

SYMMETRY GROUP SOLUTIONS TO DIFFERENTIAL EQUATIONS—A HISTORICAL PERSPECTIVE

JACOB HARRY BURKMAN

My thanks to Professor George Wilkens for putting up with me during the last year while working through this project and its background material. I would also like to thank Professor Matt Richey of St. Olaf College, Northfield, MN for first introducing me to the joys of differential equations.

ABSTRACT. In this project we will be looking at Sophus Lie's desire (his so called *Idée Fixe*) to apply Contact Transformations (what would eventually develop into the modern idea of a Lie Algebra) in order to arrive at symmetries of differential equations, and thus certain solutions. Our goal—as well as Lie's—is to develop a more universal method for solving differential equations than the familiar cook-book methods we learn in an introductory ordinary or partial differential equations class. We answer three questions. What was the historical underpinning of Sophus Lie's theory? How do we find the symmetry Lie algebras? How do we use the symmetry Lie algebras to find solutions to the differential equation? (In order to answer these questions we will need to fill in some background material and our answers will also result in a novel derivation of the “Fundamental Source Solution.”) Our second objective will be to establish a connection between solvability in Galois Theory and in Differential Equations. We will assume a familiarity with certain algebraic concepts from Abstract Algebra.

Date: April 2, 2007.

Master's Plan B Project: In partial fulfillment of the the degree of Master of Arts in Mathematics.

1. ALGEBRAIC BACKGROUND

The desire to solve algebraic equations is something that has dated back to the time of the ancient Greeks and even before. However, like almost everything (except somewhat superstitious theology) algebra—and mathematics—didn't advance very far in western Europe during the Dark Ages. There was, however, a great deal of advancement in the Arab world, especially in Algebra. Following a pattern similar to many academic pursuits, mathematics came back into vogue (at least with some in academia) by the 18th century in Europe. This development was spurred on by the rethinking that took place during the Renaissance. Aiding this was the discovery of the work of the Arab mathematicians of the preceding centuries as well as the work of the Greeks—which had essentially been lost to the western world.

We later deal with other historical influences on Lie's theory (c.f. Section (5)). Here, we focus on some of the algebraic material and influences. We begin this with the work of Evariste Galois, culminating in his Fundamental Theorem of Galois Theory (see [Hun80] p. 245-246).

Theorem 1. Fundamental Theorem of Galois Theory. *If F is a finite dimensional Galois extension of K , then there is a one-to-one correspondence between the set of all intermediate fields of the extension and the set of subgroups of the Galois group $\text{Aut}_K F$ (given by $E \mapsto E' = \text{Aut}_E F$) such that:*

- (1) *the relative dimension of two intermediate fields is equal to the relative index of the corresponding subgroups; in particular, $\text{Aut}_K F$ has order $[F : K]$;*
- (2) *F is Galois over every intermediate field E , but E is Galois over K if and only if the corresponding subgroup $E' = \text{Aut}_E F$ is normal in $G = \text{Aut}_K F$; in this case G/E' is (isomorphic to) the Galois group $\text{Aut}_K E$ of E over K .*

In a normal course,¹ this theorem receives much more attention than we are giving it here. However, our intention here is not to cover Abstract Algebra, but merely to point out the connection that was sought between the polynomial algebra related to Galois Theory (in particular that it is possible to find some sort of correspondence between

¹I would like to thank Professor Ron Brown for his encouragement and for two wonderful years of Algebra.

intermediate fields of a field extension $K \subset F$ and the subgroups of the Galois group $\text{Aut}_K F$, and the desired method of dealing with differential equations. See Section (8) for a connection between the two ideas.

An important Abstract Algebra topic is the concept of a *group action*. We define various results from this below (see [Hun80] Chapter II section 4).

Definition 1. An *action* of a group G on a set S is a function $G \times S \rightarrow S$ denoted by $(g, x) \mapsto gx$ such that for all $x \in S$ and $g_1, g_2 \in G$:

$$ex = x \quad \text{and} \quad (g_1 g_2)x = g_1(g_2 x)$$

(where e is the identity element of the group). When such an action is given, we say that G **acts on the set** S .

Now that we have a group action, the key part of this, for us, is the concept of an *orbit*. (We are, at this point, not particularly interested in the other important results such as the Class Equation that arise in an Abstract Algebra sense.)

Definition 2. Suppose that G is a group that acts on a set S . Then the relation defined on S by

$$x \sim x' \Leftrightarrow gx = x' \quad \text{for some} \quad g \in G$$

is an equivalence relation. The equivalence classes of this equivalence relation are called the **orbits** of G on S .

Perhaps the most important connection between Galois theory and theory regarding symmetries of differential equations is the concept of a *solvable group*. Again, we will mention its importance in Algebra, and realize this connection at the end of Section (7). We again use Hungerford (see [Hun80] p. 107-8) for our algebraic definition.

Definition 3. A **subnormal series** of a group G is a chain of subgroups $G = G_0 > G_1 > \cdots > G_n$ such that G_{i+1} is normal in G_i for $0 \leq i \leq n$. The **factors** of the series are the quotient groups G_i/G_{i+1} . If $G_n = \langle e \rangle$ and each of the factors is abelian, the series is called **solvable**.

From this, we have a theorem that states “A group G is *solvable* if and only if it has a solvable series.” We take this as our definition.

Another, obviously equivalent, definition to this relates to the concept of a *commutator subgroup* (see [Hun80] p. 102).

Definition 4. Let G be a group. The subgroup of G generated by the set $\{aba^{-1}b^{-1} \mid a, b \in G\}$ is called the **commutator subgroup** and is

denoted by G' . Now, if we let $G^{(1)} = G'$, then for each $i \geq 1$, define $G^{(i)} = (G^{(i-1)})'$ (where $G^{(i)}$ is called the i^{th} **derived subgroup** of G). We then say that G is **solvable** if $G^{(n)} = \langle e \rangle$ for some n .

The final algebraic connection we need is contained in the following concepts from Abstract Algebra, which we state without proof. For proofs, see [Hun80] Chapter V, Section 9.

- Definition 5.**
- (1) Let K be a field. The **Galois group of a polynomial** $f \in K[x]$ is the group $\text{Aut}_K F$, where F is a splitting field of f over K .
 - (2) An extension field F of a field K is a **radical extension** of K if $F = K(u_1, \dots, u_n)$, some power of u_1 lies in K and for each $i \geq 2$, some power of u_i lies in $K(u_1, \dots, u_{i-1})$.
 - (3) Let K be a field and $f \in K[x]$. The equation $f(x) = 0$ is **solvable by radicals** if there exists a radical extension F of K and a splitting field E of f over K such that $K \subset E \subset F$.

Theorem 2. *If F is a radical extension field of K and E is an intermediate field, then $\text{Aut}_K E$ is a solvable group.*

Theorem 3. *Let K be a field and $f \in K[x]$. If $f(x)$ is solvable by radicals, then the Galois group of f is a solvable group.*

The previous theorem is half of our desired result. However, we need a technical proposition in order to get the converse. (Alternatively, we could define a radical extension in a slightly different way. We will not pursue that here; Hungerford covers it in exercise 2 p. 309.)

Theorem 4. *Let E be a finite dimensional Galois extension field of K with solvable Galois group $\text{Aut}_K E$. Assume that the char K does not divide $[E : K]$. Then there exists a radical extension F of K , such that $K \subset E \subset F$.*

The following theorem is the result we really want, for it is this theorem that has a very important parallel in the case of differential equations (see Theorem 20).

Theorem 5. *Let K be a field and $f \in K[x]$ a polynomial of degree $n > 0$, where char K does not divide $n!$ (which is always true when char $K = 0$). Then the equation $f(x) = 0$ is solvable by radicals if and only if the Galois group of f is solvable.*

2. SOPHUS LIE BACKGROUND

Marius Sophus Lie was born on December 17, 1842 in Nordfjordeide, Norway to a Lutheran minister, or perhaps a farmer (see [Hel90]), Johann Herman Lie. (It would not be beyond the scope of possibilities if he were both.) Lie attended the University of Christiania (later the city was renamed Kristiania and then finally Oslo) studying a broad science course. During his time at the University, Lie showed relatively little interest in mathematics (see [OR]). What little interest in mathematics Lie showed was in Geometry (as we shall see, his geometric interpretation permeated much of his work). There was, however, an experience at the University that turned—albeit somewhat slowly—Lie’s direction in life toward our interest. Lie attended some lectures by Ludwig Sylow in 1862-1863. The topic of these lectures was the aforementioned Galois Theory—see Section (1). However, it was not until considerably later (1870) after consultation with his long-time friend and colleague Felix Klein that Lie decided it would be possible to develop an analogue for differential equations to Galois’s theory for algebraic equations (see [Hel90]).

I feel that I would be remiss in my duties if I failed to mention a very interesting event that took place in Lie’s life. Klein (a German) and Lie had moved to Paris in the spring of 1870 (they had earlier been working in Berlin). However, in July 1870, the Franco-Prussian war broke out. Being a German alien in France, Klein decided that it would be safer to return to Germany; Lie also decided to go home to Norway. However (in a move that I think questions his geometric abilities), Lie decided that to go from Paris to Norway, he would walk to Italy (and then presumably take a ship to Norway). The trip did not go as Lie had planned. On the way, Lie ran into some trouble—first some rain, and he had a habit of taking off his clothes and putting them in his backpack when he walked in the rain (so he was walking to Italy in the nude). Second, he ran into the French military (quite possibly while walking in the nude) and they discovered in his sack (in addition to his hopefully dry clothing) letters written to Klein in German containing the words “lines” and “spheres” (which the French interpreted as meaning “infantry” and “artillery”). Lie was arrested as a (insane) German spy. However, due to intervention by Gaston Darboux, he was released four weeks later and returned to Norway to finish his doctoral dissertation (see [Haw00] p. 26).

3. DIFFERENTIAL EQUATIONS AS USUALLY ENCOUNTERED

Before we proceed much farther, I wish to introduce a definition and some notation that may be helpful (note that this will apply to both ordinary and partial differential equations) (see [Olv93] p. 96).

Definition 6. The *order* of a differential equation (be it ordinary or partial) is the highest order derivative that occurs in the equation.

Notation 1. A differential equation Δ having p independent variables and q dependent variables is of the form $\Delta(x, u^{(n)}) = 0$ where $x = (x^1, x^2, \dots, x^p)$, $u = (u^1, u^2, \dots, u^q)$ (each u^j can be a function of x), and contained in $u^{(n)}$ is the set of all partial derivatives of u up to the n^{th} order.

All students of mathematics encounter differential equations early in their mathematical careers—usually by the second semester of Calculus. In fact, when we introduce the concept of an antiderivative, we are really solving a differential equation. This introduces us to the first basic method of solving an ordinary differential equation—by integrating. It does not (hopefully) take calculus students very long to discover that the differential equation by itself has an infinite number of solutions, each differing from the other by an additive constant. They also realize fairly quickly that, if a suitable number of initial conditions are added, the solution is unique. This is an example of the Uniqueness Theorem of solutions to linear differential equations, which we do not prove here.

Often the above is the extent of students' encounters with differential equations in an introductory calculus course. However, if they continue their studies, they begin to realize that there are a large number of categories into which differential equations fit. Namely are the equations ordinary or partial; linear or non-linear; constant coefficient or variable coefficient? We initially encounter ordinary linear differential equations of the first order (that is ones that contain no partial derivatives), the most general form being

$$(1) \quad \frac{du}{dx} + p(x)u = g(x).$$

We are told (see [BD65] p. 19-20) that in order to solve this equation, we need to find an *integrating factor*, namely

$$(2) \quad \mu(x) = e^{\int p(x)dx}$$

and arrive at the *general solution* to Equation (1):

$$(3) \quad u = \frac{\int \mu(x)g(x) dx + c}{\mu(x)}.$$

Of course, anyone who has experience with integration can see that evaluating the numerator of Equation (3) may not be the most direct thing to do. However, if we adopt the convention that the differential equation is solved once we have reduced it to quadrature, we are finished with this solution. But, if we insist that we need to actually evaluate this antiderivative, we may be in trouble since an elementary antiderivative of $\mu(x)g(x)$ may not exist.

Another first order equation that is “easy” to solve, especially if we again solve only to reduction to quadrature, is a *separable* equation, that is, if we can write the equation in the form (see [BD65] p. 33)

$$(4) \quad M(x) + N(u)\frac{du}{dx} = 0.$$

Again, Boyce (see [BD65] p. 34-35) provides a “cookbook” method for solving the separable equation. Provided that we can actually evaluate the antiderivatives $\int M(x) dx$ and $\int N(y) dy$, we have the implicit solution

$$(5) \quad H_1(x) + H_2(u) = c,$$

where c is an arbitrary constant and H_1, H_2 are arbitrary antiderivatives of M, N respectively.

Example 1.² Solve the differential equation

$$(6) \quad 2x + u^2 + 2xuu' = 0.$$

Solution: Note that this equation is not linear nor is it separable (it is actually what is known as an *exact* equation), so the methods that we learned to solve linear and separable equations will not work. We need the method to solve exact equations. In particular, note that if we define a function $\psi(x, u) = x^2 + xu^2$, ψ satisfies the conditions that

$$2x + u^2 = \frac{\partial \psi}{\partial x}, \quad \text{and} \quad 2xu = \frac{\partial \psi}{\partial u}.$$

Hence we can rewrite Equation (6) as

$$(7) \quad \frac{\partial}{\partial x}(x^2 + xu^2) + \frac{\partial}{\partial u}(x^2 + xu^2)\frac{du}{dx} = 0.$$

²See [BD65] p. 83.

But in order for this example to really have any meaning, we would want u to be a function of x , so by the chain rule, we can rewrite Equation (7) as

$$\frac{d}{dx}(x^2 + xu^2) = 0.$$

which leads us to the conclusion that $x^2 + xu^2 = c$ for some arbitrary constant c is a solution of Equation (6). \blacklozenge

Notice in the above example it was necessary to find a suitable function ψ ; this is the crux of solving exact differential equations. Nonlinear equations such as the exact equation, can become messy (exact equations are actually very nice nonlinear equations). Perhaps this is the reason that when we initially encounter differential equation (especially nonlinear ones), the equations are very contrived examples that fit into these various categories. We discuss several other methods and categories below. Beside the method of solving exact equations above, another means that can possibly be used to find solutions to a nonlinear equation

$$M(x, u)dx + N(x, u)du = 0$$

is to consider integrating factors as was done in the case of linear equations. This is not really our emphasis, so we continue with another example of a nonlinear equation that also fits well into our box of well-behaved equations.

Definition 7. An equation of the form $\frac{du}{dx} = f(x, u)$ is said to be **homogeneous** whenever the function f does not depend on x and u separately, but only on their ratios; that is, f is actually a function of $\frac{x}{u}$ or $\frac{u}{x}$. Thus homogeneous equations are of the form $\frac{du}{dx} = F\left(\frac{u}{x}\right)$.

The solution to solving homogeneous systems involves a change of variable, as we consider in the following example:

Example 2.³ Solve the differential equation

$$(8) \quad \frac{du}{dx} = \frac{u^2 + 2xu}{x^2}.$$

Solution: By isolating the denominator, we can see that we can really write Equation (8) as

$$\frac{du}{dx} = \left(\frac{u}{x}\right)^2 + 2\frac{u}{x}$$

³See [BD65] p. 91-92.

which shows that the equation is actually homogeneous. While it is homogeneous, it is not linear, separable, nor exact. Thus we must find another method for solving the equation. Here, we consider the substitution $u = xv$, and with this, we can rewrite Equation (8) as

$$(9) \quad x \frac{dv}{dx} = v^2 + v.$$

Notice that Equation (9) clearly has the constant solutions $v = 0, -1$. These give us solutions $u = 0, -x$ of Equation (8). Therefore let us assume that $v \neq 0, v \neq -1$, in which case, we can rewrite Equation (9) using partial fractions as

$$\frac{dx}{x} = \left(\frac{1}{v} - \frac{1}{v+1} \right) dv.$$

This is separable, so we solve it by integrating, arriving at

$$\ln |x| + \ln |c| = \ln |v| - \ln |v+1|$$

where c is the arbitrary constant of integration. Using rules of logarithms and by taking the exponential of each side, we arrive at

$$cx = \frac{v}{v+1}.$$

But, recall that $v = \frac{y}{x}$, so we get the implicit solution

$$cx = \frac{\frac{u}{x}}{\frac{u}{x} + 1} = \frac{u}{u+x}$$

to Equation (8). However, the above is actually nice enough that we can solve for u , namely the general solution

$$(10) \quad u = \frac{cx^2}{1 - cx}.$$

Just as a point of reference, notice that the solutions $u = 0$ and $u = -x$ are included in Equation (10); the former when c is chosen to be 0, and the latter as a limit when $c \rightarrow \pm\infty$. \blacklozenge

We now do a foray into partial differential equations. Very often (as is the case in Bleecker & Csordas [BC96]),⁴ the first named partial differential equation that we see is the heat equation: $u_t = u_{xx}$.⁵ For Bleecker & Csordas this happens at the beginning of Chapter 3. After a slight physical interpretation introduction to the heat equation (where

⁴My thanks to Professor Bleecker for agreeing to be on my committee, and for Professor Csordas for his invigorating teaching.

⁵If one goes precisely off what Bleecker & Csordas have as the heat equation, there is a positive parameter k such that $u_t = ku_{xx}$. By a change of coordinates, we assume without loss of generality that $k = 1$ in our calculations.

the authors actually arrive at the heat equation), we, the readers, are told that the bizarre function

$$u(x, t) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}}, \quad t > 0, -\infty < x < \infty$$

is a special solution to the heat equation called the “Fundamental Source Solution” (see [BC96] p. 124). The authors, and presumably the readers, then show that this is a solution by using logarithmic differentiation. Bleecker & Csordas later derive (see [BC96] p. 460-461) the “Fundamental Source Solution” solution using the method of Fourier transforms.⁶ We also derive this solution in Section (5) using the nonstandard method of symmetry groups. (However, once we’ve established the background, this derivation is really no longer, nor more complicated, than that using the more standard method of Fourier transforms.)

Following this, Bleecker & Csordas introduce us to what is presumably (at least from reading them) the most common analytical way of solving partial differential equations, and that is to use the method of *Separation of Variables* by seeking product solutions. In separation of variables we seek solutions of the form $u(x, t) = X(x)T(t)$. Applying this to the heat equation, we get the result that

$$\frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)}.$$

But, since the left hand side is a function of t and the right hand side is a function of x , the only way this can happen is if they are a constant, c . We are then left with two ordinary differential equations: $T'(t) - cT(t) = 0$, and $X''(x) - cX(x) = 0$, which we solve using ordinary differential equation methods, giving us the following solutions.

- (1) If $c = -\lambda^2 < 0$, then the product solution is of the form:

$$u(x, t) = e^{-\lambda^2 t} (c_1 \sin(\lambda x) + c_2 \cos(\lambda x)).$$

- (2) If $c = \lambda^2 > 0$, then the product solution is of the form:

$$u(x, t) = e^{\lambda^2 t} (c_1 e^{\lambda x} + c_2 e^{-\lambda x}).$$

⁶In Exercise 13 p. 139, Bleecker/Csordas also describe a way by which the “Fundamental Source Solution” can be derived using a probabilistic method similar to a random walk; however, this method is rather complicated—it even involves using Sterling’s formula. There are additional means by which this can be derived; however, our goal here is neither to discover all means by which one can derive the “Fundamental Source Solution,” nor the easiest.

