_2 = v^3_2v^3_3 \) can be expressed in terms of \( I'_2, I'_3 \) as follows

\[
J_1 = v^3_2 + v^3_3 = -27I'_3 \\
J_2 = v^3_2v^3_3 = -27I'_2^3
\] (1.40)

The resolvent quadratic equation whose solution provides \( v^3_2, v^3_3 \) is

\[
x^2 - (-27I'_3)x + (-27I'_2^3) = 0 \] (1.41)

The two solutions to this resolvent quadratic equation are

\[
v^3_2, v^3_3 = -\frac{27}{2}I'_3 \pm \frac{1}{2} \left[ (27I'_3)^2 + 4 \cdot 27I'_2^3 \right]^{1/2} \] (1.42)

The roots \( v_2 \) and \( v_3 \) are obtained by taking cube roots of \( v^3_2 \) and \( v^3_3 \).

\[
\left(\frac{v_2}{v_3}\right) = \left\{ -\frac{27}{2}I'_3 \pm \frac{1}{2} \left[ (27I'_3)^2 + 4 \cdot 27I'_2^3 \right]^{1/2} \right\}^{1/3}
\]

Finally, the roots \( s_1, s_2, s_3 \) are linearly related to \( v_1, v_2, v_3 \) by

\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & \omega & \omega^2 \\
1 & \omega^2 & \omega
\end{bmatrix}
\begin{bmatrix}
s_1 \\
s_2 \\
s_3
\end{bmatrix} =
\begin{bmatrix}
v_1 \\
v_2 \\
v_3
\end{bmatrix}
\] (1.43)

Again, determination of the roots is accomplished by solving a set of simultaneous linear equations

\[
\begin{bmatrix}
s_1 \\
s_2 \\
s_3
\end{bmatrix} = \frac{1}{3}
\begin{bmatrix}
1 & 1 & 1 \\
1 & \omega & \omega^2 \\
1 & \omega^2 & \omega
\end{bmatrix}
\begin{bmatrix}
I_1 \\
v_2 \\
v_3
\end{bmatrix} = \frac{1}{3}
\begin{bmatrix}
v_1 + v_2 + v_3 \\
v_1 + \omega v_2 + \omega v_3 \\
v_1 + \omega^2 v_2 + \omega^2 v_3
\end{bmatrix}
\] (1.44)

1.7 Solution of the Quartic Equation

The general quartic equation has the form

\[
(z - t_1)(z - t_2)(z - t_3)(z - t_4) = z^4 - I_1z^3 + I_2z^2 - I_3z + I_4 = 0
\]
\begin{align*}
I_1 &= t_1 + t_2 + t_3 + t_4 \\
I_2 &= t_1t_2 + t_1t_3 + t_1t_4 + t_2t_3 + t_2t_4 + t_3t_4 \\
I_3 &= t_1t_2t_3 + t_1t_2t_4 + t_1t_3t_4 + t_2t_3t_4 \\
I_4 &= t_1t_2t_3t_4
\end{align*} (1.45)

For later convenience we will construct the auxiliary quartic by shifting the origin of coordinates through the Tschirnhaus transformation \( z = z' + \frac{1}{4}I_1 \)

\[(z' - t_1)(z' - t_2)(z' - t_3)(z' - t_4) = z'^4 - I'_1z'^3 + I'_2z'^2 - I'_3z' + I'_4 = 0 \]

\begin{align*}
I'_1 &= 0 \\
I'_2 &= I_2 - \frac{3}{8}I_1^2 \\
I'_3 &= I_3 - \frac{1}{18}I_2I_1 + \frac{1}{8}I_3^3 \\
I'_4 &= I_4 - \frac{1}{144}I_3I_1 + \frac{1}{18}I_2I_1^2 - \frac{3}{32}I_4^4
\end{align*} (1.46)

Fig. 1.5. Group chain for the Galois group \( S_4 \) of the general quartic equation.

