32 Lie Groups and Differential Equations

Lie groups and Lie algebras, because of their manifold—and therefore, differentiability—structure, find very natural applications in areas of physics and mathematics in which symmetry and differentiability play important roles. Lie himself started the subject by analyzing the symmetry of differential equations in the hope that a systematic method of solving them could be discovered. Later, Emmy Noether applied the same idea to variational problems involving symmetries and obtained one of the most beautiful pieces of mathematical physics: the relation between symmetries and conservation laws. More recently, generalizing the gauge invariance of electromagnetism, Yang and Mills have considered nonabelian gauge theories in which gauge invariance is governed by a nonabelian Lie group. Such theories have been successfully built for three of the four fundamental interactions: electromagnetism, weak nuclear, and strong nuclear. Furthermore, it has been possible to cast the fourth interaction, gravity—as described by Einstein's general theory of relativity—in a language very similar to the other three interactions with the promise of unifying all four interactions into a single force. This chapter is devoted to a treatment of the first topic, application of Lie groups to DEs. The second topic, the calculus of variations and conservation laws, will be discussed in the next chapter. The third topic, that of gauge theories, is treated under the more general setting of fiber bundles in the last part of the book.

32.1 Symmetries of Algebraic Equations

The symmetry group of a system of DEs is a transformation group that acts on both the independent and dependent variables and transforms solutions of the system to other solutions. In order to understand this symmetry group, we shall first tackle the simpler question of the symmetries of a system of *algebraic* equations.

Definition 32.1.1 Let G be a local Lie group of transformations acting on a manifold M. A subset $\mathcal{S} \subset M$ is called G-**invariant** and G is called a **symmetry group** of $\mathcal{S}$ if whenever $g \cdot P$ is defined for $P \in \mathcal{S}$ and $g \in G$, then $g \cdot P \in \mathcal{S}$.

G-invariance and symmetry group defined

S. Hassani, *Mathematical Physics*, DOI 10.1007/978-3-319-01195-0_32,

Example 32.1.2 Let $M = \mathbb{R}^2$.

(a) Let $G = \mathbb{R}^+$ be the abelian multiplicative group of real numbers. Let it act on M componentwise: $r \cdot (x, y) = (rx, ry)$. Then any line going through the origin is a G-invariant subset of $\mathbb{R}^2$.

(b) If $G = SO(2)$ and it acts on M as usual, then any circle is a G-invariant subset of $\mathbb{R}^2$.

system of algebraic equations and their symmetry group

A *system of algebraic equations* is a system of equations

$$F_\nu(x) = 0, \quad \nu = 1, 2, \ldots, n,$$

in which $F_\nu : M \to \mathbb{R}$ is smooth. A *solution* is a point $x \in M$ such that $F_\nu(x) = 0$ for $\nu = 1, \ldots, n$. The *solution set* of the system is the collection of all solutions. A Lie group G is called a **symmetry group of the system** if the solution set is G-invariant.

invariant map

Definition 32.1.3 Let G be a local Lie group of transformations acting on a manifold M. A map $F : M \to N$, where N is another manifold, is called a G-**invariant map** if for all $P \in M$ and all $g \in G$ such that $g \cdot P$ is defined, $F(g \cdot P) = F(P)$. A real-valued G-invariant function is called simply an **invariant**.

The crucial property of Lie group theory is that locally the group and its algebra "look alike". This allows the complicated nonlinear conditions of invariance of subsets and functions to be replaced by the simpler linear conditions of invariance under infinitesimal actions. From Definition 29.1.30, we obtain the following proposition.

Proposition 32.1.4 *Let G be a local group of transformations acting on a manifold M. A smooth real-valued function $f : M \to \mathbb{R}$ is G-invariant if and only if*

$$\boldsymbol{\xi}_M|_P(f) = 0 \quad \text{for all } P \in M \tag{32.1}$$

and for every infinitesimal generator $\boldsymbol{\xi} \in \mathfrak{g}$.

Example 32.1.5 The infinitesimal generator for $SO(2)$ is $\boldsymbol{\xi}_M = x\partial_y - y\partial_x$. Any function of the form $f(x^2 + y^2)$ is an $SO(2)$-invariant. To see this, we apply Proposition 32.1.4:

$$(x\partial_y - y\partial_x) f\left(x^2 + y^2\right) = x(2y)f' - y(2x)f' = 0,$$

where f' is the derivative of f.

The criterion for the invariance of the solution set of a system of equations is a little more complicated, because now we are not dealing with functions themselves, but with their solutions. The following theorem gives such a criterion (for a proof, see [Olve 86, pp. 82–83]):

Theorem 32.1.6 *Let G be a local Lie group of transformations acting on an m-dimensional manifold M. Let $F : M \to \mathbb{R}^n$, $n \le m$, define a system of algebraic equations $\{F_\nu(x) = 0\}_{\nu=1}^n$, and assume that the Jacobian matrix $(\partial F_\nu / \partial x^k)$ is of rank n at every solution x of the system. Then G is a symmetry group of the system if and only if*

$$\boldsymbol{\xi}_M|_x(F_\nu) = 0 \quad \forall \nu \text{ whenever } F_\nu(x) = 0 \tag{32.2}$$

for every infinitesimal generator $\boldsymbol{\xi} \in \mathfrak{g}$.

Note that Eq. (32.2) is required to hold *only for solutions x of the system*.

Example 32.1.7 Let $M = \mathbb{R}^2$ and $G = SO(2)$. Consider $F : M \to \mathbb{R}$ defined by $F(x, y) = x^2 + y^2 - 1$. The Jacobian matrix is simply the gradient,

$$(\partial F/\partial x, \partial F/\partial y) = (2x, 2y),$$

and is of rank 1 for all points of the solution set, because it never vanishes at the points where $F(x, y) = 0$, i.e., the unit circle. It follows from Theorem 32.1.6 that G is a symmetry group of the equation $F(x, y) = 0$ if and only if $\boldsymbol{\xi}_M|_{\mathbf{r}}(F) = 0$ whenever $\mathbf{r} \in S^1$. But

$$\boldsymbol{\xi}_M|_{\mathbf{r}}(F) = (x\partial_y - y\partial_x)F|_{\mathbf{r}} = 2xy - 2yx = 0.$$

This is a proof of the obvious fact that $SO(2)$ takes points of S^1 to other points of S^1.

As a less trivial example, consider the function $F : \mathbb{R}^2 \to \mathbb{R}$ given by

$$F(x, y) = x^2y^2 + y^4 + 2x^2 + y^2 - 2.$$

The infinitesimal action of the group yields

$$\boldsymbol{\xi}_M(F) = (x\partial_y - y\partial_x)F = 2x^3y + 2xy^3 - 2xy = 2xy\big(x^2 + y^2 - 1\big).$$

The reader may check that $\boldsymbol{\xi}_M(F) = 0$ whenever $F(x, y) = 0$. The Jacobian matrix of the "system" of equations is the gradient

$$\nabla F = \big(2xy^2 + 4x, 2x^2y + 4y^3 + 2y\big),$$

which vanishes only when $x = 0 = y$, which does not belong to the solution set. Therefore, the rank of the Jacobian matrix is 1. We conclude that the solution set of $F(x, y) = 0$ is a rotationally invariant subset of $\mathbb{R}^2$. Indeed, we have

$$F(x, y) = x^2y^2 + y^4 + 2x^2 + y^2 - 2 = \big(y^2 + 2\big)\big(x^2 + y^2 - 1\big),$$

and the solution set is just the unit circle. Note that although the solution set of $F(x, y) = 0$ is G-invariant, the function itself is not.

We now discuss how to find invariants of a given group action. Start with a one-parameter group and write

$$\mathbf{v} \equiv \boldsymbol{\xi}_M = X^i \frac{\partial}{\partial x^i}$$

for the infinitesimal generator of the group in some local coordinates. A local invariant $F(x)$ of the group is a solution of the linear, homogeneous PDE

$$\mathbf{v}(F) = X^1(x)\frac{\partial F}{\partial x^1} + \cdots + X^n(x)\frac{\partial F}{\partial x^n} = 0. \tag{32.3}$$

It follows that the gradient of F is perpendicular to the vector $\mathbf{v}$. Since the gradient of F is the normal to the hypersurface of constant F, we may consider the solution of Eq. (32.3) as a surface $F(x) = c$ whose normal is perpendicular to $\mathbf{v}$. Each normal determines one hypersurface, and since there are $n-1$ linearly independent vectors perpendicular to $\mathbf{v}$, there must be $n-1$ different hypersurfaces that solve (32.3). Let us write these hypersurfaces as

$$F^j\left(x^1, \ldots, x^n\right) = c^j, \quad j = 1, 2, \ldots, n-1, \tag{32.4}$$

and note that

$$\Delta F^j \approx \sum_{i=1}^n \frac{\partial F^j}{\partial x^i}\Delta x^i = 0, \quad j = 1, 2, \ldots, n-1.$$

A solution to this equation is suggested by (32.3):

$$\Delta x^i = \alpha X^i \quad \Rightarrow \quad \frac{\Delta x^i}{X^i} = \alpha.$$

For $\Delta x^i \to dx^i$, we obtain the following set of ODEs, called the **characteristic system** of the original PDE,

characteristic system of a PDE

$$\frac{dx^1}{X^1(x)} = \frac{dx^2}{X^2(x)} = \cdots = \frac{dx^n}{X^n(x)}, \tag{32.5}$$

whose solutions determine $\{F^j(x)\}_{j=1}^{n-1}$. To find these solutions,

Box 32.1.8 *Take the equalities of (32.5) one at a time, solve the first order DE, write the solution in the form of (32.4), and read off the functions.*

The reader may check that any function of the F^j's is also a solution of the PDE. In fact, it can be shown that *any solution* of the PDE is a function of these F^j's (see [Olve 86, pp. 86–90]).

Example 32.1.9 Once again, let us consider $SO(2)$, whose infinitesimal generator is $-y\partial_x + x\partial_y$. The characteristic "system" of equations is

$$\frac{dx}{-y} = \frac{dy}{x} \quad \Rightarrow \quad x\,dx + y\,dy = 0 \quad \Rightarrow \quad x^2 + y^2 = c.$$

Thus, $F(x,y)=x^2+y^2$, or any function thereof, is an invariant of the rotation group in two dimensions.

As a less trivial example, consider the vector field

$$\mathbf{v}=-y\frac{\partial}{\partial x}+x\frac{\partial}{\partial y}+\sqrt{a^2-z^2}\frac{\partial}{\partial z}$$

where a is a constant. The characteristic system of ODEs is

$$\frac{dx}{-y}=\frac{dy}{x}=\frac{dz}{\sqrt{a^2-z^2}}.$$

The first equation was solved above, giving the invariant $F_1(x,y,z)=\sqrt{x^2+y^2}=r$. To find the other invariant, solve for x and substitute in the second equation to obtain

$$\frac{dy}{\sqrt{r^2-y^2}}=\frac{dz}{\sqrt{a^2-z^2}}.$$

The solution to this DE is

$$\underbrace{\arcsin\frac{y}{r}}_{\alpha}=\underbrace{\arcsin\frac{z}{a}}_{\beta}+C \quad\Rightarrow\quad \arcsin\frac{y}{r}-\arcsin\frac{z}{a}=C.$$

Hence, $F_2(x,y,z)=\arcsin(y/r)-\arcsin(z/a)$ is a second invariant. By taking the sine of F_2, we can come up with an invariant that is algebraic (rather than trigonometric) in form:

$$\begin{aligned}s&=\sin F_2=\sin(\alpha-\beta)=\sin\alpha\cos\beta-\cos\alpha\sin\beta\\&=\frac{y}{r}\sqrt{1-\frac{z^2}{a^2}}-\sqrt{1-\frac{y^2}{r^2}}\frac{z}{a}=\frac{y\sqrt{a^2-z^2}-xz}{ra}.\end{aligned}$$

Any function of r and s is also an invariant.

When the dimension of the Lie group is larger than one, the computation of the invariants can be very complicated. If $\{\mathbf{v}_k\}_{k=1}^r$ form a basis for the infinitesimal generators, then the invariants are the joint solutions of the system of first order PDEs

$$\mathbf{v}_k(F)=\sum_{j=1}^n X_k^j(x)\frac{\partial F}{\partial x^j},\quad k=1,\dots,r.$$

To find such a solution, one solves the first equation and finds all its invariants. Since any function of these invariants is also an invariant, it is natural to express F as a function of the invariants of $\mathbf{v}_1$. One then writes the remaining equations in terms of these new variables and proceeds inductively.

Example 32.1.10 Consider the vector fields

$$\mathbf{u}=-y\partial_x+x\partial_y,$$

$$\mathbf{v} = \left(\frac{x^3z + xy^2z + a^3x}{\sqrt{x^2+y^2}}\right)\partial_x + \left(\frac{x^2yz + y^3z + a^3y}{\sqrt{x^2+y^2}}\right)\partial_y$$
$$-\left(\sqrt{x^2+y^2}z^2 + b^3\right)\partial_z$$

where a and b are constants. The invariants of $\mathbf{u}$ are functions of $r = \sqrt{x^2+y^2}$ and z. If we are to have a nontrivial solution, the invariant of $\mathbf{v}$ as well as its PDE should be expressible in terms of $r = \sqrt{x^2+y^2}$ and z. The reader may verify that

$$\mathbf{v}(F) = \left(r^2z + a^3\right)\frac{\partial F}{\partial r} - \left(rz^2 + b^3\right)\frac{\partial F}{\partial z} = 0$$

with the characteristic equation

$$\frac{dr}{r^2z + a^3} = -\frac{dz}{rz^2 + b^3}.$$

This is an exact first-order DE whose solutions are given by

$$\frac{1}{2}r^2z^2 + a^3z + b^3r = c$$

with c an arbitrary constant. Therefore, $F = \frac{1}{2}r^2z^2 + a^3z + b^3r$, or

$$F(x,y,z) = \frac{1}{2}\left(x^2+y^2\right)z^2 + a^3z + b^3\sqrt{x^2+y^2},$$

or a function thereof, is the single invariant of this group.