(3) If $c = 0$, then the product solution is of the form:

$$u(x, t) = c_1x + c_2.$$

As the reader progresses through Bleecker & Csordas, we begin to see the work of Fourier applied to partial differential equations. In this process, one combines the use of product solutions and the superposition principle (that a linear combination of solutions are solutions) in order to arrive at solutions that can be represented by a finite sum of product solutions. In most applications, we really want to know what happens when there are certain specified initial conditions. In the case previously mentioned, the initial conditions of the finite sum of product solutions are trigonometric polynomials. Now suppose that we have an initial condition that is an arbitrary function (it need not even be continuous). Fourier's goal was to show that this arbitrary function could be approximated by a sum (not necessarily finite) of these trigonometric polynomials. This in fact works very well; so much so that we can actually represent an arbitrary function (serving as an initial condition) by a Fourier series (which is an infinite sum of the trigonometric polynomials). We can then show that a sequence of solutions coming from partial sums converges to an actual solution (this is done merely formally). As we proceed into the later part of Bleecker & Csordas, we get to perhaps the most direct analytical means to solve partial differential equations and that is to use Fourier transforms (as mentioned earlier, it is by means of Fourier transforms that one can derive the "Fundamental Source Solution").

4. LIE GROUPS AND DIFFERENTIAL GEOMETRY

While it is not our emphasis here, we need to do a slight overview of some topics in Differential Geometry and Lie Groups, since these will be important to our understanding of the culmination of Lie's interpretation of how to solve differential equations. The basic structure is a manifold (see [Olv93] p. 3).

Definition 8. An *m -dimensional manifold* is a nonempty set M , together with a countable collection of subsets $U_\alpha \subset M$, called *coordinate charts*, and one-to-one functions $\chi_\alpha : U_\alpha \rightarrow V_\alpha$ onto open connected subsets $V_\alpha \subset \mathbb{R}^m$, called *local coordinate maps* which satisfy the following properties.

- (1) The coordinate charts cover M ; that is

$$\bigcup_{\alpha} U_{\alpha} = M.$$

- (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 (meaning infinitely differentiable) function.

- (3) If $x \in U_\alpha, y \in U_\beta$ are distinct points in M , then there exist open subsets $W \subset V_\alpha, Y \subset V_\beta$, with $\chi_\alpha(x) \in W, \chi_\beta(y) \in Y$, satisfying

$$\chi_\alpha^{-1}(W) \cap \chi_\beta^{-1}(Y) = \emptyset.$$

Once we have manifolds (which we will, by means of the coordinate maps, consider to be essentially some subset of Euclidean space, and we also consider to be "smooth"), we can define (see [Olv93] p. 7-8) the following.

Definition 9. If M and N are smooth manifolds a map $F : M \rightarrow N$ is said to be *smooth* if its local coordinate expression is a smooth map in every coordinate chart.

Definition 10. Let $F : M \rightarrow N$ be a smooth mapping from an m -dimensional manifold M to an n -dimensional manifold N . The *rank* of F at a point $x \in M$ is the rank of the $n \times m$ Jacobian matrix $(\partial F^i / \partial x^j)$ at x , where $y = F(x)$ is expressed in any local coordinates near x . The mapping F is of *maximal rank* on a subset $S \subset M$ if

for each $x \in S$, the rank of F is as large as possible (i.e. the minimum of m and n).

Note that the rank mentioned in Definition 10 is independent of the choice of coordinates on the manifolds.

We now move on to the important concept of vectors in Differential Geometry. Connected with each point $x \in M$ are tangent vectors (basically in the usual calculus sense). We consider these in the following definitions.

Definition 11. The collection of all tangent vectors to all possible curves passing through a given point $x \in M$ is called the *tangent space to M at x* .

Notation 2. We denote the tangent space to M at x by $TM|_x$.

Note that if M is an m -dimensional manifold, then $TM|_x$ is an m -dimensional vector space with basis $\{\partial/\partial x^1, \dots, \partial/\partial x^m\}$ in the local coordinates.

Definition 12. The set

$$TM := \bigcup_{x \in M} TM|_x$$

is the *tangent bundle of M* . This set has a natural manifold structure of dimension $2m$.

One of the most important concepts that we will need is that of a *vector field*.

Definition 13. A *vector field \mathbf{v} on M* is a function that assigns a tangent vector $\mathbf{v}|_x \in TM|_x$ to each point $x \in M$, where $\mathbf{v}|_x$ varies smoothly from point to point.

We will follow Olver's (see [Olv93] p. 26) notation for a vector field, namely if we have local coordinates (x^1, x^2, \dots, x^m) , the vector field has the form

$$\mathbf{v}|_x = \xi^1(x) \frac{\partial}{\partial x^1} + \xi^2(x) \frac{\partial}{\partial x^2} + \dots + \xi^m(x) \frac{\partial}{\partial x^m}$$

where each $\xi^i(x)$ is a smooth function of x (and we are using superscripts as indices).

Proposition 1. *Vector fields are first order linear differential operators on smooth functions.*

One of the primary operations on vector fields that we will use is known as the Lie Bracket.

Definition 14. If \mathbf{v}, \mathbf{w} are vector fields on M , then their *Lie Bracket*, which we denote by $[\mathbf{v}, \mathbf{w}]$, is the unique vector field satisfying

$$[\mathbf{v}, \mathbf{w}] \cdot (f) = \mathbf{v}(\mathbf{w} \cdot (f)) - \mathbf{w}(\mathbf{v} \cdot (f))$$

for all smooth functions $f : M \rightarrow \mathbb{R}$.

Proposition 2. *The Lie Bracket satisfies the following properties.*

- (1) Bilinearity: *If $\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_1, \mathbf{w}_2$ are vector fields and c_1, c_2, c_3, c_4 are constants, then*

$$[c_1\mathbf{v}_1 + c_2\mathbf{v}_2, c_3\mathbf{w}_1 + c_4\mathbf{w}_2] = c_1c_3[\mathbf{v}_1, \mathbf{w}_1] + c_1c_4[\mathbf{v}_1, \mathbf{w}_2] + c_2c_3[\mathbf{v}_2, \mathbf{w}_1] + c_2c_4[\mathbf{v}_2, \mathbf{w}_2].$$

- (2) Skew-Symmetry: *If \mathbf{v}, \mathbf{w} are vector fields, then*

$$[\mathbf{v}, \mathbf{w}] = -[\mathbf{w}, \mathbf{v}].$$

- (3) Jacobi Identity: *If $\mathbf{v}, \mathbf{w}, \mathbf{u}$ are vector fields, then*

$$[\mathbf{u}, [\mathbf{v}, \mathbf{w}]] + [\mathbf{w}, [\mathbf{u}, \mathbf{v}]] + [\mathbf{v}, [\mathbf{w}, \mathbf{u}]] = 0.$$

Now suppose that we have two vector fields, $a(x, y)\frac{\partial}{\partial x}$ and $b(x, y)\frac{\partial}{\partial y}$, then we can find their Lie Bracket by

$$(11) \quad \left[a(x, y)\frac{\partial}{\partial x}, b(x, y)\frac{\partial}{\partial y} \right] = a(x, y)\frac{\partial}{\partial x}b(x, y)\frac{\partial}{\partial y} - b(x, y)\frac{\partial}{\partial y}a(x, y)\frac{\partial}{\partial x}.$$

Sometimes it may be helpful to consider vector fields and tangent bundles in the sense of category theory (this may be helpful to those who are more familiar with the algebraic structure of categories and less familiar with the differential geometry structures). In the sense of category theory, we consider a covariant functor, T , from the category whose objects are manifolds and whose morphisms are smooth maps between manifolds to the category whose objects are bundles and whose morphisms are bundle maps. The application of T to $\phi : M \rightarrow N$ is indicated in the following commutative diagram.

$$\begin{array}{ccc} TM & \xrightarrow{T(\phi)} & TN \\ \pi_M \downarrow & & \downarrow \pi_N \\ M & \xrightarrow{\phi} & N \end{array}$$

We can then define a vector field to be a special function, $\mathbf{v} : M \rightarrow TM$ such that the $\pi_M \circ \mathbf{v}$ is the identity map on M . (We may also call \mathbf{v} a *section of the bundle*.) Furthermore, the map $T(\phi)$ is the *differential* of the map ϕ .

One important theorem in vector fields and coordinates is what is known as the “Flow Box Theorem” (see [Olv93] p. 30).

Theorem 6. Flow Box Theorem. *Let*

$$\mathbf{v}|_x = \xi^1(x) \frac{\partial}{\partial x^1} + \xi^2(x) \frac{\partial}{\partial x^2} + \cdots + \xi^n(x) \frac{\partial}{\partial x^n}$$

be a vector field. Let $m \in M$ (where M , of course, is a manifold) such that $\mathbf{v}|_x(m) \neq 0$. Then there is a local coordinate chart $y = (y^1, y^2, \dots, y^n)$ at m such that in these coordinates, we have that $\mathbf{v} = \frac{\partial}{\partial y^1}$.

What this theorem does is enable us to make a suitable change of coordinates and make the vector field much easier to deal with; in effect, we can straighten out the flow. (We haven’t defined flow here, but one can essentially think of it as the path a particle would take through a vector field as encountered in an introductory differential equations course.)

Now that we have at least some definitions available we can finally define a Lie Group (see [Olv93] p. 15).

Definition 15. An *r -parameter Lie group* is a group G which also carries the structure of an r -dimensional manifold in such a way as both the group operation

$$m : G \times G \rightarrow G, \quad m(g, h) = g \cdot h, \quad g, h \in G,$$

and the inversion

$$i : G \rightarrow G, \quad i(g) = g^{-1}, \quad g \in G,$$

are smooth maps between manifolds.

Almost all properties we will be interested in are actually local. Hence we are usually interested only in a local Lie Group (see [Olv93] p. 18).

Definition 16. An *r -parameter local Lie group* consists of open subsets $V_0 \subset V \subset \mathbb{R}^r$ containing the origin 0, and smooth maps

$$m : V \times V \rightarrow \mathbb{R}^r,$$

defining the group operation, and

$$i : V_0 \rightarrow V$$

defining the group inversion which has the following properties.

(1) *Associativity*: If $x, y, z \in V$, and $m(x, y), m(y, z) \in V$, then

$$m(x, m(y, z)) = m(m(x, y), z).$$

(2) *Identity Element*: For all $x \in V$, $m(0, x) = x = m(x, 0)$.

(3) *Inverses*: For each $x \in V_0$, $m(x, i(x)) = 0 = m(i(x), x)$.

The previous two definitions have to do with a modern interpretation. However when Lie would consider a group, he actually would be considering what we could call a Lie Algebra, not what has today become known as a Lie Group. We will adopt the more modern terminology. Throughout this, let G be a Lie Group with right multiplication map $R_g : G \rightarrow G$ defined by $R_g(h) = hg$ for all $h \in G$. We then say that \mathbf{v} is a *right invariant vector field* if $T(R_g) \circ \mathbf{v} \circ R_{g^{-1}} = \mathbf{v}$ for all $g \in G$.

Proposition 3. *The Lie Bracket of two right invariant vector fields is right invariant.*

The proof of this proposition follows directly from appropriate applications of definitions. But, once we have it, we can define a Lie Algebra.

Definition 17. A *Lie algebra* is a vector space \mathfrak{g} together with the operation of the Lie Bracket (c.f. Definition 14) satisfying conditions 1, 2, and 3 of Proposition 2.

An example of a Lie Algebra is the vector space of all right invariant vector fields on G with the operation of the Lie Bracket. One thing to note is that we often do not consider \mathfrak{g} as we have defined it. The way we consider the Lie Algebra is actually that $\mathfrak{g} = TG|_e$, that is, it is the tangent space to the identity. One considers the set of vector fields to be the infinitesimal generators of a Lie Algebra, so suppose that $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ forms a basis for \mathfrak{g} . Then by closure, $[\mathbf{v}_i, \mathbf{v}_j] \in \mathfrak{g}$.

Definition 18. From this, we arrive at some very important constants, c_{ij}^k , $i, j, k = 1, \dots, r$, the so called *structure constants*. They satisfy the relationship

$$[\mathbf{v}_i, \mathbf{v}_j] = \sum_{k=1}^r c_{ij}^k \mathbf{v}_k, \quad i, j = 1, \dots, r.$$

Proposition 4. *From the properties of the Lie Bracket, we can clearly see that the structure constants also satisfy the following properties.*

(1) *Skew Symmetry*: $c_{ij}^k = -c_{ji}^k$, and

(2) Jacobi Identity:

$$\sum_{k=1}^r (c_{ij}^k c_{kl}^m + c_{li}^k c_{kj}^m + c_{jl}^k c_{ki}^m) = 0 \quad m = 1, \dots, r.$$

Definition 19. A Lie Algebra \mathfrak{g} is a *abelian* if $[\mathbf{v}, \mathbf{w}] = 0$ for every $\mathbf{v}, \mathbf{w} \in \mathfrak{g}$.

In fact, we can actually completely reduce our study to any one of the following: a Lie Group G , a Lie Algebra \mathfrak{g} , or a set of structure constants. For instance a set of constants satisfy the results of Proposition 4 if and only if they are the structure constants for some Lie Algebra \mathfrak{g} . Similarly given any Lie Group, we can find a Lie Algebra, and given any Lie Algebra, \mathfrak{h} , we can find a Lie Group that has \mathfrak{h} as its Lie Algebra.⁷

Before we start talking about how Lie saw differential equations, we need more concepts. The first of these is *prolongation*. While prolongation is actually something that can occur with diffeomorphisms, vector fields, group actions and flows, we will, for the sake of some brevity, primarily focus on the prolongation of a vector field. Since a vector field is really a function, we first introduce the idea of the prolongation of a function. We then extend this to the idea of the prolongation of a vector field. However, we need to build up some background before we can get to prolongation. Here, we consider the manifold to simply be Euclidean space.

Let $f : X \rightarrow U$ be a smooth function from $X \cong \mathbb{R}^p$ to $U \cong \mathbb{R}^q$. (Notice how this is related to our general differential equation in Notation 1.) Now let $J = (j_1, \dots, j_k)$ be an unordered k -tuple of integers with entries $1 \leq j_k \leq p$ indicating which derivatives are being taken. The order, k , indicates how many derivatives are being taken. Since f has p independent variables there are

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

different k^{th} order partial derivatives of f .

Notation 3. We denote these derivatives by

$$u_J^\alpha = \partial_J f^\alpha(x) = \frac{\partial^k f^\alpha(x)}{\partial x^{j_1} \partial x^{j_2} \dots \partial x^{j_k}}$$

where $\alpha = 1, \dots, q$.

⁷This is a slight simplification of matters for there are topological concerns primarily dealing with the need for the Lie group to be simply-connected. See **Theorem 1.54** in [Olv93] p. 47 and comments in [Olv93] p. 50.

Note that there are $q \cdot p_k$ such derivatives.

Notation 4. Let us denote the Euclidean space $\mathbb{R}^{q \cdot p_k}$ by U_k .

Furthermore, let us endow the space U_k with coordinates u_j^α which represent the derivatives in Notation 3.

Notation 5. Set $U^{(n)} := U \times U_1 \times \cdots \times U_n$ to be the cartesian product space. The coordinates of the space $U^{(n)}$ represent all the derivatives of functions $u = f(x)$ of all orders from 0 to n .

Note that $U^{(n)}$ is a Euclidean space with dimension $q \binom{p+n}{n} \equiv qp^{(n)}$.

Notation 6. We will denote by $u^{(n)}$ a typical point in the space $U^{(n)}$; the components of $u^{(n)}$ are the aforementioned u_j^α where $0 \leq k \leq n$.

Definition 20. Let $u = f(x)$ be a smooth function from X to U . Then there is an induced function, $u^{(n)} = \text{pr}^{(n)} f(x)$, called the n^{th} **prolongation of f** , which is defined by the equations $u_j^\alpha = \partial_j f^\alpha(x)$.

Example 3. Suppose we have the heat equation

$$(12) \quad u_t - u_{xx} = 0.$$

Here, note that there are two independent variables, x, t , one dependent variable, u , so we have the underlying space $X \times U \cong \mathbb{R}^3$ with coordinates (x, t, u) . Since Equation (12) contains a second order partial derivative, we will want to consider the prolongation to the 2-jet, namely to the set $X \times U^{(2)}$, which has coordinates $(x, t, u, u_x, u_t, u_{xx}, u_{xt}, u_{tt})$. We return to solutions involving this example later. \blacklozenge

Definition 21. The set $X \times U^{(n)}$ is the n^{th} **order jet space** of the underlying space $X \times U$.

In considering prolongations of functions, the key idea is to *identify the function u with its graph in $X \times U$* . We can now generalize the idea to that of a vector field, \mathbf{v} . Note that a vector field will operate on the space $X \times U \cong M$, and it will prolong up to a function operating on $X \times U^{(n)} \cong M^{(n)}$.

$$\begin{array}{ccc} M^{(n)} & \xrightarrow{\text{pr}^{(n)} \mathbf{v}} & TM^{(n)} \\ \pi \downarrow & & \downarrow T(\pi) \\ M & \xrightarrow{\mathbf{v}} & T \end{array}$$

The prolongation is a lifting of the vector field in the sense that the above diagram commutes, i.e. $\mathbf{v} \circ \pi = T(\pi) \circ \text{pr}^{(n)} \mathbf{v}$.

Here we shall calculate a simple prolongation of a vector field. (We do another calculation in Example 6, where there are more independent variables.)

Example 4. Find the most general 1st prolongation of a vector field over a $X \times U$ where there is only one independent variable, in other words, find $\text{pr}^{(1)} \mathbf{v}$ where

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

To solve this, first note that the prolongation will be a vector field

$$\text{pr}^{(1)} \mathbf{v} = \xi(x, u) \frac{\partial}{\partial x} + \phi(x, u) \frac{\partial}{\partial u} + \phi^x(x, u, u_x) \frac{\partial}{\partial u_x}.$$

By the lifting and covering mentioned above, all we need to do is to find the coefficient function ϕ^x (we assume that we are given the functions ξ and ϕ). For notation, we will follow Olver and have the superscripts of the function ϕ match with the basis element of the coordinate map. We can solve this several ways. Here we find the prolongation using Lie Derivatives and the invariance of a certain 1-form. (We use an alternative method in Example 6.)

Definition 22. Let \mathbf{v} be a vector field on M with flow φ_ϵ and σ a vector field or differential form defined on M . The **Lie Derivative** of σ with respect to \mathbf{v} is the object whose value at $x \in M$ is

$$(\mathcal{L}_{\mathbf{v}}\sigma)|_x = \lim_{\epsilon \rightarrow 0} \frac{\varphi_\epsilon^*(\sigma|_{\varphi_\epsilon(x)}) - \sigma|_x}{\epsilon}.$$

This definition is even more ungainly than the corresponding definition for regular derivatives which nearly every Calculus 1 student has difficulty grasping. And, as in the case of Calculus 1, where soon the student learns the various rules for differentiating a function of a real variable, we skip ahead to some important properties of the Lie Derivative.