The Galois group is \( S_4 \). This has the subgroup chain shown in Fig. 1.5. The alternating group \( A_4 \) consists of the twelve group operations that have determinant +1 in the permutation matrix representation. The\( \text{fourgroup (vierergruppe, Klein group, Klein four-group)} \) \( V_4 \) is \( \{ I, (12)(34), (13)(24), (14), (23) \} \). The chain

\[ S_4 \supset A_4 \supset V_4 \supset I \]

satisfies both conditions of Galois’ theorem. In particular

\begin{enumerate}
\item \( A_4 \) is invariant in \( S_4 \) and \( S_4/A_4 = S_2 \)
\item \( V_4 \) is invariant in \( A_4 \) and \( A_4/V_4 = C_3 = \{ I, (234), (432) \} \)
\end{enumerate}
(iii) \( I \) is invariant in \( V_4 \) and \( V_4/I = V_4 = \{ I, (12)(34), (13)(24), (14), (23) \} \).

We again begin at the end of the chain with the commutative group \( V_4 \) whose character table is

\[
\begin{array}{c|cccc|c}
\Gamma & (12)(34) & (13)(24) & (14)(23) & \text{Basis Functions} \\
\hline
\Gamma^1 & 1 & 1 & 1 & 1 & w_1 = t_1 + t_2 + t_3 + t_4 \\
\Gamma^2 & 1 & 1 & -1 & -1 & w_2 = t_1 + t_2 - t_3 - t_4 \\
\Gamma^3 & 1 & -1 & 1 & -1 & w_3 = t_1 - t_2 + t_3 - t_4 \\
\Gamma^4 & 1 & -1 & -1 & 1 & w_4 = t_1 - t_2 - t_3 + t_4 \\
\end{array}
\]

The linear combinations of these roots that transform under each of the irreducible representations are

\[
\begin{bmatrix}
w_1 \\
w_2 \\
w_3 \\
w_4 \\
\end{bmatrix} = \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 1 & -1 & -1 \\
1 & -1 & 1 & -1 \\
1 & -1 & -1 & 1 \\
\end{bmatrix} \begin{bmatrix}
t_1 \\
t_2 \\
t_3 \\
t_4 \\
\end{bmatrix} = \begin{bmatrix}
t_1 + t_2 + t_3 + t_4 \\
t_1 + t_2 - t_3 - t_4 \\
t_1 - t_2 + t_3 - t_4 \\
t_1 - t_2 - t_3 + t_4 \\
\end{bmatrix}
\]

These basis vectors are symmetric under \( I \) but the basis vectors \( w_2, w_3, w_4 \) are not symmetric under \( V_4 \).

We now advance to the next group-subgroup pair: \( A_4 \supset V_4 \). It is a simple matter to construct from these linear combinations functions that are

(i) Symmetric under \( V_4 \)
(ii) Permuted among themselves by \( A_4 \) and the group \( A_4/V_4 \).

These functions are \( w_1 = I_1 \) and \( w_2^2, w_3^2, w_4^2 \). In the coordinate system in which the sum of the roots is zero, the three functions \( w_2^2, w_3^2, w_4^2 \) are

\[
w_2^2 = (t'_1 + t'_2 - t'_3 - t'_4)^2 = 2^2(t'_1 + t'_2)^2 = -4(t'_1 + t'_2)(t'_3 + t'_4)
\]

\[
w_3^2 = (t'_1 - t'_2 + t'_3 - t'_4)^2 = 2^2(t'_1 + t'_3)^2 = -4(t'_1 + t'_3)(t'_2 + t'_4)
\]

\[
w_4^2 = (t'_1 - t'_2 - t'_3 + t'_4)^2 = 2^2(t'_1 + t'_4)^2 = -4(t'_1 + t'_4)(t'_2 + t'_3)
\]