32.2 Symmetry Groups of Differential Equations

Let $\mathcal{S}$ be a system of partial differential equations involving p independent variables $x = (x^1, \dots, x^p)$, and q dependent variables $u = (u^1, \dots, u^q)$. The solutions of the system are of the form $u = f(x)$, or, in component form, $u^\alpha = f^\alpha(x^1, \dots, x^p)$, $\alpha = 1, \dots, q$. Let $X = \mathbb{R}^p$ and $U = \mathbb{R}^q$ be the spaces of independent and dependent variables with coordinates $\{x^i\}$ and $\{u^\alpha\}$, respectively. Roughly speaking, a symmetry group of the system $\mathcal{S}$ will be a local group of transformations that map solutions of $\mathcal{S}$ into solutions of $\mathcal{S}$.

Historical Notes

Marius Sophus Lie (1842–1899) was the youngest son of a Lutheran pastor in Norway. He studied mathematics and science at Christiania (which became Kristiania, then Oslo in 1925) University where he attended Sylow's lectures on group theory. There followed a few years when he could not decide what career to follow. He tutored privately after his graduation and even dabbled a bit in astronomy and mechanics.

A turning point came in 1868 when he read papers on geometry by Poncelet and Plücker from which originated the inspiration in the topic of creating geometries by using elements other than points in space, and provided the seed for the rest of Lie's career, prompting him to call himself a student of Plücker, even though the two had never met.

Marius Sophus Lie
1842–1899

Lie's first publication won him a scholarship to work in Berlin, where he met Klein, who had also been influenced by Plücker's papers. The two had quite different styles—Lie always pursuing the broadest generalization, while Klein could become absorbed in a charming special case—but collaborated effectively for many years. However, in 1892 the lifelong friendship between Lie and Klein broke down, and the following year Lie publicly attacked Klein, saying, "I am no pupil of Klein, nor is the opposite the case, although this might be closer to the truth." Lie and Klein spent a summer in Paris, then parted for some time before resuming their collaboration in Germany. While in Paris, Lie discovered the *contact transformation*, which, for instance, maps lines into spheres. During the Franco-Prussian War, Lie decided to hike to Italy. On the way, however, he was arrested as a German spy and his mathematics notes were assumed to be coded messages. Only after the intervention of Darboux was Lie released, and he decided to return to Christiania. In 1871 Lie became an assistant at Christiania and obtained his doctorate.

After a short stay in Germany, he again returned to Christiania University, where a chair of mathematics was created for him. Several years later Lie succeeded Klein at Leipzig, where he was stricken with a condition, then called neurasthenia, resulting in fatigue and memory loss and once thought to result from exhaustion of the nervous system. Although treatment in a mental hospital nominally restored his health, the once robust and happy Lie became ill-tempered and suspicious, despite the recognition he received for his work. To lure him back to Norway, his friends at Christiania created another special chair for him, and Lie returned in the fall of 1898. He died of anemia a few months later. Lie had started examining partial differential equations, hoping that he could find a theory that was analogous to Galois's theory of equations. He examined his contact transformations considering how they affected a process due to Jacobi of generating further solutions from a given one. This led to combining the transformations in a way that Lie called a group, but is today called a *Lie algebra*. At this point he left his original intention of examining partial differential equations and examined Lie algebras. Killing was to examine Lie algebras quite independently of Lie, and Cartan was to publish the classification of semisimple Lie algebras in 1900. Much of the work on transformation groups for which Lie is best known was collected with the aid of a postdoctoral student sent to Christiania by Klein in 1884. The student, F. Engel, remained nine months with Lie and was instrumental in the production of the three volume work *Theorie der Transformationsgruppen*, which appeared between 1888 and 1893. A similar effort to collect Lie's work in contact transformations and partial differential equations was sidetracked as Lie's coworker, F. Hausdorff, pursued other topics.

The transformation groups now known as **Lie groups** provided a very fertile area for research for decades to come, although perhaps not at first. When Killing tried to classify the simple Lie groups, Lie considered his efforts so poor that he admonished one of his departing students with these words: "Farewell, and if ever you meet that s.o.b., kill him." Lie's work was continued (somewhat in isolation) by Cartan, but it was the papers of Weyl in the early 1920s that sparked the renewal of strong interest in Lie groups. Much of the foundation of the quantum theory of fundamental processes is built on Lie groups. In 1939, Wigner showed that application of Lie algebras to the Lorentz transformation required that all particles have the intrinsic properties of mass and spin.

To make precise the above statement, we have to clarify the meaning of the action of G on a function $u = f(x)$. We start with identifying the function f (i.e., a map) with its graph (see Chap. 1),

transform the *graph* of a function to find the function's transform!

$$\Gamma_f \equiv \left\{\bigl(x, f(x)\bigr) \mid x \in \Omega\right\} \subset X \times U,$$

where $\Omega \subset X$ is the domain of definition of f. If the action of $g \in G$ on Γ_f is defined, then the transform of Γ_f by g is

$$g \cdot \Gamma_f = \left\{(\tilde{x}, \tilde{u}) = g \cdot (x, u) \mid (x, u) \in \Gamma_f\right\}.$$

transform of a function by a group element

In general, $g \cdot \Gamma_f$ may not represent the graph of a function—in fact, it may not be even a function at all. However, by choosing g close to the identity of G and shrinking the size of Ω, we can ensure that $g \cdot \Gamma_f = \Gamma_{\tilde{f}}$, i.e., that $g \cdot \Gamma_f$ is indeed the graph of a function $\tilde{u} = \tilde{f}(\tilde{x})$. We write $\tilde{f} = g \cdot f$ and call $\tilde{f}$ the **transform** of f by g.

Example 32.2.1 Let $X = \mathbb{R} = U$, so that we are dealing with an ODE. Let $G = SO(2)$ be the rotation group acting on $X \times U = \mathbb{R}^2$. The action is given by

$$(\tilde{x}, \tilde{u}) = \theta \cdot (x, u) = (x \cos\theta - u \sin\theta, x \sin\theta + u \cos\theta). \tag{32.6}$$

If $u = f(x)$ is a function, the group $SO(2)$ acts on its graph Γ_f by rotating it. This process can lead to a rotated graph $\theta \cdot \Gamma_f$, which may not be the graph of a single-valued function. However, if we restrict the interval of definition of f, and make θ small enough, then $\theta \cdot \Gamma_f$ will be the graph of a well-defined function $\tilde{u} = \tilde{f}(\tilde{x})$ with $\Gamma_{\tilde{f}} = \theta \cdot \Gamma_f$. If we substitute $f(x)$ for u, we obtain

$$(\tilde{x}, \tilde{u}) = \theta \cdot \big(x, f(x)\big) = \big(x \cos\theta - f(x) \sin\theta, x \sin\theta + f(x) \cos\theta\big),$$

or

$$\begin{aligned} \tilde{x} &= x \cos\theta - f(x) \sin\theta, \\ \tilde{u} &= x \sin\theta + f(x) \cos\theta. \end{aligned} \tag{32.7}$$

Eliminating x from these two equations yields $\tilde{u}$ in terms of $\tilde{x}$, from which the function $\tilde{f}$ can be deduced.

As a specific example, consider $f(x) = kx^2$. Then, the first equation of (32.7) gives

$$(k \sin\theta)x^2 - \cos\theta x + \tilde{x} = 0 \quad \Rightarrow \quad x = \frac{\cos\theta - \sqrt{\cos^2\theta - 4k\tilde{x}\sin\theta}}{2k \sin\theta},$$

where we kept the root of the quadratic equation that gives a finite answer in the limit $\theta \to 0$. Inserting this in the second equation of (32.7) and simplifying yields

$$\tilde{u} = \tilde{f}(\tilde{x}) = \frac{\cos\theta - \sqrt{\cos^2\theta - 4k\tilde{x}\sin\theta}}{2k \sin^2\theta} - \tilde{x} \cot\theta.$$

We write this as

$$\tilde{f}(x) \equiv (\theta \cdot f)(x) = \frac{\cos\theta - \sqrt{\cos^2\theta - 4kx\sin\theta}}{2k \sin^2\theta} - x \cot\theta.$$

This equation defines the function $\tilde{f} = \theta \cdot f$.

The foregoing example illustrates the general procedure for finding the transformed function $\tilde{f} = g \cdot f$:

Box 32.2.2 *If the rule of transformation of $g \in G$ is given by*

$$(\tilde{x}, \tilde{u}) = g \cdot (x, u) = \big(\Psi_g(x, u), \Phi_g(x, u)\big),$$

then the graph $\Gamma_{\tilde{f}} = g \cdot \Gamma_f$ of $g \cdot f$ is given parametrically *by*

$$\tilde{x} = \Psi_g\big(x, f(x)\big), \qquad \tilde{u} = \Phi_g\big(x, f(x)\big). \tag{32.8}$$

In principle, we can solve the first equation for x in terms of $\tilde{x}$ and substitute in the second equation to find $\tilde{u}$ in terms $\tilde{x}$, and consequently $\tilde{f}$.

For some special but important cases, the transformed functions can be obtained explicitly. If G is **projectable**, i.e., if the action of G on x does not depend on u, then Eq. (32.8) takes the special form $\tilde{x} = \Psi_g(x)$ and $\tilde{u} = \Phi_g(x, f(x))$ in which Ψ_g is a diffeomorphism of X with inverse $\Psi_{g^{-1}}$. If Γ_f is the graph of a function f, then its transform $g \cdot \Gamma$ is *always* the graph of some function. In fact,

projectable group

$$\tilde{u} = \tilde{f}(\tilde{x}) = \Phi_g\big(x, f(x)\big) = \Phi_g\big(\Psi_{g^{-1}}(\tilde{x}), f\big(\Psi_{g^{-1}}(\tilde{x})\big)\big). \tag{32.9}$$

In particular, if G transforms *only the independent variables*, then

$$\tilde{u} = \tilde{f}(\tilde{x}) = f(x) = f\big(\Psi_{g^{-1}}(\tilde{x})\big) \quad \Rightarrow \quad \tilde{f} = f \circ \Psi_{g^{-1}}. \tag{32.10}$$

For example, if G is the group of translations $x \mapsto x + a$, then the transform of f will be defined by $\tilde{f}(x) = f(x - a)$.

Definition 32.2.3 A **symmetry group** of a system of DEs $\mathcal{S}$ is a local group of transformations G acting on an open subset M of $X \times U$ with the property that whenever $u = f(x)$ is a solution of $\mathcal{S}$ and $\tilde{f} \equiv g \cdot f$ is defined for $g \in G$, then $u = \tilde{f}(x)$ is also a solution of $\mathcal{S}$.

symmetry group of a system of DEs

The importance of knowing the symmetry group of a system of DEs lies in the property that from one solution we may be able to obtain a family of other solutions by applying the group elements to the given solution. To find such symmetry groups, we have to be able to "prolong" the action of a group to derivatives of the dependent variables as well. This is obvious because to test a symmetry, we have to substitute not only the transformed function $u = \tilde{f}(x)$, but also its derivatives in the DE to verify that it satisfies the DE.

32.2.1 Prolongation of Functions

Given a function $f : \mathbb{R}^p \supset X \to \mathbb{R}$, there are

$$p_k \equiv \binom{p + k - 1}{k}$$

different derivatives of order k of f. We use the multi-index notation

$$\partial_{J^{(k)}} f(x) \equiv \frac{\partial^k f(x)}{\partial x^{j_1} \partial x^{j_2} \cdots \partial x^{j_k}}$$

for these derivatives, where $J^{(k)} \equiv (j_1, \dots, j_k) \in \mathbb{N}^k$ is an *unordered* k-tuple of nonnegative integers with $1 \le j_k \le p$ (see also Sect. 21.1).[1] The *order* of the multi-index $J^{(k)}$, denoted by $|J^{(k)}|$, is the sum of its components and indicates the order of differentiation. So, in the derivative above, $|J^{(k)}| = j_1 + \cdots + j_k = k$. For a smooth map $f : X \to U$, we have $f(x) = (f^1(x), \dots, f^q(x))$, so that we need $q \cdot p_k$ numbers to represent all k-th order derivatives $\partial_{J^{(k)}} f^\alpha(x)$ of all components of f. We include the case of $k = 0$, in which case $\partial_{J^{(0)}} f^\alpha(x) = f^\alpha(x)$.

To geometrize the treatment of DEs (and thus facilitate the study of their invariance), we need to construct a space in which derivatives of all orders up to a certain number n participate. Since derivatives need functions to act on, we arrive at the space of functions whose derivatives share certain common properties. To be specific, me make the following definition.

n-equivalence of functions

Definition 32.2.4 Let f and h be functions defined on a neighborhood of a point $a \in X$ with values in U. We say that f and h are n-**equivalent** at a if $\partial_{J^{(k)}} f^\alpha(a) = \partial_{J^{(k)}} h^\alpha(a)$ for all α and $k = 0, 1, \dots, n$. The collection of all U-valued functions defined on a neighborhood of a will be denoted by $\Gamma_a(X \times U)$, and all functions n-equivalent to f at a by $j_a^n f$.

A convenient representative of such equivalent functions is the Taylor polynomial of order n (the terms in the Taylor series up to nth order) about a. Now collect all $j_a^n f$ for all a and f, and denote the result by $J^n(X \times U)$, so that

$$J^n(X \times U) \equiv \left\{ j_a^n f \mid \forall a \in X \text{ and } \forall f \in \Gamma_a(X \times U) \right\}. \tag{32.11}$$

nth jet space of U

$J^n(X \times U)$ is called the nth **prolongation** of U, or the nth **jet space** of U. It turns out that $J^n(X \times U)$ is a manifold (see [Saun 89, pp. 98 and 199]).