- (1) The Lie Derivative is linear over the reals (as any good derivative should be).
- (2) If f is a function (note that a function is a differential form), then $\mathcal{L}_{\mathbf{v}}f = \mathbf{v} \cdot f$; that is, apply the vector field to the function f .
- (3) If df is the differential of a function f , then $\mathcal{L}_{\mathbf{v}}df = d(\mathcal{L}_{\mathbf{v}}f) = d(\mathbf{v} \cdot f)$.

Now, suppose that $u = f(x)$ is a function with graph $\gamma(x) = (x, u(x))$. Then $(\text{pr}^{(1)} \gamma) \cdot (x) = (x, u(x), u_x(x))$. Consider the 1-form $du - u_1 dx$ (where $u_1 = \frac{du}{dx}$). Let $\alpha(x) = (x, u(x), u_x(x))$ be a graph in the 1-jets

for an arbitrary α . Then let γ^* denote the restriction of α to the graph γ . Consider

$$\begin{aligned}\gamma^*(du - u_1 dx) &= d(u(x)) - \frac{du}{dx} dx \\ &= \frac{du}{dx} dx - \frac{du}{dx} dx \\ &= 0.\end{aligned}$$

We can, therefore, see that the differential form $du - u_1 dx$ is “special” in some way. In fact, we want to consider the $\text{Span}\{du - u_1 dx\}$, which is known as the *contact system*. One important property of *prolongation* is that it preserves this contact system. Hence, to find the value of ϕ^x we examine

$$\mathcal{L}_{\text{pr}^{(1)} \mathbf{v}}(du - u_1 dx) \equiv 0 \pmod{du - u_1 dx}.$$

Consider the following.

$$\begin{aligned}\mathcal{L}_{\text{pr}^{(1)} \mathbf{v}}(du - u_1 dx) &= \mathcal{L}_{\text{pr}^{(1)} \mathbf{v}}(du) - \mathcal{L}_{\text{pr}^{(1)} \mathbf{v}}(u_1 dx) \\ &= \mathcal{L}_{\text{pr}^{(1)} \mathbf{v}}(du) - \left(u_1 \mathcal{L}_{\text{pr}^{(1)} \mathbf{v}}(dx) + dx \mathcal{L}_{\text{pr}^{(1)} \mathbf{v}}(u_1) \right) \\ &= d\mathcal{L}_{\text{pr}^{(1)} \mathbf{v}}(u) - \left(u_1 d\mathcal{L}_{\text{pr}^{(1)} \mathbf{v}}(x) + dx \mathcal{L}_{\text{pr}^{(1)} \mathbf{v}}(u_1) \right) \\ &= d(\text{pr}^{(1)} \mathbf{v} \cdot u) - u_1 d(\text{pr}^{(1)} \mathbf{v} \cdot x) - dx(\text{pr}^{(1)} \mathbf{v} \cdot u_1).\end{aligned}$$

We must now perform some side calculations.

$$\text{pr}^{(1)} \mathbf{v} \cdot u = \xi \frac{\partial u}{\partial x} + \phi \frac{\partial u}{\partial u} + \phi^x \frac{\partial u}{\partial u_x} = \phi.$$

Similarly, we have that $\text{pr}^{(1)} \mathbf{v} \cdot x = \xi$, $\text{pr}^{(1)} \mathbf{v} \cdot u_1 = \phi^x$. Hence we have that

$$\mathcal{L}_{\text{pr}^{(1)} \mathbf{v}}(du - u_1 dx) = d\phi - u_1 d\xi - \phi^x dx.$$

Note that $d\phi = \frac{\partial \phi}{\partial x} dx + \frac{\partial \phi}{\partial u} du$ and $d\xi = \frac{\partial \xi}{\partial x} dx + \frac{\partial \xi}{\partial u} du$, and making the modulus substitution $du = u_1 dx$, we can rewrite

$$0 \equiv \mathcal{L}_{\text{pr}^{(1)} \mathbf{v}}(du - u_1 dx) \equiv \left(\frac{\partial \phi}{\partial x} dx + u_1 \frac{\partial \phi}{\partial u} dx \right) - u_1 \left(\frac{\partial \xi}{\partial x} dx + u_1 \frac{\partial \xi}{\partial u} dx \right) - \phi^x dx$$

which, solving we get

$$(13) \quad \phi^x(x, u) = \frac{\partial \phi}{\partial x} + u_1 \frac{\partial \phi}{\partial u} - u_1 \frac{\partial \xi}{\partial x} - u_1^2 \frac{\partial \xi}{\partial u}. \quad \blacklozenge$$

The calculation of prolongations can become quite tedious, so we do not compute more of them here. However, the prolongations of higher order are found recursively; we will not go through that process yet (see Example 6 where we calculate the second prolongation). Here we merely note several important coefficient functions of prolongations. (For these calculations, see [Olv93] p. 114.) Here we look at the case where there are two independent variables, namely we have a vector field over $X \times U \cong \mathbb{R}^2 \times \mathbb{R}$ which has the general form

$$(14) \quad \mathbf{v} = \xi(x, t, u) \frac{\partial}{\partial x} + \tau(x, t, u) \frac{\partial}{\partial t} + \phi(x, t, u) \frac{\partial}{\partial u}.$$

Recalling what happens when we make prolongations, we can see directly that (here, for the sake of space, we remove the arguments from the coefficient functions, which is a practice we will continue to follow provided it does not cause confusion)

$$(15) \quad \text{pr}^{(1)} \mathbf{v} = \xi \frac{\partial}{\partial x} + \tau \frac{\partial}{\partial t} + \phi \frac{\partial}{\partial u} + \phi^x \frac{\partial}{\partial u_x} + \phi^t \frac{\partial}{\partial u_t}$$

and that

$$(16) \quad \text{pr}^{(2)} \mathbf{v} = \xi \frac{\partial}{\partial x} + \tau \frac{\partial}{\partial t} + \phi \frac{\partial}{\partial u} + \phi^x \frac{\partial}{\partial u_x} + \phi^t \frac{\partial}{\partial u_t} + \phi^{xx} \frac{\partial}{\partial u_{xx}} + \phi^{xt} \frac{\partial}{\partial u_{xt}} + \phi^{tt} \frac{\partial}{\partial u_{tt}}.$$

In Equations (15) and (16), we have that

$$(17) \quad \phi^x = \phi_x + (\phi_u - \xi_x)u_x - \tau_x u_t - \xi_u u_x^2 - \tau_u u_x u_t,$$

$$(18) \quad \phi^t = \phi_t - \xi_t u_x + (\phi_u - \tau_t)u_t - \xi_u u_x u_t - \tau_u u_t^2,$$

$$(19) \quad \begin{aligned} \phi^{xx} = & \phi_{xx} + (2\phi_{xu} - \xi_{xx})u_x - \tau_{xx}u_t + (\phi_{uu} - 2\xi_{xx})u_x^2 - 2\tau_{xu}u_x u_t - \xi_{uu}u_x^3 \\ & - \tau_{uu}u_x^2 u_t + (\phi_u - 2\xi_x)u_{xx} - 2\tau_x u_{xt} - 3\xi_u u_x u_{xx} - \tau_u u_t u_{xx} - 2\tau_u u_x u_{xt}. \end{aligned}$$

We will return to these when we discuss the heat equation.

5. SOPHUS LIE AND DIFFERENTIAL EQUATIONS

Recall from Section (2) that Sophus Lie was at heart a geometer, and it was through this lens that he viewed much of his work. We had previously mentioned (again see Section (2)) that Lie had a close association with the German mathematician Felix Klein (who was a student of Julius Plücker), and it was through studies of Plücker’s line geometry that Klein and Lie’s association began.⁸ While it is not completely relevant to the topic, we will spend a little bit of time dealing with this initial topic of Klein and Lie’s since we see a similar approach later in Lie’s approach to differential equations.

Let Δ denote a fixed tetrahedron, with faces determined by planes $\pi_1, \pi_2, \pi_3, \pi_4$. Then we clearly have that each ℓ in $\mathbb{P}^3(\mathbb{C})$ meets Δ in four places. A *tetrahedral line complex* \mathcal{T} is the set of lines ℓ for which the cross ratio of these points is the same.

Notation 7. Let \mathcal{B} denote the totality of all projective transformations of the space that fixes the vertices of the tetrahedron Δ . We essentially think of the elements of \mathcal{B} as coordinate changes.

What was unique (and for our sake informative) about Lie’s approach to these tetrahedral line complexes was (see [Haw00] p. 6) that to him a tetrahedral complex was the orbit (in the sense of a group action) of some fixed line under the transformations of \mathcal{B} . For us, this is informative since he wanted to solve differential equations in a similar way.

Lie’s study of differential equations in this sense became his *Idée Fixe*. To get into this, let us suppose that we have a point p in our space. From this point p , we can arrive at the *complex cone with vertex* p , which we will denote by $\mathcal{C}(p)$. This is the set of all lines of \mathcal{T} that pass through the point p .

Problem: Determine all surfaces S with the property that at each point $p \in S$, the cone $\mathcal{C}(p)$ touches S at p . In other words, the tangent plane to S meets the cone $\mathcal{C}(p)$ in exactly one straight line.

⁸One important connection was Klein’s Erlanger Programm which “consisted in implicitly regarding all the geometries or geometrical modes of treatment...[as being] completely determined by specifying a manifold of elements and a distinguished set of transformations acting on the elements of the manifold and defining the permissible geometrical operations” (see [Haw00] p. 35). One then considers those elements of the manifold that are invariant under the operations.

Since the geometrical condition on S can be translated into an equation of the form

$$(20) \quad f(x, y, z, p, q) = 0, \quad p = \frac{\partial z}{\partial x}, \quad q = \frac{\partial z}{\partial y},$$

the solution to this problem (according to Lie) (see [Haw00] p. 21-22), amounts to solving a first order partial differential equation. We can clearly note that geometrically any $T \in \mathcal{B}$ will take one solution surface into another; Lie would later describe this as saying that *the differential equation admits the transformations of \mathcal{B}* . Lie's actual solution to the problem involved the maps

$$X = \log x, \quad Y = \log y, \quad Z = \log z.$$

But since the differential equation noted admits all elements of \mathcal{B} , Lie concluded that the differential equation could actually be reduced to one of the form $f(p, q) = 0$ which could be solved directly. His reason for this is contained in Theorem 7.

When Lie shared this result with Klein (who apparently had more algebraic knowledge or possibly interest), he noticed that there was an analogy between Lie's method of integrating differential equations admitting a 3-parameter group of commuting transformations (as in the problem above) and in Abel's work on abelian polynomial equations. Recall that in polynomial Galois theory the elements of the Galois group permute the roots of the associated polynomial; that is, they take solutions into solutions. The above problem and Lie's solution to it seemed to indicate a similar result may happen here: knowing something about the group admitted yields information about the solution.

We can see another example of this in the case of Equation (8), the homogeneous equation example mentioned in Section (3), which admits the 1-parameter group of commuting transformations $x' = ax$ and $y' = ay$. (That is, the homogeneous equation admits a scaling symmetry—see Example 7.) Lie felt this was the reason we could solve Equation (8) by using a change of variables. With these in hand, we see that Lie felt that “knowledge of transformations, ‘finite’ or infinitesimal, admitted by a differential equation was the true basis for its solution or at least for its reduction to a simpler equation...It thus seemed plausible that a bona fide theory of differential equations admitting known, commuting, infinitesimal transformations could be developed that would serve a function analogous to that served by Galois's theory for polynomial equations.”⁹ To this end, Lie came up with the following theorem (see [Haw00] p. 24).

⁹See [Haw00] p. 24.

Theorem 7. Lie, ca. 1870.

- (1) *If $f(x, y, z, p, q) = 0$ admits three commuting infinitesimal projective transformations, then it can be transformed into $F(P, Q) = 0$.*
- (2) *If $f(x, y, z, p, q) = 0$ admits two commuting infinitesimal projective transformations, then it can be transformed into $F(Z, P, Q) = 0$.*
- (3) *If $f(x, y, z, p, q) = 0$ admits one commuting infinitesimal projective transformation, then it can be transformed into an equation without the third coordinate $F(X, Y, P, Q) = 0$.*

As is the case with much of Lie's work, his proof to this theorem was not very rigorous (such as obtaining an unlimited number of differential equations for which it applied). The theorem also has a number of problems with its lack of generality. Lie's *idée fixe* was that the results of this theorem could be expanded and developed on a far greater scale and generality than implied in this theorem. However, because of the vastness of differential equations (compared to polynomial equations) the full extent of Lie's *idée fixe* has yet to be realized. But Lie's interest in differential equations was kindled, and many mathematicians since Lie have attempted to expand on his *idé fixe*.

As we saw above, Lie became fascinated with differential equations. We would like to take a little bit of time to go over other motivations of his as well as knowledge that he would have had of differential equations. We had previously mentioned the geometric impetus given to Lie by Felix Klein. Much of Lie's formative knowledge on differential equations came from an essay written by V. G. Imschenetsky in 1869. Imschenetsky was inspired by the work of one of the greatest mathematicians of all time, Carl Gustav Jacob Jacobi. Jacobi, who did his work primarily in the 1830's, did not publish many of his results, for the results were published posthumously by Clebsch in the 1860's. It was Jacobi, "the quintessential analyst, [who] provided the analytical framework for Lie's geometrical ideas."¹⁰ As a basis for Jacobi's (and hence Lie's) work, we start with a result of Lagrange (see [Haw00] p. 45) that is essentially the Flow Box Theorem (c.f. Theorem 6).

Theorem 8. Lagrange's Theorem. *The problem of determining all solutions $f = f(x_1, \dots, x_n)$ to a linear homogeneous equation*

$$\xi_1 \frac{\partial f}{\partial x_1} + \dots + \xi_n \frac{\partial f}{\partial x_n} = 0, \quad \xi_i = \xi_i(x_1, \dots, x_n)$$

¹⁰See [Haw00] p. 46.

is equivalent to the problem of completely integrating (that is, to solve with arbitrary initial conditions) the system of ordinary differential equations

$$\frac{dx_i}{dt} = \xi_i(x_1, \dots, x_n), \quad i = 1, \dots, n.$$

Moreover, every solution is of the form $f = \Theta(f_1, \dots, f_{n-1})$ where f_1, \dots, f_{n-1} are functionally independent solutions, and Θ is arbitrary.

Jacobi further expanded this, and reintroduced the concept of a Poisson (or Possion-Jacobi) bracket, in the following theorem (see [Haw00] p. 47) which deals with the abelian case (c.f. Definition 19).

Theorem 9. *A partial differential equation $F_1(x, p) = 0$ can be integrated if $n - 1$ functions F_2, \dots, F_n of the $2n$ variables $(x, p) = (x_1, \dots, x_n, p_1, \dots, p_n)$ can be determined such that F_1, \dots, F_n are functionally independent and satisfy the relations $(F_i, F_j) = 0$ for all i, j where for any functions G, H of (x, p) ,*

$$(21) \quad (G, H) = \sum_{i=1}^n \left(\frac{\partial G}{\partial p_i} \frac{\partial H}{\partial x_i} - \frac{\partial G}{\partial x_i} \frac{\partial H}{\partial p_i} \right).$$

(Note that in the above, we essentially want to consider the p_i to be a derivative of a function that is dependent on the x_i , but the differential equation is not dependent on this function.)

Jacobi, by means of Theorem 9, was able to eliminate the problem of finding all information at once. His method, which enabled one to find the F_i 's by successively solving systems of linear homogeneous partial differential equations, simplified the classical problem. "The problem of integrating $F_1(x, p) = 0$ reduces to the problem of determining (in the worst possible case) one solution to each of the following systems of ordinary differential equation: 1 system of order $2n - 2$, 2 systems of order $2n - 4$, 3 systems of equations of order $2n - 6, \dots, n - 1$ systems of order 2."¹¹ This result was later refined by Weiler in 1863 and Clebsch in 1866 to "The integration of $F(x, p) = 0$ can be reduced to obtaining one solution to one system of order $2n - 2$ and two systems each of order $2n - 4, 2n - 6, \dots, 2$."¹²

The result was made even better by Adolph Mayer and Lie in 1872.

Theorem 10. Mayer & Lie 1872 (separately). *The integration of $F(x, p) = 0$ can be reduced to obtaining one solution to one system of orders $2n - 2, 2n - 4, \dots, 2$.*

¹¹See [Haw00] p. 48.

¹²See [Haw00] p. 48.

Adolph Mayer (whose specialty was in differential equations) was much more of an analyst than Lie. Mayer arrived at the above result using analytical means; Lie, by his somewhat typically inspirational means, arrived at the result geometrically. It was, however, through Mayer that Lie gained the Jacobian analytical view of differential equations.

The above mentioned results were critical in Lie’s realization of his *idée fixe*: These theorems, particularly the Mayer/Lie Theorem, made more specific the manner in which knowledge that a differential equation admits a group of transformations should yield information about its integration by translating into the number and order of systems of ordinary differential equations needed to integrate the original partial differential equation. In this sense, we are finally beginning to see the connection between the way Lie saw differential equations and the results for Galois theory of polynomials. “*The system of ordinary differential equations played somewhat the role of the auxiliary polynomial equations associated to the composition series of the Galois group of a polynomial equation and utilized to analyze its algebraic resolution.*”¹³ Lie needed a little bit more before he could really do this—he needed to translate group related information into auxiliary systems.

Jacobi, as mentioned, failed to publish his later results (largely due to failing health), and also failed to provide justification for his results. Many did not believe his claims regarding the general case of Theorem 9—it is a generalization since it removes the abelian condition—resulting in what is known as “Jacobi’s Problem” (see [Haw00] p. 51).

Jacobi’s Problem. Suppose that $f = F_1, \dots, F_r$ are r functionally independent solutions to $(F_1, f) = 0$ with the property that the bracketing produces no more solutions in the sense that for all i, j , (F_i, F_j) is functionally dependent upon F_1, \dots, F_r , so that functions Ω_{ij} of r variables exist for which

$$(F_i, F_j) = \Omega_{ij}(F_1, \dots, F_r), \quad i, j = 1, \dots, r.$$

How does knowledge of F_1, \dots, F_r simplify the problem of solving $F_1(x, p) = 0$?

While several mathematicians between Jacobi and Lie had dealt with this problem, it was Lie who “was the first to tackle the general problem, inspired by the fact that for him, it was an expression of his *idée fixe*.”¹⁴ In fact, we can even express the problem in terms of the *idée*

¹³See [Haw00] p. 50.

¹⁴See [Haw00] p. 51.

fixe: “Given that the partial differential equation $F_1(x, p) = 0$ admits the group $\mathfrak{g}_{\mathcal{F}}$, how does knowledge of $\mathfrak{g}_{\mathcal{F}}$ simplify the problem of solving $F_1(x, p) = 0$?”¹⁵

In realizing a solution similar to what Lie would have envisioned, we need several key theorems and definitions (see [Olv93] p. 93, 100-101).

