Lie Groups and PDE

Carl Öhrnell
Abstract

We introduce some central concepts of Lie groups and algebras to a reader without a background in topology and differentiable manifolds. Presenting the jet space and prolongation of vector fields, we then find Lie groups of transformations that leave differential equations invariant, called symmetries. These are used to find solutions using coordinate changes, transformations of known solutions and invariant solutions. A non-linear ODE and the heat equation (PDE) are used to exemplify the methods.
Contents

1 Introduction 2

2 Lie Theory 4
   2.1 Manifolds ............................................. 4
   2.2 Lie Groups ............................................. 6
   2.3 Vector Fields ........................................... 8
   2.4 The Lie Algebra ........................................ 11

3 Differential Equations and the Application of Lie Groups 14
   3.1 Geometrical Setting .................................... 14
   3.2 Symmetry Groups ....................................... 15
   3.3 Solving ODE and PDE Using Symmetries ................. 18
      3.3.1 Canonical Coordinates ............................. 18
      3.3.2 Transformation of a Known Solution .............. 19
      3.3.3 Group-invariant Solutions ....................... 19

4 Examples 20
   4.1 A Non-linear ODE ...................................... 20
   4.2 The Heat Equation ..................................... 22
   4.3 The Kolmogorov Equation ............................... 26
Chapter 1

Introduction

Differential equations are equations that contain one or more derivatives of an unknown function. They appear in just about every field of science, wherever a state or quantity can be related to its rate of change. The mathematical study of these equations removed from direct applicability is also a thriving field in and of itself. The diversity of different equations naturally leads to classification attempts to inform on what properties to expect and how to attempt to solve them.

The first distinction is made by considering the number of variables the unknown function is dependent on: only one gives an ordinary differential equation (ODE), while more than one results in a partial differential equation (PDE). Generally speaking, the latter are more difficult to solve. Complexity usually also increases with the order of the equation, given by the highest order derivative present. Another source of complexity is the presence of non-linearity in the equation.

The natural progression in the curriculum follows this increase in difficulty, usually beginning with linear first order ODE, to successively progress to higher order linear ODE, the occasional non-linear ODE, then onto linear PDE, etc. The methods taught are often tailored to the specific type of equation, and in some instances the underlying principles are somewhat obscure.

This paper is dedicated to present methods that have many merits: they are applicable to a wide variety of differential equations and they have a geometric foundation making them intuitive. They build on the theory of Lie groups, so the ambition is to introduce some concepts central to this theory while simultaneously giving an intuition for its geometry and later its application to differential equations.

In Chapter 2 we introduce the concept of a smooth manifold, Lie groups as manifolds with group operations and their action on manifolds. To find linear models we introduce vector fields and their flows which correspond to one-parameter subgroups. We conclude by defining the Lie algebra corresponding to a Lie group and its application in finding symmetry groups of equations. Throughout the chapter we exemplify using the circle group.

Chapter 3 describes how the introduced Lie group machinery can be used
to solve or simplify differential equations. We start by introducing the geometric extensions needed to apply Lie group actions to differential equations, most importantly the jet space and the prolongation of group actions generating vector fields. We then present specific methods: using canonical variables to solve ODE, finding new solutions to PDE by transforming a known solution and finally finding group-invariant solutions of a PDE.

In Chapter 4 we give two extensive examples to demonstrate the methods in practice. The first is a non-linear ODE to which we find a symmetry, an invariant to that symmetry and finally canonical coordinates which let us solve the equation by quadrature. The second is the heat equation, a PDE in two dimensions. We solve the determining equation to find the Lie algebra of symmetries, transform the constant solution using one of the found symmetries, and finally find translation invariant travelling wave solutions. We then conclude by giving a brief history of the Kolmogorov equation, a PDE still under theoretical investigation, finding the translations and dilations that define the Lie group structure needed for further study of the equation.
Chapter 2

Lie Theory

Most of the definitions in this chapter are adapted from the first chapter of Olver [11], which gives an introduction to Lie groups to then focus on applications in later chapters, as well as his lectures [12]. For a more comprehensive introduction to differentiable manifolds, vector fields and Lie groups, the first four chapters of Boothby [2] are recommended.

2.1 Manifolds

The natural setting for our development is on $m$-dimensional smooth ($C^\infty$) manifolds, which can be thought of as smooth deformations of $\mathbb{R}^m$ with no self-intersections. Thus $m = 1$ gives a curve, $m = 2$ a surface, and $m \geq 3$ a hyper-surface.

**Definition 2.1.** An $m$-dimensional manifold is a set $M$, together with a countable collection of subsets $U_\alpha \subset M$, called coordinate charts, and injective functions $\chi_\alpha : U_\alpha \rightarrow V_\alpha$ onto connected open subsets $V_\alpha$ of $\mathbb{R}^m$ called local coordinate maps, which satisfy the following properties:

1. The coordinate charts cover $M$:
   \[ \cup_\alpha U_\alpha = M. \]

2. On the overlap of any pair of coordinate charts $U_\alpha \cap U_\beta$ the composite map
   \[ \chi_\beta \circ \chi_\alpha^{-1} : \chi_\alpha(U_\alpha \cap U_\beta) \rightarrow \chi_\beta(U_\alpha \cap U_\beta) \]
   is a smooth function.

3. If $x \in U_\alpha$, $\hat{x} \in U_\beta$ are distinct points of $M$, then there exist open subsets $W$ of $\chi_\alpha(x)$ in $V_\alpha$ and $\hat{W}$ of $\chi_\beta(\hat{x})$ in $V_\beta$ such that
   \[ \chi_\alpha^{-1}(W) \cap \chi_\beta^{-1}(\hat{W}) = \emptyset. \]

**Example 2.2.** The most obvious example of an $m$-dimensional manifold is $\mathbb{R}^m$ itself, or any open subset of $\mathbb{R}^m$, using the identity map as coordinate map.
Example 2.3. An example of a one-dimensional manifold that we are familiar with is the unit circle

\[ S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \]

We cover it with the charts

\[ U_1 = S^1 \setminus \{(0,1)\}, \quad U_2 = S^1 \setminus \{(0,-1)\}, \]

i.e. the unit circle with deleted north and south pole, respectively. We then let the coordinate maps

\[ \chi_1 : U_1 \to \mathbb{R}, \quad \chi_2 : U_2 \to \mathbb{R}, \]

be the stereographic projections from the north respectively south pole,

\[ \chi_1(x, y) = \left( \frac{x}{1-y} \right), \quad \chi_2(x, y) = \left( \frac{x}{1+y} \right). \]

We then check that the composite map on the overlap,

\[ \chi_1 \circ \chi_2^{-1} : \mathbb{R}\setminus\{(0)\} \to \mathbb{R}\setminus\{(0)\}, \]

is a smooth map. Solving for \( t \) in \( \chi_2(x, y) = t \) using that \( x^2 + y^2 = 1 \) on the unit circle we get

\[ \chi_2^{-1}(t) = \left( \frac{2t}{t^2+1}, \frac{1-t^2}{t^2+1} \right), \]

which composed with \( \chi_1 \) gives us

\[ \chi_1 \circ \chi_2^{-1}(t) = \frac{1}{t}, \]

which is clearly a smooth map away from the origin. The third criterion concerning the separation of points, called the Hausdorff separation property, is inherited from \( \mathbb{R}^2 \).

Note that this collection of charts and maps is not unique. Another common atlas for \( S^1 \) uses four charts where the maps are projections onto the \( x \) and \( y \) axes.

In the following, we will often use the Cartesian product of two manifolds to create new manifolds. If \( M \) is an \( m \)-dimensional manifold and \( N \) is an \( n \)-dimensional manifold, then \( M \times N \) is an \( (m+n) \)-dimensional manifold. E.g., two open subsets of \( \mathbb{R}^m \) and \( \mathbb{R}^n \), respectively, gives an open subset of \( \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n} \), and the Cartesian product of two circles, \( S^1 \times S^1 \), is the two-dimensional torus \( T^2 \), which has the shape of a doughnut.

We conclude with two useful definitions regarding maps between manifolds.

Definition 2.4. The rank of a map \( F : M \to N \) at a point \( x \in M \) is defined to be the rank of the \( n \times m \) Jacobian matrix \( (\partial F^i/\partial x^j) \) of any local coordinate expression for \( F \) at the point \( x \). The map \( F \) is called regular if its rank is constant.

Definition 2.5. A smooth \( n \)-dimensional immersed submanifold of a manifold \( M \) is a subset \( N \subset M \) parametrized by a smooth, one-to-one map \( F : \tilde{N} \to N \subset M \) whose domain \( \tilde{N} \), the parameter space, is a smooth \( n \)-dimensional manifold, and such that \( F \) is everywhere regular, of maximal rank \( n \).