Theorem 32.2.5 *$J^n(X \times U)$ is a manifold with **natural coordinate functions** $(x^i, \{u^\alpha_{J^{(k)}}\}_{k=1}^n) \equiv (x^i, u^\alpha_J)$ defined by*

$$x^i\big(j_a^n f\big) = a^i, \qquad u^\alpha_{J^{(k)}}\big(j_a^n f\big) = \partial_{J^{(k)}} f^\alpha(a), \quad k = 0, 1, \dots, n.$$

Note that the "points" of $J^n(X \times U)$ are U-valued functions!

The natural coordinate functions allow us to identify the space of the derivatives with various powers of $\mathbb{R}$. Let $U_k \equiv \mathbb{R}^{qp_k}$ denote the set of coordinates $u^\alpha_{J^{(k)}}$, and let $U^{(n)} \equiv U \times U_1 \times \cdots \times U_n$ be the Cartesian product space[2] whose coordinates represent all the derivatives $u^\alpha_{J^{(k)}}$ of all orders

[1] We shall usually omit the superscript (k) in $J^{(k)}$ when it is understood that all orders of $J^{(k)}$ up to a certain given number are involved.

[2] Note that U, identified with the space of zeroth derivative, is a factor in $U^{(n)}$.

from 0 to n. The dimension of $U^{(n)}$ is

$$q + qp_1 + \cdots + qp_n = q\binom{p+n}{n} \equiv qp^{(n)}.$$

A typical point in $U^{(n)}$ is denoted by $u^{(n)}$, which has $qp^{(n)}$ different components $\{u^\alpha_{J^{(k)}}\}_{\alpha=1}^q$, where $J^{(k)}$ runs over all unordered multi-indices $J^{(k)} = (j_1, \ldots, j_k)$ with $1 \le j_k \le p$ and $0 \le k \le n$. The nth jet space $J^n(X \times U)$ can now be identified with $X \times U^{(n)}$. From now on, we shall use $X \times U^{(n)}$ in place of $J^n(X \times U)$.

Example 32.2.6 Let $p = 3$ and $u = 1$, i.e., $X = \mathbb{R}^3$ and $U = \mathbb{R}$. The coordinates of X are (x, y, z) and that of U is u. The coordinates of U_1 are (u_x, u_y, u_z), where the subscript denotes the variable of differentiation. Similarly, the coordinates of U_2 are

$$(u_{xx}, u_{xy}, u_{xz}, u_{yy}, u_{yz}, u_{zz})$$

and those of $U^{(2)} \equiv U \times U_1 \times U_2$ are

$$(u; u_x, u_y, u_z; u_{xx}, u_{xy}, u_{xz}, u_{yy}, u_{yz}, u_{zz}),$$

which shows that $U^{(2)}$ is 10-dimensional.

Definition 32.2.7 Given a smooth map $f : X \supset \Omega \to U$, we define a map $\mathrm{pr}^{(n)} f : \Omega \to U^{(n)}$ whose components $(\mathrm{pr}^{(n)} f)^\alpha_J$ are given by

prolongation of a function

$$\left(\mathrm{pr}^{(n)} f\right)^\alpha_J(x) \equiv \partial_J f^\alpha(x) \equiv \left(\left\{\partial_{J^{(k)}} f^\alpha(x)\right\}_{k=0}^n\right).$$

This map is called the nth **prolongation** of f.

Thus, for each $x \in X$, $\mathrm{pr}^{(n)} f(x)$ is a vector in $\mathbb{R}^{qp^{(n)}}$ whose components are the values of f and all its derivatives up to order n at the point x. For example, in the case of $p = 3$, $q = 1$ discussed above, $\mathrm{pr}^{(2)} f(x, y, z)$ has components

$$\left(f; \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}; \frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial x \partial z}, \frac{\partial^2 f}{\partial y^2}, \frac{\partial^2 f}{\partial y \partial z}, \frac{\partial^2 f}{\partial z^2}\right).$$

When the underlying space is an open subset M of $X \times U$, its corresponding jet space is

$$M^{(n)} \equiv M \times U_1 \times \cdots \times U_n,$$

which is a subspace of $X \times U^{(n)} \cong J^n(X \times U)$. If the graph of $f : X \to U$ lies in M, the graph of $\mathrm{pr}^{(n)} f$ lies in $M^{(n)}$.

Prolongation allows us to turn a system of DEs into a system of *algebraic* equations: Given a system of l DEs

$$\Delta_\nu\left(\left\{x^i\right\}, \left\{u^\alpha\right\}, \left\{\partial_i u^\alpha\right\}, \left\{\partial_i \partial_j u^\alpha\right\}, \ldots, \left\{\partial_{i_1} \ldots \partial_{i_n} u^\alpha\right\}\right) = 0, \quad \nu = 1, \ldots, l,$$

one can define a map $\Delta : M^{(n)} \to \mathbb{R}^l$ and identify the system of DEs with

$$\mathcal{S}_\Delta \equiv \left\{ \left(x, u^{(n)} \right) \in M^{(n)} \mid \Delta\left(x, u^{(n)} \right) = 0 \right\}.$$

By identifying the system of DEs with the subset $\mathcal{S}_\Delta$ of the jet space, we have translated the abstract relations among the derivatives of u into a geometrical object $\mathcal{S}_\Delta$, which is more amenable to symmetry operations.

Definition 32.2.8 Let Ω be a subset of X and $f : \Omega \to U$ a smooth map. Then f is called a **solution** of the system of DEs $\mathcal{S}_\Delta$ if

$$\Delta\left(x, \mathrm{pr}^{(n)} f(x) \right) = 0 \quad \forall x \in \Omega.$$

Just as we identified a function with its graph, we can identify the solution of a system of DEs with the graph of its prolongation $\mathrm{pr}^{(n)} f$. This graph, which is denoted by $\Gamma_f^{(n)}$, will clearly be a subset of $\mathcal{S}_\Delta$:

$$\Gamma_f^{(n)} \equiv \left\{ \left(x, \mathrm{pr}^{(n)} f(x) \right) \right\} \subset \mathcal{S}_\Delta.$$

Box 32.2.9 *An nth order system of differential equations is taken to be a subset $\mathcal{S}_\Delta$ of the jet space $J^n(X \times U)$, and a solution to be a smooth map $f : \Omega \to J^n(X \times U)$ the graph of whose nth prolongation $\mathrm{pr}^{(n)} f$ is contained in $\mathcal{S}_\Delta$.*

Example 32.2.10 Consider Laplace's equation

$$\nabla^2 u = u_{xx} + u_{yy} + u_{zz} = 0$$

with $p = 3$, $q = 1$, and $n = 2$. The total jet space is the 13-dimensional Euclidean space $X \times U^{(2)}$, whose coordinates are taken to be

$$(x, y, z; u; u_x, u_y, u_z; u_{xx}, u_{xy}, u_{xz}, u_{yy}, u_{yz}, u_{zz}).$$

In this 13-dimensional Euclidean space, Laplace's equation defines a 12-dimensional subspace $\mathcal{S}_\Delta$ consisting of all points in the jet space whose eighth, eleventh, and thirteenth coordinates add up to zero. A solution $f : \mathbb{R}^3 \supset \Omega \to U \subset \mathbb{R}$ must satisfy

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0 \quad \forall (x, y, z) \in \Omega.$$

This is the same as requiring the graph $\Gamma_f^{(2)}$ to lie in $\mathcal{S}_\Delta$. For example, if

$$f(\mathbf{r}) \equiv f(x, y, z) = x^3 yz - y^3 xz + y^3 - 3yz^2,$$

then, collecting each section of $\mathrm{pr}^{(2)} f$ with fixed α (see Definition 32.2.7) separately, we have

$$\left(\mathrm{pr}^{(2)} f \right)^1(\mathbf{r}) = x^3 yz - y^3 xz + y^3 - 3yz^2,$$

$$\left(\mathrm{pr}^{(2)} f\right)^2(\mathbf{r}) = \left(3x^2yz - y^3z, x^3z - 3y^2xz + 3y^2 - 3z^2, x^3y - y^3x - 6yz\right),$$

$$\left(\mathrm{pr}^{(2)} f\right)^3(\mathbf{r}) = \left(6xyz, 3x^2z - 3y^2z, 3x^2y - y^3,\right.$$
$$\left. - 6xyz + 6y, x^3 - 3xy^2 - 6z, -6y\right)$$

which lies in $\mathcal{S}_\Delta$ because the sum of the eighth, the eleventh, and the thirteenth coordinates of $(x, y, z; \mathrm{pr}^{(2)} f(x, y, z))$ is $6xyz - 6xyz + 6y - 6y = 0$.

32.2.2 Prolongation of Groups

Suppose G is a group of transformations acting on $M \subset X \times U$. It is possible to prolong this action to the n-jet space $M^{(n)}$. The resulting group that acts on $M^{(n)}$ is called the nth **prolongation** of G and denoted by $\mathrm{pr}^{(n)} G$ with group elements $\mathrm{pr}^{(n)} g$, for $g \in G$. This prolongation is defined naturally: The derivatives of a function f with respect to x are transformed into derivatives of $\tilde{f} = g \cdot f$ with respect to $\tilde{x}$. More precisely,

nth prolongation of a group action

$$\boxed{\mathrm{pr}^{(n)} g \cdot \left(j_x^n f\right) \equiv \left(\tilde{x}, \mathrm{pr}^{(n)} \tilde{f}(\tilde{x})\right) \equiv \left(\tilde{x}, \mathrm{pr}^{(n)} (g \cdot f)(\tilde{x})\right).} \tag{32.12}$$

For $n = 0$, Eq. (32.12) reduces to the action of G on M as given by Eq. (32.8). That the outcome of the action of the prolongation of G in the defining equation (32.12) is independent of the representative function f follows from the chain rule and the fact that only derivatives up to the nth order are involved. The following example illustrates this.

Example 32.2.11 Let X, U, and G be as in Example 32.2.1. In this case, we have

$$\mathrm{pr}^{(1)}\theta \cdot \left(j_x^1 f\right) \equiv \mathrm{pr}^{(1)}\theta \cdot (x, u, u_1) \equiv (\tilde{x}, \tilde{u}, \tilde{u}_1).$$

We calculated $\tilde{x}$ and $\tilde{u}$ in that example. They are

$$\begin{aligned} \tilde{x} &= x\cos\theta - u\sin\theta, \\ \tilde{u} &= x\sin\theta + u\cos\theta. \end{aligned} \tag{32.13}$$

To find $\tilde{u}_1$, we need to differentiate the second equation with respect to $\tilde{x}$ and express the result in terms of the original variables. Thus

$$\tilde{u}_1 \equiv \frac{d\tilde{u}}{d\tilde{x}} = \frac{d\tilde{u}}{dx}\frac{dx}{d\tilde{x}} = \left(\sin\theta + \frac{du}{dx}\cos\theta\right)\frac{dx}{d\tilde{x}} = (\sin\theta + u_1\cos\theta)\frac{dx}{d\tilde{x}};$$

$dx/d\tilde{x}$ is obtained by differentiating the first equation of (32.13):

$$1 = \frac{dx}{d\tilde{x}}\cos\theta - \frac{du}{d\tilde{x}}\sin\theta = \frac{dx}{d\tilde{x}}\cos\theta - \frac{du}{dx}\frac{dx}{d\tilde{x}}\sin\theta = (\cos\theta - u_1\sin\theta)\frac{dx}{d\tilde{x}},$$

or

$$\frac{dx}{d\tilde{x}} = \frac{1}{\cos\theta - u_1\sin\theta}.$$

It therefore follows that

$$\tilde{u}_1 = \frac{\sin\theta + u_1 \cos\theta}{\cos\theta - u_1 \sin\theta}$$

and

$$\mathrm{pr}^{(1)}\theta \cdot (x, u, u_1) = \left(x\cos\theta - u\sin\theta, x\sin\theta + u\cos\theta, \frac{\sin\theta + u_1\cos\theta}{\cos\theta - u_1\sin\theta} \right).$$

We note that the RHS involves derivatives up to order one. Therefore, the transformation is independent of the representative function. So, if we had chosen $j_x^1 h$ where h is 1-equivalent to f, we would have obtained the same result. This holds for derivatives of all orders. Therefore, the prolongation of the action of the group G is well-defined.

In many cases, it is convenient to choose the nth-order Taylor polynomial as the representative of the class of n-equivalent functions, and, if possible, write the transformed function $\tilde{f}$ explicitly in terms of $\tilde{x}$, and differentiate it to obtain the transformed derivatives (see Problem 32.3).

Example 32.2.11 illustrates an important property of the prolongation of G. We note that the first prolongation $\mathrm{pr}^{(1)}G$ acts on the original coordinates (x, u) in exactly the same way that G does. This holds in general:

Box 32.2.12 *The effect of the nth prolongation $\mathrm{pr}^{(n)}G$ to derivatives up to order $m \le n$ is exactly the same as the effect of $\mathrm{pr}^{(m)}G$. If we already know the action of the mth-order prolonged group $\mathrm{pr}^{(m)}G$, then to compute $\mathrm{pr}^{(n)}G$ we need only find how the derivatives u_J^α of order higher than m transform, because the lower-order action is already determined.*

32.2.3 Prolongation of Vector Fields

The geometrization of a system of DEs makes it possible to use the machinery of differentiable manifolds, Lie groups, and Lie algebras to unravel the symmetries of the system. At the heart of this machinery are the infinitesimal transformations, which are directly connected to vector fields. Therefore, it is necessary to find out how a vector field defined in $M \subset X \times U$ is prolonged. The most natural way to prolong a vector field is to prolong its integral curve—which is a one-parameter group of transformations of M—to a curve in $M^{(n)}$ and then calculate the tangent to the latter curve.

nth prolongation of a vector field

Definition 32.2.13 Let M be an open subset of $X \times U$ and $\mathbf{X} \in \mathcal{X}(M)$. The nth **prolongation** of $\mathbf{X}$, denoted by $\mathrm{pr}^{(n)}\mathbf{X}$, is a vector field on the nth jet space $M^{(n)}$ defined by

$$\text{pr}^{(n)}\mathbf{X}\big|_{(x,u^{(n)})} = \frac{d}{dt}\text{pr}^{(n)}\big[\exp(t\mathbf{X})\big]\cdot\big(x,u^{(n)}\big)\big|_{t=0}$$

for any $(x, u^{(n)}) \in M^{(n)}$.