Definition 23. Let \mathcal{S} be a system of differential equations. A **symmetry group** of the system \mathcal{S} is a local group of transformations G acting on an open subset M of the space of independent and dependent variables for the system with the property that whenever $u = f(x)$ is a solution of \mathcal{S} , and whenever $g \cdot f$ is defined for $g \in G$, then $u = g \cdot f(x)$ is also a solution of the system.

Notation 8. Let Δ be an n^{th} order differential equation. Denote by \mathcal{S}_{Δ} the subvariety of $X \times U^{(n)}$ that is $\Delta^{-1}(0)$.

Theorem 11. *Let M be an open subset of $X \times U$ and suppose that $\Delta(x, u^{(n)}) = 0$ is an n^{th} order system of differential equations defined over M , with corresponding sub-variety $\mathcal{S}_{\Delta} \subset M^{(n)}$. Suppose that G is a local group of transformations acting on M whose prolongation leaves \mathcal{S}_{Δ} invariant, meaning that whenever $(x, u^{(n)}) \in \mathcal{S}_{\Delta}$, we have $\text{pr}^{(n)} g \cdot (x, u^{(n)}) \in \mathcal{S}_{\Delta}$ for all $g \in G$ such that this is defined. Then G is a symmetry group of the system of differential equations in the sense of definition (23).*

The theorem that we will primarily use is the following (usually with $l = 1$) (see [Olv93] p. 104).

Theorem 12. *Suppose that $\Delta_v(x, u^{(n)}) = 0$ for $v = 1, \dots, l$ is a system of differential equations of maximal rank defined over $M \subset X \times U$. If G is a local group of transformations acting on M and $\text{pr}^{(n)} \mathbf{v}(\Delta_v(x, u^{(n)})) = 0$, $v = 1, \dots, l$ whenever $\Delta(x, u^{(n)}) = 0$ for every infinitesimal generator \mathbf{v} of G , then G is a symmetry group of the system.*

We can also interpret Theorem 12 in terms of the Lie Algebra.

Theorem 13. *G is a symmetry group for Δ if and only if $\mathfrak{g}^{(n)}$ is tangent to \mathcal{S}_{Δ} .*

¹⁵See [Haw00] p. 59.

6. THE HEAT AND WAVE EQUATIONS

Example 5.¹⁶ **The Heat Equation.** Recall that the heat equation (c.f Equation (12)) is $u_t - u_{xx} = 0$. Since there are two independent variables and the differential equation is second order, in order to represent the equation, we want to find the sub-variety of $X \times U^{(2)}$ that is determined by where $\Delta(x, t, u^{(2)}) = u_t - u_{xx}$ vanishes.

Let us now find the symmetry groups for the heat equation, recalling the second prolongation formula for a general vector field on $X \times U$, namely, equation (16),

$$\text{pr}^{(2)} \mathbf{v} = \xi \frac{\partial}{\partial x} + \tau \frac{\partial}{\partial t} + \phi \frac{\partial}{\partial u} + \phi^x \frac{\partial}{\partial u_x} + \phi^t \frac{\partial}{\partial u_t} + \phi^{xx} \frac{\partial}{\partial u_{xx}} + \phi^{xt} \frac{\partial}{\partial u_{xt}} + \phi^{tt} \frac{\partial}{\partial u_{tt}}.$$

We find the infinitesimal condition by means of Theorem 12; that is, we evaluate Equation (16), or in other words we let Equation (16) differentiate, the differential equation $\Delta = u_t - u_{xx}$. What happens here is that u_t picks off the coefficient function, ϕ^t of $\frac{\partial}{\partial u_t}$ in Equation (16), and similarly, u_{xx} picks off the coefficient function ϕ^{xx} of $\frac{\partial}{\partial u_{xx}}$. Hence by Theorem 12, we have that

$$(22) \quad \phi^t - \phi^{xx} = 0 \quad \text{whenever} \quad u_t - u_{xx} = 0.$$

(For future reference, we may want to consider the equivalent condition that $\phi^t = \phi^{xx}$ whenever $u_t = u_{xx}$.)

We are fortunate, for we have equations for ϕ^t and ϕ^{xx} , Equations (18) and (19) respectively. Our task is then to equate coefficients of monomials of partials of u , and use this to come up with various information about the functions ξ, τ, ϕ .

Since anything that satisfies the heat equation must satisfy $u_t = u_{xx}$, we make our restriction in this case by replacing u_t by u_{xx} in Equations (18) and (19), arriving at (where the results are simplified)

$$\phi^t = \phi_t - \xi_t u_x + (\phi_u - \tau_t) u_{xx} - \xi_u u_x u_{xx} - \tau_u u_{xx}^2$$

and

$$\begin{aligned} \phi^{xx} &= \phi_{xx} + (2\phi_{xu} - \xi_{xx}) u_x + (\phi_u - 2\xi_x - \tau_{xx}) u_{xx} + (\phi_{uu} - 2\xi_{xu}) u_x^2 \\ &- (2\tau_{xu} + 3\xi_u) u_x u_{xx} - \xi_{uu} u_x^3 - \tau_{uu} u_x^2 u_{xx} - 2\tau_x u_{xt} - \tau_u u_{xx}^2 - 2\tau_u u_x u_{xt}. \end{aligned}$$

Then, by Equation (22) we have that these two are actually equal. Hence, by equating the coefficients of the partials of u , we have the following partial differential equations:

¹⁶See [Olv93] p. 117-9. (I'm expanding it.)

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

The goal is now to solve these elementary partial differential equations and by so doing, simplify the remaining equations, and thus determine equations for the unknown functions ξ, τ, ϕ . This will enable us to find the Lie algebra of infinitesimal symmetries, and from this the symmetry groups, and finally, given that u is a solution, we can easily determine other solutions. Furthermore, recall that knowing ξ, τ, ϕ is all that is needed in order to determine the coefficient functions of $\text{pr}^{(2)} \mathbf{v}$. (In actuality, we really are not that interested in them specifically; they are just the means by which we build up to the desired solutions of the differential equation.)

We solve as follows. Note that conditions (a) and (b) require τ to be a function only of t . Therefore, in (e), we have that $-2\tau_{xu} = 0$ which implies that $\xi_u = 3\xi_u$ and this is possibly only if ξ is not a function of u , i.e $\xi_u = 0$. Also using the fact that $\tau(t)$, we know that $-\tau_{xx} = 0$, so therefore, from (f), we have that $\tau_t = 2\xi_x$, which we solve by integration with respect to x , giving us the relation that

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

where $\sigma(t)$ is an arbitrary function of t (recall that when doing this type of problem we arrive at an arbitrary function of the non-integration variable). Furthermore since we have that ξ is not a function of u , equation (h) tells us that ϕ is at most linear in u , so there exist functions α, β such that

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

Hence, by applying condition (j) (and remembering that ξ is at most linear in x so that $\xi_{xx} = 0$), we have that $\xi_t = -2\beta_x$. If we differentiate $\xi_t = -2\beta_x$ twice with respect to x , we have that $0 = \xi_{txx} = -2\beta_{xxx}$. Hence we must have that β is at most quadratic in x . In fact, by using the fact that $\xi(x, t) = \frac{1}{2}\tau_t x + \sigma(t)$ and $\xi_t = -2\beta_x$ (by means of

differentiating ξ with respect to t and then integrating with respect to x) we actually arrive at

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

(again $\rho(t)$ is an arbitrary function of integration).

Now, condition (k) says that ϕ satisfies the heat equation; therefore, we must also have that α and β (in the definition of ϕ) satisfy the heat equation. Since β satisfies the heat equation, we have that

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

So, by equating coefficients of x , we arrive at the facts that $\tau_{ttt} = 0$, $\sigma_{tt} = 0$, $\rho_t = -\frac{1}{4}\tau_{tt}$. Hence we have the following general equations:

$$\tau(t) = at^2 + bt + c, \quad \sigma(t) = dt + e, \quad \rho(t) = -\frac{1}{4}(2a)t + k$$

where a, b, c, d, e, k are arbitrary constants. By substituting these into the already established general formulas for ξ and ϕ , we have

$$(23) \quad \tau(t) = at^2 + bt + c, \quad \xi(x, t) = atx + \frac{1}{2}bx + dt + e,$$

$$(24) \quad \phi(x, t, u) = \left(-\frac{1}{4}ax^2 - \frac{1}{2}dx - \frac{1}{2}at + k\right)u + \alpha(x, t).$$

In the interest of making things a little simpler for future calculations, we'll now make a slight adjustment to the arbitrary constants above (but note if we compare the coefficients in the expressions to each other, the relation remains the same, for instance the coefficient of x is ξ is half the coefficient of t in τ) arriving at the following equations:

$$(25) \quad \xi(x, t) = c_1 + c_4x + 2c_5t + 4c_6xt,$$

$$(26) \quad \tau(t) = c_2 + 2c_4t + 4c_6t^2,$$

$$(27) \quad \phi(x, t, u) = (c_3 - c_5x - 2c_6t - c_6x^2)u + \alpha(x, t).$$

The purpose of renumbering the constants (in this perhaps unnatural way) will become clear in the following. We wish to find a spanning set (it'll actually be an infinite dimensional basis) of the Lie algebra of the infinitesimal symmetries. In order to find this set, we successively set each constant equal to 1 and the rest equal to 0. For instance, for the first, we will set $c_1 = 1$ and $c_2, \dots, c_6, \alpha = 0$ (note that this implies

that $\xi = c_1$, $\tau = 0$, $\phi = 0$). We then evaluate Equation (14) at these values giving us the vector fields:

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

Now that we've worked our way to this point, we have almost achieved the desired effect. We have one more result, the one-parameter subgroups G_i (which are, by the way they are defined, symmetry groups of the differential equation) that are generated by each of the \mathbf{v}_i . We find these by evaluating $e^{\epsilon\mathbf{v}_i}(x, t, u) = (\tilde{x}, \tilde{t}, \tilde{u})$ (note that it is not necessary that ϵ be positive—as is often the case in analysis—nor that it be small; although it must be small enough so that everything is defined—remember we are in essence only sure of local groups):

$$\begin{aligned}G_1 &: (x + \epsilon, t, u), \\ G_2 &: (x, t + \epsilon, u), \\ G_3 &: (x, t, e^\epsilon u), \\ G_4 &: (e^\epsilon x, e^{2\epsilon} t, u), \\ G_5 &: (x + 2\epsilon t, t, ue^{-\epsilon x - \epsilon^2 t}), \\ G_6 &: \left(\frac{x}{1 - 4\epsilon t}, \frac{t}{1 - 4\epsilon t}, u\sqrt{1 - 4\epsilon t} \cdot \exp\left(\frac{-\epsilon x^2}{1 - 4\epsilon t}\right) \right), \\ G_\alpha &: (x, t, u + \epsilon\alpha(x, t)) \quad \text{where } \alpha \text{ is a solution to the heat equation.}\end{aligned}$$

We have now arrived at the fruits of our labor. Since the above groups G_i are in fact symmetry groups, we can arrive directly at a large number of solutions to the heat equation. The above gives us several important invariants (of which we should already have been aware), but in a new light.

- (1) G_1 tells us that a solution of the heat equation is invariant about translations in x .
- (2) G_2 gives the analogous result for translations in t .

- (3) G_3 tells us that at least any positive multiple of a solution is also a solution.
- (4) G_4 gives us a well known scaling symmetry.
- (5) G_5 is something that may not be entirely recognizable. (It is a Galilean boost to a moving coordinate frame).¹⁷
- (6) G_6 is unnatural (being a full-fledged local group of transformations), but has many nice consequences including the possibly unintuitive “Fundamental Source Solution.”
- (7) G_α is the superposition principle.



In Bleeker & Csordas Exercise 1 p. 135, we are asked to show that given a solution $u(x, t)$ of the heat equation, various other values are solutions. While these solutions can be shown in a straight forward manner using differentiation, they can also be shown by making use of the above symmetry groups.

We now derive the solution

$$u(x, t) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}}, \quad t > 0, \quad -\infty < x < \infty$$

of Equation (12) using the above results (see [Olv93] p. 120).

Proof. From G_6 , we see that if an arbitrary function $u = f(x, t)$ is a solution of the heat equation, so is

$$v = g(x, t) = \frac{1}{\sqrt{1 + 4\epsilon t}} e^{-\frac{\epsilon x^2}{1 + 4\epsilon t}} f\left(\frac{x}{1 + 4\epsilon t}, \frac{t}{1 + 4\epsilon t}\right)$$

for any ϵ in a proper neighborhood of the origin. Now clearly $u = f(x, t) = c$ is a solution of the heat equation for any constant c . Hence we have that

$$v_1 = \frac{c}{\sqrt{1 + 4\epsilon t}} e^{-\frac{\epsilon x^2}{1 + 4\epsilon t}}$$

is a solution of the heat equation. But this holds for any value of c , so in particular, it holds if we set $c = \sqrt{\frac{\epsilon}{\pi}}$, we get the solution

$$v_2 = \sqrt{\frac{\epsilon}{\pi(1 + 4\epsilon t)}} e^{-\frac{\epsilon x^2}{1 + 4\epsilon t}}.$$

¹⁷This symmetry group gives the standard product solutions if we let ϵ be complex, see [BC96] exercise 7 and 8 p.137

We can see that v_2 is algebraically equivalent to

$$v_3 = \frac{1}{\sqrt{4\pi \left(t + \frac{1}{4\epsilon}\right)}} e^{\frac{-x^2}{4\left(t + \frac{1}{4\epsilon}\right)}}$$

by multiplying surreptitiously by $\frac{1}{4\epsilon}$. We now apply G_2 and translate v_3 “right” $\frac{1}{4\epsilon}$ in t . This gives us the Fundamental Source Solution to the heat equation. \square

In a previous example (c.f. Example 4), we had derived the function ϕ^x using Lie Derivatives. The method of Lie Derivatives has certain advantages, but it has one major disadvantage for finding the prolongation formulas—it is necessary to have foreknowledge of the differential form that remains invariant. There is a method that does not actually require this foreknowledge; for the next example, we will use this method in computing the prolongation formulas (in the interest of space and not including too many excess calculations, only a complete derivation for x is included). The example will be the wave equation, but before we get into that we introduce the notion of a *total derivative*.

Definition 24. Let $P(x, u^{(n)})$ be a smooth function of x, u and the derivatives of u up to order n , defined on an open subset $M^{(n)} \subset X \times U^{(n)}$. The ***total derivative*** of P with respect to x^i is the unique smooth function $D_i P(x, u^{(n+1)})$ defined on $M^{(n+1)}$ and depending on derivatives of u up to order $n + 1$, with the property that if $u = f(x)$ is any smooth function, $D_i P(x, \text{pr}^{(n+1)} f(x)) = \frac{\partial}{\partial x^i} \left(P(x, \text{pr}^{(n)} f(x)) \right)$.

There are several comments that should be made in regards to this definition. First remember that by X we are referring to the entire *space* of independent variables. Second, if ϕ is a function on an n -jet, then $D_x \phi$ is actually a function on an $(n + 1)$ -jet. Finally, the means by which the definition describes the computation of the total derivative is ambiguous. To clear this up, suppose we have a function $\xi(x, y, t, u)$ defined on the 0-jet (which is locally diffeomorphic to the space $\mathbb{R}^3 \times \mathbb{R}$), then

$$D_x \xi = \frac{\partial \xi}{\partial x} + u_x \frac{\partial \xi}{\partial u}.$$

(This can be extended naturally to any of the independent variables and functions under consideration.)

If we consider $u = u(x, y, t)$, then the calculation becomes even simpler. In this case $D_x u = u_x$; in other words, any time we are taking a total derivative of the function (or derivatives of the function) that

is only dependent on the independent variables, we simply add an extra partial derivative. We see this in the following example where we calculate symmetry groups of the wave equation.

Example 6.¹⁸ **The Wave Equation.** For our purpose we will consider the wave equation in two spacial dimensions:

$$(28) \quad u_{tt} - u_{xx} - u_{yy} = 0.$$

We have that a general vector field on the 0-jet space $\mathbb{R}^3 \times \mathbb{R}$ is

$$(29) \quad \mathbf{v} = \xi \frac{\partial}{\partial x} + \eta \frac{\partial}{\partial y} + \tau \frac{\partial}{\partial t} + \phi \frac{\partial}{\partial u},$$

and noting that our equation is of order 2, we need to prolong up to the space of 2-jets giving us the prolonged vector field

$$(30) \quad \begin{aligned} \text{pr}^{(2)} \mathbf{v} = \mathbf{v} &+ \phi^x \frac{\partial}{\partial u_x} + \phi^y \frac{\partial}{\partial u_y} + \phi^t \frac{\partial}{\partial u_t} + \phi^{xx} \frac{\partial}{\partial u_{xx}} + \phi^{xy} \frac{\partial}{\partial u_{xy}} + \phi^{xt} \frac{\partial}{\partial u_{xt}} \\ &+ \phi^{yy} \frac{\partial}{\partial u_{yy}} + \phi^{yt} \frac{\partial}{\partial u_{yt}} + \phi^{tt} \frac{\partial}{\partial u_{tt}}. \end{aligned}$$

Recall that one way we can represent the differential equation $\Delta : \mathbb{R}^3 \times \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^6 \rightarrow \mathbb{R}$ such that

$$\Delta(x, y, t, u, u_x, u_y, u_t, u_{xx}, u_{xy}, u_{xt}, u_{yy}, u_{yt}, u_{tt}) = u_{tt} - u_{xx} - u_{yy}.$$

is as the zero set of the function Δ . The subvariety \mathcal{S}_Δ is then $\Delta^{-1}(0)$. We again find the infinitesimal condition by making use of Theorem 12, letting $\text{pr}^{(2)} \mathbf{v}$ differentiate Δ . As in the last example each partial derivative picks off the corresponding coefficient function of the vector field. Since Theorem 12 tells us that $\text{pr}^{(2)} \mathbf{v} \cdot \Delta$ vanishes whenever Δ vanishes, we essentially “factor” Δ out of $\text{pr}^{(2)} \mathbf{v} \cdot \Delta$, giving us the following result

$$(31) \quad \phi^{tt} - \phi^{xx} - \phi^{yy} = (u_{tt} - u_{xx} - u_{yy})Q = Qu_{tt} - Qu_{xx} - Qu_{yy}$$

where Q is some function on the space of 2-jets.

We will now, as promised, calculate the prolongations. But first we need a definition.

Definition 25. The *characteristic* of \mathbf{v} (c.f. Equation (29)) in the case of the wave equation (suitable changes can easily be made depending on the number of independent variables) is the function $\phi - \xi u_x - \eta u_y - \tau u_t$.

¹⁸Some parts of this can be found in [Olv93] p. 123-125; however, his solution is only sketched.