Remark 2.6. In the following we will assume that \( M \) is a smooth manifold of dimension \( m \), all manifolds are connected, and maps are regular.
2.2 Lie Groups

We begin by recalling the properties and some examples of abstract groups.

**Definition 2.7.** A group is a set $G$ with an operation $G \times G \rightarrow G$, $(x, y) \mapsto xy$ such that

1. The set $G$ is closed under the operation: $x, y \in G \implies xy \in G$,
2. The operation is associative: $(xy)z = x(yz)$ for all $x, y, z \in G$,
3. There exists an identity element $e$ such that $xe = ex = x$ for all $x \in G$,
4. All elements have inverses: $\forall x \in G \exists x^{-1} \in G$ such that $x^{-1}x = xx^{-1} = e$.

**Example 2.8.** Some examples of groups are

1. $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{R}^n$ with addition,
2. $\mathbb{Q}^*$, $\mathbb{R}^*$ or $\mathbb{C}^*$ with multiplication,
3. The circle group, i.e. $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ with multiplication.

Now, merging the algebraic properties of a group with the topological and differentiable properties of a smooth manifold, we get a Lie group.

**Definition 2.9.** A Lie group is a smooth manifold $G$ with a group structure such that the operation $\mu : G \times G \rightarrow G$, $(x, y) \mapsto xy$ and the inversion $\iota : G \rightarrow G$, $x \mapsto x^{-1}$ are smooth maps.

It turns out that many of the examples of algebraic groups listed above are also examples of Lie groups, specifically all involving some set of $\mathbb{R}$ or $\mathbb{C}$.

**Example 2.10.** An interesting example is the general linear group in $n$ dimensions, $\text{GL}(n, \mathbb{R})$. It can be identified with $n \times n$ invertible matrices under matrix multiplication as group operation. The set of such matrices form an open subset of the space $\text{M}_{n \times n}$ of all $n \times n$ matrices which is isomorphic to $\mathbb{R}^{n^2}$. Further, matrix multiplication is smooth since it is polynomial in the elements, and since the determinant is non-zero so is the inverse. Thus $\text{GL}(n, \mathbb{R})$ is a Lie group.

We will primarily be interested in Lie groups acting as transformations on manifolds, sometimes only locally. As such the group action is not always defined for the entire manifold or all the elements of the group.

**Definition 2.11.** Let $M$ be a smooth manifold. A local group of transformations acting on $M$ is given by a (local) Lie group $G$, an open subset $U$, with

$$\{e\} \times M \subset U \subset G \times M,$$

which is the domain of definition of the group action, and a smooth map $\psi : U \rightarrow M$, denoted $\psi(g, x) = g \cdot x$, with the following properties:

(a) If $(h, x) \in U$, $(g, h \cdot x) \in U$, and also $(gh, x) \in U$, then

$$g \cdot (h \cdot x) = (gh) \cdot x.$$
(b) For all \( x \in M \),
\[ e \cdot x = x. \]

(c) If \((g, x) \in U\), then \((g^{-1}, g \cdot x) \in U\) and
\[ g^{-1} \cdot (g \cdot x) = x. \]

**Remark 2.12.** In the following we will not always distinguish between global and local groups of transformations. Statements will concern the group elements for which the action is defined.

**Example 2.13.** We mentioned the circle group as an example of a Lie group. Regarded as a transformation group acting on \( \mathbb{R}^2 \) it is called the special orthogonal group, \( \text{SO}(2) \), or the rotation group, since its elements rotate points around the origin by an angle \( \varepsilon \),
\[
g_{\varepsilon} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \varepsilon & -\sin \varepsilon \\ \sin \varepsilon & \cos \varepsilon \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \cos \varepsilon - y \sin \varepsilon \\ x \sin \varepsilon + y \cos \varepsilon \end{pmatrix}.
\]

We include the matrix form to give a glimpse of the fact that \( \text{SO}(2) \) consists of orthogonal matrices with determinant 1, and is a subgroup of \( \text{GL}(2) \).

**Definition 2.14.** Starting at a point \( x \) in a manifold \( M \), the set of points that can be reached by sequences of group transformations is called the orbit of \( G \) through \( x \). For a regular group action, the orbits are submanifolds of \( M \).

In the example above, since the transformation depends on one parameter, \( \varepsilon \), the orbits are curves in \( \mathbb{R}^2 \). Specifically, starting with a point \( x \) a distance \( c \) from the origin, the orbit will be the circle \( x^2 + y^2 = c \). The circle as a set will remain unchanged by all the rotations of the group; we say that it is invariant under the group \( \text{SO}(2) \). Defining such a circle as the solution set of a real-valued function \( f(x, y) = x^2 + y^2 - c \), we have an example of the following definition.

**Definition 2.15.** Let \( G \) be a group acting on \( M \), and \( f : M \to \mathbb{R} \). The group \( G \) is called a symmetry (group) of an equation \( f(x) = 0 \) if and only if the solution set of \( f \)
\[ S_f = \{ x : f(x) = 0 \} \]
is invariant under \( G \), that is \( g \cdot x \in S_f \) whenever \( g \in G \) and \( x \in S_f \).

Given that the equation is of maximal rank at all points satisfying \( f(x) = 0 \), the solution set is a submanifold of \( M \) called the solution surface. Thus the transformations of \( G \) move points along the solution surface, meaning it maps solutions of \( f \) to other solutions. This property will be essential when we study symmetry groups of differential equations in the next chapter.

We could also make a similar definition regarding all the values of a function, not only its solution set.

**Definition 2.16.** A function \( f : M \to \mathbb{R} \) is an invariant of a group of transformations \( G \) if for all \( x \in M \) and all \( g \in G \),
\[ f(g \cdot x) = f(x). \quad (2.1) \]

Put differently, the function \( f \) is constant on all the orbits of \( G \).
2.3 Vector Fields

A very useful property of connected Lie groups is that the whole group can be recovered from the tangent space at the identity element. To show this we need to introduce the notion of tangents and vector fields on manifolds. One way is to define the tangent of a curve lying on the manifold.

**Definition 2.17.** Let $C$ be a smooth curve on a manifold $M$, parametrized by $\gamma: I \rightarrow M$, where $I \subset \mathbb{R}$. In local coordinates $x = (x^1, \ldots, x^m)$, $C$ is given by $m$ smooth functions $\gamma(\varepsilon) = (\gamma^1(\varepsilon), \ldots, \gamma^m(\varepsilon))$ of the real variable $\varepsilon$. At each point $x = \gamma(\varepsilon)$ of $C$ the curve has a **tangent vector**, namely the componentwise derivative $\dot{\gamma}(\varepsilon) = \frac{d\gamma}{d\varepsilon}$. We write
\[
\frac{d\gamma}{d\varepsilon} = \dot{\gamma}^1(\varepsilon) \frac{\partial}{\partial x^1} + \ldots + \dot{\gamma}^m(\varepsilon) \frac{\partial}{\partial x^m}
\]
for the tangent vector to $C$ at $x = \gamma(\varepsilon)$.

**Example 2.18.** One parametrization of the unit circle in $\mathbb{R}^2$, $\gamma(\varepsilon) = (\cos \varepsilon, \sin \varepsilon)$, with coordinates $(x, y)$, has tangent vector
\[
\dot{\gamma}(\varepsilon) = -\sin \varepsilon \frac{\partial}{\partial x} + \cos \varepsilon \frac{\partial}{\partial y} = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y}
\]
at the point $(x, y) = \gamma(\varepsilon) = (\cos \varepsilon, \sin \varepsilon)$.

The collection of all such possible tangent vectors forms the **tangent space** to $M$ at the point $x$, denoted $T_xM$. For each point the tangent space is a vector space of the same dimension as $M$. Choosing one tangent from $T_xM$ for each $x \in M$ we get a vector field.

**Definition 2.19.** A smooth **vector field** $X$ is an assignment of a tangent vector $X_x \in T_xM$ to each point $x$ in $M$, varying smoothly. In local coordinates
\[
X = \xi^1(x) \frac{\partial}{\partial x^1} + \ldots + \xi^m(x) \frac{\partial}{\partial x^m},
\]
where the $\xi^i(x)$ are smooth functions.

**Definition 2.20.** An **integral curve** of a vector field $X$ is a smooth parametrized curve $x = \gamma(\varepsilon)$ whose tangent vector is the same as the vector field $X$ at every point:
\[
\dot{\gamma}(\varepsilon) = X_{\gamma(\varepsilon)}
\]
for all $\varepsilon$.

