1

Introduction

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 nineteenth 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 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 nth 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. Nonsemisimple 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 nonsemisimple 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 classical functions of mathematical physics are matrix elements of simple Lie groups, in particular matrix representations. There is a very rich connection between Lie groups and special functions that is still evolving.

1.2 A result of Galois

In 1830 Galois developed machinery that allowed mathematicians to 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 the following (Lang, 1984; Stewart, 1989).

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 $G = G_0 \supset G_1 \supset \cdots \supset G_\omega = I$ with the properties:

(i) G_{i+1} is an invariant subgroup of G_i;
(ii) 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 nth 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

$$\begin{aligned} (z - z_1)(z - z_2) \cdots (z - z_n) &= 0 \\ z^n - I_1 z^{n-1} + I_2 z^{n-2} + \cdots + (-1)^n I_n &= 0 \end{aligned} \tag{1.1}$$

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

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \longrightarrow \begin{bmatrix} z_{i_1} \\ z_{i_2} \\ \vdots \\ z_{i_n} \end{bmatrix} \tag{1.2}$$

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 $6 = 3!$ 3×3 matrices for the permutation representation of S_3 are

$$\begin{aligned} I &\to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (123) \to \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad (321) \to \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\ (12) &\to \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (23) \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (13) \to \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{aligned} \tag{1.3}$$

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

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \xrightarrow{(123)} \begin{bmatrix} z_2 \\ z_3 \\ z_1 \end{bmatrix} \tag{1.4}$$

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

$$\begin{bmatrix} z_2 \\ z_3 \\ z_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \tag{1.5}$$

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

$$\begin{array}{ccccc} g_i & \cdot & g_j & = & g_i \cdot g_j \\ \downarrow & \downarrow & \downarrow & & \downarrow \\ \Gamma(g_i) & \times & \Gamma(g_j) & = & \Gamma(g_i \cdot g_j) \end{array} \tag{1.6}$$

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, $G \to \Gamma(G)$, 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 3×3 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:

$$\begin{aligned} S_2(12) &= \{I, (12)\} \\ S_2(23) &= \{I, (23)\} \\ S_2(13) &= \{I, (13)\} \end{aligned}$$

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.

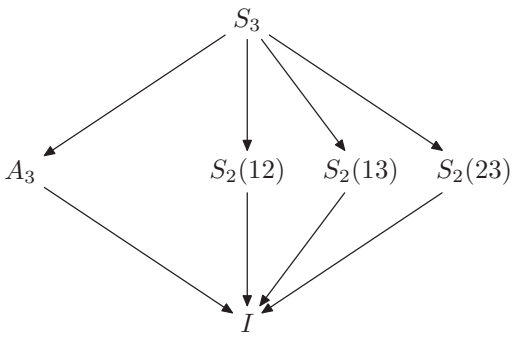

Figure 1.1. Subgroups of S_3.

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.

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

$$g_i \cdot g_j = g_j \cdot g_i \tag{1.7}$$

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

$$\begin{aligned} (12)(23) &= (321) \\ & \qquad\qquad (123) \neq (321) \\ (23)(12) &= (123) \end{aligned} \tag{1.8}$$

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

$$g H_1 g^{-1} = H_2 \tag{1.9}$$

Example The subgroups $S_2(12)$ and $S_2(13)$ are conjugate in S_3 since

$$(23)S_2(12)(23)^{-1} = (23)\{I, (12)\}(23)^{-1} = \{I, (13)\} = S_2(13) \tag{1.10}$$

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:

$$(23)A_3(23)^{-1} = (23)\{I, (123), (321)\}(23)^{-1} = \{I, (321), (123)\} = A_3 \tag{1.11}$$

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.

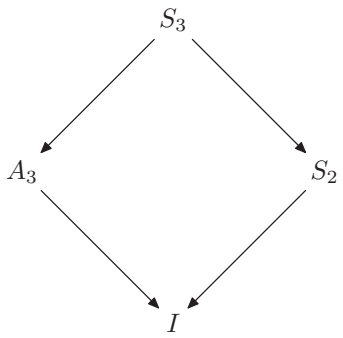

Figure 1.2. Subgroups of S_3, combining conjugate subgroups.