In the prolongation formula (see [Olv93] p. 110), we are told that, in the wave equation,

$$\phi^x = D_x(\phi - \xi u_x - \eta u_y - \tau u_t) + \xi u_{xx} + \eta u_{xy} + \tau u_{xt}.$$

Remember that ϕ^x is a function on the 1-jets, and thus should only be dependent up to first order derivatives; however, also remember that D pushes the function up the the next jet. We perform the above calculation, remembering the above comments on how D works and realizing that it obeys linearity and the product rule.

$$\begin{aligned} \phi^x &= D_x(\phi - \xi u_x - \eta u_y - \tau u_t) + \xi u_{xx} + \eta u_{xy} + \tau u_{xt} \\ &= D_x\phi - (\xi D_x u_x + u_x D_x \xi) - (\eta D_x u_y + u_y D_x \eta) \\ &\quad - (\tau D_x u_t + u_t D_x \tau) + \xi u_{xx} + \eta u_{xy} + \tau u_{xt} \\ &= D_x\phi - \xi u_{xx} - u_x D_x \xi - \eta u_{xy} - u_y D_x \eta - \tau u_{xt} - u_t D_x \tau \\ &\quad + \xi u_{xx} + \eta u_{xy} + \tau u_{xt} \\ &= D_x\phi - u_x D_x \xi - u_y D_x \eta - u_t D_x \tau \\ &= \left(\frac{\partial \phi}{\partial x} + u_x \frac{\partial \phi}{\partial u} \right) - u_x \left(\frac{\partial \xi}{\partial x} + u_x \frac{\partial \xi}{\partial u} \right) - u_y \left(\frac{\partial \eta}{\partial x} + u_x \frac{\partial \eta}{\partial u} \right) \\ &\quad - u_t \left(\frac{\partial \tau}{\partial x} + u_x \frac{\partial \tau}{\partial u} \right) \\ &= \phi_x + \phi_u u_x - \xi_x u_x - \xi_u u_x^2 - \eta_x u_y - \eta_u u_x u_y - \tau_x u_t - \tau_u u_x u_t \\ &= \phi_x + (\phi_u - \xi_x) u_x - \eta_x u_y - \tau_x u_t - \xi_u u_x^2 - \eta_u u_x u_y - \tau_u u_x u_t. \end{aligned}$$

Notice that indeed we do get that ϕ^x is really only a function on the 1-jet (despite it looking like it may not be originally). We will wish to consider all of these prolongation formulas as “polynomials” with indeterminants being the partial derivatives of u .

The remaining 1st prolongation formulas are

$$\begin{aligned} \phi^y &= D_y(\phi - \xi u_x - \eta u_y - \tau u_t) + \xi u_{xy} + \eta u_{yy} + \tau u_{yt} \\ &= \phi_y - \xi_y u_x + (\phi_u - \eta_y) u_y - \tau_y u_t - \xi_u u_x u_y - \eta_u u_y^2 - \tau_u u_y u_t \end{aligned}$$

and

$$\begin{aligned} \phi^t &= D_t(\phi - \xi u_x - \eta u_y - \tau u_t) + \xi u_{xt} + \eta u_{yt} + \tau u_{tt} \\ &= \phi_t - \xi_t u_x - \eta_t u_y + (\phi_u - \tau_t) u_t - \xi_u u_x u_t - \eta_u u_t u_y - \tau_u u_t^2. \end{aligned}$$

We now calculate the second prolongation, ϕ^{xx} , recalling the recursive nature of prolongation.

(32)

$$\begin{aligned}
\phi^{xx} &= D_x(\phi^x - \xi u_{xx} - \eta u_{xy} - \tau u_{xt}) + \xi u_{xxx} + \eta u_{xxy} + \tau u_{xxt} \\
&= D_x \phi^x - (\xi D_x u_{xx} + u_{xx} D_x \xi) - (\eta D_x u_{xy} + u_{xy} D_x \eta) \\
&\quad - (\tau D_x u_{xt} + u_{xt} D_x \tau) + \xi u_{xxx} + \eta u_{xxy} + \tau u_{xxt} \\
&= D_x \phi^x - \xi u_{xxx} - u_{xx} D_x \xi - \eta u_{xxy} - u_{xy} D_x \eta - \tau u_{xxt} - u_{xt} D_x \tau \\
&\quad + \xi u_{xxx} + \eta u_{xxy} + \tau u_{xxt} \\
&= D_x \phi^x - u_{xx} D_x \xi - u_{xy} D_x \eta - u_{xt} D_x \tau \\
&= \left(\frac{\partial \phi^x}{\partial x} + u_x \frac{\partial \phi^x}{\partial u} \right) - u_{xx} \left(\frac{\partial \xi}{\partial x} + u_x \frac{\partial \xi}{\partial u} \right) - u_{xy} \left(\frac{\partial \eta}{\partial x} + u_x \frac{\partial \eta}{\partial u} \right) \\
&\quad - u_{xt} \left(\frac{\partial \tau}{\partial x} + u_x \frac{\partial \tau}{\partial u} \right) \\
&= \phi_x^x + \phi_u^x u_x - \xi_x u_{xx} - \xi_u u_x u_{xx} - \eta_x u_{xy} - \eta_u u_x u_{xy} - \tau_x u_{xt} - \tau_u u_x u_{xt} \\
&= \left(\phi_{xx} + \phi_u u_{xx} - \xi_x u_{xx} + \phi_{xu} u_x - \xi_{xx} u_x - \eta_x u_{xy} - \eta_{xx} u_y - \tau_x u_{xt} \right. \\
&\quad - \tau_{xx} u_t - 2\xi_u u_x u_{xx} - \xi_{xu} u_x^2 - \eta_u u_x u_{xy} - \eta_{uu} u_y u_{xx} - \eta_{uu} u_y u_{xx} \\
&\quad - \eta_{xu} u_x u_y - \tau_u u_x u_{xt} - \tau_{uu} u_t u_{xx} - \tau_{xu} u_x u_t \left. \right) + \left(\phi_{xu} u_x + \phi_{uu} u_x^2 \right. \\
&\quad - \xi_{xu} u_x^2 - \eta_{xu} u_x u_y - \tau_{xu} u_x u_t - \xi_{uu} u_x^3 - \eta_{uu} u_x^2 u_y - \tau_{uu} u_x^2 u_t \left. \right) \\
&\quad - \xi_x u_{xx} - \xi_u u_x u_{xx} - \eta_x u_{xy} - \tau_x u_{xt} - \tau_u u_x u_{xt} \\
&= \phi_{xx} + (2\phi_{xu} - \xi_{xx}) u_x - \eta_{xx} u_y - \tau_{xx} u_t + (\phi_{uu} - 2\xi_{xu}) u_x^2 - 2\eta_{xu} u_x u_y \\
&\quad - 2\tau_{xu} u_x u_t + (\phi_u - 2\xi_x) u_{xx} - 2\tau_x u_{xt} - 2\eta_x u_{xy} - \xi_{uu} u_x^3 - \eta_{uu} u_x^2 u_y \\
&\quad - \tau_{uu} u_x^2 u_t - 3\xi_u u_x u_{xx} - \eta_u u_y u_{xx} - \tau_{uu} u_{xx} - 2\eta_u u_x u_{xy} - 2\tau_u u_x u_{xt}.
\end{aligned}$$

Since the calculations for ϕ^{yy} and ϕ^{tt} are essentially identical to that for ϕ^{xx} above, we will not include these values here. Also, there is a pattern in the calculation of these prolongations, so once one discovers this pattern, actually computing them is unnecessary. However, we are interested in the left hand side of Equation (31), so we include that here (even though it is very ungainly).

$$\begin{aligned}
(33a) \quad & \phi^{tt} - \phi^{xx} - \phi^{yy} = (\phi_{tt} - \phi_{xx} - \phi_{yy}) + (-\xi_{tt} - 2\phi_{xu} + \xi_{xx} + \xi_{yy})u_x \\
(33b) \quad & + (-\eta_{tt} + \eta_{xx} - 2\phi_{yu} + \eta_{yy})u_y + (2\phi_{tu} - \tau_{tt} + \tau_{xx} + \tau_{yy})u_t \\
(33c) \quad & - (\phi_{uu} - 2\xi_{xu})u_x^2 - (\phi_{uu} - 2\eta_{yu})u_y^2 + (\phi_{uu} - 2\tau_{tu})u_t^2 \\
(33d) \quad & + (2\eta_{xu} + 2\xi_{yu})u_xu_y + (\tau_{xu} - 2\xi_{tu})u_xu_t + (2\tau_{yu} - 2\eta_{tu})u_yu_t \\
(33e) \quad & - (\phi_u - 2\xi_x)u_{xx} - (\phi_u - 2\eta_y)u_{yy} + (\phi_u - 2\tau_t)u_{tt} \\
(33f) \quad & + (2\tau_x - 2\xi_t)u_{xt} + (2\eta_x + 2\xi_y)u_{xy} + (2\tau_y - 2\eta_t)u_{yt} \\
(33g) \quad & + \xi_{uu}u_x^3 + \eta_{uu}u_y^3 - \tau_{uu}u_t^3 \\
(33h) \quad & + \eta_{uu}u_x^2u_y + \tau_{uu}u_x^2u_t + \xi_{uu}u_xu_y^2 + \tau_{uu}u_tu_y^2 - \eta_{uu}u_yu_t^2 - \xi_{uu}u_xu_t^2 \\
(33i) \quad & + 3\xi_uu_xu_{xx} + 3\eta_uu_yu_{yy} - 3\tau_uu_tu_{tt} + \eta_uu_yu_{xx} + \tau_uu_tu_{xx} \\
(33j) \quad & + \xi_uu_xu_{yy} + \tau_uu_tu_{yy} - \eta_uu_yu_{tt} - \xi_uu_xu_{tt} + 2\eta_uu_xu_{xy} + 2\tau_uu_xu_{xt} \\
(33k) \quad & + 2\xi_uu_yu_{xy} + 2\tau_uu_yu_{yt} - 2\eta_uu_tu_{yt} - 2\xi_uu_tu_{xt}.
\end{aligned}$$

Recalling that $\phi^{tt} - \phi^{xx} - \phi^{yy} = (u_{tt} - u_{xx} - u_{yy})Q = Qu_{tt} - Qu_{xx} - Qu_{yy}$, we eventually see that almost all of the terms in Equation (33) are actually 0. When we identified the symmetry Lie algebra for the heat equation in the last example, we dealt with it as a whole. In this case, we will do what is more commonly the practice in dealing with this type of problem, namely to “tackle the solution of the symmetry equations in stages, first extracting information from higher order derivatives appearing in them, and then using this information to simplify the prolongation formulas.”¹⁹ This is often the most efficient way to analyze such problems (for the wave equation, as far as differential equations go, is fairly straightforward), and we already have an equation that is very long, and attempting to write down all the corresponding partial differential equations we would need to solve would be even longer. In this case, we first consider the coefficients of the second order derivatives, $u_{xy}, u_{xt}, u_{yt}, u_{xx}, u_{yy}, u_{tt}$ (these are in lines (33f), (33j) and (33k) of Equation (33)). Since everything on the right hand side of Equation (31) needs to contain at least u_{xx}, u_{yy}, u_{tt} , the coefficient functions of u_{xy}, u_{xt}, u_{yt} must be 0. With these, we see

¹⁹See [Olv93] p. 121.

$$\begin{aligned}
u_{xt} &: (2\tau_x - 2\xi_t) + 2\tau_u u_x - 2\xi_u u_t = 0, \\
u_{xy} &: (2\eta_x + 2\xi_y) + 2\eta_u u_x + 2\xi_u u_y = 0, \\
u_{yt} &: (2\tau_y - 2\eta_t) + 2\tau_u u_y - 2\eta_u u_t = 0.
\end{aligned}$$

If we consider the above to be functions on the 1-jets, we notice that they are linear “polynomials.” Since they are equal to 0, we can see from basic polynomial facts each of the coefficients are 0 (the only way to have the 0 polynomial is for all of its coefficients to be 0). In particular, for the constant terms, we have

$$(34) \quad \tau_x - \xi_t = 0, \quad \eta_x + \xi_y = 0, \quad \tau_y - \eta_t = 0.$$

If we look at the other terms, we see that $\tau_u = 0$, $\xi_u = 0$, $\eta_u = 0$. This implies that none of these are a function of u , hence we have that $\xi = \xi(x, y, t)$, $\eta = \eta(x, y, t)$, $\tau = \tau(x, y, t)$. Note that this result means that all the terms in lines (33g), (33h), (33i), (33j), and (33k) of Equation (33) are 0. This is a major simplification of Equation (33), and that is our (and Lie’s) goal in using this method to understand differential equations.

We can also see that the coefficients of u_{xx} , u_{yy} , u_{tt} imply that

$$Q = \phi_u - 2\xi_x = \phi_u - 2\eta_y = \phi_u - 2\tau_t.$$

So clearly we have the result that

$$(35) \quad \tau_t = \xi_x = \eta_y.$$

The reader may recognize that Equations (34) and (35) are the equations for an infinitesimal conformal transformation on \mathbb{R}^3 with Lorentz metric $dt^2 - dx^2 - dy^2$ (see [Olv93] p. 123).

Using substitutions from Equations (34) and (35) and taking derivatives, we can see that (the first is from Olver, we do two other calculations)

$$\begin{aligned}
\xi_{xxx} &= \eta_{xxy} = -\xi_{xyy} = -\tau_{yyt} = -\eta_{ytt} = -\xi_{xtt} = -\tau_{xxt} = -\xi_{xxx}, \\
\eta_{yyy} &= \xi_{xyy} = -\eta_{xxy} = -\tau_{xxt} = -\xi_{xtt} = -\eta_{ytt} = -\tau_{yyt} = -\eta_{yyy}, \\
\xi_{yyy} &= -\eta_{xyy} = -\tau_{xyt} = -\xi_{tty} = \eta_{ttx} = \tau_{ytx} = \eta_{yyx} (= -\xi_{yyy}).
\end{aligned}$$

The last equality in the third line is included only to continue the pattern, for by saying that $-\eta_{xyy} = \eta_{yyx}$ we already have that the entire string, including ξ_{yyy} , is equal to 0.

Doing similar calculations, i.e. showing that all the possible third order derivatives of ξ, τ, η are actually equal to their opposite, we can

see that they are all 0. Hence ξ, τ, η are at most quadratic in each of x, y, t . Hence, by using the relations in Equations (34) and (35), we can actually find that

$$\begin{aligned}\xi &= c_1 + c_4x - c_5y + c_6t + c_8(x^2 - y^2 + t^2) + 2c_9xy + 2c_{10}xt, \\ \eta &= c_2 + c_5x + c_4y + c_7t + 2c_8xy + c_9(-x^2 + y^2 + t^2) + 2c_{10}yt, \\ \tau &= c_3 + c_6x + c_7y + c_4t + 2c_8xt + 2c_9yt + c_{10}(x^2 + y^2 + t^2),\end{aligned}$$

where c_1, \dots, c_{10} are arbitrary constants.

Now since all of the indeterminants u_x^2, u_y^2, u_t^2 do not appear in the right hand side of Equation (31), we must have that the coefficients of u_x^2, u_y^2, u_t^2 in Equation (33) must be 0 (that is all the coefficients in (33c) are 0); since we have already shown that none of ξ, η, τ depend on u , it follows that $\phi_{uu} = 0$. Hence ϕ is at most linear in u , so we can write

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

We then finally examine the constant terms and the first order derivatives of Equation (33) (they're in line (33a)), finding the following relations:

$$\begin{aligned}2\beta_x &= \xi_{xx} + \xi_{yy} - \xi_{tt}, \\ 2\beta_y &= \eta_{xx} + \eta_{yy} - \eta_{tt}, \\ 2\beta_t &= \tau_{tt} - \tau_{xx} - \tau_{yy}, \\ \alpha_{tt} - \alpha_{xx} - \alpha_{yy} &= 0.\end{aligned}$$

Hence we can see that α is an (arbitrary) solution to the wave equation, and by making use of our previous definitions of ξ, η, τ , we can see that

$$\beta = c_{11} - c_8x - c_9y - c_{10}t.$$

We find the infinitesimal symmetries in the same way as we did in the example of the heat equation—by successively setting the constants equal to 1 and the rest equal to 0. The first 10 generate the conformal algebra for \mathbb{R}^3 with the given Lorentz metric.

$$\begin{array}{ll}
\partial_x, \partial_y, \partial_t, & \text{translations: } \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \\
\mathbf{r}_{xy} = -y\partial_x + x\partial_y, & \text{rotation: } \mathbf{v}_5 \\
\mathbf{r}_{xt} = t\partial_x + x\partial_t, \mathbf{r}_{yt} = t\partial_y + y\partial_t, & \text{hyperbolic rotations: } \mathbf{v}_6, \mathbf{v}_7 \\
\mathbf{d} = x\partial_x + y\partial_y + t\partial_t, & \text{dilations: } \mathbf{v}_4 \\
\mathbf{i}_x = (x^2 - y^2 + t^2)\partial_x + 2xy\partial_y + 2xt\partial_t - xu\partial_u, & \text{inversion } x: \mathbf{v}_8 \\
\mathbf{i}_y = 2xy\partial_x + (y^2 - x^2 + t^2)\partial_y + 2yt\partial_t - yu\partial_u, & \text{inversion } y: \mathbf{v}_9 \\
\mathbf{i}_t = 2xt\partial_x + 2yt\partial_y + (x^2 + y^2 + t^2)\partial_t - tu\partial_u, & \text{inversion } t: \mathbf{v}_{10}
\end{array}$$

and the additional vector fields

$$\mathbf{v}_{11} = u\partial_u, \quad \mathbf{v}_\alpha = \alpha(x, y, t)\partial_u$$

showing the invariance of dilation in u , i.e. a constant multiple of a solution is also a solution, and the superposition principle due to the linearity of the wave equation.

One can understand $\mathbf{v}_1, \dots, \mathbf{v}_7$ in terms of linear algebra. Recall that any planar rotation transformation can be described by

$$\left. \frac{d}{d\theta} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right|_{\theta=0}$$

and that a hyperbolic rotation can be described by

$$\left. \frac{d}{d\theta} \begin{pmatrix} \cosh \theta & \sinh \theta \\ \sinh \theta & \cosh \theta \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} \right|_{\theta=0}.$$

In this light, we can see that indeed we do get the translations, dilation and rotation that we claimed would be present. Consider the following.

$$\begin{bmatrix} \xi \\ \eta \\ \tau \\ 1 \end{bmatrix} = \left[\begin{array}{ccc|c} c_4 & -c_5 & c_6 & c_1 \\ c_5 & c_4 & c_7 & c_2 \\ c_6 & c_7 & c_4 & c_3 \\ \hline 0 & 0 & 0 & 1 \end{array} \right] \cdot \begin{bmatrix} x \\ y \\ t \\ 1 \end{bmatrix}$$

Here again, we successively set each constant equal to 0, and we arrive at the same result. However, it is clearer to see the transformation that is occurring. When each of $c_1, c_2, c_3 = 1$ and the rest are 0, we can see that we have a translation with respect to x, y, t . If $c_4 = 1$, we have the matrix that we recognize from linear algebra as a dilation. Similarly, if $c_5 = 1$ we get the aforementioned 90° rotation; and if $c_6, c_7 = 1$ we get the matrix corresponding to a hyperbolic rotation. \blacklozenge

7. SYMMETRIES WITH ORDINARY DIFFERENTIAL EQUATIONS

In the previous section, our focus was to find the symmetry Lie algebra, and not really spend much time actually finding solutions. The symmetry vector fields, and hence groups, did give us information on how once we had a solution to arrive at an entire family of solutions to the partial differential equation.