Since $(x, u^{(n)}) \in M^{(n)}$ form a coordinate system on $M^{(n)}$, any vector field in $M^{(n)}$ can be written as a linear combination of $\partial/\partial x^i$ and $\partial/\partial u_J^\alpha$ with coefficients being, in general, functions of *all* coordinates x^i and u_J^α. For a prolonged vector, however, we have

$$\text{pr}^{(n)}\mathbf{X} = \sum_{i=1}^{p} X^i \frac{\partial}{\partial x^i} + \sum_{\alpha=1}^{q}\sum_J X_J^\alpha \frac{\partial}{\partial u_J^\alpha}, \tag{32.14}$$

where X^i and X_0^α are functions only of x^k and u. This is due to the remark made in Box 32.2.12. For the same reason, the coefficients $X_{J^{(m)}}^\alpha$ corresponding to derivatives of order m will be independent of coordinates $u_{J^{(k)}}^\alpha$ for $k > m$. Thus, it is possible to construct various prolongations of a given vector field recursively.

Example 32.2.14 Let us consider our recurrent example of $X \cong U \cong \mathbb{R}$, $G = SO(2)$. Given the infinitesimal generator $\mathbf{v} = -u\partial_x + x\partial_u$, one can solve the DE of its integral curve to obtain[3]

$$\exp(t\mathbf{v})(x,u) = (x\cos t - u\sin t, x\sin t + u\cos t).$$

Example 32.2.11 calculated the first prolongation of $SO(2)$. So

$$\begin{aligned}&\text{pr}^{(1)}\exp(t\mathbf{v})\cdot(x,u,u_1)\\ &\quad= \left(x\cos t - u\sin t, x\sin t + u\cos t, \frac{\sin t + u_1\cos t}{\cos t - u_1\sin t}\right).\end{aligned}$$

Differentiating the components with respect to t at $t = 0$ gives

$$\begin{aligned}\frac{\partial}{\partial t}(x\cos t - u\sin t)\bigg|_{t=0} &= -u,\\ \frac{\partial}{\partial t}(x\sin t + u\cos t)\bigg|_{t=0} &= x,\\ \frac{\partial}{\partial t}\left(\frac{\sin t + u_1\cos t}{\cos t - u_1\sin t}\right)\bigg|_{t=0} &= 1 + u_1^2.\end{aligned}$$

Therefore,

$$\text{pr}^{(1)}\mathbf{v} = -u\frac{\partial}{\partial x} + x\frac{\partial}{\partial u} + \big(1+u_1^2\big)\frac{\partial}{\partial u_1}.$$

Note that the first two terms in $\text{pr}^{(1)}\mathbf{v}$ are the same as those in $\mathbf{v}$ itself, in agreement with Box 32.2.12.

[3] One can, of course, also write the finite group element directly.

32.3 The Central Theorems

We are now in a position to state the first central theorem of the application of Lie groups to the solution of DEs. This theorem is the exact replica of Theorem 32.1.6 in the language of prolongations:

Theorem 32.3.1 *Let* $\{\Delta_\nu(x, u^{(n)}) = 0\}_{\nu=1}^{l}$ *be a system of DEs defined on* $M \subset X \times U$ *whose Jacobian matrix*

$$\mathsf{J}_\Delta\big(x, u^{(n)}\big) \equiv \left(\frac{\partial \Delta_\nu}{\partial x^i}, \frac{\partial \Delta_\nu}{\partial u_J^\alpha}\right)$$

has rank l *for all* $(x, u^{(n)}) \in \mathcal{S}_\Delta$. *If* G *is a local Lie group of transformations acting on* M, *and*

$$\mathsf{pr}^{(n)}\boldsymbol{\xi}_M\big|_{(x,u^{(n)})}\Delta_\nu = 0, \quad \nu = 1, \dots, l, \text{ whenever } \Delta\big(x, u^{(n)}\big) = 0$$

for every infinitesimal generator $\boldsymbol{\xi}$ *of* G, *then* G *is the symmetry group of the system.*

Example 32.3.2 Consider the first order (so, $n = 1$) ordinary DE

$$\Delta(x, u, u_1) = \big(u - x^3 - u^2 x\big)u_1 + x + x^2 u + u^3 = 0,$$

so that $X \cong \mathbb{R}$ and $U \cong \mathbb{R}$. We first note that

$$\begin{aligned}\mathsf{J}_\Delta &= \left(\frac{\partial \Delta}{\partial x}, \frac{\partial \Delta}{\partial u}, \frac{\partial \Delta}{\partial u_1}\right)\\ &= \big(\big(-3x^2 - u^2\big)u_1 + 1 + 2xu, (1 - 2ux)u_1 + x^2 + 3u^2, u - x^3 - u^2 x\big),\end{aligned}$$

which is of rank 1 everywhere. Now let us apply the first prolongation of the generator of $SO(2)$—calculated in Example 32.2.14—to Δ. We have

$$\begin{aligned}\mathsf{pr}^{(1)}\mathbf{v}(\Delta) &= -u\frac{\partial \Delta}{\partial x} + x\frac{\partial \Delta}{\partial u} + \big(1 + u_1^2\big)\frac{\partial \Delta}{\partial u_1}\\ &= -u\big[\big(-3x^2 - u^2\big)u_1 + 1 + 2xu\big] + x\big[(1 - 2ux)u_1 + x^2 + 3u^2\big]\\ &\quad + \big(1 + u_1^2\big)\big(u - x^3 - u^2 x\big)\\ &= u_1\big[\big(u - x^3 - u^2 x\big)u_1 + x + x^2 u + u^3\big] = u_1\Delta.\end{aligned}$$

It follows that $\mathsf{pr}^{(1)}\mathbf{v}(\Delta) = 0$ whenever $\Delta = 0$, and that $SO(2)$ is a symmetry group of the DE. Thus, rotations will change solutions of the DE into other solutions. In fact, the reader may verify that in polar coordinates, the DE can be written in the incredibly simple form

$$\frac{dr}{d\theta} = r^3,$$

and the symmetry of the DE conveys the fact that adding a constant to θ does not change the polar form of the DE.

Theorem 32.3.1 reduces the invariance of a system of DEs to a criterion involving the prolongation of the infinitesimal generators of the symmetry group. The urgent task in front of us is therefore to construct an explicit formula for the prolongation of a vector field. In order to gain insight into this construction, we first look at the simpler case of $U \cong \mathbb{R}$ and a group G that transforms only the independent variables. Furthermore, we restrict ourselves to the first prolongation. An infinitesimal generator of such a group will be of the form

$$\mathbf{v} = \sum_{i=1}^{p} X^i(x) \frac{\partial}{\partial x^i},$$

which is assumed to act on the space $M \subset X \times U$. The integral curve of this vector field is $\exp(t\mathbf{v})$ which acts on points of M as follows:

$$(\tilde{x}, \tilde{u}) = \exp(t\mathbf{v}) \cdot (x, u) \equiv \big(\Psi_t(x), u\big) \equiv \big(\Psi_t^i(x), u\big).$$

By the construction of the integral curves in general, we have

$$\left.\frac{d\Psi_t^i(x)}{dt}\right|_{t=0} = X^i(x). \tag{32.15}$$

Denote the coordinates of the first jet space $M^{(1)}$ by (x^i, u, u_k), where $u_k(j_x^1 f) \equiv \partial f/\partial x^k$. By the definition of the action of the prolonged group,

$$\mathrm{pr}^{(1)} \exp(t\mathbf{v}) \cdot \big(j_x^1 f\big) = \big(\Psi_t(x), u, \tilde{u}_j\big), \tag{32.16}$$

where $u = f(x)$ and $\tilde{u}_j = \partial \tilde{f}/\partial \tilde{x}^j$. Once we find $\tilde{u}_j$, we can differentiate Eq. (32.16) with respect to t at $t = 0$ to obtain the prolonged vector field. Using Eq. (32.10) and commas to indicate differentiation,[4] we obtain

$$\begin{aligned}\tilde{u}_j = \tilde{f}_{,j}(\tilde{x}) = (f \circ \Psi_{-t})_{,j}(\tilde{x}) &= \sum_{i=1}^{p} f_{,i}\big(\underbrace{\Psi_{-t}(\tilde{x})}_{=x}\big)\Psi_{-t,j}^i(\tilde{x}) \\ &= \sum_{i=1}^{p} u_i \Psi_{-t,j}^i(\tilde{x}).\end{aligned}$$

[4]We have found it exceedingly convenient to use commas to indicate differentiation with respect to the x's. The alternative, i.e., the use of partials, makes it almost impossible to find one's way in the maze of derivatives involving x^i, $\tilde{x}^j$, and t, with $\tilde{x}^j$ depending on t. The reader will recall that the index after the comma is to be thought of as a "position holder", and the argument as a substitution. Thus, for example,

$$\tilde{f}_{,j}(\tilde{x}) \equiv \left.\frac{\partial \tilde{f}(r_1, \ldots, r_p)}{\partial r_j}\right|_{r=\tilde{x}} \equiv \left.\frac{\partial \tilde{f}(s_1, \ldots, s_p)}{\partial s_j}\right|_{s=\tilde{x}} \equiv \left.\frac{\partial \tilde{f}(\heartsuit_1, \ldots, \heartsuit_p)}{\partial \heartsuit_j}\right|_{\heartsuit=\tilde{x}}.$$

Since $\mathbf{v}$ does not have any component along $\partial/\partial u$, its prolongation will not have such components either. The components U_j along $\partial/\partial u_j$ are obtained by differentiating $\tilde{u}_j$ with respect to t:

$$\begin{aligned} U_j(x,u,u_j) &= \sum_{i=1}^{p} u_i \frac{d}{dt}\Psi^i_{-t,j}(\tilde{x})\bigg|_{t=0} \\ &= \sum_{i=1}^{p} u_i \frac{\partial \Psi^i_{-t,j}}{\partial t}(\tilde{x})\bigg|_{t=0} + \sum_{i=1}^{p} u_i \sum_{k=1}^{p}\left(\Psi^i_{-t,jk}(\tilde{x})\frac{\partial \tilde{x}^k}{\partial t}\right)\bigg|_{t=0}. \end{aligned} \tag{32.17}$$

The derivative of the first term in the sum can be evaluated as follows:

$$\begin{aligned} \frac{\partial \Psi^i_{-t,j}}{\partial t}\big(\tilde{x}(t)\big)\bigg|_{t=0} &= -\frac{\partial \Psi^i_{s,j}}{\partial s}\big(\tilde{x}(-s)\big)\bigg|_{s=0} \equiv -\frac{\partial \Psi^i_{,j}}{\partial s}\big(s,\tilde{x}(-s)\big)\bigg|_{s=0} \\ &= -\frac{\partial \Psi^i_{,j}}{\partial s}\big(s,\tilde{x}(0)\big)\bigg|_{s=0} = -\frac{\partial \Psi^i_{,j}}{\partial s}(s,x)\bigg|_{s=0} \\ &= -\frac{\partial \Psi^i_{s,j}}{\partial s}(x)\bigg|_{s=0} = -\left(\frac{\partial \Psi^i_s}{\partial s}\bigg|_{s=0}\right)_{,j}(x) = -X^i_{,j}(x), \end{aligned}$$

where we have emphasized the dependence of $\tilde{x}$ on t (or s), treated s as the first independent variable, and in the second line substituted $s=0$ in all $\tilde{x}$'s before differentiation. This is possible because we are taking the *partial* derivative with respect to the first variable holding all others constant. The derivative of the second term in Eq. (32.17) can be calculated similarly:

$$\Psi^i_{-t,jk}\big(\tilde{x}(t)\big)\big|_{t=0} = \Psi^i_{0,jk}\big(\tilde{x}(0)\big) = \frac{\partial^2 x^i}{\partial x^j \partial x^k} = 0$$

because $\Psi^i_0 = x^i$. We therefore have

$$U_j(x,u,u_k) = -\sum_{i=1}^{p} \frac{\partial X^i}{\partial x^j} u_i \tag{32.18}$$

and

$$\mathrm{pr}^{(1)}\mathbf{v} = \sum_{i=1}^{p}\left(X^i \frac{\partial}{\partial x^i} + U_i \frac{\partial}{\partial u_i}\right) = \mathbf{v} - \sum_{i,k=1}^{p} \frac{\partial X^k}{\partial x^i} u_k \frac{\partial}{\partial u_i}. \tag{32.19}$$

It is also instructive to consider the case in which U is still $\mathbb{R}$, but G acts only on the dependent variable. Then $\mathbf{v} = U(x,u)\partial_u$, and

$$(\tilde{x},\tilde{u}) = \big(x, \Phi_t(x,u)\big), \quad \text{with} \quad U(x,u) = \frac{d\Phi_t(x,u)}{dt}\bigg|_{t=0}.$$

The reader may check that in this case, the prolongation of $\mathbf{v}$ is given by

$$\mathrm{pr}^{(1)}\mathbf{v} = \mathbf{v} + \sum_{j=1}^{p} U_j\big(x, u^{(1)}\big)\frac{\partial}{\partial u_j}, \qquad U_j\big(x, u^{(1)}\big) = \frac{\partial U}{\partial x^j} + u_j\frac{\partial U}{\partial u}. \tag{32.20}$$

The second equation in (32.20) can also be written as

$$U_j\big(x, \mathrm{pr}^{(1)} f(x)\big) = \frac{\partial U}{\partial x^j} + \frac{\partial U}{\partial u}\frac{\partial f}{\partial x^j} = \frac{\partial}{\partial x^j}\big[U\big(x, f(x)\big)\big].$$

In other words, $U_j(x, u^{(1)})$ is obtained from $U(x, u)$ by differentiation with respect to x^j, while treating u as a function of x. This leads us to the definition of the total derivative.