Thus in local coordinates, $x = \gamma(\varepsilon)$ must be a solution to the autonomous system of ordinary differential equations
\[
\frac{dx^i}{d\varepsilon} = \xi^i(x), \quad i = 1, \ldots, m,
\]
(2.2)
where the $\xi^i(x)$ are the coefficients of $X$ at $x$. For a smooth vector field the existence and uniqueness theorems for systems of ODE guarantee there is a unique solution to the system for each set of initial data $\gamma(0) = x_0$. So through each $x$ in $M$ passes a unique, maximal integral curve.
Example 2.21. In example 2.18 we found the tangents of a parametrized circle. Conversely, given the vector field $X = -y \partial_x + x \partial_y$, and initial condition $\gamma(0) = (1, 0)$, we find the integral curve by solving the system

$$\frac{dx}{d\varepsilon} = -y, \quad \frac{dy}{d\varepsilon} = x.$$  

The solution is as expected the curve $\gamma(\varepsilon) = (\cos \varepsilon, \sin \varepsilon)$, see figure 2.1.

![Figure 2.1](image)

Figure 2.1: The vector field $X = -y \partial_x + x \partial_y$ and a (non-maximal) integral curve with $\gamma(0) = (1, 0)$. (The vectors are scaled for visibility)

We can picture the vector field $X$ as the velocity field of a fluid, carrying particles along in its flow. Starting at $x$ for $\varepsilon = 0$, letting the particle follow the flow for a time $\varepsilon$, the trajectory of the particle is an integral curve $\gamma(\varepsilon)$.

When viewed as tangent vectors, the $\partial / \partial x^i$ are seen as basis vectors in the tangent space. Another way to define vector fields is as derivations on real-valued functions $f : M \to \mathbb{R}$. This reflects our notation in a natural way, giving

$$X(f) = \sum_{i=1}^m \xi^i(x) \frac{\partial f}{\partial x^i},$$

which is itself a smooth function given any smooth $f \in C^\infty(M)$. Since we have local coordinates in $\mathbb{R}^m$ with basis vectors $e_i$, we can think of a derivation as a directional derivative in the direction $\sum_{i=1}^m \xi^i e_i$.

For maps between manifolds, and Lie groups in particular, we want to know how tangent vectors get mapped between the tangent spaces. This is described by the induced linear map, the differential.

Definition 2.22. Let $F : M \to N$ be a smooth map, $x = \gamma(\varepsilon)$ a curve in $M$ and $F(x) = F(\gamma(\varepsilon)) \in N$. The differential $dF$ of $F$ is a linear map between the tangent spaces, $dF : T_x M \to T_{F(x)} N$, where

$$dF\left(\frac{d}{d\varepsilon} \gamma(\varepsilon)\right) = \frac{d}{d\varepsilon} F(\gamma(\varepsilon)).$$

In local coordinates, for $X_x = \sum_{i=1}^n \xi^i(x) \partial / \partial x^i,$

$$dF(X_x) = \sum_{i=1}^n \left( \sum_{i=1}^m \xi^i \frac{\partial F_j}{\partial x^i}(x) \frac{\partial}{\partial y^j} \right) = \sum_{i=1}^n X(F^i(x)) \frac{\partial}{\partial y^i}.$$
So, when \( F \) maps a smooth curve \( \gamma \in M \) to smooth curve \( F(\gamma) \in N \), the differential \( dF \) maps the tangent to \( \gamma \) at \( x = \gamma(\varepsilon) \) to the tangent of \( F(\gamma) \) at \( F(x) = F(\gamma(\varepsilon)) \). In local coordinates, \( dF \) is the Jacobian matrix of \( F \) at \( x \).

**Definition 2.23.** Let \( X \) and \( Y \) be vector fields on a manifold \( M \), then their Lie bracket or commutator \( [X,Y] \) is the unique vector field satisfying

\[
[X,Y](f) = X(Y(f)) - Y(X(f))
\]

for all smooth functions \( f : M \to \mathbb{R} \). In local coordinates, if

\[
X = \sum_{i=1}^{m} \xi^i(x) \frac{\partial}{\partial x^i}, \quad Y = \sum_{i=1}^{m} \eta^i(x) \frac{\partial}{\partial x^i},
\]

then

\[
[X,Y] = \sum_{i=1}^{m} (X(\eta^i) - Y(\xi^i)) \frac{\partial}{\partial x^i} = \sum_{i=1}^{m} \sum_{j=1}^{m} \left( \xi^j \frac{\partial \eta^i}{\partial x^j} - \eta^j \frac{\partial \xi^i}{\partial x^j} \right) \frac{\partial}{\partial x^i}. \tag{2.3}
\]

**Proposition 2.24.** The Lie bracket has the following properties:

1. Bilinearity. For constants \( a, b \),

\[
[aX + bY,Z] = a[X,Z] + b[Y,Z], \quad [X,aY + bZ] = a[X,Y] + b[X,Z],
\]

2. Skew-symmetry

\[
[X,Y] = -[Y,X],
\]

3. Jacobi identity

\[
[Z,[X,Y]] + [Y,[Z,X]] + [X,[Y,Z]] = 0.
\]

It is convenient to collect the commutators in a commutator table, where the intersection of the \( X_i \) row and the \( X_j \) column is the commutator \( [X_i,X_j] \). Since the commutator is anti-symmetric, we only need the cells above the diagonal. We illustrate with an example. Note that we will write \( \partial_x \) to mean \( \partial/\partial x \) when motivated to save space.

**Example 2.25.** In the next chapter we will investigate vector fields associated with the heat equation, \( u_t = u_{xx} \). Six vector fields will be of special importance since they generate the symmetry groups of the equation:

\[
X_1 = \partial_t, \quad X_2 = x \partial_x + 2t \partial_t, \quad X_3 = 4xt \partial_x + 4t^2 \partial_t - (x^2 + 2t)u \partial_u, \\
X_4 = \partial_x, \quad X_5 = 2t \partial_x - xu \partial_u, \quad X_6 = u \partial_u.
\]

The commutators are given in the following table:

<table>
<thead>
<tr>
<th></th>
<th>( X_1 )</th>
<th>( X_2 )</th>
<th>( X_3 )</th>
<th>( X_4 )</th>
<th>( X_5 )</th>
<th>( X_6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_1 )</td>
<td>0</td>
<td>2( X_1 )</td>
<td>4( X_2 ) - 2( X_6 )</td>
<td>0</td>
<td>2( X_4 )</td>
<td>0</td>
</tr>
<tr>
<td>( X_2 )</td>
<td>0</td>
<td>2( X_3 )</td>
<td>( X_4 )</td>
<td>( X_5 )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>0</td>
<td>2( X_5 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( X_4 )</td>
<td>0</td>
<td>0</td>
<td>-( X_6 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( X_5 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( X_6 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
For instructive purposes we calculate one of the table entries using (2.3):

\[ [X_4, X_6] = \left( (x \partial_x + 2t \partial_t)(4xt) - (4xt \partial_x + 4t^2 \partial_t - (x^2 + 2t)u \partial_u)(x) \right) \partial_x \]
\[ + \left( (x \partial_x + 2t \partial_t)(4t^2) - (4xt \partial_x + 4t^2 \partial_t - (x^2 + 2t)u \partial_u)(2t) \right) \partial_t \]
\[ + \left( (x \partial_x + 2t \partial_t)(-x^2u - 2tu) - (4xt \partial_x + 4t^2 \partial_t - (x^2 + 2t)u \partial_u)(0) \right) \partial_u \]
\[ = \left( (4xt + 8xt) - 4xt \right) \partial_x + \left( 16t^2 - 8t^2 \right) \partial_t + \left( (-2x^2u - 4tu) - 0 \right) \partial_u \]
\[ = 2 \left( 4xt \partial_x + 4t^2 \partial_t - (x^2 - 2t)u \partial_u \right) = 2X_6. \]

Note that we consider \((x,t,u)\) as independent differential variables, and therefore \(\partial_x u = \partial_t u = 0\).

### 2.4 The Lie Algebra

In order to work with a linear algebra instead of the full complexity of a Lie group, we now wish to establish a correspondence between vector fields and Lie groups. We start by defining a property of vector fields called left invariance.

**Definition 2.26.** For a Lie group \((G, \cdot)\) and any element \(x \in G\) we define the \textit{left translation}

\[ L_x : G \to G, \quad L_x(y) = x \cdot y, \]

which is a diffeomorphism with inverse

\[ (L_x)^{-1}(y) = L_{x^{-1}}(y) = x^{-1} \cdot y. \]

**Definition 2.27.** A vector field \(X\) is called \textit{left invariant} if it is preserved under left translations, meaning that

\[ dL_x(X) = X, \]

for all \(x \in G\). Equivalently, if we consider the value of \(X\) at two distinct points \(y \in G\) and \(z = xy \in G\), they must be related by the linear map \(dL_x : T_yG \to T_zG\)

\[ X_z = X_{xy} = dL_x(X_y). \]

This means that a left invariant vector field is completely determined by its value at the identity, \(X_e\), since the value of the tangent at any point \(x \in G\) is obtained by a left translation

\[ X_x = X_{xe} = dL_x(X_e). \quad (2.4) \]

Now, linear combinations of left invariant vector fields are also left invariant, meaning the set of them form a vector space. Using the above observation, we can identify this vector space with the tangent space at the identity. We state this as a theorem, a proof of which is found in van den Ban [13, p. 15].