We now move on to considering ordinary differential equations. By contrast, in this section, our emphasis is no longer going to be arriving at the symmetry groups, but rather focusing on how these symmetry groups enable us to find solutions to the ODE. In many of the examples, it is clear once we see the symmetry group that it is in fact a symmetry group. We will follow a classical convention which is that the differential equation will be *solved* once we have *reduced it to quadrature*; that is, once we are to the point of finding an antiderivative.

There are three methods that we will use. The first involves the Flow Box Theorem (c.f. Theorem 6) which enables us find a change of coordinates reducing the vector field to a single coordinate (this is also called using *canonical coordinates*). The second involves using an integrating factor in order to make the differential equation exact (c.f. Section (3) for a different view of this method). The final method is the use of differential invariants.

In the following, we are going to have a differential equation of only one independent and one dependent variable

$$(36) \quad \frac{du}{dx} = F(x, u).$$

In aligning this with our previous notions, we can define $\Delta : X \times U^{(1)} \rightarrow \mathbb{R}$ where $\Delta(x, u, u_x) = u_x - F(x, u)$. Note that the solution to Equation (36) is the graph of a function, and it prolongs up to a graph of a function on the surface $\Delta = 0$ in \mathbb{R}^3 .

A general vector field on $X \times U$ is $\mathbf{v} = \xi(x, u) \frac{\partial}{\partial x} + \phi(x, u) \frac{\partial}{\partial u}$, which extends up to

$$(37) \quad \text{pr}^{(1)} \mathbf{v} = \xi \frac{\partial}{\partial x} + \phi \frac{\partial}{\partial u} + \phi^x \frac{\partial}{\partial u_x}.$$

Letting Equation (37) differentiate Δ , we get by Theorem 12 the infinitesimal condition

$$(38) \quad \phi^x - (\xi F_x + \phi F_u) = 0 \quad \text{whenever} \quad u_x = F(x, u).$$

But, the additional condition $u_x = F(x, u)$ provides no additional constraints (unlike the constraints in the heat and wave equations). As mentioned earlier, sometimes the group may seem to come out of the

blue. This is because “solving Equation (38) is often much more difficult than solving Equation (36). However, led on by inspired-guess work, or geometric intuition, we may be able to ascertain a particular solution to Equation (38) which will allow us to integrate Equation (36). Herein lies the art of Lie’s method.”²⁰ As we have seen earlier, Lie was someone who would often make some leaps; here we have to do the same.

We now specifically consider the result using the Flow Box Theorem (c.f. Theorem 6). According to this theorem, we can find a change of coordinates (in this case the independent variable $y = \eta(x, u)$ and the dependent variable $w = \zeta(x, u)$) so that the vector field \mathbf{v} also changes coordinates into $\mathbf{v}^* = \frac{\partial}{\partial w}$. Since the prolongations involve derivatives (in this case total derivatives) and the coefficient functions of $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial w}$ in \mathbf{v}^* are 0 and 1 respectively, we see that \mathbf{v}^* actually prolongs to itself, i.e. $\text{pr}^{(1)} \mathbf{v}^* = \frac{\partial}{\partial w}$. We can further re-express Equation (36) in terms of the new variables $\Delta^* = w_y - H(y, w)$, and letting the prolongation differentiate this function, we find the infinitesimal condition

$$\text{pr}^{(1)} \mathbf{v}^* \cdot \Delta^* = -\frac{\partial H}{\partial w}(y, w) = 0 \quad \text{whenever} \quad w_y = H(y, w).$$

This, of course, implies that H does not depend on w ; thus we can reduce the differential equation to

$$\frac{dw}{dy} = H(y) \implies w = \int H(y) dy.$$

We now look at this in an example.

Example 7.²¹ Recall from Section (3) (where we do an alternative treatment) that a homogeneous equation is a differential equation for which F is dependent only on the ratio $\frac{u}{x}$, i.e.

$$\frac{du}{dx} = F\left(\frac{u}{x}\right).$$

By inspection, we can see that the 1-parameter group of scaling transformations $G_\lambda : (x, u) \mapsto (\lambda x, \lambda u)$ (where $\lambda > 0$) is a symmetry group (let G act on the differential equation, which will replace x with λx and u with λu , and see that the λ 's cancel leaving the equation unchanged). Let λ_0 denote the parameter corresponding to the identity

²⁰See [Olv93] p. 131.

²¹See [Olv93] p. 133.

element of the group. (In our example, $\lambda_0 = 1$.) We can then find the symmetry vector field that corresponds to this symmetry group by

$$\mathbf{v}(x, u) = \left. \frac{d}{d\lambda} G_\lambda(\lambda x, \lambda u) \right|_{\lambda=\lambda_0}.$$

In our specific case,

$$\mathbf{v}(x, u) = \left. \frac{d}{d\lambda}(\lambda x, \lambda u) \right|_{\lambda=1} = x \frac{\partial}{\partial x} + u \frac{\partial}{\partial u}.$$

We then use the flow box theorem to reduce \mathbf{v} to having a single coordinate. To do this, we must find the change of coordinates η, ζ (these are the canonical coordinates) (see [Olv93] p. 132).

$$(39) \quad \mathbf{v} \cdot \eta = \xi \frac{\partial \eta}{\partial x} + \phi \frac{\partial \eta}{\partial u} = 0,$$

$$(40) \quad \mathbf{v} \cdot \zeta = \xi \frac{\partial \zeta}{\partial x} + \phi \frac{\partial \zeta}{\partial u} = 1.$$

We can find η by solving the associated characteristic ordinary differential equation

$$\frac{dx}{\xi(x, u)} = \frac{du}{\phi(x, u)}.$$

However, much of the time, we come up with the change of variables largely by inspection by looking at the form of the differential equation. A logical choice would be to choose $\eta = y = \frac{u}{x}$ for this would make the function of the differential equation just a function of y . The choice of $\zeta = w$ is in some sense arbitrary, provided it is independent of y , in particular that the Jacobian determinant $\frac{\partial(y,w)}{\partial(x,u)} \neq 0$. In this case, we choose $\zeta = w = \ln x$ (these do indeed satisfy the above equations). We then re-express the homogenous differential equation in terms of y, w .

$$F(y) = \frac{du}{dx} = \frac{du/dy}{dx/dy} = \frac{x(1 + yw_y)}{xw_y} = \frac{1 + yw_y}{w_y}.$$

Solving this for w_y , we get

$$\frac{dw}{dy} = \frac{1}{F(y) - y} \implies w = \int \frac{dy}{F(y) - y} + c.$$

We can then re-substitute and find a solution u implicitly (recall we often are only able to come up with the implicit solution). \blacklozenge

A differential equation we often see, especially when we start dealing with symmetry groups, is a Ricatti equation. In most cases, we will not solve the Ricatti equation, but merely consider the problem solved if we have reduced the problem to that of finding a solution to a Ricatti equation. Various sources have a variety of equations that on

the surface appear different (but in many senses are very similar) that they call Ricatti equations. Bluman has perhaps the simplest form, $\frac{du}{dx} = A + Bu + Cu^2$ (see [BK89] p. 105). Olver relaxes the condition on constant coefficients and calls the expression $z_x = -z^2 - p(x)z - q(x)$ a Ricatti equation (see [Olv93] p. 139). We then would consider that a Ricatti equation is one with $\frac{du}{dx}$ being quadratic in u . The following two examples look at solving specific Ricatti equations using canonical coordinates.

We earlier concluded symmetry groups must act on solutions by sending them to solutions. We will adopt Hydon's notation (see [Hyd00] p. 8), which is slightly different than Olver's, for parts of these next examples. For instance, we will denote by (\hat{x}, \hat{u}) the image of (x, u) under the symmetry map. The reader can clearly see that the results are the same as our earlier approaches. We have previously talked about symmetry conditions of Equation (36) and other differential equations, so here we will skip some of the preliminaries; however, note that the previously mentioned methods get us to the equivalent condition (see [Hyd00] p. 8)

$$(41) \quad F(\hat{x}, \hat{u}) = \frac{\hat{u}_x + F(x, u)\hat{u}_u}{\hat{x}_x + F(x, u)\hat{x}_u}.$$

Example 8. ²² Solve the Ricatti equation

$$(42) \quad \frac{du}{dx} = xu^2 - \frac{2u}{x} - \frac{1}{x^3}, \quad (x \neq 0).$$

Equation (42) has a scaling symmetry

$$(43) \quad (\hat{x}, \hat{u}) = (e^\epsilon x, e^{-2\epsilon}u).$$

This is easy to see by either using Equation (41), or by simply substituting Equation (43) into Equation (42) and seeing that it is unchanged after simplification. Noting our previously mentioned methods for determining the vector field corresponding to the symmetry (find the value of the parameter that is the identity, and then differentiate this with respect to the parameter, and then set the parameter equal to the identity), we find that the vector field for equation (43) is

$$\mathbf{v} = x \frac{\partial}{\partial x} - 2u \frac{\partial}{\partial u}.$$

As an aside, we can then find what Hydon calls the *reduced characteristic* (see [Hyd00] p. 21),

$$(44) \quad \overline{Q} = (\phi - u_x \xi)|_{(u_x=F(x,u))} = \phi - F\xi.$$

²²See [Hyd00]. This is a running example in Chapter 2.

In the case of Equation (42), we can see that it has reduced characteristic

$$\overline{Q} = \frac{1}{x^2} - x^2 u^2.$$

The values of u making $\overline{Q} = 0$ are solutions to Equation (42). We can clearly see that this occurs when $u = \pm x^{-2}$. However, these are not the only solutions to Equation (42). We will now compute the complete set of solutions for Equation (42). From the above, we already have the two solutions $u = \pm x^{-2}$. Furthermore, we can get canonical coordinates

$$(\eta, \zeta) = (x^2 u, \ln |x|)$$

for Equation (42) (c.f. Equations (39) and (40)).

We can then use these to change the coordinates, in line with the Flow Box Theorem, of Equation (42) to the simpler form

$$\frac{d\zeta}{d\eta} = \frac{1}{\eta^2 - 1} \implies \zeta = \int \frac{1}{\eta^2 - 1} d\eta.$$

Evaluating this is straightforward, and after re-substituting back to x and u , we arrive at the general solution

$$u = \frac{c + x^2}{x^2(c - x^2)}.$$

(In this solution, we see that the solutions of $u = \pm x^{-2}$ are present; $u = x^{-2}$ is the limit as $c \rightarrow \infty$, and $u = -x^{-2}$ corresponds to the solution when $c = 0$. \blacklozenge)

Hydon also includes other example of a Ricatti equation that we solve in the next example.

Example 9. ²³ Solve the Ricatti Equation

$$(45) \quad \frac{du}{dx} = \frac{u+1}{x} + \frac{u^2}{x^3}, \quad (x \neq 0).$$

We claim that one of the symmetries of Equation (45) is

$$(46) \quad (\hat{x}, \hat{u}) = \left(\frac{x}{1 - \epsilon x}, \frac{u}{1 - \epsilon x} \right).$$

This is directly proven by substituting Equation (46) and Equation (45) into Equation (41), and is purely computation, so we do not include it here. (Of course, how it was determined that this was in fact a symmetry is a little more mysterious.) One thing of interest to note

²³See [Hyd00] p. 10-11, 28.

is that this symmetry Equation (45) is indeed local—it is not defined for $\epsilon > \frac{1}{|x|}$. This is similar to the case of the inversions that we have encountered previously (like in the wave equation example).

The tangent vector field corresponding to Equation (46) is

$$\mathbf{v} = x^2 \frac{\partial}{\partial x} + xu \frac{\partial}{\partial u}.$$

Once we have established the vector field, we can find the canonical coordinates

$$(47) \quad (\eta, \zeta) = \left(\frac{u}{x}, -\frac{1}{x} \right), \quad (x \neq 0)$$

for Equation (45) (c.f. Equations (39) and (40)).

Using this, we can again change coordinates, reducing Equation (45) to the ODE

$$\frac{d\zeta}{d\eta} = \frac{1}{1 + \eta^2}.$$

This now has an obvious solution (we would want our Calculus 1 students to recognize it—although many do not)

$$\zeta = \tan^{-1} \eta + c,$$

or in terms of x, u :

$$u = -x \tan \left(\frac{1}{x} + c \right).$$

◆

In Section (3) we looked at an example of an exact equation. We now do a slightly more general interpretation of this. We first consider a *total* differential equation, that is, an equation of the form

$$(48) \quad P(x, u)dx + Q(x, u)du = 0.$$

Equation (48) is *exact* if there is some function $g(x, u)$ so that $dg = P(x, u)dx + Q(x, u)du = 0$. This will also be the case if $\frac{\partial P}{\partial u} = \frac{\partial Q}{\partial x}$. Noting that partial derivatives commute, this implies that

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial u} = \frac{\partial^2 g}{\partial x \partial u} - \frac{\partial^2 g}{\partial u \partial x} = 0.$$

This will clearly imply that the function $g(x, u) = g(x, u(x)) = c$. Therefore $\frac{\partial g}{\partial x} + u'(x) \frac{\partial g}{\partial u} = 0$.

We now consider the case of using an integrating factor. For this, we have the following theorem (see [Olv93] p. 135-136).

Theorem 14. *Suppose that the equation $Pdx + Qdu = 0$ has a one-parameter symmetry group with infinitesimal generator $\mathbf{v} = \xi \frac{\partial}{\partial x} + \phi \frac{\partial}{\partial u}$. Then $\mathbf{v} \lrcorner (Pdx + Qdu) = P\xi + Q\phi$. Furthermore, $R = \frac{1}{P\xi + Q\phi}$ is an integrating factor; hence the equation $\frac{Pdx + Qdu}{P\xi + Q\phi} = 0$ is exact.*

Proof. The first claim that $\mathbf{v} \lrcorner (Pdx + Qdu) = P\xi + Q\phi$ is computational and direct from the definition of interior product. Recalling this, we see that $\langle \mathbf{v} \lrcorner (Pdx + Qdu) \rangle = \langle (Pdx + Qdu); \mathbf{v} \rangle = P\xi + Q\phi$. (Note that if we have a one-form $\omega = h_1(x)dx^1 + \dots + h_m(x)dx^m$ and vector field $\mathbf{v} = \xi^1(x) \frac{\partial}{\partial x^1} + \dots + \xi^m(x) \frac{\partial}{\partial x^m}$, then $\langle \omega; \mathbf{v} \rangle = \sum_{i=1}^m h_i(x)\xi^i(x)$.) Hence we have the desired result.

Let us note that once we show that R is an integrating factor, the last statement follows from the definition of integrating factor; that is, it is the function that is multiplied to the equation to make it exact.

By Equation (38), we can see that \mathbf{v} is a symmetry vector field of Equation (48) if and only if

$$\begin{aligned} \left(\xi \frac{\partial P}{\partial x} + \phi \frac{\partial P}{\partial u} \right) Q - \left(\xi \frac{\partial Q}{\partial x} + \phi \frac{\partial Q}{\partial u} \right) P + \frac{\partial \phi}{\partial x} Q^2 \\ - \left(\frac{\partial \phi}{\partial u} - \frac{\partial \xi}{\partial x} \right) PQ - \frac{\partial \xi}{\partial u} P^2 = 0. \end{aligned}$$

Now R is an integrating factor if

$$\frac{\partial}{\partial u}(RP) = \frac{\partial}{\partial x}(RQ).$$

Substituting this into the equation for R , this condition becomes

$$\begin{aligned} R^2 \left(\phi \left(Q \frac{\partial P}{\partial u} - P \frac{\partial Q}{\partial u} \right) - \frac{\partial \xi}{\partial u} P^2 - \frac{\partial \phi}{\partial u} PQ \right) \\ = R^2 \left(\xi \left(P \frac{\partial Q}{\partial x} - Q \frac{\partial P}{\partial x} \right) - \frac{\partial \xi}{\partial x} PQ - \frac{\partial \phi}{\partial x} Q^2 \right). \end{aligned}$$

The theorem is proved by comparing this result with the earlier infinitesimal criterion. \square

In the previous examples, we have looked at using symmetry groups to find solutions to ODEs in a different way than would be presented in an introduction to ordinary differential equations class; however, many of the problems could be solved using the methods of introductory ordinary differential equations. For those that could be solved using traditional methods, the method of symmetry groups is often more complicated, and one would wonder why bother trying to solve differential equations using symmetry methods. In the examples of the

Riccati equations as well as the following example, we see the reason—it is often possible to obtain (once we’ve divined a symmetry group) solutions to a differential equation that at best would be difficult to solve using traditional methods.

Example 10. ²⁴ Consider the ODE

$$(49) \quad \frac{du}{dx} = \frac{u^3 + u - 3x^2u}{3x^3u + x - x^3}.$$

The reader familiar with ODE’s can see that this is not reasonably solvable by any traditional method, such as we’ve mentioned above and in Section (3).

However, note that Equation (49) has an infinitesimal symmetry vector field

$$\mathbf{v} = (u^3 + u - 3x^2u) \frac{\partial}{\partial x} + (x^3 - x - 3xu^2) \frac{\partial}{\partial u}.$$

We then find that the characteristic equation is

$$(50) \quad \frac{du}{dx} = \frac{x^3 - x - 3xu^2}{u^3 + u - 3x^2y}.$$

This does not appear to be any more easily solvable than Equation (49). However, we can use an integrating factor (see Theorem 14), noting that

$$\int \frac{du - Fdx}{\phi - F\xi} = c.$$

Applying this to Equation (50), we get after doing some calculations

$$\int \frac{(x^3 - x - 3xu^2)du + (u^3 + u - 3x^2u)dx}{(u^2 + x^2)(u^2 + (x + 1)^2)(u^2 + (x - 1)^2)} = c,$$

which, finally can be evaluated relatively easily, using partial fractions, to yield the result

$$\frac{1}{2} \tan^{-1} \left(\frac{u}{x + 1} \right) + \frac{1}{2} \tan^{-1} \left(\frac{u}{x - 1} \right) - \tan^{-1} \left(\frac{u}{x} \right) = c.$$

We can use trigonometric identities to reduce the above to the simpler form

$$\frac{(u^2 + x^2)^2 + u^2 - x^2}{xu} = k$$

where k is a constant. (We would have been extremely hard pressed to come up with this by traditional methods.) \blacklozenge

²⁴See [Hyd00] p. 37.

As we have seen in these examples, using symmetry groups to solve equations really does have advantages—one can solve certain equations that would have been very difficult or impossible with traditional methods.

Another important method for integrating differential equations is to make use of the concept of a *differential invariant* which we define below, but to understand what is going on with these differential invariants, let us first look at the invariants of group actions.