Definition 32.3.3 Let $S : M^{(n)} \to \mathbb{R}$ be a smooth function of x, u, and all derivatives of u up to nth order. The **total derivative** of S with respect to x^i, denoted by $D_i S$, is a smooth function $D_i S : M^{(n+1)} \to \mathbb{R}$ defined by

total derivative

$$D_i S\big(j_x^{n+1} f\big) = \frac{\partial}{\partial x^i}\big[S\big(x, \mathrm{pr}^{(n)} f(x)\big)\big];$$

i.e., $D_i S$ is obtained from S by differentiating S with respect to x^i, treating u and all the u_J^α's as functions of x.

The following proposition, whose proof is a straightforward application of the chain rule, gives the explicit formula for calculating the total derivative:

Proposition 32.3.4 *The ith total derivative of $S : M^{(n)} \to \mathbb{R}$ is of the form*

$$D_i S = \frac{\partial S}{\partial x^i} + \sum_{\alpha=1}^{q}\sum_J u_{J,i}^\alpha \frac{\partial S}{\partial u_J^\alpha}$$

where, for $J = (j_1, \ldots, j_k)$,

$$u_{J,i}^\alpha \equiv \frac{\partial u_J^\alpha}{\partial x^i} = \frac{\partial^{k+1} u^\alpha}{\partial x^i \partial x^{j_1} \ldots \partial x^{j_k}}$$

and the sum over J includes derivatives of all orders from 0 *to n.*

An immediate consequence of this proposition is

$$\begin{aligned} D_i u_J^\alpha &= u_{J,i}^\alpha = \frac{\partial u_J^\alpha}{\partial x^i} \quad \forall J, \alpha, \\ D_i(ST) &= T D_i S + S D_i T. \end{aligned} \tag{32.21}$$

Higher-order total derivatives are defined in analogy with partial derivatives: If I is a multi-index of the form $I = (i_1, \ldots, i_k)$, then the Ith total derivative is

$$D_I S = D_{i_1} D_{i_2} \cdots D_{i_k} S. \tag{32.22}$$

As in the case of partial derivatives, the order of differentiation is immaterial.

We are now ready to state the second central theorem of the application of Lie groups to the solution of DEs (for a proof, see [Olve 86, pp. 113–115]).

Theorem 32.3.5 *Let*

$$\mathbf{v} = \sum_{i=1}^{p} X^i(x,u)\frac{\partial}{\partial x^i} + \sum_{\alpha=1}^{q} U^\alpha(x,u)\frac{\partial}{\partial u^\alpha}$$

be a vector field on an open subset $M \subset X \times U$. *The nth prolongation of* $\mathbf{v}$, *i.e.,* $\mathrm{pr}^{(n)}\mathbf{v} \in \mathfrak{X}(M^{(n)})$, *is*

$$\mathrm{pr}^{(n)}\mathbf{v} = \mathbf{v} + \sum_{\alpha=1}^{q}\sum_{J} U_J^\alpha\big(x, u^{(n)}\big)\frac{\partial}{\partial u_J^\alpha},$$

where for $J = (j_1, \ldots, j_k)$, *the inner sum extends over* $1 \le |J| \le n$ *and the coefficients* U_J^α *are given by*

$$U_J^\alpha\big(x, u^{(n)}\big) = D_J\left(U^\alpha - \sum_{i=1}^{p} X^i \frac{\partial u^\alpha}{\partial x^i}\right) + \sum_{i=1}^{p} X^i \frac{\partial u_J^\alpha}{\partial x^i}$$

and the higher-order derivative D_J *is as given in Eq.* (32.22).

Example 32.3.6 Let $p = 2$, $q = 1$, and consider the case in which G acts only on the independent variables (x, y). The general vector field for this situation is

$$\mathbf{v} = \xi(x,y)\frac{\partial}{\partial x} + \eta(x,y)\frac{\partial}{\partial y}.$$

We are interested in the first prolongation of this vector field. Thus, $n = 1$, $X^1 = \xi$, $X^2 = \eta$, and J has only one component, which we denote by j (also written as x or y). Theorem 32.3.5 gives

$$U_J^\alpha \equiv U_j = -D_j \sum_{i=1}^{2} X^i \frac{\partial u}{\partial x^i} + \sum_{i=1}^{2} X^i \frac{\partial u_j}{\partial x^i} = -\sum_{i=1}^{2} \frac{\partial X^i}{\partial x^j}\frac{\partial u}{\partial x^i},$$

and using the notation $u_x = \partial u/\partial x$ and $u_y = \partial u/\partial y$, we obtain

$$\mathrm{pr}^{(1)}\mathbf{v} = \mathbf{v} + U_x \frac{\partial}{\partial u_x} + U_y \frac{\partial}{\partial u_y}, \tag{32.23}$$

where

$$U_x = -u_x \frac{\partial \xi}{\partial x} - u_y \frac{\partial \eta}{\partial x} \quad \text{and} \quad U_y = -u_x \frac{\partial \xi}{\partial y} - u_y \frac{\partial \eta}{\partial y}.$$

In particular, if $G = SO(2)$, so that $\mathbf{v} = -y\partial_x + x\partial_y$, then $U_x = -u_y$ and $U_y = u_x$. It then follows that

$$\mathrm{pr}^{(1)}\mathbf{v} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} - u_y\frac{\partial}{\partial u_x} + u_x\frac{\partial}{\partial u_y}.$$

Example 32.3.7 Let $p=1, q=1$, and $G=SO(2)$. The general vector field for this situation is

$$\mathbf{v}=-u\frac{\partial}{\partial x}+x\frac{\partial}{\partial u}.$$

For the first prolongation of this vector field $n=1$, $X^1=-u$, and J has only one component, which we denote by x. Theorem 32.3.5 gives

$$U_J^\alpha \equiv U_x = D_x\left(x-X^1u_x\right)+X^1u_{xx}=1-\frac{\partial X^1}{\partial x}u_x=1+u_x^2.$$

It follows that

$$\mathrm{pr}^{(1)}\mathbf{v}=-u\frac{\partial}{\partial x}+x\frac{\partial}{\partial u}+\left(1+u_x^2\right)\frac{\partial}{\partial u_x},$$

which is the result obtained in Example 32.2.14.

The second prolongation can be obtained as well. Once again we use Theorem 32.3.5 with obvious change of notation:

$$\begin{aligned}U_{xx}&=D_xD_x\left(x-X^1u_x\right)+X^1u_{xxx}=D_x\left(1+u_x^2+uu_{xx}\right)-uu_{xxx}\\&=3u_xu_{xx}.\end{aligned}$$

Then

$$\mathrm{pr}^{(2)}\mathbf{v}=-u\frac{\partial}{\partial x}+x\frac{\partial}{\partial u}+\left(1+u_x^2\right)\frac{\partial}{\partial u_x}+3u_xu_{xx}\frac{\partial}{\partial u_{xx}}.$$

Using Theorem 32.3.1, we note that the DE $u_{xx}=0$ has $SO(2)$ as a symmetry group, because with $\Delta(x,u,u_x,u_{xx})\equiv u_{xx}$,

$$\mathrm{pr}^{(2)}\mathbf{v}(\Delta)=-u\frac{\partial\Delta}{\partial x}+x\frac{\partial\Delta}{\partial u}+\left(1+u_x^2\right)\frac{\partial\Delta}{\partial u_x}+3u_xu_{xx}\frac{\partial\Delta}{\partial u_{xx}}=3u_xu_{xx},$$

which vanishes whenever $\Delta(x,u,u_x,u_{xx})$ vanishes. This is the statement that rotations take straight lines to straight lines.

32.4 Application to Some Known PDEs

We have all the tools at our disposal to compute (in principle) the most general symmetry group of almost any system of PDEs. The coefficients U_J^α of the prolonged vector field $\mathrm{pr}^{(n)}\mathbf{v}$ will be functions of the partial derivatives of the coefficients X^i and U^α of $\mathbf{v}$ with respect to both x and u. The infinitesimal criterion of invariance as given in Theorem 32.3.1 will involve x, u, and the derivatives of u with respect to x, as well as X^i and U^α and their partial derivatives with respect to x and u. Using the system of PDEs, we can obtain some of the derivatives of the u's in terms of the others. Substituting these relations in the equation of infinitesimal criterion, we get an equation involving u's and powers of its derivatives that are to be treated as independent. We then equate the coefficients of these powers of partial derivatives of u to zero. This will result in a large number of elementary PDEs for the defining equations for the symmetry group of a system of PDEs

coefficient functions X^i and U^α of $\mathbf{v}$, called the **defining equations** for the symmetry group of the given system of PDEs. In most applications, these defining equations can be solved, and the general solution will determine the most general infinitesimal symmetry of the system. The symmetry group itself can then be calculated by exponentiation of the vector fields, i.e., by finding their integral curves. In the remaining part of this section, we construct the symmetry groups of the heat and the wave equations.

32.4.1 The Heat Equation

The one-dimensional heat equation $u_t = u_{xx}$ corresponds to $p = 2$, $q = 1$, and $n = 2$. So it is determined by the vanishing of $\Delta(x, t, u^{(2)}) = u_t - u_{xx}$. The most general infinitesimal generator of symmetry appropriate for this equation can be written as

$$\mathbf{v} = \xi(x,t,u)\frac{\partial}{\partial x} + \tau(x,t,u)\frac{\partial}{\partial t} + \phi(x,t,u)\frac{\partial}{\partial u}, \tag{32.24}$$

which, as the reader may check (see Problem 32.11), has a second prolongation of the form

$$\mathrm{pr}^{(2)}\mathbf{v} = \mathbf{v} + \phi^x\frac{\partial}{\partial u_x} + \phi^t\frac{\partial}{\partial u_t} + \phi^{xx}\frac{\partial}{\partial u_{xx}} + \phi^{xt}\frac{\partial}{\partial u_{xt}} + \phi^{tt}\frac{\partial}{\partial u_{tt}},$$

where, for example,

$$\begin{aligned}
\phi^t &= \phi_t - \xi_t u_x + (\phi_u - \tau_t)u_t - \xi_u u_x u_t - \tau_u u_t^2 \\
\phi^{xx} &= \phi_{xx} + (2\phi_{xu} - \xi_{xx})u_x - \tau_{xx}u_t + (\phi_{uu} - 2\xi_{xu})u_x^2 \\
&\quad - 2\tau_{xu}u_x u_t - \xi_{uu}u_x^3 - \tau_{uu}u_x^2 u_t + (\phi_u - 2\xi_x)u_{xx} \\
&\quad - 2\tau_x u_{xt} - 3\xi_u u_x u_{xx} - \tau_u u_t u_{xx} - 2\tau_u u_x u_{xt},
\end{aligned} \tag{32.25}$$

and subscripts indicate partial derivatives. Theorem 32.3.1 now gives

$$\mathrm{pr}^{(2)}\mathbf{v}(\Delta) = \phi^t - \phi^{xx} = 0 \quad \text{whenever } u_t = u_{xx} \tag{32.26}$$

as the infinitesimal criterion. Substituting (32.25) in (32.26), replacing u_t with u_{xx} in the resulting equation, and equating to zero the coefficients of the monomials in derivatives of u, we obtain a number of equations involving ξ, τ, and ϕ. These equations as well as the monomials of which they are coefficients are given in Table 32.1. Complicated as the defining equations may look, they are fairly easy to solve. From (d) and (f) we conclude that τ is a function of t only. Then (c) shows that ξ is independent of u, and (e) gives $2\xi_x = \tau_t$, or $\xi(x,t) = \frac{1}{2}\tau_t x + \eta(t)$, for some arbitrary function η. From (h) and the fact that ξ is independent of u we get $\phi_{uu} = 0$, or

$$\phi(x,t,u) = \alpha(x,t)u + \beta(x,t)$$

Table 32.1 The defining equations of the heat equation and the monomials that give rise to them

Monomial	Coefficient equation	
u_{xx}^2	$0=0$	(a)
$u_x^2 u_{xx}$	$\tau_{uu}=0$	(b)
$u_x u_{xx}$	$2\xi_u+2\tau_{xu}=0$	(c)
$u_x u_{xt}$	$2\tau_u=0$	(d)
u_{xx}	$2\xi_x+\tau_{xx}-\tau_t=0$	(e)
u_{xt}	$2\tau_x=0$	(f)
u_x^3	$\xi_{uu}=0$	(g)
u_x^2	$2\xi_{xu}-\phi_{uu}=0$	(h)
u_x	$\xi_{xx}-2\phi_{xu}-\xi_t=0$	(i)
1	$\phi_t-\phi_{xx}=0$	(j)

for some as yet undetermined functions α and β. Since ξ is linear in x, $\xi_{xx}=0$, and (i) yields $\xi_t=-2\phi_{xu}=-2\alpha_x$, or

$$\frac{\partial\alpha}{\partial x}=-\frac{1}{2}\xi_t=-\frac{1}{4}\tau_{tt}x-\frac{1}{2}\eta_t \quad\Rightarrow\quad \alpha(x,t)=-\frac{1}{8}\tau_{tt}x^2-\frac{1}{2}\eta_t x+\rho(t).$$