**Theorem 2.28.** The space of all left invariant vector fields on a Lie group \(G\) is isomorphic to the tangent space at the identity \(T_eG\).
Recalling that an algebra is a vector field with a multiplication operation, we define the Lie algebra corresponding to a Lie group using the Lie bracket from definition 2.23 as multiplication.

**Definition 2.29.** The Lie algebra $\mathfrak{g}$ of a Lie group $G$ is the space of all left invariant vector fields on $G$ equipped with the Lie bracket.

Returning to the flow generated by a vector field, we will establish a correspondence between one-parameter subgroups of a Lie group, and the left-invariant vector fields of the Lie algebra.

**Definition 2.30.** The integral curve that passes through $x \in M$ at $\varepsilon = 0$, given by the flow generated by the vector field $X$, is written $\gamma_X(\varepsilon) = \exp(\varepsilon X)x$.

**Proposition 2.31.** The integral curve in definition 2.30 satisfies the following properties:

1. $\exp(\varepsilon X)x = x$ for $\varepsilon = 0$
2. $\exp(\varepsilon X) \exp(\delta X)x = \exp((\varepsilon + \delta)X)x$
3. $\exp(\varepsilon X)^{-1}x = \exp(-\varepsilon X)x$
4. $\frac{d}{d\varepsilon} \exp(\varepsilon X)x = X_{\exp(\varepsilon X)}x$

**Proposition 2.32.** Let $X$ be a left invariant vector field on a Lie group $G$. Then the flow generated by $X$ through the identity

$$g_\varepsilon = \exp(\varepsilon X)e \equiv \exp(\varepsilon X)$$

is defined for all $\varepsilon \in \mathbb{R}$ and forms a one-parameter subgroup of $G$, with

$$g_{\varepsilon + \delta} = g_\varepsilon \cdot g_\delta, \quad g_0 = e, \quad g_\varepsilon^{-1} = g_{-\varepsilon}.$$

Conversely, any connected one-dimensional subgroup of $G$ is generated by such a left invariant vector field.

**Proof.** A more general statement is proved in Olver [11, p. 45].

We now have a correspondence between transformation groups and vector fields. A one-parameter group of transformations on a manifold can be generated by a vector field, obtained by using the fourth property of proposition 2.31 setting $\varepsilon = 0$. Conversely, given a vector field $X$ we find the transformation group by solving the system (2.2). We will therefore refer to $X$ as the (infinitesimal) generator of the corresponding group.

In definition 2.15 we introduced a condition for being a symmetry of an equation. In practice, finding a symmetry group is much simpler on the level of generating vector fields using the following theorem.

**Theorem 2.33.** A connected Lie group $G$ is a symmetry group of the regular equation $f(x) = 0$ if and only if

$$X(f(x)) = 0, \quad \text{when } f(x) = 0,$$

for every infinitesimal generator $X \in \mathfrak{g}$ of $G$. 

12
Proof. A proof can be found in Olver [11, p. 80].

This result is central in that it allows us to find the vector fields and thus the one-parameter groups under which an equation is invariant. In the next chapter we will introduce geometrical constructs that allow us to apply the theorem to differential equations.

We finish the chapter by defining properties of a Lie algebra that can be leveraged when solving or reducing the order of differential equations [9].

Definition 2.34. Let \( g \) be a Lie algebra. A subspace \( h \subseteq g \) is called a subalgebra of the Lie algebra \( g \) if it is closed under commutation:

\[
[X,Y] \in h \quad \text{for all} \quad X,Y \in h.
\]

The subalgebra \( h \subseteq g \) is called an ideal of \( g \) if

\[
[X,Y] \in h \quad \text{for all} \quad X \in h, Y \in g.
\]

A finite Lie algebra \( g \) is said to be solvable if there is a sequence of subalgebras

\[
g_1 \subseteq \ldots \subseteq g_{m-1} \subseteq g_m = g
\]

of dimensions \( 1, \ldots, m-1, m \) such that \( g_k \) is an ideal in \( g_{k+1} \) for \( k = 1, \ldots, m-1 \).
Chapter 3

Differential Equations and the Application of Lie Groups

3.1 Geometrical Setting

The geometrical interpretation of differential equations explored in this section is adapted from Cicogna [4] and Olver [12].

Consider an equation $f(x) = 0$ for $x$ in $\mathbb{R}^n$, as discussed in section 2.2. For a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ of maximal rank the solution set $S_f = \{x : f(x) = 0\}$ is a submanifold of $\mathbb{R}^n$. We want to regard solutions of differential equations in an analogous way: as the solution set defined by the vanishing of functions in the space of all variables as well as relevant derivatives, the jet space. For clarity we will only consider a single scalar differential equation at a time, but the extension to vector valued and systems of equations is straightforward.

A general $n$-th order differential equation

$$F(x, u^{(n)}) = 0$$

is a relation between $p$ independent variables $x = (x^1, \ldots, x^p)$, the dependent variable $u$ and partial derivatives up to order $n$, here collectively denoted $u^{(n)}$. The differential equation is thus defined by the vanishing of a differential function $F : M^{(n)} \to \mathbb{R}$ defined on the $n$-th jet space $M^{(n)}$, defined as follows:

**Definition 3.1.** Let $x \in \Omega \subset \mathbb{R}^p$, $u \in U \subset \mathbb{R}$, and let $U_k$ be the space of all $k$-th order derivatives of $u$ with respect to $x^i$, i.e. $\partial u/\partial x^i$, $\partial^2 u/\partial x^i \partial x^j$, etc., up to order $k$, for $i, j = 1, \ldots, p$. Then the jet space of order $n$ is

$$M^{(n)} = \Omega \times U \times U_1 \times \cdots \times U_n,$$

also called the $n$-th prolongation of the space of variables $M \simeq \Omega \times U$.

The jet space is locally isomorphic to a real space with one important distinction: there is a contact structure in it that sets the relations among the derivatives. This is best illustrated by example.
Example 3.2. If we consider the case with one independent variable $x$ and one dependent variable $y$, then we have $\Omega \simeq \mathbb{R}$, $U \simeq \mathbb{R}$, $M \simeq \mathbb{R}^2$ and the first jet space $M^{(1)} \simeq \mathbb{R}^3$ with coordinates $(x, y, y_x)$. Now, if we have a curve $\gamma \in M$ corresponding to a function $y(x)$ with derivative $dy/dx = y_x(x)$, and we then consider its prolongation $\gamma^{(1)} \in M^{(1)}$ given by $(x, y(x), y_x(x))$, then $\gamma^{(1)}$ is not an arbitrary curve in $\mathbb{R}^3$, for it must at all points satisfy the condition $y_x(x) = dy/dx$.

So, just like the solution set of a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ defines a manifold in $\mathbb{R}^n$, the solution of a differential equation defines a submanifold

$$S_F = \{ (x, u^{(n)} : F(x, u^{(n)}) = 0 \}$$

(3.2)

in the jet space $M^{(n)}$, called the solution manifold of the equation.

Given a function $u = f(x), f : \Omega \to U$, we define its graph $\Gamma_f$ in $M$

$$\Gamma_f = \{ (x, u) \in M : u = f(x) \} \subset M.$$ 

Since we know all the partial derivatives of a given $f$ up to order $n$, we can define its $n$-th prolongation $f^{(n)}$, and identify it with its graph

$$\Gamma_{f^{(n)}} = \{ (x, u^{(n)}) \in M^{(n)} : u^{(n)} = f^{(n)}(x) \} \subset M^{(n)}.$$ 

So, for $u = f(x)$ to be a solution to the differential equation (3.1), the graph of $f^{(n)}$ must be entirely contained in the solution manifold: $\Gamma_{f^{(n)}} \subset S_F$. This geometric reformulation is equivalent to the classical requirement that a solution must satisfy $F(x, f^{(n)}) = 0$, but it will enable us to find solutions through prolonged groups of transformations acting on submanifolds of $M^{(n)}$.

3.2 Symmetry Groups

This section is based on material from Hydon [6] and Ibragimov [8], both of which present a large number of symmetry methods for ODE and PDE, as well as Olver [11, 12].