It is clear that the three \( w_j^2 \) (\( j = 2, 3, 4 \)) are permuted among themselves by the factor group \( C_3 = A_4/V_4 \), which is a subgroup of the Galois group of a resolvent cubic equation whose three roots are \( w_2^2, w_3^2, w_4^2 \):

\[
(y - w_2^2)(y - w_3^2)(y - w_4^2) = y^3 - J_1y^2 + J_2y - J_3 = 0
\]
Introduction

\[ J_1 = w_2^2 + w_3^2 + w_4^2 \]
\[ J_2 = w_2^2 w_3^2 + w_2^2 w_4^2 + w_3^2 w_4^2 \] (1.50)
\[ J_3 = w_2^2 w_3^2 w_4^2 \]

Since the three \( J_k \) are invariant under \( C_3 \), they can be expressed in terms of the symmetric functions (coefficients) of the original quartic equation (1.45) or (1.46). We find by direct calculation

\[ J_1 = (-4)^1 (2I'_2) \]
\[ J_2 = (-4)^2 (I'_2^2 - 4I'_4) \] (1.51)
\[ J_3 = (-4)^3 (-I'_3) \]

This cubic equation is solved by proceeding to the first group-subgroup pair in the chain: \( S_4 \supset A_4 \), with \( S_4/A_4 = S_2 \). The cubic is solved by introducing the resolvent quadratic, as described in the previous section.

If the three solutions of the resolvent cubic equation are called \( y_2, y_3, y_4 \), then the functions \( w_2, w_3, w_4 \) are

\[ w_2 = \pm \sqrt{y_2} \]
\[ w_3 = \pm \sqrt{y_3} \] (1.52)
\[ w_4 = \pm \sqrt{y_4} \]

A simple computation shows that \( w_2 w_3 w_4 = 8I'_3 \). The signs \( \pm \sqrt{y_j} \) are chosen so that their product is \( 8I'_3 \). The simple linear relation between the roots \( t_i \) and the invariants \( I_1 \) and functions \( w_j(I') \) is easily inverted:

\[
\begin{bmatrix}
  t_1 \\
  t_2 \\
  t_3 \\
  t_4
\end{bmatrix} = \frac{1}{4} \begin{bmatrix}
  1 & 1 & 1 & 1 \\
  1 & 1 & -1 & -1 \\
  1 & -1 & 1 & -1 \\
  1 & -1 & -1 & 1
\end{bmatrix} \begin{bmatrix}
  I_1 \\
  w_2 \\
  w_3 \\
  w_4
\end{bmatrix} \] (1.53)

where the \( w_j \) are square roots of the solutions of the resolvent cubic equation whose coefficients are functions (1.51) of the auxiliary quartic equation.

1.8 The Quintic Cannot be Solved

To investigate whether the typical quintic equation is solvable (and if so, how), it is sufficient to study the structure of its Galois group \( S_5 \). The alternating subgroup \( A_5 \) of order 60 is an invariant subgroup. \( S_5 \) has no invariant subgroups except \( A_5 \) and \( I \). Further, \( A_5 \) has only \( I \) as
an invariant subgroup. The only chain of invariant subgroups in $S_5$ is

$$S_5 \supset A_5 \supset I$$

(1.54)

Although $S_5/A_5 = S_2$ is commutative, $A_5/I = A_5$ is not. Therefore the quintic equation does not satisfy the conditions of Galois’ theorem, so cannot be solved by radicals. General polynomial equations of degree greater than 5 also cannot be solved by radicals.

1.9 Example

To illustrate the solution of a polynomial equation by radicals using the machinery introduced above, we begin with a quartic equation whose roots are: $-2, -1, 2, 5$. We will carry out the algorithm on the corresponding quartic equation. As we proceed through the algorithm, we indicate the numerical values of the functions present. Those values that would not be available at each stage of the computation are indicated by arrows.