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.

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:

$$\begin{array}{ccccc} g_i & \cdot & g_j & = & g_i \cdot g_j \\ \downarrow & \downarrow & \downarrow & & \downarrow \\ f(g_i) & \times & f(g_j) & = & f(g_i \cdot g_j) \end{array} \tag{1.12}$$

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

$$\begin{array}{ccc} S_3 & \xrightarrow{f} & S_2 \\ I, (123), (321) & \longrightarrow & I \\ (12), (23), (31) & \longrightarrow & (12) \end{array} \tag{1.13}$$

is a homomorphism.

Example The 1:1 mapping of S_3 onto the six 3×3 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 $|G/H| = |G|/|H|$. For example, for subgroups $H = A_3 = \{I, (123), (321)\}$ and $H = S_2(23) = \{I, (23)\}$ we have

$$\begin{array}{rcl} G/H & \cdot \quad H & = \quad G \\ \{I, (12)\} & \cdot \{I, (123), (321)\} & = \{I, (123), (321), (12), (13), (23)\} \\ \{I, (12), (321)\} & \cdot \quad \{I, (23)\} & = \{I, (23), (12), (123), (321), (13)\} \end{array} \tag{1.14}$$

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 $G/H = S_3/A_3 = \{I, (13)\}$ or $\{I, (23)\}$; for $S_2(23)$ we could equally well have chosen $G/H = S_3/S_2(23) = \{I, (123), (321)\}$.

In general, it is not possible to choose the group elements in G/H so that they form a subgroup of G. However, if H is an invariant subgroup of G, it is always possible to choose the group elements in the quotient space G/H in such a way that they form a subgroup in G. This group is called the **factor group**, also denoted G/H. Since A_3 is an invariant subgroup of S_3, the coset S_3/A_3 is a group, and this group is isomorphic to S_2. More generally, if H is an invariant subgroup of G, then the group G is the **direct product** of the invariant subgroup H with the factor group G/H: $G = G/H \times H$.

1.4 Approach to solving polynomial equations

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

$$\begin{aligned} (z - z_1)(z - z_2)\cdots(z - z_n) &= z^n - I_1 z^{n-1} + I_2 z^{n-2} - \cdots + (-)^n I_n = 0 \\ I_1 &= \sum_{i=1}^{n} z_i = z_1 + z_2 + \cdots + z_n \\ I_2 &= \sum_{i<j}^{n} z_i z_j = z_1 z_2 + z_1 z_3 + \cdots + z_1 z_n + z_2 z_3 + \cdots + z_{n-1} z_n \\ &\vdots \; \vdots \; \vdots \\ I_n &= \sum_{i<j<\cdots<k}^{n} z_i z_j \cdots z_k = z_1 z_2 \cdots z_n \end{aligned} \tag{1.15}$$

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 (see Eq. (1.15)). 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 (Lang, 1984; Stewart, 1989)

$$G = G_0 \supset G_1 \supset \cdots \supset G_\omega = I \tag{1.16}$$

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.

Figure 1.3. Group chain for the Galois group S_2 of the general quadratic equation.

1.5 Solution of the quadratic equation

The general quadratic equation has the form

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

The Galois group is S_2 with subgroup chain shown in Fig. 1.3.

The character table for the commutative group S_2 is

	I	(12)	Basis functions
Γ^1	1	1	$u_1 = r_1 + r_2$
Γ^2	1	-1	$u_2 = r_1 - r_2$

(1.18)

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

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} r_1 + r_2 \\ r_1 - r_2 \end{bmatrix} \tag{1.19}$$

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

$$(r_1 - r_2) \begin{cases} \xrightarrow{I} +(r_1 - r_2) \\ \xrightarrow{(12)} (r_2 - r_1) = -(r_1 - r_2) \end{cases} \tag{1.20}$$

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

$$\begin{aligned} (r_1 - r_2)^2 &= r_1^2 - 2r_1 r_2 + r_2^2 \\ &= r_1^2 + 2r_1 r_2 + r_2^2 - 4r_1 r_2 = I_1^2 - 4I_2 = D \end{aligned} \tag{1.21}$$