In the previous chapter we introduced Lie groups and their action as (local) groups of transformations. We also introduced the correspondence between Lie groups and vector fields, giving rise to the associated Lie algebra. We shall now apply these concepts to manipulate and solve differential equations.

We consider smooth, invertible transformations $g_\varepsilon : M \to M$,

$$\begin{align*}
(x, u) \mapsto g_\varepsilon \cdot (x, u) = (\tilde{x}, \tilde{u}) = \big( \tilde{x}(x, u, \varepsilon), \tilde{u}(x, u, \varepsilon) \big),
\end{align*}$$

(3.3)

where $\varepsilon$ is a real parameter. For transformations such that

1. $g_0$ is the identity transformation, i.e. $(\tilde{x}, \tilde{u}) = (x, u)$ for $\varepsilon = 0$,

2. $g_\varepsilon^{-1} = g_{-\varepsilon}$, and
3. \( g_\varepsilon g_\delta = g_{\varepsilon + \delta} \) for every \( \varepsilon, \delta \) sufficiently close to zero, \( g_\varepsilon \) is a one-parameter Lie group.

We write the infinitesimal generator, the vector field that generates a flow equal to the transformations,

\[
X = \sum_{i=1}^{p} \xi^i(x, u) \frac{\partial}{\partial x^i} + \eta(x, u) \frac{\partial}{\partial u},
\]

(3.4)

where

\[
\xi^i(x, u) = \frac{\partial \tilde{x}^i(x, u, \varepsilon)}{\partial \varepsilon}, \quad \eta(x, u) = \frac{\partial \tilde{u}(x, u, \varepsilon)}{\partial \varepsilon}.
\]

At the heart of applying Lie theory to the solution of differential equations is finding symmetries, transformations that take solutions to solutions. We could define a symmetry only in terms of this property as follows.

**Definition 3.3.** A group of transformations \( G \) acting on \( M \) is a symmetry of the differential equation (3.1) if whenever \( u = f(x) \) is a solution, then \( \tilde{u} = \tilde{f}(\tilde{x}) \) is also a solution.

However, we want to make use of our introduced geometrical setting for a definition from which we can construct solutions. In definition 2.15 we defined a symmetry for the solutions of a function in \( \mathbb{R}^n \). Now that we have defined the jet space \( M^{(n)} \), we can adapt this definition to the solution manifold of a differential equation.

**Definition 3.4.** Let \( G \) be a group acting on \( M \), and \( F : M^{(n)} \to \mathbb{R} \). Then \( G \) is a symmetry of \( F(x, u^{(n)}) = 0 \) if the solution manifold

\[
S_F = \{(x, u^{(n)}) : F(x, u^{(n)}) = 0\} \subset M^{(n)}
\]

is invariant under the action of the prolonged actions of \( G^{(n)} \) on \( M^{(n)} \), which is given in terms of its generator in the following (see proposition 3.6).

One question arises: a given group of transformations only defines an action on the independent and dependent variables, but how are the derivative variables that we have introduced affected? We need to *prolong* the transformations to act on the jet space, so that we extend our transformations from being maps \( g : M \to M \) to being prolonged maps \( g^{(n)} : M^{(n)} \to M^{(n)} \). To this end we use the total derivative, which treats \( u \) and its derivatives as functions of the independent variables. We emphasise that we consider \( u_i = \partial u(x)/\partial x^i \), \( u_{ij} = \partial^2 u(x)/\partial x^i \partial x^j \), etc., as variables and coordinates in the jet space \( M^{(n)} \).

**Definition 3.5.** For a differential function \( F(x, u^{(n)}) : M^{(n)} \to \mathbb{R} \),

\[
D_1 F = \frac{\partial F}{\partial x^i} + u_i \frac{\partial F}{\partial u} + \sum_{j=1}^{p} u_{ij} \frac{\partial F}{\partial u_j} + \ldots
\]

is the *total derivative* of \( F \) with respect to \( x^i \).

The total derivative treats \( u \) and all its derivatives as functions of \( x^i \), and so captures all of \( F \)'s dependence on \( x^i \). We can now use it to deduce how the transformations on \( M \) act on the derivatives.
Proposition 3.6. For a one-parameter group of transformations $G$ acting on $M \simeq \mathbb{R}^p \times \mathbb{R}$, with generator

$$X = \sum_{i=1}^{p} \xi^i(x,u) \frac{\partial}{\partial x^i} + \eta(x,u) \frac{\partial}{\partial u},$$

the first prolongation of the generator of $G^{(1)}$, acting on $M^{(1)}$, is

$$X^{(1)} = X + \sum_{i=1}^{p} \eta^i(x, u^{(1)}) \frac{\partial}{\partial u^i},$$

the second prolongation of the generator of $G^{(2)}$, acting on $M^{(2)}$, is

$$X^{(2)} = X^{(1)} + \sum_{j=1}^{p} \sum_{i=1}^{p} \eta^{ij}(x, u^{(2)}) \frac{\partial}{\partial u^{ij}},$$

and so on. The $\eta^i$ and $\eta^{ij}$ are found recursively by the formulas

$$\eta^i = D_i(\eta) - \sum_{k=1}^{p} u_k D_i(\xi^k), \quad (3.5)$$

$$\eta^{ij} = D_j(\eta^i) - \sum_{k=1}^{p} u_{ik} D_j(\xi^k). \quad (3.6)$$

Proof. A proof of a more general proposition can be found in [11, p. 110].

Now that we have introduced the machinery of prolongations we can apply theorem 3.7 to differential equations. We reformulate it in this context.

Theorem 3.7. A connected Lie group $G$ is a symmetry group of an $n$-th order differential equation $F(x, u^{(n)}) = 0$ if and only if

$$X^{(n)}(F(x, u^{(n)})) = 0, \quad \text{whenever } F(x, u^{(n)}) = 0, \quad (3.7)$$

for every infinitesimal generator $X \in \mathfrak{g}$ of $G$.

Geometrically we want to find a transformation that leaves the solution manifold invariant. If we imagine our solution set as a manifold embedded in the jet space, we seek a transformation whose prolonged generating vector field is always parallel to it.

Equation (3.7) leads to a system of linear, over-determined system of PDEs for the coefficients $\xi^i$ and $\eta$ of $X$. These are called the determining equations, and they can often be explicitly solved.

Example 3.8. Let us consider a first order ODE on the form $y'(x) = f(x,y)$. Viewed in the jet space $M^{(1)}$, with coordinates $(x, y, y')$, the solution surface is defined by $F(x, y, y') = f(x, y) - y' = 0$. To find a group of transformations $G$ acting on the $(x, y)$-plane with generator $X = \xi(x,y) \partial_x + \eta(x,y) \partial_y$ that is a symmetry of the ODE, we consider its first prolongation,

$$X^{(1)} = \xi(x,y) \frac{\partial}{\partial x} + \eta(x,y) \frac{\partial}{\partial y} + \eta^2(x,y,y') \frac{\partial}{\partial y'}.$$
to find its effect on the derivative coordinate $y'$. By (3.5), we have
$$\eta^x = D_x(\eta) - y_x D_x(\xi) = \eta_x + y' \eta_y - y'(\xi_x + y' \xi_y),$$
and so by theorem 3.7, $G$ is a symmetry of the ODE when
$$X^{(1)}(F) = \xi(x, y) \frac{\partial F}{\partial x} + \eta(x, y) \frac{\partial F}{\partial y} + \eta^x(x, y, y') \frac{\partial F}{\partial y'} = \xi f_x + \eta f_y - (\eta_x + y' \eta_y - \xi_x + (y')^2 \xi_y) = 0,$$
where the last equality follows from (3.8). Since $y'(x) = f(x, y)$, we can substitute $f$ for $y'$ above to get
$$\xi f_x + \eta f_y - (\eta_x + f(\eta_y - \xi_x) + (f)^2 \xi_y) = 0,$$
which for a given function $f(x, y)$ is a system of equations to be solved for $\xi$ and $\eta$. See chapter 4.1 for a concrete example of an ODE on this form.

### 3.3 Solving ODE and PDE Using Symmetries

Solving the determining equations we get a number of generators for one-parameter groups that are symmetries of our differential equation. How they can be used in solving or exploring it depends on the type of equation. We give but a brief summary of three methods. For a more thorough treatment we recommend the accessible textbooks by Hydon [6] and Ibragimov [8].

#### 3.3.1 Canonical Coordinates

A first order ODE with one symmetry can be solved by quadrature (integration). We will present the method of finding invariants of the symmetry and using them to find a suitable change of coordinates.