The fourth degree equation is

$$(z + 2)(z + 1)(z - 2)(z - 5) = z^4 - 4z^3 - 9z^2 + 16z + 20 = 0$$

$$I_1 = 4$$
$$I_2 = -9$$
$$I_3 = -16$$
$$I_4 = 20$$

(1.55)

We now center the roots by making a Tschirnhaus transformation

$$z = z' + \frac{1}{4}I_1 = z' + 1$$

The new roots are $-3, -2, 1, 4$ and the auxiliary quartic equation is

$$(z' + 1)^4 - 4(z' + 1)^3 - 9(z' + 1)^2 + 16(z' + 1) + 20 =$$

$$(z' + 3)(z' + 2)(z' - 1)(z' + 4) = z'^4 - 15z'^2 - 10z' + 24 = 0$$

$$I'_1 = 0$$
$$I'_2 = -15$$
$$I'_3 = 10$$
$$I'_4 = 24$$

(1.56)
Next, we introduce linear combinations of the four roots $t_1' = -3, t_2' = -2, t_3' = 1, t_4' = 4$

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} t_1' \\ t_2' \\ t_3' \\ t_4' \end{bmatrix} \rightarrow \begin{bmatrix} 0 \\ -10 \\ -4 \\ 2 \end{bmatrix} \quad (1.57)$$

Observe at this stage that $w_2w_3w_4 = 8I'_3$.

Now we compute the squares of these numbers

$$w_2^2 = y_2 \rightarrow (-10)^2 = 100$$
$$w_3^2 = y_3 \rightarrow (-4)^2 = 16$$
$$w_4^2 = y_4 \rightarrow (+2)^2 = 4 \quad (1.58)$$

From the auxiliary quartic (1.56) the resolvent cubic equation can be constructed

$$y^3 - J_1y^2 + J_2y - J_3 = 0$$

$$J_1 = (-4)^3[2I'_2] = (-4)(-30) = 120$$
$$J_2 = (-4)^2[I'_2^2 - 4I'_4] = 16(225 - 4 \cdot 24) = 2064$$
$$J_3 = (-4)^3[-I'_4^2] = (-64)(-100) = 6400 \quad (1.59)$$

Note that these are the coefficients of the equation

$$(y - 2^2)(y - 4^2)(y - 10^2) = y^3 - 120y^2 + 2064y - 6400 = 0 \quad (1.60)$$

Now we construct the cubic equation auxiliary to this cubic. This is done by defining $y = y' + \frac{1}{3}J_1 = y' + \frac{1}{3}(4 + 16 + 100) = y' + 40$. The roots are now

$$y'_1 = y_1 - 40 \rightarrow 4 - 40 = -36$$
$$y'_2 = y_2 - 40 \rightarrow 16 - 40 = -24$$
$$y'_3 = y_3 - 40 \rightarrow 100 - 40 = 60 \quad (1.61)$$

The auxiliary cubic is

$$y'^3 - J'_1y'^2 + J'_2y' - J'_3 = 0$$

$$J'_1 = 0$$
$$J'_2 = -2736$$
$$J'_3 = 51840 \quad (1.62)$$

We note that these are the coefficients of the equation

$$(y' + 36)(y' + 24)(y' - 60) = 0 \quad (1.63)$$
1.9 Example

These coefficients are obtained directly from the coefficients of the resolvent cubic, in principle without knowledge of the values of the roots.

Next we construct the functions \( v_1, v_2, v_3 \)

\[
\begin{bmatrix}
  v_1 \\
  v_2 \\
  v_3
\end{bmatrix}
= \begin{bmatrix}
  1 & 1 & 1 \\
  1 & \omega & \omega^2 \\
  1 & \omega^2 & \omega
\end{bmatrix}
\begin{bmatrix}
  s_1 \\
  s_2 \\
  s_3
\end{bmatrix}
\begin{bmatrix}
  -24 \\
  -36 \\
  60
\end{bmatrix}
\rightarrow
\begin{bmatrix}
  0 \\
  -36 - i48\sqrt{3} \\
  -36 + i48\sqrt{3}
\end{bmatrix}
\]

(1.64)

We can express \( v_2^3 + v_3^3, v_2 v_3^3 \) in terms of \( J'_2, J'_3 \):

\[
\begin{align*}
  v_2^3 + v_3^3 &= 27J'_3 = 27 \times 518400 = 1399680 \\
  v_2^3 v_3^3 &= -27J'_2 = -27 \times (-2736)^3 = 552983334912
\end{align*}
\]

(1.65)