Finally, with $\phi_t=\alpha_t u+\beta_t$ and $\phi_{xx}=\alpha_{xx}u+\beta_{xx}$ (recall that when taking partial derivatives, u is considered independent of x and t), the last defining equation (j) gives $\alpha_t=\alpha_{xx}$ and $\beta_t=\beta_{xx}$, i.e., that α and β are to satisfy the heat equation. Substituting α in the heat equation, we obtain

$$-\frac{1}{8}\tau_{ttt}x^2-\frac{1}{2}\eta_{tt}x+\rho_t=-\frac{1}{4}\tau_{tt},$$

which must hold for all x and t. Therefore,

$$\tau_{ttt}=0,\qquad \eta_{tt}=0,\qquad \rho_t=-\frac{1}{4}\tau_{tt}.$$

These equations have the solution

$$\tau=c_1t^2+c_2t+c_3,\qquad \rho=-\frac{1}{2}c_1t+c_4,\qquad \eta=c_5t+c_6.$$

It follows that $\alpha(x,t)=-\frac{1}{4}c_1x^2-\frac{1}{2}c_5x-\frac{1}{2}c_1t+c_4$ and

$$\begin{aligned}\xi(x,t)&=\frac{1}{2}(2c_1t+c_2)x+c_5t+c_6,\\ \tau(t)&=c_1t^2+c_2t+c_3,\\ \phi(x,t,u)&=\left(-\frac{1}{4}c_1x^2-\frac{1}{2}c_5x-\frac{1}{2}c_1t+c_4\right)u+\beta(x,t).\end{aligned}$$

Inserting in Eq. (32.24) yields

$$\begin{aligned}\mathbf{v} &= \left[\frac{1}{2}(2c_1t+c_2)x+c_5t+c_6\right]\frac{\partial}{\partial x}+\left(c_1t^2+c_2t+c_3\right)\frac{\partial}{\partial t}\\ &\quad+\left[\left(-\frac{1}{4}c_1x^2-\frac{1}{2}c_5x-\frac{1}{2}c_1t+c_4\right)u+\beta(x,t)\right]\frac{\partial}{\partial u}\\ &= c_1\left[xt\frac{\partial}{\partial x}+t^2\frac{\partial}{\partial t}-\frac{1}{4}(2t+x^2)u\frac{\partial}{\partial u}\right]+c_2\left(\frac{1}{2}x\frac{\partial}{\partial x}+t\frac{\partial}{\partial t}\right)\\ &\quad+c_3\frac{\partial}{\partial t}+c_4u\frac{\partial}{\partial u}+c_5\left(t\frac{\partial}{\partial x}-\frac{1}{2}xu\frac{\partial}{\partial u}\right)+c_6\frac{\partial}{\partial x}+\beta(x,t)\frac{\partial}{\partial u}.\end{aligned}$$

Thus the Lie algebra of the infinitesimal symmetries of the heat equation is spanned by the six vector fields

$$\begin{aligned}&\mathbf{v}_1=\partial_x, \qquad \mathbf{v}_2=\partial_t, \qquad \mathbf{v}_3=u\partial_u, \qquad \mathbf{v}_4=x\partial_x+2t\partial_t,\\ &\mathbf{v}_5=2t\partial_x-xu\partial_u, \qquad \mathbf{v}_6=4tx\partial_x+4t^2\partial_t-\left(x^2+2t\right)u\partial_u\end{aligned} \tag{32.27}$$

and the infinite-dimensional subalgebra

$$\mathbf{v}_\beta=\beta(x,t)\partial_u,$$

where β is an arbitrary solution of the heat equation.

The one-parameter groups G_i generated by the $\mathbf{v}_i$ can be found by solving the appropriate DEs for the integral curves. We show a sample calculation and leave the rest of the computation to the reader. Consider $\mathbf{v}_5$, whose integral curve is given by the set of DEs

$$\frac{dx}{ds}=2t, \qquad \frac{dt}{ds}=0, \qquad \frac{du}{ds}=-xu.$$

The second equation shows that t is not affected by the group. So, $t=t_0$, where t_0 is the initial value of t. The first equation now gives

$$\frac{dx}{ds}=2t_0 \quad\Rightarrow\quad x=2t_0s+x_0,$$

and the last equation yields

$$\frac{du}{ds}=-(2t_0s+x_0)u \quad\Rightarrow\quad \frac{du}{u}=-(2t_0s+x_0)ds \quad\Rightarrow\quad u=u_0e^{-t_0s^2-x_0s}.$$

Changing the transformed coordinates to $\tilde{x}^i$ and removing the subscript from the initial coordinates, we can write

$$\exp(\mathbf{v}_5s)\cdot(x,t,u)=(\tilde{x},\tilde{t},\tilde{u})=\left(x+2ts,t,ue^{-sx-s^2t}\right).$$

Table 32.2 gives the result of the action of $\exp(\mathbf{v}_is)$ to (x,t,u).

The symmetry groups G_1 and G_2 reflect the invariance of the heat equation under space and time translations. G_3 and G_β demonstrate the linearity of the heat equation: We can multiply solutions by constants and add solutions to get new solutions. The scaling symmetry is contained in G_4, which shows that if you scale time by the square of the scaling of x, you obtain

Table 32.2 The transformations caused by the symmetry group of the heat equation

Group element	Transformed coordinates $(\tilde{x}, \tilde{t}, \tilde{u})$
$G_1 = \exp(\mathbf{v}_1 s)$	$(x + s, t, u)$
$G_2 = \exp(\mathbf{v}_2 s)$	$(x, t + s, u)$
$G_3 = \exp(\mathbf{v}_3 s)$	$(x, t, e^s u)$
$G_4 = \exp(\mathbf{v}_4 s)$	$(e^s x, e^{2s} t, u)$
$G_5 = \exp(\mathbf{v}_5 s)$	$(x + 2ts, t, ue^{-sx-s^2 t})$
$G_6 = \exp(\mathbf{v}_6 s)$	$(\frac{x}{1-4st}, \frac{t}{1-4st}, u\sqrt{1-4st}\exp[\frac{-sx^2}{1-4st}])$
$G_\beta = \exp(\mathbf{v}_\beta s)$	$(x, t, u + s\beta(x, t))$

a new solution. G_5 is a Galilean boost to a moving frame. Finally, G_6 is a transformation that cannot be obtained from any physical principle. Since each group G_i is a one-parameter group of symmetries, if f is a solution of the heat equation, so are the functions $f_i \equiv G_i \cdot f$ for all i. These functions can be obtained from Eq. (32.8). As an illustration, we find f_6. First note that for $u = f(x, t)$, we have

$$\tilde{x} = \frac{x}{1-4st}, \qquad \tilde{t} = \frac{t}{1-4st},$$
$$\tilde{u} \equiv \tilde{f}(\tilde{x}, \tilde{t}) = f(x, t)\sqrt{1-4st}\exp\left[\frac{-sx^2}{1-4st}\right].$$

Next solve the first two equations above for x and t in terms of $\tilde{x}$ and $\tilde{t}$:

$$x = \frac{\tilde{x}}{1+4s\tilde{t}}, \qquad t = \frac{\tilde{t}}{1+4s\tilde{t}}.$$

Finally, substitute in the last equation to get

$$\tilde{f}(\tilde{x}, \tilde{t}) = f\left(\frac{\tilde{x}}{1+4s\tilde{t}}, \frac{\tilde{t}}{1+4s\tilde{t}}\right)\sqrt{\frac{1}{1+4s\tilde{t}}}\exp\left[\frac{-s\tilde{x}^2}{1+4s\tilde{t}}\right],$$

or, changing $\tilde{x}$ to x and $\tilde{t}$ to t,

$$f_6(x, t) = \frac{1}{\sqrt{1+4st}}\exp\left[\frac{-sx^2}{1+4st}\right] f\left(\frac{x}{1+4st}, \frac{t}{1+4st}\right).$$

The other transformed functions can be obtained similarly. We simply list these functions:

$$f_1(x, t) = f(x - s, t), \qquad f_2(x, t) = f(x, t - s),$$
$$f_3(x, t) = e^s f(x, t), \qquad f_4(x, t) = f\left(e^{-s}x, e^{-2s}t\right),$$
$$f_5(x, t) = e^{-sx+s^2 t} f(x - 2st, t), \qquad f_\beta(x, t) = f(x, t) + s\beta(x, t),$$
$$f_6(x, t) = \frac{1}{\sqrt{1+4st}}\exp\left[\frac{-sx^2}{1+4st}\right] f\left(\frac{x}{1+4st}, \frac{t}{1+4st}\right). \tag{32.28}$$

We can find the *fundamental solution* to the heat equation very simply as follows. Let $f(x,t)$ be the trivial constant solution c. Then

$$u = f_6(x,t) = \frac{c}{\sqrt{1+4st}} e^{-sx^2/(1+4st)}$$

is also a solution. Now choose $c = \sqrt{s/\pi}$ and translate t to $t - 1/4s$ (an allowed operation due to the invariance of the heat equation under time translation G_2). The result is

$$u = \frac{1}{\sqrt{4\pi t}} e^{-x^2/4t},$$

which is the fundamental solution of the heat equation [see (22.45)].

32.4.2 The Wave Equation

As the next example of the application of Lie groups to differential equations, we consider the wave equation in two dimensions. This equation is written as

$$u_{tt} - u_{xx} - u_{yy} = 0, \quad \text{or} \quad \eta^{ij} u_{ij} = 0 \quad \text{and} \quad \Delta = \eta^{ij} u_{ij}, \tag{32.29}$$

where $\boldsymbol{\eta} = \mathrm{diag}(1, -1, -1)$, and subscripts indicate derivatives with respect to coordinate functions $x^1 = t$, $x^2 = x$, and $x^3 = y$. With $p = 3$ and $q = 1$, a typical generator of symmetry will be of the form

$$\mathbf{v} = \sum_{i=1}^{3} X^{(i)} \frac{\partial}{\partial x^i} + U \frac{\partial}{\partial u}, \tag{32.30}$$

where $\{X^{(i)}\}_{i=1}^3$ and U are functions of t, x, y, and u to be determined. The second prolongation of such a vector field is

$$\mathrm{pr}^{(2)}\mathbf{v} = \mathbf{v} + \sum_{i=1}^{3} U^{(i)}\big(x, u^{(2)}\big) \frac{\partial}{\partial u_i} + \sum_{i,j=1}^{3} U^{(ij)}\big(x, u^{(2)}\big) \frac{\partial}{\partial u_{ij}},$$

where by Theorem 32.3.5,

$$U^{(i)} = D_i\left(U - \sum_{k=1}^{3} X^{(k)} u_k\right) + \sum_{k=1}^{3} X^{(k)} u_{ik} = D_i U - \sum_{k=1}^{3} \big(D_i X^{(k)}\big) u_k,$$

$$\begin{aligned} U^{(ij)} &= D_i D_j\left(U - \sum_{k=1}^{3} X^{(k)} u_k\right) + \sum_{k=1}^{3} X^{(k)} u_{ijk} \\ &= D_i D_j U - \sum_{k=1}^{3} \big[\big(D_i D_j X^{(k)}\big) u_k + u_{ik}\big(D_j X^{(k)}\big) + u_{jk}\big(D_i X^{(k)}\big)\big], \end{aligned}$$

and we have used Eq. (32.21). Using (32.21) further, the reader may show that

$$\begin{aligned}U^{(ij)} &= U_{ij} + u_k\big(\delta_{jk}U_{iu} + \delta_{ik}U_{ju} - X_{ij}^{(k)}\big)\\ &\quad + u_l u_k\big(\delta_{il}\delta_{jk}U_{uu} - X_{iu}^{(k)}\delta_{jl} - X_{ju}^{(k)}\delta_{il}\big)\\ &\quad + u_{kl}\big(\delta_{il}\delta_{jk}U_u - X_i^{(k)}\delta_{jl} - X_j^{(k)}\delta_{il}\big)\\ &\quad - u_k u_{lm}\big(X_u^{(k)}\delta_{il}\delta_{jm} + X_u^{(m)}\delta_{il}\delta_{jk} + X_u^{(m)}\delta_{ik}\delta_{jl}\big) - u_i u_j u_k X_{uu}^{(k)},\end{aligned} \tag{32.31}$$

where a sum over repeated indices is understood.