A symmetry is a bijective transformation that maps solutions to other solutions. The generating vector field in general depends on all variables, but with the right invertible coordinate change $r = r(x, y)$, $s = s(x, y)$, called canonical coordinates, we transform $X_{(x,y)} = \xi(x, y) \partial_x + \eta(x, y) \partial_y$ into the constant vector field $X_{(r,s)} = \partial_s$. The flow it generates is a translation in $s$, and so the ODE will not depend on $s$ but be of the form $s'(r) = f(r)$, which can be integrated.

**Proposition 3.9.** We find these canonical coordinates by solving
$$X(r) = \xi(x, y) \frac{\partial r}{\partial x} + \eta(x, y) \frac{\partial r}{\partial y} = 0,$$
$$X(s) = \xi(x, y) \frac{\partial s}{\partial x} + \eta(x, y) \frac{\partial s}{\partial y} = 1.$$  

**Proof.** This is immediate by the chain rule, since in the new coordinates
$$X_{(r(x,y), s(x,y))} = \xi(x, y) \left( \frac{\partial r}{\partial x} \frac{\partial}{\partial r} + \frac{\partial r}{\partial y} \frac{\partial}{\partial s} \right) + \eta(x, y) \left( \frac{\partial s}{\partial x} \frac{\partial}{\partial r} + \frac{\partial s}{\partial y} \frac{\partial}{\partial s} \right)$$
$$= X(r) \frac{\partial}{\partial r} + X(s) \frac{\partial}{\partial s}.$$ 

\[\square\]
Equation (3.10) dictates that \( r(x, y) \) is an invariant of the symmetry, constant on the orbits of \( X \). It is found by the method of characteristics. Essentially, we solve (2.2) to get the group action, and then eliminate the parameter and solve for the constant of integration to get an invariant function \( \psi(x, y) = C \), also called a *first integral*. We then set \( r = \psi(x, y) \), or a function thereof.

This can also be done by solving the parameter independent characteristic equation

\[
\frac{dx}{\xi(x, y)} = \frac{dy}{\eta(x, y)}.
\]

Similarly, we find \( s = s(x, y) \) that satisfies (3.11) by solving

\[
\frac{dx}{\xi(x, y)} = \frac{dy}{\eta(x, y)} = ds.
\]

See chapter 4.1 for an example.

This technique can also be used on higher order ODE. For each one-parameter group the order can be lowered by one as long as the Lie algebra is solvable, see e.g. Hydon [6, Ch. 4.1] and Olver [11, Ch. 2.5].

### 3.3.2 Transformation of a Known Solution

In the case of PDE, we can use the knowledge of one solution to generate whole families of solutions using the symmetries we have found. A one-parameter group of transformations gives rise to a one-parameter family of solutions. So if \( u = f(x) \) is a solution to a PDE, then given a symmetry \((\tilde{x}, \tilde{u}) = g_\epsilon \cdot (x, u)\)

\[
u = g_\epsilon \cdot f(x)
\]

is also a solution for those \( \epsilon \) where it is defined. For an example of this, see in chapter 4.2.

### 3.3.3 Group-invariant Solutions

This method uses invariants of a group \( G \) to reformulate the differential equation in terms of these, which will result in an equation with one less independent variable. Say we have a PDE \( F(x, t, u^{(n)}) = 0 \) in two independent variables \( x \) and \( t \) with a symmetry generated by

\[
X = \xi(x, t, u) \frac{\partial}{\partial x} + \tau(x, t, u) \frac{\partial}{\partial t} + \eta(x, t, u) \frac{\partial}{\partial u},
\]

with corresponding characteristic equations

\[
\frac{dx}{\xi} = \frac{dt}{\tau} = \frac{du}{\eta}.
\]

We get two invariants \( r(x, t, u) \) and \( s(x, t, u) \) (in general one less than the number of variables [3]). Letting one play the role of dependent variable, \( s = f(r) \), we determine the derivatives of \( u \) in terms of \( s \) with respect to \( r \) (and perhaps \( x \) or \( t \) which will later be redundant). Substituting back into the PDE, the result is an ODE, which we solve to find the solutions that are invariant under the given group action. An example is given at the end of chapter 4.2.
Chapter 4

Examples

4.1 A Non-linear ODE

Let us consider the ODE

\[ y'(x) = \frac{x^2y - x + y^3 - y}{x^3 + xy^2 - x + y} \quad (4.1) \]

Seeking a symmetry we use equation (3.7),

\[ X^{(1)}(F) = \xi(x, y) \frac{\partial F}{\partial x} + \eta(x, y) \frac{\partial F}{\partial y} + \eta^s(x, y, y') \frac{\partial F}{\partial y'} = 0, \]

where \( F = \frac{x^2y - x + y^3 - y}{x^3 + xy^2 - x + y} - y'(x) \), to determine \( \xi \) and \( \eta \). Using the prolongation formula (3.8) for \( \eta \) and (4.1) to replace \( y' \), after simplification we get

\[
\begin{align*}
\xi (-x^4y + 2x^3 - 2x^2y^3 + 2x^2y + 2xy^2 - y^5 + 2y^3 - 2y) \\
+ \eta (x^5 + 2x^3y^2 - 2x^3 + 2x^2y + xy^4 - 2xy^2 + 2x + 2y^3) \\
+ \eta_x (-x^6 - 2x^4y^2 + 2x^4 - 2x^3y - x^2y^4 + 2x^2y^2 - x^2 - 2xy^3 + 2xy - y^2) \\
+ (\eta_y - \xi_y) (-x^5y + x^4 - 2x^3y^3 + 2x^3y - x^2 - xy^5 + 2xy^3 - y^4 + y^2) \\
+ \xi_y (x^4y^2 - 2x^3y + 2x^2y^4 - 2x^2y^2 + x^2 - 2xy^3 + 2xy + y^6 - 2y^4 + y^2) = 0.
\end{align*}
\]

To solve this unwieldy equation we make an ansatz that the coefficient functions are polynomial. Constant polynomials \( \xi(x, y) = c_1 \) and \( \eta(x, y) = c_2 \) give no non-trivial solutions, and so the ODE has no translation symmetry in the \((x, y)\)-plane. Using first degree polynomials \( \xi(x, y) = c_1 x + c_2 y + c_3 \) and \( \eta(x, y) = c_4 x + c_5 y + c_6 \) gives non-trivial solutions. We get an over-determined system of equations, with one equation for each monomial, of which only 5 linearly independent, namely

\[
\begin{align*}
c_3 &= 0 \\
c_6 &= 0 \\
2(c_1 + c_5) &= 0 \\
c_1 + c_2 + c_4 - c_5 &= 0 \\
2(-c_1 + c_2 + c_4 + c_3) &= 0.
\end{align*}
\]
This gives a one-dimensional space of solutions, namely \( c_2 = -c_4 \) with all other \( c_i = 0 \). Choosing \( c_2 = -1, c_4 = 1 \) gives us the generator \( X = -y\partial_x + x\partial_y \) of the rotation group \( \text{SO}(2) \).

Next we wish to use this symmetry to find canonical coordinates in which the ODE is integrable by quadrature. Since the transformation under which the equation is invariant is rotation around the origin, it might be clear that polar coordinates are the right choice. Generally, the choice of coordinates is not obvious and we therefore demonstrate how to find them using the method of characteristics.

We use (3.10) to determine the independent variable \( r \):

\[
X(r) = -y\frac{\partial r}{\partial x} + x\frac{\partial r}{\partial y} = 0,
\]

which corresponds to the characteristic equation

\[
\frac{dx}{-y} = \frac{dy}{x}.
\]

It is integrated to give the first integral \( x^2 + y^2 = c_1 \), and so the general solution is \( r = F(x^2 + y^2) \). We choose \( r = \sqrt{x^2 + y^2} \). The dependent variable is found using (3.11),

\[
X(s) = -y\frac{\partial s}{\partial x} + x\frac{\partial s}{\partial y} = 1,
\]

which corresponds to the characteristic equations

\[
\frac{dx}{-y} = \frac{dy}{x} = ds.
\]

For \( x > 0 \) we substitute \( x = \sqrt{r^2 - y^2} \) giving

\[
ds = \frac{dy}{\sqrt{r^2 - y^2}},
\]

with solution \( s = \arcsin(\frac{y}{r}) + c_2 \). Using \( \arcsin(\frac{y}{r}) = \arctan(\frac{y}{\sqrt{r^2 - y^2}}) \), and substituting back for \( r \) we get \( s = \arctan(\frac{y}{x}) + c_2 \). We can therefore confirm that polar coordinates

\[
r = \sqrt{x^2 + y^2}, \quad s = \arctan(\frac{y}{x}),
\]

are canonical coordinates for (4.1) and will render it translation invariant and thus solvable by quadrature. Using the chain rule and the inverse coordinate change \( x = r \cos s, y = r \sin s \) we get

\[
\frac{ds}{dr} = \frac{s_x + s_y y_x}{r_x + r_y y_x} = \frac{1}{\sqrt{x^2 + y^2 (1 - x^2 - y^2)}} = \frac{1}{r(1 - r^2)}.
\]

which can be integrated (using partial fractions) to yield

\[
s(r) = \int \frac{dr}{r(1 - r^2)} = \frac{1}{2} \ln \left| \frac{r^2}{1 - r^2} \right| + C. \quad (4.2)
\]
Changing back to \((x, y)\)-coordinates we get an implicit expression for the solutions where \(x > 0\), namely
\[
\arctan\left(\frac{y}{x}\right) = \ln\left|\frac{x^2 + y^2}{1 - x^2 - y^2}\right| + C.
\]

The solution \((4.2)\) in polar coordinates is more illuminating, revealing e.g. that \(\lim_{r \to \infty} s(r) = C\), i.e. the constant \(C\) is an asymptote for the angle \(s\). See figure \(4.1\) for the graph of solutions for two choices of the constant \(C\). It is clear to see that one solution is attained from the other by rotating \(\pi\) radians, or, equivalently, following the flow of the vector field \(X_{(r,s)} = \partial_s\) for \(\varepsilon = \pi\).