The quadratic resolvent for the auxiliary cubic is

\[
x^2 - 1399680x + 552983334912 = 0
\]

\[
K_1 = 1399680 \\
K_2 = 552983334912
\]

(1.66)

A Tschirnhaus transformation \( x = x' + \frac{1}{2}K_1 \) produces the auxiliary quadratic

\[
x'^2 + 63207309312 = 0
\]

\[
K'_1 = 0 \\
K'_2 = 63207309312
\]

(1.67)

The square of the difference between the two roots of this equation is easily determined:

\[
x'^1 - x'^2 = x_1 - x_2 = \pm 2\sqrt{-K_2} = \pm 2i\sqrt{K_2}
\]

\[
= \pm 2i \times 145152\sqrt{3} = \pm i \times 290304\sqrt{3}
\]

(1.68)

Now we work backwards. The solutions of the resolvent quadratic are given by the linear equation

\[
\begin{bmatrix}
  x_1 \\
  x_2
\end{bmatrix}
= \frac{1}{2}
\begin{bmatrix}
  1 & 1 \\
  1 & -1
\end{bmatrix}
\begin{bmatrix}
  K_1 = 1399680 \\
  2\sqrt{-K_2} = i \times 290304\sqrt{3}
\end{bmatrix}
= 699840 \pm i \times 145152\sqrt{3}
\]

(1.69)

These solutions are the values of \( v_2^3 \) and \( v_3^3 \):

\[
\begin{align*}
  v_2^3 &= 699840 + i \times 145152\sqrt{3} \\
  v_3^3 &= 699840 - i \times 145152\sqrt{3}
\end{align*}
\]

(1.70)
Next, we take cube roots of these quantities. These are unique up to a factor of $\omega$

\[
v_2 = -36 + 48\sqrt{3} \\
v_3 = -36 - 48\sqrt{3}
\]  

(1.71)

The values $y_1, y_2, y_3$ of the resolvent cubic are complex linear combinations of $v_2, v_3$

\[
\begin{bmatrix}
y_1 \\
y_2 \\
y_3
\end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega^2 & \omega \\ 1 & \omega & \omega^2
\end{bmatrix} \begin{bmatrix}
J_1 = 120 \\
v_2 = -36 + i48\sqrt{3} \\
v_3 = -36 - i48\sqrt{3}
\end{bmatrix} = \begin{bmatrix} 16 \\ 100 \\ 4
\end{bmatrix}
\]

(1.72)

\[
w_2^2 = y_1 \\
w_3^2 = y_2 \\
w_4^2 = y_3
\]

\[
w_2 = w_3 = \pm 4 \\
w_3 = w_4 = \pm 10 \\
w_4 = w_4 = \pm 2
\]

(1.73)

Since $w_2w_3w_4 = 8I_3 = 80$, an even number of these signs must be negative. The simplest choice is to take all signs positive. This is different from the results shown in (1.57); this choice of signs serves only to permute the order of the roots. In the final step, the roots of the original quartic are linear combinations of $w_2, w_3, w_4$ and the linear symmetric function $w_1 = I_1$

\[
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1
\end{bmatrix} \begin{bmatrix}
I_1 = 4 \\
w_2 = 4 \\
w_3 = 10 \\
w_4 = 2
\end{bmatrix} = \begin{bmatrix}
20/4 = +5 \\
-4/4 = -1 \\
8/4 = +2 \\
-8/4 = -2
\end{bmatrix}
\]

(1.74)

We have recovered the four roots of the original quartic equation using Galois’ algorithm, based on the structure of the invariance group $S_4$ of the quartic equation.

1.10 Conclusion

One of the many consequences of Galois’ study of algebraic equations and the symmetries that leave them invariant is the proof that an algebraic equation can be solved by radicals if and only if its invariance group has a certain structure. This proof motivated Lie to search for analogous results involving differential equations and their symmetry groups, now called Lie groups. We have described in this chapter how the structure of the discrete symmetry group (Galois group) of a polynomial equation determines whether or not that equation can be solved by
radicals. If the answer is ‘yes,’ we have shown how the structure of the Galois group determines the structure of the algorithm for constructing solutions. This algorithm has been developed for the cubic and quartic equations, and illustrated by example for a quartic equation.