Applying $\mathrm{pr}^{(2)}\mathbf{v}$ to Δ, we obtain the infinitesimal criterion

$$U^{(tt)} = U^{(xx)} + U^{(yy)} \quad \text{or} \quad \eta^{ij}U^{(ij)} = 0.$$

Multiplying Eq. (32.31) by η^{ij} and setting the result equal to zero yields

$$\begin{aligned}0 &= \eta^{ij}U^{(ij)}\\ &= \eta^{ij}U_{ij} + u_k\big(2\eta^{ik}U_{iu} - \eta^{ij}X_{ij}^{(k)}\big) + u_l u_k\big(\eta^{kl}U_{uu} - 2X_{iu}^{(k)}\eta^{il}\big)\\ &\quad - 2u_{kl}X_i^{(k)}\eta^{il} - 2u_k u_{lm}X_u^{(m)}\eta^{kl} - u_i u_j u_k X_{uu}^{(k)}\eta^{ij},\end{aligned} \tag{32.32}$$

where we have used the wave equation, $\eta^{kl}u_{kl} = 0$. Equation (32.32) must hold for all derivatives of u and powers thereof (treated as independent) modulo the wave equation. Therefore, the coefficients of such "monomials" must vanish. For example, since all the terms involving $u_k u_{lm}$ are independent (even after substituting $u_{xx} + u_{yy}$ for u_{tt}), we have to conclude that $X_u^{(m)} = 0$ for all m, i.e., that $X^{(i)}$ are independent of u. Setting the coefficient of $u_k u_l$ equal to zero and noting that $X_{iu}^{(k)} = \partial X_u^{(k)}/\partial x^i = 0$ yields

$$U_{uu} = 0 \quad \Rightarrow \quad U(x, y, t, u) = \alpha(x, y, t)u + \beta(x, y, t). \tag{32.33}$$

Let us concentrate on the functions $X^{(i)}$. These are related via the term linear in u_{kl}. After inserting the wave equation in this term, we get

$$\begin{aligned}u_{kl}X_i^{(k)}\eta^{il} &= u_{12}\big(X_1^{(2)} - X_2^{(1)}\big) + u_{13}\big(X_1^{(3)} - X_3^{(1)}\big) - u_{23}\big(X_2^{(3)} + X_3^{(2)}\big)\\ &\quad + u_{22}\big(X_1^{(1)} - X_2^{(2)}\big) + u_{33}\big(X_1^{(1)} - X_3^{(3)}\big).\end{aligned}$$

The u_{ij} in this equation are all independent; so, we can set their coefficients equal to zero:

$$\begin{aligned}&X_1^{(2)} = X_2^{(1)}, \qquad X_1^{(3)} = X_3^{(1)}, \qquad X_2^{(3)} + X_3^{(2)} = 0,\\ &X_1^{(1)} = X_2^{(2)} = X_3^{(3)}.\end{aligned} \tag{32.34}$$

The reader may verify that these relations imply that $X_{jkl}^{(i)} = 0$ for any i, j, k, and l. For example,

Table 32.3 The generators of the conformal group for $\mathbb{R}^3$, part of the symmetry group of the wave equation in two dimensions

Infinitesimal generator	Transformation
$\mathbf{v}_1 = \partial_t, \mathbf{v}_2 = \partial_x, \mathbf{v}_3 = \partial_y$	Translation
$\mathbf{v}_4 = x\partial_t + t\partial_x, \mathbf{v}_6 = -y\partial_x + x\partial_y, \mathbf{v}_7 = y\partial_t + t\partial_y$	Rotation/Boost
$\mathbf{v}_5 = t\partial_t + x\partial_x + y\partial_y$	Dilatation
$\mathbf{v}_8 = (t^2+x^2+y^2)\partial_t + 2xt\partial_x + 2yt\partial_y - tu\partial_u$, $\mathbf{v}_9 = 2xt\partial_t + (t^2+x^2-y^2)\partial_x + 2xy\partial_y - xu\partial_u$, $\mathbf{v}_{10} = 2yt\partial_t + 2xy\partial_x + (t^2-x^2+y^2)\partial_y - yu\partial_u$	Inversions

$$\begin{aligned} X^{(2)}_{222} = X^{(1)}_{122} &= \underbrace{X^{(3)}_{322}}_{\equiv X^{(3)}_{223}} = -\underbrace{X^{(2)}_{323}}_{\equiv X^{(2)}_{233}} = -\underbrace{X^{(1)}_{133}}_{\equiv X^{(1)}_{313}} = -\underbrace{X^{(3)}_{113}}_{\equiv X^{(3)}_{311}} \\ &= -\underbrace{X^{(2)}_{211}}_{\equiv X^{(2)}_{121}} = -\underbrace{X^{(1)}_{221}}_{\equiv X^{(1)}_{122}} = -X^{(2)}_{222}. \end{aligned} \tag{32.35}$$

So, the first link of this chain is equal to its negative. Therefore, all the third derivatives in the chain of Eq. (32.35) vanish. It follows that all $X^{(i)}$'s are mixed polynomials of at most degree two. Writing the most general such polynomials for the three functions $X^{(1)}$, $X^{(2)}$, and $X^{(3)}$ and having them satisfy Eq. (32.34) yields

$$\begin{aligned} X^{(1)} &= a_1 + a_4x + a_7y + a_5t + a_8\left(x^2+y^2+t^2\right) + 2a_9xt + 2a_{10}yt, \\ X^{(2)} &= a_2 + a_5x - a_6y + a_4t + a_9\left(x^2-y^2+t^2\right) + 2a_{10}xy + 2a_8xt, \\ X^{(3)} &= a_3 + a_6x + a_5y + a_7t + a_{10}\left(-x^2+y^2+t^2\right) + 2a_9xy + 2a_8yt. \end{aligned} \tag{32.36}$$

Setting the coefficient of u_k and $\eta^{ij}U_{ij}$ equal to zero and using Eq. (32.33) gives

$$\begin{aligned} 2\alpha_x &= X^{(2)}_{xx} + X^{(2)}_{yy} - X^{(2)}_{tt}, & 2\alpha_y &= X^{(3)}_{xx} + X^{(3)}_{yy} - X^{(2)}_{tt}, \\ 2\alpha_t &= X^{(1)}_{tt} - X^{(1)}_{xx} - X^{(1)}_{yy}, & \beta_{tt} - \beta_{xx} - \beta_{yy} &= 0. \end{aligned}$$

It follows that β is any solution of the wave equation, and

$$\alpha(x,y,t) = a_{11} - a_8t - a_9x - a_{10}y.$$

By inserting the expressions found for $X^{(i)}$ and U in (32.30) and writing the result in the form $\sum_i a_i\mathbf{v}_i$, we discover that the generators of the symmetry group consist of the ten vector fields given in Table 32.3 as well as the vector fields

$$u\partial_u, \quad \mathbf{v}_\beta = \beta(x,y,t)\partial_u$$

for β an arbitrary solution of the wave equation. The ten vector fields of Table 32.3 comprise the generators of the conformal group in three dimensions whose generalization to m dimensions will be studied in Sect. 37.2.

32.5 Application to ODEs

The theory of Lie groups finds one of its most rewarding applications in the integration of ODEs. Lie's fundamental observation was that if one could come up with a sufficiently large group of symmetries of a system of ODEs, then one could integrate the system. In this section we outline the general technique of solving ODEs once we know their symmetries. The following proposition will be useful (see [Warn 83, p. 40]):

Proposition 32.5.1 *Let M be an n-dimensional manifold and $\mathbf{v} \in \mathfrak{X}(M)$. Assume that $\mathbf{v}|_P \neq 0$ for some $P \in M$. Then there exists a local chart, i.e., local set of coordinate functions, $(w^1, \dots, w^n)$ at P such that $\mathbf{v} = \partial/\partial w^1$.*

32.5.1 First-Order ODEs

The most general first-order ODE can be written as

$$\frac{du}{dx} = F(x,u) \quad \Rightarrow \quad \Delta(x,u,u_x) \equiv u_x - F(x,u) = 0. \tag{32.37}$$

A typical infinitesimal generator of the symmetry group of this equation is[5] $\mathbf{v} = X\partial_x + U\partial_u$, whose prolongation is

$$\mathrm{pr}^{(1)} = \mathbf{v} + U^{(x)}\frac{\partial}{\partial u_x}, \quad \text{where}$$
$$U^{(x)} \equiv U_x + (U_u - X_x)u_x - X_u u_x^2, \tag{32.38}$$

as the reader may verify. The infinitesimal criterion for the one-parameter group of transformations G to be a symmetry group of Eq. (32.37) is $\mathrm{pr}^{(1)}\mathbf{v}(\Delta) = 0$, or

$$\frac{\partial U}{\partial x} + \left(\frac{\partial U}{\partial u} - \frac{\partial X}{\partial x}\right)F - \frac{\partial X}{\partial u}F^2 = X\frac{\partial F}{\partial x} + U\frac{\partial F}{\partial u}. \tag{32.39}$$

Any solution (X, U) of this equation generates a 1-parameter group of transformations. The problem is that a *systematic* procedure for solving (32.39) is *more difficult* than solving the original equation. However, in most cases, one can *guess* a symmetry transformation (based on physical, or other, grounds), and that makes Lie's method worthwhile.

Suppose we have found a symmetry group G with infinitesimal generator $\mathbf{v}$ that does not vanish at $P \in M \subset X \times U$. Based on Proposition 32.5.1, we can introduce new coordinates

$$w = \xi(x,u), \qquad y = \eta(x,u) \tag{32.40}$$

[5] The reader is warned against the unfortunate coincidence of notation: X and U represent *both* the components of the infinitesimal generator and the spaces of independent and dependent variables!

in a neighborhood of P such that $\mathbf{v} = \partial/\partial w$, whose prolongation is also $\partial/\partial w$ [see (32.38)]. This transforms the DE of (32.37) into[6] $\tilde{\Delta}(y, w, w_y) = 0$, and the infinitesimal criterion into

$$\mathrm{pr}^{(1)}\mathbf{v}(\tilde{\Delta}) = \frac{\partial \tilde{\Delta}}{\partial w} = 0.$$

It follows that $\tilde{\Delta}$ is independent of w. The transformed DE is therefore $\tilde{\Delta}(y, w_y) = 0$, whose normal form, obtained by implicitly solving for dw/dy, is

$$\frac{dw}{dy} = H(y) \quad \Rightarrow \quad w = \int_a^y H(t)\, dt - w(a)$$

for some function H of y alone and some convenient point $y = a$. Substituting this expression of w as a function of y in Eq. (32.40) and eliminating y between the two equations yields u as a function of x.

Thus our task is to find the change of variables (32.40). For this, we use $\mathbf{v}(w) = 1$ and $\mathbf{v}(y) = 0$, and express them in terms of x and u:

$$\begin{aligned} \mathbf{v}(w) = \mathbf{v}(\xi) = X\frac{\partial \xi}{\partial x} + U\frac{\partial \xi}{\partial u} = 1, \\ \mathbf{v}(y) = \mathbf{v}(\eta) = X\frac{\partial \eta}{\partial x} + U\frac{\partial \eta}{\partial u} = 0. \end{aligned} \tag{32.41}$$

The second equation says that η is an invariant of the group generated by $\mathbf{v}$. We therefore use the associated characteristic ODE [see (32.3) and (32.5)] to find y (or η):

$$\frac{dx}{X(x, u)} = \frac{du}{U(x, u)}. \tag{32.42}$$

To find w (or ξ), we introduce $\chi(x, u, v) = v - \xi(x, u)$ and note that an equivalent relation containing the same information as the first equation in (32.41) is

$$X\frac{\partial \chi}{\partial x} + U\frac{\partial \chi}{\partial u} + \frac{\partial \chi}{\partial v} = 0,$$

which has the characteristic ODE

$$\frac{dx}{X(x, u)} = \frac{du}{U(x, u)} = \frac{dv}{1}, \tag{32.43}$$

for which we seek a solution of the form $v - \xi(x, u) = c$ to read off $\xi(x, u)$.

The reader may wonder whether it is sane to go through so much trouble only to replace the original single ODE with *two* ODEs such as (32.42) and (32.43)! The answer is that *in practice*, the latter two DEs are much easier to solve than the original ODE.

---

[6]Here we are choosing w to be the dependent variable. This choice is a freedom that is always available to us.

Example 32.5.2 The homogeneous FODE $du/dx = F(u/x)$ is invariant under the scaling transformation $(x, u) \mapsto (sx, su)$ whose infinitesimal generator is $\mathbf{v} = x\partial_x + u\partial_u$. The first prolongation of this vector is the same as the vector itself (reader, verify!).

To find the new coordinates w and y, first use Eq. (32.42) with $X(x, u) = x$ and $U(x, u) = u$:

$$\frac{dx}{x} = \frac{du}{u} \quad \Rightarrow \quad \frac{u}{x} = c_1 \quad \Rightarrow \quad y = \frac{u}{x} \qquad \text{(see Box 32.1.8)}.$$

Next, we note that (32.43) yields

$$\frac{dx}{x} = \frac{du}{u} = dv \quad \Rightarrow \quad \ln u = v + \ln c_2 \quad \Rightarrow \quad v = \ln(u/c_2).$$

Substituting from the previous equation, we obtain

$$v = \ln(c_1 x/c_2) = \ln x + \underbrace{\ln(c_1/c_2)}_{\equiv c} \quad \Rightarrow \quad v - \ln x = c \quad \Rightarrow \quad w = \ln x.$$

The chain rule gives $du/dx = (1 + yw_y)/w_y$, so that the DE becomes

$$\frac{1 + yw_y}{w_y} = F(y) \quad \Rightarrow \quad \frac{dw}{dy} = \frac{1}{F(y) - y},$$

which can be integrated to give $w = H(y)$ or $\ln x = H(y) = H(u/x)$, which defines u as an implicit function of x.

32.5.2 Higher-Order ODEs

The same argument used in the first order ODEs can be used for higher-order ODEs to reduce their orders.

Proposition 32.5.3 *Let*

$$\Delta\big(x, u^{(n)}\big) = \Delta(x, u, u_1, \ldots, u_n) = 0, \qquad u_k \equiv \frac{d^k u}{dx^k},$$

be an nth order ODE. If this ODE has a one-parameter symmetry group, then there exist variables $w = \xi(x, u)$ and $y = \eta(x, u)$ such that

$$\Delta\big(x, u^{(n)}\big) = \tilde{\Delta}\left(y, \frac{dw}{dy}, \ldots, \frac{d^n w}{dy^n}\right) = 0,$$

i.e., in terms of w and y, the ODE becomes of $(n-1)$st order in w_y.

Proof The proof is exactly the same as in the first-order case. The only difference is that one has to consider $\mathrm{pr}^{(n)}\mathbf{v}$, where $\mathbf{v} = \partial/\partial w$. But Problem 32.7 shows that $\mathrm{pr}^{(n)}\mathbf{v} = \mathbf{v}$, as in the first-order case. □

Example 32.5.4 Consider a second-order DE $\Delta(u, u_x, u_{xx}) = 0$, which does not depend on x explicitly. The fact that $\partial\Delta/\partial x = 0$ suggests $w = x$. So, we switch the dependent and independent variables and write $w = x$, and $y = u$. Then, using the chain rule, we get

$$\frac{du}{dx} = \frac{1}{w_y}, \qquad \frac{d^2u}{dx^2} = -\frac{w_{yy}}{w_y^3}.$$

Substituting in the original DE, we obtain

$$\tilde{\Delta}(y, w_y, w_{yy}) \equiv \Delta\left(y, \frac{1}{w_y}, -\frac{w_{yy}}{w_y^3}\right) = 0,$$

which is of first order in w_y.