![Figure 4.1: Two solutions of \((4.1)\) for \(C_1 = 0, C_2 = \pi\).](image)

### 4.2 The Heat Equation

Let us consider the heat equation in one space dimension,
\[
u_t = u_{xx},
\]
(4.3)
elaborating the treatment given in Olver [11, Examples 2.41 and 3.3], filling in some computational details.

We seek generating vector fields acting on the space of independent and dependent variables, \(M \simeq X \times U\), of the form
\[
X = \xi(x, t, u) \frac{\partial}{\partial x} + \tau(x, t, u) \frac{\partial}{\partial t} + \eta(x, t, u) \frac{\partial}{\partial u}.
\]
(4.4)

Since the equation involves \(u_{xx}\) the solution manifold lies in the second jet space \(M^{(2)} \simeq X \times U^{(2)} \simeq \mathbb{R}^2 \times \mathbb{R}^6\), and we therefore require the second prolongation,
\[
X^{(2)} = X + \eta^x \frac{\partial}{\partial u_x} + \eta^t \frac{\partial}{\partial u_t} + \eta^{xx} \frac{\partial}{\partial u_{xx}} + \eta^{xt} \frac{\partial}{\partial u_{xt}} + \eta^{tt} \frac{\partial}{\partial u_{tt}},
\]
to account for the action on all derivatives involved. A generator gives a symmetry transformation when the vector field is everywhere tangent to the solution.
manifold, by theorem 3.7 when

\[ X^{(2)}(F) = \eta^t - \eta^{xx} = 0, \]  

where \( F = u_t - u_{xx} \). The coefficients \( \eta^t \) and \( \eta^{xx} \) are found using (3.5) and (3.6):

\[
\begin{align*}
\eta^t &= D_t(\eta) - u_x D_t(\xi) - u_t D_t(\tau) \\
&= (\eta_t + u_t \eta_u) - u_x (\xi_t - u_t \xi_u) - u_t (\tau_t - u_t \tau_u) \\
&= \eta_t + \xi_t u_x + (\eta_u - \tau_t) u_x - \xi_u u_x u_t - \tau_u u_t^2,
\end{align*}
\]

\[
\begin{align*}
\eta^{xx} &= D_x(\eta) - u_x D_x(\xi) - u_t D_x(\tau) \\
&= (\eta_x + u_x \eta_u) - u_x (\xi_x - u_x \xi_u) - u_t (\tau_x - u_x \tau_u) \\
&= \eta_x + (\eta_u - \xi_x) u_x - \tau_x u_t - \xi_u u_x^2 - \tau_u u_x u_t,
\end{align*}
\]

Substituting the above expressions into equation (4.5) and substituting \( u_{xx} \) for \( u_t \) we get the determining equations, one equation for each monomial \((u_x, u_t, u_x^2, \text{etc.})\),

\[
\begin{align*}
&u_x u_{xt} - 0 = -2\tau_u \\
u_{xt} - 0 = -2\tau_x \\
u_x^2 - \tau_x = -\tau_u \\
u_x^2 u_{xx} - 0 = -\tau_{uu} \\
u_{xx} u_{xt} - \xi_u = -2\tau_{xx} - 3\xi_u \\
u_{xx} - \eta_u - \tau_x = -\tau_{xx} + \eta_u \\
u_x^3 - 0 = -\xi_{uu} \\
u_x^2 - 0 = \eta_{uu} - 2\xi_{uu} \\
u_x - \xi_t = 2\eta_{xx} - \xi_{xx} \\
1 - \eta_t = -\eta_{xx}
\end{align*}
\]

Equations (4.6) and (4.7) imply that \( \tau = \tau(t) \), i.e. \( \tau \) is a function only of \( t \), and then (4.10) implies that \( \xi = \xi(x, t) \). We can then use (4.11) to conclude

\[ \xi(x, t) = \frac{1}{2} \tau_t(t) x + \sigma(t), \]  

where \( \sigma \) is an arbitrary function of \( t \). Equation (4.13) is now \( \eta_{uu} = 0 \), so \( \eta \) is linear in \( u \), i.e.

\[ \eta(x, t, u) = \beta(x, t) u + \alpha(x, t), \]

for arbitrary functions \( \alpha \) and \( \beta \). Now (4.14) gives \( 2\beta_x(x, t) = 2\eta_{uu} = -\xi_t = -\frac{1}{2}\tau_{tt}(t) x + \sigma_t(t) \), so

\[ \beta(x, t) = \frac{1}{8}\tau_{tt}(t) x^2 - \frac{1}{2}\sigma_t(t) x + \rho(t). \]  

(4.17)
Also, (4.15) requires $\beta_t(x,t)u + \alpha_t(x,t) = \beta_{xx}(x,t)u + \alpha_{xx}(x,t)$, so $\alpha_t(x,t) = \alpha_{xx}(x,t)$ which means $\alpha(x,t)$ must be a solution of the heat equation. Secondly, together with (4.17), we get

$$-\frac{1}{8}\tau_{ttt}(t)x^2 - \frac{1}{2}\sigma_{tt}(t)x - \rho_t(t) = -\frac{1}{4}\tau_{tt}(t)$$

so $\tau(t)_{ttt} = 0$, which implies $\tau$ is a second degree polynomial in $t$, i.e.

$$\tau(t) = c_1 + c_2t + c_3t^2.$$  

Equation (4.18) also implies $\rho(t) = -\frac{1}{4}\tau_t$ and $\sigma(t)_{tt} = 0$, dictating that $\rho(t) = -\frac{1}{4}(2c_3t + c_6)$, and $\sigma(t) = c_4 + c_5t$, so

$$\beta(x,t) = -\frac{1}{8}2c_3x^2 - \frac{1}{2}c_5x - \frac{1}{4}(c_6 + 2c_3t).$$  

Then by (4.10) and (4.19) we conclude

$$\xi(x,t) = \frac{1}{2}(c_2 + 2c_3t)x + c_4 + c_5t,$$

and finally, using (4.20),

$$\eta(x,t,u) = (-\frac{1}{2}c_6 - \frac{1}{2}c_5x - \frac{1}{2}c_3t - \frac{1}{4}c_3x^2)u + \alpha(x,t).$$

So we have six arbitrary constants and degrees of freedom. By setting all $c_i$ but one to zero, recalling the form of our vector field from (4.4), we arrive at the following six vector fields spanning the Lie algebra:

$$X_1 = \partial_t \quad (c_1)$$
$$X_2 = x\partial_x + 2t\partial_t \quad (c_2)$$
$$X_3 = 4xt\partial_x + 4t^2\partial_t - (x^2 + 2t)u\partial_u \quad (c_3)$$
$$X_4 = \partial_x \quad (c_4)$$
$$X_5 = 2t\partial_x - xu\partial_u \quad (c_5)$$
$$X_6 = u\partial_u \quad (c_6)$$

and an infinite-dimensional subalgebra $X_\alpha = \alpha(x,t)\partial_u$. The commutator table is given in example 2.25. Exponentiating these give us the most general symmetry group of (4.3), and the corresponding group actions are, given as $(\tilde{x}, \tilde{t}, \tilde{u}) = \exp(\varepsilon X_i)(x,t,u)$:

$$G_1 : (x,t + \varepsilon, u)$$
$$G_2 : (e^\varepsilon x, e^{2\varepsilon}t, u)$$
$$G_3 : (\frac{x}{1 - 4\varepsilon t}, \frac{t}{1 - 4\varepsilon t}, u\sqrt{1 - 4\varepsilon t} \exp\left(\frac{-\varepsilon x^2}{1 - 4\varepsilon t}\right))$$
$$G_4 : (x + \varepsilon, t, u)$$
$$G_5 : (x + 2\varepsilon t, t, u \exp(-\varepsilon x - \varepsilon t^2))$$
$$G_6 : (x, t, e^\varepsilon u)$$
$$G_\alpha : (x, t, u + \varepsilon \alpha(x,t))$$

Some of these symmetries are easily interpreted: $G_1$ and $G_4$ show that the equation is time- and space-invariant, and $G_\alpha$ and $G_6$ show it is linear; we can add
solutions and multiply them by constants. These are also quite easily deduced from the equation (4.3) itself. The symmetry group $G_3$ on the other hand is not at all obvious, and we use it below to find an important solution.