1.11 Problems

1. Compute $S_4/A_4$, $A_4/V_4$, $V_4$ and show that they are commutative.

2. Construct the group $V_8$ with the property $S_4 \supset V_8 \supset V_4$ (cf. Fig. 1.5). Hint: include a cyclic permutation.

3. For the cubic equation $z^3 - 7z + 6 = 0 \ (z - 1)(z - 2)(z + 3) = 0$

$$I_1 = 0 \quad J_1 = 162$$

$$I_2 = -7 \quad J_2 = 9261$$

$$I_3 = -6$$

Show that the resolvent equation for $v_2^3, v_3^3$ is $(x - v_2^3)(x - v_3^3) = x^2 - 162x + 9261 = 0$. Solve this quadratic to find $v_2^3, v_3^3 = 81 \pm 30\sqrt{3}$, so that $v_2, v_3 = \frac{1}{2}(3 \pm i5\sqrt{3})$. Invert Equ. (1.43) to determine the three roots of the original equation: $(1, 2, -3)$.

4. Ruler and compass can be used to construct an orthogonal pair of axes in the plane (Euclid). A compass is used to establish a unit of length $(1)$. Then by ruler and compass it is possible to construct intervals of length $x$, where $x$ is integer. From there it is possible to construct intervals of lengths $x + y$, $x - y$, $x \times y$ and $x/y$ using ruler and compass. It is also possible to construct intervals of length $\sqrt{x}$ by these means. The set of all numbers that can be constructed from integers by addition, subtraction, multiplication, division, and extraction of square roots is called the set of constructable numbers. This forms a subset of the numbers $x + iy = (x, y)$ in the complex plane. If a number is (is not) constructable the point representing that number can (cannot) be constructed by ruler and compass alone. Since repeated square roots can be taken, a constructable number satisfies an algebraic equation of degree $K$ with integer coefficients, where $K = 2^n$ must be some power of two.

The three geometry problems of antiquity are:
Introduction

**a: Square a circle?** For the circle of radius 1 the area is $\pi$. Squaring a circle means finding an interval of length $x$, where $x^2 - \pi = 0$. This is of degree 2 but $\pi$ is not rational (not even algebraic). Argue that it is impossible to square the circle by ruler and compass alone.

**b: Double the cube?** A cube with edge length 1 has volume $1^3 = 1$. A cube with twice the volume has edge length $x$, where $x$ satisfies $x^3 - 2 = 0$. Although the coefficients are integers this equation is of degree $3 \neq 2^n$ for any integer $n$. Argue that it is impossible to double the volume of a cube by ruler and compass alone.

**c: Trisect an angle?** If $3\theta$ is some angle, the trigonometric functions of $3\theta$ and $\frac{1}{3}(3\theta) = \theta$ are related by

$$e^{i3\theta} = (e^{i\theta})^3$$

$$\cos(3\theta) + i\sin(3\theta) = (\cos^3(\theta) - 3\cos(\theta)\sin^2(\theta)) + i(3\cos^2(\theta)\sin(\theta) - \sin^3(\theta))$$

In particular

$$\cos(3\theta) = 4\cos^3(\theta) - 3\cos(\theta)$$

Whether $\cos(3\theta)$ is rational or irrational, the equation for $\cos(\theta)$:

$$4\cos^3(\theta) - 3\cos(\theta) - \cos(3\theta) = 0$$

is cubic. Argue that it is impossible to trisect an angle unless $\cos(3\theta)$ is such that the cubic factors into the form $(x^2 + ax + b)(x + c) = 0$, where $a, b, c$ are rational. For example, if $\cos(3\theta) = 0$, $c = 0$ so that $a = 0$ and $b = -3/4$. Then $\cos(\theta) = 0$ or $\pm \sqrt{3}/2$ for $3\theta = \pi/2$ (0), $3\pi/2$ (0), or $5\pi/2$ (-).