Example 32.5.5 The order of the SOLDE $u_{xx} + p(x)u_x + q(x)u = 0$ can be reduced by noting that the DE is invariant under the scaling transformation $(x, u) \mapsto (x, su)$, whose infinitesimal generator is $\mathbf{v} = u\partial_u$. With this vector field, Eqs. (32.42) and (32.43) give

$$\frac{dx}{0} = \frac{du}{u}, \qquad \frac{dx}{0} = \frac{du}{u} = dv.$$

For the first equation to make sense, we have to have

$$dx = 0 \quad \Rightarrow \quad x = c_1 \quad \Rightarrow \quad y = x \quad \text{(by Box 32.1.8)}.$$

The second equation in u gives

$$v = \ln u + c \quad \Rightarrow \quad v - \ln u = c \quad \Rightarrow \quad w = \ln u \quad \Rightarrow \quad u = e^w.$$

Using the chain rule, we obtain

$$u_x = \frac{dw}{dx}e^w = \frac{dw}{dy}e^w \quad \text{and} \quad u_{xx} = \left(w_{yy} + w_y^2\right)e^w.$$

Riccati equation

By inserting this in the original DE and writing $z = w_y$, we obtain

$$\frac{dz}{dy} = -z^2 - p(y)z - q(y),$$

which is the well-known first-order **Riccati equation**.

32.5.3 DEs with Multiparameter Symmetries

We have seen that 1-parameter symmetries reduce the order of an ODE by 1. It is natural to suspect that an r-parameter symmetry will reduce the order by r. Although this suspicion is correct, it turns out that in general, one cannot reconstruct the solution of the original equation from those of the reduced $(n-r)$th-order equation. (See [Olve 86, pp. 148–158] for a thorough discussion of this problem.) However, the special, but important, case of second-order DEs is an exception. The deep reason behind this is the exceptional structure of 2-dimensional Lie algebras given in Box 29.2.5. We cannot afford to go into details of the reasoning, but simply quote the following important theorem.

Theorem 32.5.6 *Let $\Delta(x, u^{(n)}) = 0$ be an nth-order ODE invariant under a 2-parameter group. Then there is an $(n-2)$nd-order ODE $\tilde{\Delta}(y, w^{(n-2)}) = 0$ with the property that the general solution to Δ can be found by integrating the general solution to $\tilde{\Delta}$. In particular, a second-order ODE having a 2-parameter group of symmetries can be solved by integration.*

Let us analyze the case of a second-order ODE in some detail. By Box 29.2.5, the infinitesimal generators $\mathbf{v}_1$ and $\mathbf{v}_2$ satisfy the Lie bracket relation

$$[\mathbf{v}_1, \mathbf{v}_2] = c\mathbf{v}_1, \quad c = 0 \quad \text{or} \quad 1.$$

We shall treat the abelian case ($c = 0$) and leave the nonabelian case for the reader. To begin with, we use s and t for the transformed variables, and at the end replace them with y and w.

By Proposition 32.5.1, we can let $\mathbf{v}_1 = \partial/\partial s$. Then $\mathbf{v}_2$ can be expressed as the linear combination

$$\mathbf{v}_2 = \alpha(s,t)\frac{\partial}{\partial s} + \beta(s,t)\frac{\partial}{\partial t}.$$

The commutation relation $[\mathbf{v}_1, \mathbf{v}_2] = 0$ gives

$$0 = [\partial_s, \alpha\partial_s + \beta\partial_t] = \frac{\partial\alpha}{\partial s}\partial_t + \frac{\partial\beta}{\partial s}\partial_s,$$

showing that α and β are independent of s. We want to simplify $\mathbf{v}_2$ as much as possible without changing $\mathbf{v}_1$. A transformation that accomplishes this is $S = s + h(t)$ and $T = T(t)$. Then, by Eq. (28.8) we obtain

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{v}_1(S)\frac{\partial}{\partial S} + \mathbf{v}_1(T)\frac{\partial}{\partial T} = \frac{\partial S}{\partial s}\frac{\partial}{\partial S} + \frac{\partial T}{\partial s}\frac{\partial}{\partial T} = \frac{\partial}{\partial S},\\
\mathbf{v}_2 &= \mathbf{v}_2(S)\frac{\partial}{\partial S} + \mathbf{v}_2(T)\frac{\partial}{\partial T} = \left(\alpha\frac{\partial S}{\partial s} + \beta\frac{\partial S}{\partial t}\right)\frac{\partial}{\partial S} + \left(\alpha\frac{\partial T}{\partial s} + \beta\frac{\partial T}{\partial t}\right)\frac{\partial}{\partial T}\\
&= \left(\alpha + \beta h'\right)\frac{\partial}{\partial S} + \beta T'\frac{\partial}{\partial T}.
\end{aligned}$$

If $\beta \neq 0$, we choose $T' = 1/\beta$ and $h' = -\alpha/\beta$ to obtain

$$\mathbf{v}_1 = \frac{\partial}{\partial s}, \qquad \mathbf{v}_2 = \frac{\partial}{\partial t}, \tag{32.44}$$

where we have substituted s for S and t for T. If $\beta = 0$, we choose $\alpha = T$, and change the notation from S to s and T to t to obtain

$$\mathbf{v}_1 = \frac{\partial}{\partial s}, \qquad \mathbf{v}_2 = t\frac{\partial}{\partial s}. \tag{32.45}$$

The next step is to decide which coordinate is the independent variable, prolong the vector fields, and apply it to the DE to find the infinitesimal criterion. For $\beta \neq 0$, the choice is immaterial. So, let $w = s$ and $y = t$. Then the prolongation of $\mathbf{v}_1$ and $\mathbf{v}_2$ will be the same as the vectors themselves, and with $\Delta(y, w, w_y, w_{yy}) \equiv w_{yy} - F(y, w, w_y)$, the infinitesimal criteria for invariance will be

$$0 = \mathrm{pr}^{(2)}\mathbf{v}_1(\Delta) = \mathbf{v}_1(\Delta) = \frac{\partial \Delta}{\partial w} = -\frac{\partial F}{\partial w},$$
$$0 = \mathrm{pr}^{(2)}\mathbf{v}_2(\Delta) = \mathbf{v}_2(\Delta) = \frac{\partial \Delta}{\partial y} = -\frac{\partial F}{\partial y}.$$

It follows that in the (y, w) system, F will be a function of w_y alone and the DE will be of the form

$$w_{yy} = F(w_y) \quad \Rightarrow \quad \frac{dw_y}{dy} = F(w_y) \quad \Rightarrow \quad \underbrace{\int^{w_y} \frac{dz}{F(z)}}_{\equiv H(w_y)}.$$

The last equation can be solved for w_y in terms of y and the result integrated.

For $\beta = 0$, choose $w = t$ and $y = s$. Then $\mathbf{v}_1$ will not prolongate, and as the reader may verify,

$$\mathrm{pr}^{(2)}\mathbf{v}_2 = \mathbf{v}_2 - w_y^2 \frac{\partial}{\partial w_y} - 3w_y w_{yy}\frac{\partial}{\partial w_{yy}} = w\frac{\partial}{\partial y} - w_y^2 \frac{\partial}{\partial w_y} - 3w_y w_{yy}\frac{\partial}{\partial w_{yy}},$$

and the infinitesimal criteria for invariance will be

$$0 = \mathrm{pr}^{(2)}\mathbf{v}_1(\Delta) = \mathbf{v}_1(\Delta) = \frac{\partial \Delta}{\partial y} = -\frac{\partial F}{\partial y},$$
$$0 = \mathrm{pr}^{(2)}\mathbf{v}_2(\Delta) = -w\underbrace{\frac{\partial F}{\partial y}}_{=0} + w_y^2\frac{\partial F}{\partial w_y} + 3w_y \underbrace{w_{yy}}_{=F}.$$

It follows that in the (y, w) system, F will be a function of w and w_y and satisfy the DE

$$w_y\frac{\partial F}{\partial w_y} = 3F,$$

whose solution is of the form $F(w, w_y) = w_y^3 \tilde{F}(w)$. The original DE now becomes

$$w_{yy} = w_y^3 \tilde{F}(w),$$

for which we use the chain rule $w_{yy} = w_y \partial w_y/\partial w$ to obtain

$$\frac{dw_y}{dw} = w_y^2 \tilde{F}(w) \quad \Rightarrow \quad -\frac{1}{w_y} = \underbrace{\int^w \tilde{F}(z)\,dz}_{\equiv H(w)} \quad \Rightarrow \quad \frac{dw}{dy} = -\frac{1}{H(w)},$$

which can be integrated. Had we chosen $w = s$ and $y = t$, F would have been a function of y and the DE would have reduced to $w_{yy} = F(y)$, which could be solved by two consecutive integrations. The nonabelian 2-dimensional Lie algebra can be analyzed similarly. The reader may verify that if $\beta = 0$, the vector fields can be chosen to be

$$\mathbf{v}_1 = \frac{\partial}{\partial s}, \qquad \mathbf{v}_2 = s\frac{\partial}{\partial s}, \tag{32.46}$$

leading to the ODE $w_{yy} = w_y \tilde{F}(y)$, and if $\beta \neq 0$, the vector fields can be chosen to be

$$\mathbf{v}_1 = \frac{\partial}{\partial s}, \qquad \mathbf{v}_2 = s\frac{\partial}{\partial s} + t\frac{\partial}{\partial t}, \tag{32.47}$$

leading to the ODE $w_{yy} = \tilde{F}(w_y)/y$. Both of these ODEs are integrable as in the abelian case.

32.6 Problems

32.1 Suppose that $\{F_i\}_{i=1}^n$ are invariants of the PDE (32.3). Show that any function $f(F_1, F_2, \ldots, F_n)$ is also an invariant of the PDE.

32.2 Find the function $\tilde{f} = \theta \cdot f$ when $f(x) = ax + b$ and θ is the angle of rotation of $SO(2)$.

32.3 Use the result of Problem 32.2 to find $\tilde{u}_1$. Hint: Note that $a = u_1$.

32.4 Transform the DE of Example 32.3.2 from Cartesian to polar coordinates to obtain $dr/d\theta = r^3$.

32.5 Using the definition of total derivative, verify Eq. (32.21).

32.6 Show that $SO(2)$ is a symmetry group of the first-order DE

$$\Delta(x, u, u_1) = (u - x)u_1 + x + u = 0$$

and write the same DE in polar coordinates.

32.7 Show that the nth prolongation of the generator of the ith translation, ∂_i, is the same as the original vector.

32.8 Find the first prolongation of the generator of scaling: $x\partial_x + u\partial_u$.

32.9 Show that when the group acts only on the single dependent variable u, the prolongation of $\mathbf{v} = U\partial_u$ is given by

$$\mathrm{pr}^{(1)}\mathbf{v} = \mathbf{v} + \sum_{j=1}^{p} U_j \frac{\partial}{\partial u_j}, \quad \text{where} \quad U_j = \frac{\partial U}{\partial x^j} + u_j \frac{\partial U}{\partial u}.$$

32.10 Show that the nth prolongation of $\mathbf{v} = X(x,u)\partial_x + U(x,u)\partial_u$ for an ordinary DE of nth order is

$$\mathrm{pr}^{(n)}\mathbf{v} = \mathbf{v} + \sum_{k=1}^{n} U^{[k]} \frac{\partial}{\partial u^{(k)}},$$

where

$$u^{(k)} \equiv \frac{\partial^k u}{\partial x^k} \quad \text{and} \quad U^{[k]} = D_x^k(U - Xu_x) + Xu^{(k+1)}.$$

32.11 Compute the second prolongation of the infinitesimal generators of the symmetry group of the heat equation.

32.12 Derive Eqs. (32.31) and (32.32).

32.13 Using Eq. (32.34) show that $X_{jkl}^{(i)} = 0$ for any i, j, k, and l.

Korteweg-de Vries equation

32.14 The **Korteweg-de Vries** equation is $u_t + u_{xxx} + uu_x = 0$. Using the technique employed in computing the symmetries of the heat and wave equations, show that the infinitesimal generators of symmetries of the Korteweg-de Vries equation are

$\mathbf{v}_1 = \partial_x,$	$\mathbf{v}_2 = \partial_t,$	translation
$\mathbf{v}_3 = t\partial_x + \partial_u,$		Galilean boost
$\mathbf{v}_4 = x\partial_x + 3t\partial_t - 2u\partial_u.$		scaling

32.15 Suppose $M(x,u)\,dx + N(x,u)\,du = 0$ has a 1-parameter symmetry group with generator $\mathbf{v} = X\partial_x + U\partial_u$. Show that the function $q(x,u) = 1/(XM + UN)$ is an integrating factor.

32.16 Show that the second prolongation of $\mathbf{v} = w\partial_y$ (with y treated as independent variable) is

$$\mathrm{pr}^{(2)}\mathbf{v} = \mathbf{v} - w_y^2 \frac{\partial}{\partial w_y} - 3w_y w_{yy} \frac{\partial}{\partial w_{yy}}.$$

32.17 Go through the case of $\beta = 0$ in the solution of the second order ODE and, choosing $w = s$ and $y = t$, show that F will be a function of y alone and the original DE will reduce to $w_{yy} = F(y)$.

32.18 Show that in the case of the nonabelian 2-dimensional Lie algebra,

(a) the vector fields can be chosen to be

$$\mathbf{v}_1 = \frac{\partial}{\partial s}, \qquad \mathbf{v}_2 = s\frac{\partial}{\partial s}$$

if $\beta = 0$.

(b) Show that these vector fields lead to the ODE $w_{yy} = w_y \tilde{F}(y)$.

(c) If $\beta \neq 0$, show that the vector fields can be chosen to be

$$\mathbf{v}_1 = \frac{\partial}{\partial s}, \qquad \mathbf{v}_2 = s\frac{\partial}{\partial s} + t\frac{\partial}{\partial t}.$$

(d) Finally, show that the latter vector fields lead to the ODE $w_{yy} = \tilde{F}(w_y)/y$.