To find solutions to (4.3), one method is to transform a known solution. The constant solution $u = c$, can be transforming it using the symmetry $G_3$ to get the solution
\[ u = \frac{c}{\sqrt{1 + 4\varepsilon t}} \exp\left(\frac{-\varepsilon x^2}{1 + 4\varepsilon t}\right). \] (4.23)

So for each $c \in \mathbb{R}$, we get a one-parameter family of solutions depending on $\varepsilon$. Setting $c = \sqrt{\varepsilon/\pi}$, we get
\[ u = \sqrt{\varepsilon} \sqrt{\pi(1 + 4\varepsilon t)} \exp\left(\frac{-\varepsilon x^2}{1 + 4\varepsilon t}\right), \]
and translating in $t$ with $-1/4\varepsilon$ using $G_1$, we get
\[ u = \frac{1}{\sqrt{4\pi t}} \exp\left(-\frac{x^2}{4t}\right), \]
which is the so called *fundamental solution*. It can be used to solve initial value problems of (4.3) using convolution, see e.g. Evans [5, Chapter 2.3.1].

Using the method of group-invariant solutions, we can choose a linear combination of the translation symmetries $X_1$ and $X_4$, i.e. we set $X = X_1 + cX_4$ giving the generator
\[ X = \partial_t + c\partial_x, \quad c \in \mathbb{R}. \]
We then use the method of characteristics to find the invariants of $X$. The characteristic equation
\[ \frac{dx}{c} = \frac{dt}{1} \]
gives $x - ct = C$, so one invariant is $r = x - ct$. The other is clearly $u$, so we set $s = u$. Letting $s$ be the dependent variable, the group-invariant solutions $s = f(r)$ will be $u = f(x - ct)$ which are waves travelling at speed $c$. The derivatives $u_t$ and $u_{xx}$ can then be found in terms of $s$ and $r$ using the chain rule:
\[ u_t = s_t = s_r r_t = -cs_r, \quad u_x = s_x = s_r r_x = s_r, \quad u_{xx} = (s_r)_x = (s_r)_x r_x = s_{rr}, \]
and these can then be substituted into the heat equation (4.3) reducing it to the ODE
\[ s''(r) = -cs'(r). \]
It can be solved by setting $w = s'(r)$ making it an elementary system of first order linear equations; the solution is
\[ s(r) = k_1 e^{-cr} + k_2, \]
where $k_i$ are constants. Substituting back to our original variables we finally get the solutions
\[ u(x,t) = k_1 e^{-c(x-ct)} + k_2, \]
a family of travelling waves of exponentials.
4.3 The Kolmogorov Equation

An equation still under theoretical investigation, and with applications ranging from the kinetic theory of gases to the pricing of options (models which are related by the random nature of the motion of particles as well as stock prices), is the Kolmogorov equation

\[ Lu = u_t - u_{xx} - xu_y = 0, \quad (4.24) \]

which is a degenerate parabolic PDE for \( t > 0 \).

The study of this equation spans the last century: in 1931 Kolmogorov publishes the paper "Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung," on analytical methods of probability, and a couple of years later "Zur Theorie der stetigen zufälligen Prozesse" on continuous random processes. Following the methods presented he then provides a fundamental solution for (4.24) in a short paper on Brownian motion [10] in 1934.

In 1967 Hörmander uses (4.24) as an example in a central work [7] establishing sufficient and necessary conditions for a differential operator to be hypoelliptic. Essentially, hypoellipticity means that if the right hand side of our equation is smooth, then so is the solution. Formally, a differential operator \( L \) with \( C^\infty(\Omega) \) coefficients (\( \Omega \) open subset of \( \mathbb{R}^n \)) is said hypoelliptic in \( \Omega \) if, for any open set \( \Omega' \subseteq \Omega \) and any distribution \( u \in D(\Omega) \), \( Lu \in C^\infty(\Omega) \implies u \in C^\infty(\Omega) \).

A contemporary application is the price of "Asian options". They satisfy a stochastic differential equation \( dS(t) = \mu_0 S(t)dt + \sigma SdW \), where the exercise (strike/purchase) price depends on an average on the history of the stock price. For certain conditions the price of the option can be shown to satisfy a PDE equivalent to the Kolmogorov equation (4.24) [3].

We will not solve the determining equations for all possible symmetries, but focus on finding the translations and dilations need to define a (homogeneous) group structure on \( \mathbb{R}^3 \). See Anceschi [1] for a recent article presenting a survey of results on Kolmogorov type equations.

We begin as before by prolonging a general vector field generator

\[ X = \xi \frac{\partial}{\partial x} + \eta \frac{\partial}{\partial y} + \tau \frac{\partial}{\partial t} + \omega \frac{\partial}{\partial u}, \quad (4.25) \]

with unknown coefficients \( \xi, \eta, \tau \) and \( \omega \) that are functions of \((x, y, t, u)\), finding the determining equations by imposing the invariant solution manifold condition from theorem X(2): \( X(u_t - u_{xx} - xu_y) = 0 \), substituting \( u_t \), and collecting terms for each monomial on which the generator does not depend. See section 4.2 for the details of the procedure for the similar example of the heat equation.
The determining equations are

\[
\begin{align*}
\xi_u &= 0 \quad (4.26) \\
\eta_u &= 0 \quad (4.27) \\
\tau_u &= 0 \quad (4.28) \\
\omega_{uu} &= 0 \quad (4.29) \\
\eta_x &= 0 \quad (4.30) \\
\tau_x &= 0 \quad (4.31) \\
-2\omega_{xu} - \xi_t + x\xi_y + \xi_{xx} &= 0 \quad (4.32) \\
-\tau_t + x\tau_y + 2\xi_x &= 0 \quad (4.33) \\
-\xi - \eta_t + x\eta_y - x\tau_t + x^2\tau_y &= 0 \quad (4.34) \\
\omega_t - x\omega_y - \omega_{xx} &= 0 \quad (4.35)
\end{align*}
\]

To find the dilations, we assume there exists a generator

\[
X_\delta = ax \frac{\partial}{\partial x} + by \frac{\partial}{\partial y} + ct \frac{\partial}{\partial t}. \quad (4.36)
\]

Using equations (4.33) and (4.34) and setting \(a = 1\) (since the dilations are only unique up to scaling), we get \(b = 3\) and \(c = 2\) corresponding to the dilation \(\delta(x, y, t) = (e^x e^y, e^{2t})\). Using a different parametrization to emphasise the scaling relationships we write

\[
\delta_\lambda(x, y, t) = (\lambda x, \lambda^3 y, \lambda^2 t), \quad \lambda > 0. \quad (4.37)
\]

For the translations, we begin with the ansätze

\[
X^x = \frac{\partial}{\partial x}, \quad X^y = \frac{\partial}{\partial y}, \quad X^t = \frac{\partial}{\partial t}. \quad (4.38)
\]

We find that \(X^y\) and \(X^t\) satisfy all determining equations (4.26)-(4.35), but \(X^x\) does not. Given our ansatz \(\xi = 1\) for \(X^x\), equation (4.34) demands \(\eta_t = -1\); choosing the simplest solution \(\eta = -t\) results in \(X^x = \partial_x - t\partial_y\). In total, a linear combination gives the vector field

\[
X = x_0 X^x + y_0 X^y + t_0 X^t = x_0 \frac{\partial}{\partial x} + (y_0 - x_0 t) \frac{\partial}{\partial y} + t_0 \frac{\partial}{\partial t}, \quad (4.39)
\]

corresponding to translations \((x, y, t) \circ (x_0, y_0, t_0) = (x + x_0, y + (y_0 - x_0 t), t + t_0)\), for any \((x, y, t)\) and \((x_0, y_0, t_0) \in \mathbb{R}^3\).

We have thus found the symmetries needed to define the Lie group structure, which is crucial for further study of the Kolmogorov equation.
Bibliography