As an example of this, consider the actions of the group $SO(2)$ on \mathbb{R}^2 . Since $SO(2)$ is a group of matrices that induces rotations on the plane, we can see that the orbit of an element $(x, u) \in \mathbb{R}^2$ will be a concentric circle about the origin with radius $r(x, u) = \sqrt{x^2 + u^2}$. Notice that on each orbit the function $r(x, u)$ is constant. That makes r an invariant of this group action. Furthermore, r separates the orbits, that is, if (x, u) and (\hat{x}, \hat{u}) are such that they are not on the same orbit, then $r(x, u) \neq r(\hat{x}, \hat{u})$. These conditions make r into a *differential invariant* of this group action.

Definition 26. Let G be a local group of transformations acting on $M \subset X \times U$. An n^{th} order *differential invariant* of G is a smooth function $\eta : M^{(n)} \rightarrow \mathbb{R}$, depending on x, u and the derivatives of u up to order n , such that η is an invariant of the prolonged group action $\text{pr}^{(n)} G$, that is

$$\eta(\text{pr}^{(n)} g \cdot (x, u^{(n)})) = \eta(x, u^{(n)}), \quad (x, u^{(n)}) \in M^{(n)}$$

(or that $\eta \circ \text{pr}^{(n)} g = \eta$) for all $g \in G$ for which $\text{pr}^{(n)} g \cdot (x, u^{(n)})$ is defined.

Dimensions worked nicely in the case of $SO(2)$ acting on \mathbb{R}^2 to the point that we didn't need to prolong the group action. However, this is not always the case. We may need to prolong the group action to the jet spaces in order to get a dimension greater than that of the group, for the orbits foliate the jet space by defining submanifolds that have the same dimension as the group. This enables us to find functions that are constant on orbits and obtains different values for different orbits.

There are several theorems which enable us to determine differential invariants. We note these here (see [Olv93] p. 140-142).

Theorem 15. Let G be a group of transformations acting on $M \subset X \times U \cong \mathbb{R}^2$. Suppose that $y = \eta(x, u^{(n)})$ and $w = \zeta(x, u^{(n)})$ are n^{th}

order differential invariants of G . Then the derivative

$$\frac{dw}{dy} = \frac{dw/dx}{dy/dx} = \frac{D_x \zeta}{D_x \eta}$$

(where D_x is the total derivative) is an $(n+1)^{\text{st}}$ order differential invariant of G .

Definition 27. A set $\{h_1, h_2, \dots, h_n\}$ forms a **complete set of functionally independent invariants** for a prolongations if they are invariants for which there does not exist a function f such that $f(h_1, h_2, \dots, h_n) = 0$ (that is they are functionally independent), and if g is any other invariant, then $g \in \text{Span}\{h_1, h_2, \dots, h_n\}$.

Theorem 16. Suppose that G is a one-parameter group of transformations acting on $M \subset X \times U \cong \mathbb{R}^2$. Let $y = \eta(x, u)$ and $w = \zeta(x, u, u_x)$ be a complete set of functionally independent invariants of the first prolongation $\text{pr}^{(1)}G$. Then the derivatives

$$y, w, \frac{dw}{dy}, \dots, \frac{d^{n-1}w}{dy^{n-1}}$$

provide a complete set of functionally independent invariants for the n^{th} prolongation $\text{pr}^{(n)}G$ for $n \geq 1$.

Example 11. As an example of how this theorem works, consider the example of $SO(2)$ acting on the plane. Then clearly the function $y = \sqrt{x^2 + u^2}$ is a first order differential invariant (that is, it is invariant on the 0-jet). Furthermore, we can construct a second differential invariant (which will be on the 1-jets) by considering that for $SO(2)$,

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

From this, we can see that the characteristic equation is

$$\frac{dx}{-u} = \frac{du}{x} = \frac{du_x}{1 + u_x^2}.$$

Which noting that we already have the relation that $x = \sqrt{r^2 - u^2}$ as an invariant, we can restate this characteristic equation as a separable equation

$$\frac{du}{\sqrt{r^2 - u^2}} = \frac{du_x}{1 + u_x^2},$$

which shows us that

$$\sin^{-1} \left(\frac{u}{r} \right) = \tan^{-1} u_x + k$$

where k is a constant. We can use right triangle trigonometry to show that we have the relation

$$\tan^{-1} u_x - \tan^{-1} \left(\frac{u}{x} \right) = k.$$

Then, we take the tangent of each side, and make use of the formula $\tan(\alpha - \beta) = \frac{\tan \alpha - \tan \beta}{1 + \tan \alpha \tan \beta}$, we see that

$$(51) \quad w := \frac{xu_x - u}{x + uu_x} = \text{constant}.$$

This shows us that the left hand side of equation (51) is actually an invariant of $\text{pr}^{(1)} SO(2)$ that is a function on the 1-jet. Furthermore, it is independent from $\sqrt{x^2 + u^2}$ since the former is clearly dependent on u_x whereas the latter is not.

Geometric intuition tells us that the curvature, κ , is also invariant under $SO(2)$. What Theorem 16 tells us is that we should be able to express the curvature,

$$\kappa = \frac{u_{xx}}{(1 + u_x^2)^{(3/2)}}$$

which is a function on the 2-jets (and hence invariant under the second prolongation), as a function $\kappa = \kappa \left(y, w, \frac{dw}{dy} \right)$. We compute this latter derivative.

$$w_y = \frac{dw}{dy} = \frac{dw/dx}{dy/dx} = \frac{\sqrt{x^2 - u^2}}{(x + uu_x)^3} \left((x^2 + u^2)u_{xx} - (1 + u_x^2)(xu_x - u) \right).$$

Again, it is clear that this is functionally independent from y, w . One can readily check that

$$\kappa = \frac{u_{xx}}{(1 + u_x^2)^{(3/2)}} = \frac{w_y}{(1 + w^2)^{(3/2)}} + \frac{w}{y(1 + w^2)^{(1/2)}}.$$

(Equivalently, we could have written w_y as a function of κ .) It is a straightforward calculation to show that κ and w_y are in fact invariants by letting $\text{pr}^{(2)} \mathbf{v}$ differentiate either of these and seeing that the result is 0.

This shows us that once we have a complete set of independent invariants for a particular prolongation (we use dimensions to determine how many there should be), *then any other invariant must be able to be written as a function of those invariants.*

For instance, the most general first order ordinary differential equation admitting $SO(2)$ is

$$\frac{xu_x - u}{x + uu_x} = H(\sqrt{x^2 + u^2})$$

where H is an arbitrary function (see [BK89] p. 107). ◆

One more theorem completes our list. This theorem deals with reduction of order, and as such, is very similar to Lie's Theorem 7; however, it includes some of the formality that was absent in that theorem.

Theorem 17. *Let G be a local group of transformations acting on $M \subset X \times U$. Assume $\text{pr}^{(n)} G$ acts semi-regularly on an open subset of $M^{(n)}$, and let $\eta^1(x, u^{(n)}), \dots, \eta^k(x, u^{(n)})$ be a complete set of functionally independent n^{th} order differential invariants. An n^{th} order differential equation $\Delta(x, u^{(n)}) = 0$ admits G as a symmetry group if and only if there is an equivalent equation*

$$\tilde{\Delta}(\eta^1(x, u^{(n)}), \dots, \eta^k(x, u^{(n)})) = 0$$

involving only the differential invariants of G . In particular, if G is a one-parameter group of transformations, any n^{th} order differential equation having G as a symmetry group is equivalent to an $(n - 1)^{\text{st}}$ order equation

$$\tilde{\Delta}\left(y, w, \frac{dw}{dy}, \dots, \frac{d^{n-1}w}{dy^{n-1}}\right) = 0$$

involving the invariants $y = \eta(x, u), w = \zeta(x, u, u_x)$ of $\text{pr}^{(1)} G$ and their derivatives. Furthermore, we can construct a solution $u = f(x)$ of $\Delta = 0$ by a single quadrature from a solution $w = h(y)$ of $\tilde{\Delta}$.

Example 12. Suppose that we have a differential equation $f(u, u_x, u_{xx}) = 0$ where the equation does not depend on the independent variable x . By inspection, this equation is clearly invariant under the infinitesimal transformation $\frac{\partial}{\partial x}$ (this corresponds to a translation in the x -direction). We can see this by comparing the original equation to one where we replace every x by $x + \epsilon$; the equation is obviously unchanged since there is no x in the equation. The typical method used to solve this equation in elementary courses is to simply "call" the dependent variable u the independent variable. We do this with a little bit more formality, but still wanting $\frac{\partial}{\partial x} = \frac{\partial}{\partial w}$ where, in our change of coordinates, y becomes the independent variable and w is the dependent. So we let

$$y = u, \quad w = x.$$

It is a direct calculation using the chain rule to show that

$$u_x = \frac{1}{w_x}$$

and that

$$u_{xx} = \frac{-w_{yy}}{w_y^3}.$$

Letting $z = w_y$, we have managed to reduce that $f(u, u_x, u_{xx}) = 0$ reduces to the equation $F(y, z, z_y) = 0$, thus illustrating the theorem.

◆

Example 13.²⁵ As another example, consider a homogeneous second order linear equation

$$u_{xx} + p(x)u_x + q(x)u = 0.$$

Again, by inspection, we see that this has a symmetry group, namely the group

$$G_\lambda : (x, u) \mapsto (x, \lambda u),$$

since replacing all the u 's by λu leaves the equation unchanged (since it is homogeneous). This scaling has an infinitesimal generator

$$\mathbf{v} = u \frac{\partial}{\partial u}.$$

Noting that the characteristic equation is

$$\frac{dx}{0} = \frac{du}{u} = d(\log u),$$

we see that x is constant, and that a suitable change of coordinates (in line with Theorem 6) is $y = x$ and $w = \log u \Rightarrow u = e^w$ (provided that $u \neq 0$, in which case the equation is trivial). With this change of coordinates, we can see that

$$\mathbf{v} = \frac{\partial}{\partial w}.$$

Furthermore, by differentiation of the change of variable results, we can see that

$$u_x = e^w w_y, \quad \text{and} \quad u_{xx} = (w_{yy} + w_y^2)e^w.$$

Now, we substitute these results back into our equation, arriving at the equation

$$e^w (w_{yy} + w_y^2) + p(y)w_y e^w + q(y)e^w = 0$$

and noting that $e^w \neq 0 \forall w$, we rewrite this as

$$w_{yy} + w_y^2 + p(y)w_y + q(y) = 0$$

which does not explicitly depend on w , like in the previous example. So, as in that example, we make use of the substitution $z = w_y$ arriving at the new equation

$$z_y + z^2 + p(y)z + q(y) = 0$$

²⁵See [Olv93] p. 139.

or equivalently, the first order Riccati equation

$$(52) \quad z_y = -z^2 - p(y)z - q(y).$$

◆

We make only a slight foray into equations that admit multi-parameter subgroups, since much of the analysis is similar to that above where the equation admits a single parameter subgroup. Recall in that case, we were able to reduce the order of the equation by one. Now suppose that $\Delta(x, u^{(n)}) = 0$ admits an r -dimensional (or parameter) symmetry group G . Our intuition tells us that we should be able to reduce the order to an $(n-r)^{th}$ order differential equation. In general, this is true; however, it may not be possible to reconstruct the solutions to the n^{th} order original equation from these by quadrature alone (remember that in order for us to consider a differential equation “solved” we must be left only with evaluating an antiderivative). In our case, we will assume that the r^{th} prolongation $\text{pr}^{(r)}G$ has r -dimensional orbits on $X \times U^{(r)}$; we make this assumption to avoid some technically complicated degenerate cases. This will imply that there are two independent r^{th} order differential invariants

$$(53) \quad y = \eta(x, u^{(r)}), \quad w = \zeta(x, u^{(r)}).$$

Furthermore, by noting that $r = \text{dimension of orbits} \leq \text{dimension } G = r$, we see that the orbits will remain r -dimensional for any further prolongations. Hence by Theorem 16, we can arrive at a complete set of independent invariants

$$y, w, \frac{dw}{dy}, \frac{d^2w}{dy^2}, \dots, \frac{d^{n-r}w}{dy^{n-r}}.$$

By Theorem 17, since $\Delta(x, u^{(n)}) = 0$ by assumption, we have that there exists an equivalent ODE $\tilde{\Delta}$ such that

$$(54) \quad \tilde{\Delta} \left(y, w, \frac{dw}{dy}, \dots, \frac{d^{n-r}w}{dy^{n-r}} \right) = 0.$$

So, here we have, in Equation (54), essentially reduced the problem to solving an $(n-r)^{th}$ order equation. However, “the principle problem at this juncture is that it is unclear how we determine the solution $u = f(x)$ of the original system from the general solution $w = h(y)$ of the reduced system (54).”²⁶ We consider this in the following example.

²⁶See [Olv93] p. 145.

Example 14. Let $SL(2)$ act on \mathbb{R} by the action of linear fractional transformation

$$(x, u) \mapsto \left(\frac{\alpha x + \beta}{\gamma x + \delta}, u \right).$$

Note, that since

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL(2),$$

$$\begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} = 1.$$

This determinant condition gives a restriction on the parameters $\alpha, \beta, \gamma, \delta$; thus we consider $SL(2)$ to be a 3-parameter group.

This action has the generators corresponding to the transformations

$$\begin{aligned} \mathbf{v}_1 &= \frac{\partial}{\partial x} && \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} && \text{translation,} \\ \mathbf{v}_2 &= 2x \frac{\partial}{\partial x} && \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} && \text{dilation,} \\ \mathbf{v}_3 &= x^2 \frac{\partial}{\partial x} && \begin{pmatrix} 1 & 0 \\ -\gamma & 1 \end{pmatrix} && \text{inversion.} \end{aligned}$$

The space $X \times U^{(3)}$ has general coordinates $(x, u, u_x, u_{xx}, u_{xxx})$, so it is five dimensional. Since $SL(2)$ is three dimensional, we have by assumption that the orbits are also three dimensional. Hence there are two invariants. (The number of invariants is the codimension of the orbits.) The invariant $y = u$ on the 0-jet is obvious. Much less obvious is the Schwarzian derivative

$$w = 2u_x^{-3}u_{xxx} - 3u_x^{-4}u_{xx}^2.$$

How this is arrived at is a bit of a mystery (it's a rather involved process that follows from the linear fractional transformation); however, it is computational (and the calculation is not particularly informative) to check that it is invariant under $\text{pr}^{(3)} \mathbf{v}_i$, $i = 1, 2, 3$. Since y, w are clearly independent, and there are only two differential invariants, it follows that these two form a complete set of functionally independent invariants for the prolongation of $\text{pr}^{(3)} SL(2)$.

Hence if $\Delta(x, u^{(n)}) = 0$ is invariant under this action, we must have that Δ is equivalent to a $(n - 3)^{rd}$ order equation

$$(55) \quad \tilde{\Delta} \left(y, w, \frac{dw}{dy}, \dots, \frac{d^{n-3}w}{dy^{n-3}} \right) = 0$$

involving only the invariants of $\text{pr}^{(n)} SL(2)$. Suppose that we have arrived at a solution $w = h(y)$ for Equation (55). Then by using the fact that the Schwarzian derivative is an invariant, we must have that

$$(56) \quad 2u_x^{-3}u_{xxx} - 3u_x^{-4}u_{xx}^2 = h(u),$$

which is equivalent to

$$(57) \quad 2u_x u_{xxx} - 3u_{xx}^2 = u_x^4 h(u).$$

In order to find a solution $u = f(x)$ of $\Delta(x, u^{(n)}) = 0$, we must find a solution, $u = f(x)$, of Equation (57). Note that Equation (57) is invariant under $SL(2)$ since it is equivalent to Equation (56) which can be written entirely in terms of the invariants. So, we can try to use this knowledge to integrate it. Letting $y = u$, $z = u_x$, we can rewrite equation (57) as

$$(58) \quad 2z \frac{d^2 z}{dy^2} - \left(\frac{dz}{dy} \right)^2 = z^2 h(y)$$

which is clearly invariant under the scaling $(y, z) \mapsto (y, \lambda z)$, and so Equation (58) can be reduced to the first order Riccati equation

$$(59) \quad 2 \frac{dv}{dy} + v^2 = h(y)$$

where $v = (\log z)_y = \frac{z_y}{z}$.

But now we are stuck. We have used \mathbf{v}_1 and \mathbf{v}_2 already, and the Riccati Equation (59) does not admit \mathbf{v}_3 ; so this symmetry is useless to us at this stage. Hence all we can really say is “that the solution of an n^{th} order differential equation invariant under the projective group can be found from the general solution of a reduced $(n - 3)^{\text{rd}}$ order equation using two quadratures and the solution of an auxiliary first order Riccati equation.”²⁷ \blacklozenge

“*This whole example is illustrative of an important point. If we reduce the order of an ordinary differential equation using only a subgroup of a full symmetry group, we may very well lose any additional symmetry properties present in the full group.*”²⁸ In order to solve this problem, we must in fact use normal subgroups (where we consider “normal” in the usual algebraic sense). This result is exemplified in the following theorem (see [Olv93] p. 148).

Theorem 18. *Let $H \subset G$ be an s -parameter normal subgroup of a Lie Group of transformations acting on $M \subset X \times U \cong \mathbb{R}^2$ such that*

²⁷See [Olv93] p. 147.

²⁸See [Olv93] p. 147.

$\text{pr}^{(s)} H$ has s dimensional orbits in $M^{(s)}$. Let $\Delta(x, u^{(n)}) = 0$ be an n^{th} order ordinary differential equation admitting H as a symmetry group, with corresponding reduced equation $\tilde{\Delta}(y, w^{(n-s)}) = 0$ for the invariants $y = \eta(x, u^{(s)}), w = \zeta(x, u^{(s)})$ of H . There is an induced action of the quotient group G/H on $\tilde{M} \subset Y \times W$ and Δ admits all of G as a symmetry group if and only if the H -reduced equation $\tilde{\Delta}$ admits the quotient group G/H as a symmetry group.

There are several problems to note with this. First, it is vital that the reduction procedure be done in the right order—doing a reduction in one order makes the solution straightforward, whereas doing it in a different order may make it impossible to solve. Second, depending on the dimension, there may not be a sufficient number of normal subgroups to continue applying Theorem 18. To solve this final problem, we return to the concept of a *solvable* group (c.f. Definition 3 for the algebraic treatment). In fact, the definition (see [Olv93] p. 151) is almost identical.²⁹

Definition 28. Let G be a Lie Group with Lie Algebra \mathfrak{g} . Then G is **solvable** if there exists a chain of Lie subgroups

$$\{e\} = G^{(0)} \subset G^{(1)} \subset \dots \subset G^{(r-1)} \subset G^{(r)} = G$$

such that for each $k = 1, \dots, r$, $G^{(k)}$ is a k -dimensional subgroup of G and $G^{(k-1)}$ is a normal subgroup of $G^{(k)}$.

There are several equivalent ways that we can express this concept, dealing with the Lie algebra and vector fields. For instance, we can equivalently say that there exists a chain of subalgebras

$$\{0\} = \mathfrak{g}^{(0)} \subset \mathfrak{g}^{(1)} \subset \dots \subset \mathfrak{g}^{(r-1)} \subset \mathfrak{g}^{(r)} = \mathfrak{g}$$

such that for each k , $\dim \mathfrak{g}^{(k)} = k$ and $\mathfrak{g}^{(k-1)}$ is a normal subalgebra of $\mathfrak{g}^{(k)}$, i.e. $[\mathfrak{g}^{(k-1)}, \mathfrak{g}^{(k)}] \subset \mathfrak{g}^{(k-1)}$. Furthermore, we could also consider it to be equivalent to the existence of a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ of \mathfrak{g} such that

$$[\mathbf{v}_i, \mathbf{v}_j] = \sum_{k=1}^{j-1} c_{ij}^k \mathbf{v}_k, \quad \text{whenever } i < j.$$

As a note, clearly we have that any abelian Lie algebra is solvable (this makes sense since any abelian group is solvable). “Also every two

²⁹Olver is using a strange numbering system in his chains. It is backwards from the way many authors—including Hungerford—chose to write them (normally, one would consider $G^{(0)} = G, G^{(r)} = \{e\}$ in a solvable series). The probable reason for this discrepancy is that Olver is trying to match his notation with the dimension of the subgroup (or subalgebra) in question.

dimensional Lie algebra is solvable. The simplest Lie algebra that is not solvable is the three-dimensional algebra $\mathfrak{sl}(2)$, the Lie algebra of $SL(2)$.”³⁰

We are about to finally arrive at the fruits of our labor. We sought a connection between the Abstract Algebra result that solvability of the Galois group leads to the polynomial equation being solvable by radicals (c.f. Theorem 5) and the result in the Lie version of Differential Equations that an equation admitting a solvable symmetry group leads to the equation being solvable by quadrature (c.f. Theorem 20). However, as in the case with the result in Abstract Algebra, we will bootstrap our way to this result, starting with a elementary proposition.

Proposition 5. *Suppose that G is a two-parameter group. There exists a basis $\{\mathbf{v}, \mathbf{w}\}$ for the Lie algebra \mathfrak{g} such that $[\mathbf{v}, \mathbf{w}] = \mathbf{v}$.*

Proof. Suppose that G is a two-parameter group. Then we can clearly find a basis for \mathfrak{g} containing two elements, $\{e_1, e_2\}$. We can find another basis $\{\hat{\mathbf{v}}, \hat{\mathbf{w}}\}$ for \mathfrak{g} by considering

$$\begin{bmatrix} \hat{\mathbf{v}} \\ \hat{\mathbf{w}} \end{bmatrix} = A \cdot \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

where A is any nonsingular matrix. In particular, let $a, b \in \mathbb{R}$, $a^2 + b^2 \neq 0$. Then choose

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

With this choice of A , we can see that

$$\begin{aligned} \hat{\mathbf{v}} &= ae_1 + be_2, \\ \hat{\mathbf{w}} &= -be_1 + ae_2. \end{aligned}$$

Which, using Equation (11), we can see that

$$[\hat{\mathbf{v}}, \hat{\mathbf{w}}] = a^2 + b^2.$$

Now, if we let $\mathbf{w} = \frac{1}{a^2+b^2}\hat{\mathbf{w}}$, $\mathbf{v} = \hat{\mathbf{v}}$, we see that indeed

$$[\mathbf{v}, \mathbf{w}] = \mathbf{v}.$$

□

We begin with the simple case (see [Olv93] p. 149)—recalling that all 2-parameter Lie groups are solvable.

³⁰See [Olv93] p. 151.

Theorem 19. *Let $\Delta(x, u^{(n)}) = 0$ be an n^{th} order ordinary differential equation invariant under a two-parameter (and hence solvable) symmetry group G . Then there is an $(n - 2)^{\text{nd}}$ order equation $\hat{\Delta}(z, v^{(n-2)}) = 0$ with the property that the general solution to Δ can be found by a pair of quadratures from the general solution to $\hat{\Delta}$.*

Proof. Let $\Delta(x, u^{(n)}) = 0$ be an n^{th} order differential equation that has a two-parameter symmetry group G . Then, by Proposition 5, we can find a basis $\{\mathbf{v}, \mathbf{w}\}$ for \mathfrak{g} with the property that

$$[\mathbf{v}, \mathbf{w}] = \mathbf{v}.$$

The one-parameter subgroup H generated by \mathbf{v} is then a normal subgroup of G , with one-parameter quotient group G/H . To effect the reduction of Δ , we begin by determining first order differential invariants $y = \eta(x, u), w = \zeta(x, u, u_x)$ for H using our earlier methods. By Theorem 17, our n^{th} order equation is equivalent to an $(n - 1)^{\text{st}}$ order equation

$$\Delta(y, w^{(n-1)}) = 0;$$

moreover, once we know the solution $w = h(y)$ of this latter equation, we can reconstruct the solution to Δ by solving

$$\zeta(x, u, u_x) = h(\eta(x, u))$$

with a single quadrature.

Since H is normal, by Theorem 18, the reduced equation $\tilde{\Delta}$ is invariant under the action of G/H on the variables (y, w) , and hence we can employ our earlier methods for one-parameter symmetry groups to reduce the order yet again by one to

$$\hat{\Delta} \left(z, v, \frac{dv}{dz}, \dots, \frac{d^{n-2}v}{dz^{n-2}} \right) = 0.$$

Now, suppose that we can find a solution $v = g(z)$ of $\hat{\Delta} = 0$. Then by Theorem 17, we can construct a solution $w = h(y)$ of $\tilde{\Delta}$ by a single quadrature from $v(y, w, w_y) = g(z(y, w))$. But as noted above, once we have a solution of $\tilde{\Delta}$, we can find a solution $u = f(x)$ for Δ by a single quadrature. Thus we can go from a solution of $\hat{\Delta}$ to a solution of Δ by two quadratures. \square

Now, suppose that we have that H is a normal abelian s -parameter subgroup of G , where G acts on the space $X \times U$. Assuming that H is a symmetry group of Δ , we can reduce the n^{th} order ODE Δ on the set $X \times U^{(n)}$ to an equivalent ODE, $\tilde{\Delta}$ on the set $Y \times W^{(n-s)}$ (where y, w are invariants of the action of H). But since $H \triangleleft G$, we have that

there is an induced action of G/H on $Y \times W$. However, Theorem 18 then ensures us that G is in fact a symmetry group of Δ .

$$\begin{array}{ccc}
 X \times U^{(n)} & \xrightarrow{\Delta} & \mathbb{R} \\
 \downarrow & & \\
 \vdots & \longrightarrow & Y \times W^{(n-s)} \xrightarrow{\tilde{\Delta}} \mathbb{R} \\
 \downarrow & & \downarrow \\
 \vdots & \longrightarrow & \vdots \\
 \downarrow & & \downarrow \\
 X \times U^{(s)} & \xrightarrow{(y,w)} & Y \times W \\
 \downarrow & & \\
 \vdots & & \\
 \downarrow & & \\
 X \times U & &
 \end{array}$$

Before we can get to the final result, we need to cover one other idea. Let G be an r -parameter Lie Group acting on an m -dimensional manifold, M . Suppose that $\Delta(x, u^{(n)}) = 0$ is an n^{th} order ODE, and that $H < G$ is an s -parameter symmetry group of Δ . Let $\{y^1, y^2, \dots, y^s\}$ be the invariants of H . We can consider this is part of a coordinate system, and, somewhat arbitrarily extend it to a whole local coordinate system, $\{y^1, y^2, \dots, y^s, \hat{x}^1, \hat{x}^2, \dots, \hat{x}^{m-s}\}$ for M (we can do this primarily by the independence of the coordinates $\{y^1, y^2, \dots, y^s\}$). Hence, we can say that each infinitesimal generator of G (or perhaps more aptly each element of \mathfrak{g}) must be of the form

$$(60) \quad \mathbf{v}_k = \sum_{i=1}^s \eta_k^i(y) \frac{\partial}{\partial y^i} + \sum_{j=1}^{m-s} \hat{\xi}_k^j(y, \hat{x}) \frac{\partial}{\partial \hat{x}^j}$$

where

$$\sum_{i=1}^s \eta_k^i(y) \frac{\partial}{\partial y^i}$$

is a well defined vector field on M/H .

We are now ready to state and prove the final major result (see [Olv93] p. 151-152).

Theorem 20. *Let $\Delta(x, u^{(n)}) = 0$ be an n^{th} order ordinary differential equation. If Δ admits a solvable r -parameter group of symmetries G such that for $1 \leq k \leq r$ the orbits of $\text{pr}^{(k)} G^{(k)}$ have dimension k , then the general solution of Δ can be found by quadratures from the general solution of an $(n - r)^{\text{th}}$ order differential equation $\tilde{\Delta}(y, w^{(n-r)}) = 0$. In particular, if Δ admits an n -parameter solvable group of symmetries, then (subject to the above technical restrictions), the general solution to Δ can be found by quadratures alone.*

Proof. The proof proceeds by induction along the chain of subalgebras in the equivalent condition of solvability of G . The initial stage of the induction argument was Theorem 19. Now assume the induction hypothesis: that the result of the theorem holds for the k^{th} stage. In other words, we have that Δ is invariant under the k -dimensional subalgebra $\mathfrak{g}^{(k)}$; thus, it can be reduced to an $(n - k)^{\text{th}}$ order equation

$$\tilde{\Delta}^{(k)}(y, w^{(n-k)}) = 0$$

in which $y, w, \frac{dw}{dy}, \dots, \frac{d^{n-k}w}{dy^{n-k}}$ form a complete set of functionally independent differential invariants for the n^{th} prolongation $\text{pr}^{(n)} G^{(k)}$; in particular, $y = \eta(x, u^{(k)}), w = \zeta(x, u^{(k)})$ form a complete set of invariants of the k^{th} prolongation of $G^{(k)}$. (Recall that we are assuming that the orbits of $G^{(k)}$ are k -dimensional.) Furthermore, we also assume that we can construct the general solution $u = f(x)$ from the general solution $w = h(y)$ of $\tilde{\Delta}^{(k)}$ by a series of k quadratures.

We must now show that the result holds for the $(k + 1)^{\text{st}}$ stage. In order to do this, consider a generator \mathbf{v}_{k+1} of $\mathfrak{g}^{(k+1)}$ which does not lie in $\mathfrak{g}^{(k)}$. Since $\mathfrak{g}^{(k)}$ is a normal subalgebra (in other words, an ideal) of $\mathfrak{g}^{(k+1)}$, $\text{pr}^{(k)} \mathbf{v}_{k+1}$ takes the form (see Equation 60)

$$\text{pr}^{(k)} \mathbf{v}_{k+1} = \text{pr}^{(k-2)} \mathbf{v}_{k+1} + \alpha(y, w) \frac{\partial}{\partial y} + \psi(y, w) \frac{\partial}{\partial w} \equiv \text{pr}^{(k-2)} \mathbf{v}_{k+1} + \tilde{\mathbf{v}}_{k+1}$$

in which $\text{pr}^{(k-2)} \mathbf{v}_{k+1}$ depends on the non-invariant coordinates x, u, \dots, u_{k-2} needed to complete y, w to a coordinate system on $M^{(k)}$. (One can further see this by assuming that the coordinates can be selected in that manner, and then seeing that all of the functions of the first $k - 2$ prolongations do not change with the change of variables. The only place there is something happening is with the coefficient functions of ∂_y, ∂_w .)

Then, Theorem 18 says (if we interpret it in terms of infinitesimals) that the original equation Δ is invariant under all of $\mathfrak{g}^{(k+1)}$ if and only if the reduced equation $\tilde{\Delta}^{(k)}$ is invariant under the reduced vector field

$\tilde{\mathbf{v}}_{k+1}$ which allows us to implement our reduction procedure for $\tilde{\Delta}^{(k)}$ using the vector field $\tilde{\mathbf{v}}_{k+1}$. Namely, we set

$$\hat{y} = \hat{\eta}(y, w), \quad \hat{w} = \hat{\zeta}(y, w, w_y)$$

to be independent invariants of the first prolongation $\text{pr}^{(1)} \tilde{\mathbf{v}}_{k+1}$. Then $\hat{y}, \hat{w}, \frac{d\hat{w}}{d\hat{y}}, \dots, \frac{d^{n-k-1}\hat{w}}{d\hat{y}^{n-k-1}}$ form a set of invariants for the $(n-k)^{\text{th}}$ prolongation $\text{pr}^{(n-k)} \tilde{\mathbf{v}}_{k+1}$. Since $\tilde{\Delta}^{(k)}$ determines an invariant subvariety of this group, there is an equivalent equation

$$\hat{\Delta}^{(k+1)}(\hat{y}, \hat{w}^{(n-k-1)}) = 0$$

depending only on the invariants of $\text{pr}^{(n-k)} \tilde{\mathbf{v}}_{k+1}$. Moreover, to reconstruct the solutions to $\tilde{\Delta}^{(k)}$ from those, $\hat{w} = \hat{h}(\hat{y})$, to $\hat{\Delta}^{(k+1)}$, we need only solve the first order equation

$$\hat{\zeta}(y, w, w_y) = \hat{h}(\hat{\eta}(y, w)).$$

This is invariant under the one-parameter group generated by $\tilde{\mathbf{v}}_{k+1}$ and hence can be integrated by quadrature. This completes the induction step, and thus the proof to the theorem. \square

We close this by considering an example.

Example 15. ³¹ Consider the third order equation

$$(61) \quad u_x^5 u_{xxx} = 3u_x^4 u_{xx}^2 + u_{xx}^3.$$

There is a three-parameter group of symmetries, generated by the vector fields

$$\mathbf{v}_1 = \frac{\partial}{\partial u}, \quad \mathbf{v}_2 = \frac{\partial}{\partial x}, \quad \mathbf{v}_3 = u \frac{\partial}{\partial x}.$$

This group is solvable since

$$[\mathbf{v}_1, \mathbf{v}_2] = 0, \quad [\mathbf{v}_1, \mathbf{v}_3] = \mathbf{v}_2, \quad [\mathbf{v}_2, \mathbf{v}_3] = 0,$$

i.e. its derived algebra is \mathbf{v}_2 which is abelian.

Hence, by Theorem 20, we have that Equation (61) can be solved by quadratures. First for $\mathfrak{g}^{(1)}$, which is generated by \mathbf{v}_1 , we have invariants $x, v = u_x$. (We need to choose the first one so that it is abelian in order for the series to be solvable.) Using these invariants, we can reduce Equation (61) to

$$(62) \quad v^5 v_{xx} = 3v^4 v_x^2 + v_x^3.$$

³¹See [Olv93] p. 152-154.

Now, the vector field \mathbf{v}_2 maintains its form $\tilde{\mathbf{v}}_2 = \frac{\partial}{\partial x}$ when written using the invariants of $\mathfrak{g}^{(1)}$, so to reduce Equation (62), for $\mathfrak{g}^{(2)} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ we need the invariants

$$y = v, \quad w = v_x, \quad w_y = \frac{v_{xx}}{v_x},$$

of $\text{pr}^{(2)} \tilde{\mathbf{v}}_2$. In terms of these, we can then reduce Equation (62) to the first order equation

$$(63) \quad y^5 w_y = 3y^4 w + w^2.$$

This is a Riccati equation. However, in contrast to Example 14, Equation (63) should still retain a symmetry corresponding the vector field \mathbf{v}_3 that we haven't yet used in any reduction. Indeed, in terms of $x, y = u_x, w = u_{xx}$,

$$\text{pr}^{(2)} \mathbf{v}_3 = u \frac{\partial}{\partial x} - y^2 \frac{\partial}{\partial y} - 3yw \frac{\partial}{\partial w}$$

and the reduced vector field

$$\tilde{\mathbf{v}}_3 = -y^2 \frac{\partial}{\partial y} - 3yw \frac{\partial}{\partial w}$$

is a symmetry of Equation (63). We now perform a change of variables in Equation (63) by setting $t = \frac{-1}{y}, z = \frac{w}{y^3}$. (Here we are essentially using Theorem 6.) Using this change of variables, we can see that $\tilde{\mathbf{v}}_3 = -\frac{\partial}{\partial t}$, and hence reduce Equation (63) to

$$\frac{dz}{dt} = z^2, \quad \Rightarrow z = \frac{1}{c - t}$$

for some constant c . Of course, what we really want is a solution of the original Equation (61). In many of the previous examples, working backwards was difficult to impossible. However, here we can. If we re-express the previous in terms of the invariants of $\tilde{\mathbf{v}}_2$,

$$w = \frac{y^4}{cy + 1}.$$

Now, going farther back to find v , we need to solve the autonomous (and separable) equation

$$\frac{dv}{dx} = \frac{v^4}{cv + 1}.$$

From this, we can find an implicit solution for $v = u_x$.

$$6(x - k)v^3 + 3cv + 2 = 6(x - k)u_x^3 + 3cu_x + 2 = 0$$

where k is another constant of integration.

We then solve the above equation for u_x . This is not direct, and may be very difficult (but remember there is a general form for a cubic equation), but once we do, we have reduced the solution of Equation (61) to quadrature. \blacklozenge

Before we close, several comments should be made about the results of the previous example. First of all, note that it is possible for a symmetry to disappear and then reappear. Note that there is no corresponding symmetry to \mathbf{v}_3 to Equation (62). Therefore, we have that “in the general reduction procedure, it is important to wait until we have the invariants for $\mathfrak{g}^{(k)}$ before trying to reduce the next vector field \mathbf{v}_{k+1} ; one cannot expect \mathbf{v}_{k+1} to naturally reduce relative to an earlier subalgebra $\mathfrak{g}^{(j)}$ if $j < k$.”³²

³²See [Olv93] p. 154.

8. DIFFERENTIAL GALOIS THEORY AND PICARD-VISSIOT EXTENSIONS

Recall, from Abstract Algebra, and from Section (1), the basics of Galois Theory for polynomial equations. Here we establish a slight parallel with differential equations. Let R denote a commutative ring (often we will actually want a field). We then can define the following (see [San98] p. 58).

Definition 29. An additive map $d : R \rightarrow R$ is called a *derivation* if d satisfies the Leibnitz rule, namely that $d(ab) = a \cdot d(b) + d(a) \cdot b$.

If R is actually an integral domain, then we can extend it to its field of quotients K and extend the map d to a map $d : K \rightarrow K$ (where we follow the traditional notation of referring to the extension by the same name) that obeys something similar to the quotient rule. Of course, if R is actually a field, R is its own field of quotients. Once we have this, we can define the following.

Definition 30. The pair (K, d) described above is called a *differential field*.

Now, let $L = Y^{(n)} + a_1 Y^{(n-1)} + \dots + a_n Y^{(0)}$ be a differential operator where the $a_j \in (F, d)$ (where F is a field).

Notation 9. We denote by $F\{Y\}$ the adjoining to F a differential indeterminate Y .

(The reader should recognize this is similar to the notation established for the polynomial ring $R[x]$.)

We are now ready to define a Picard-Vessiot extension (see [San98] p. 61-62).

Definition 31. Let (F, d) be a differential field and let $L \in F\{Y\}$ be a monic linear homogeneous differential operator. A differential field extension E of F is said to be a *Picard-Vessiot extension of F for L* , if the following are satisfied.

- (1) The field of constants of F is the same as the field of constants of E .
- (2) E is generated over F by the vector space V of all solutions of $L = 0$ in E .
- (3) E contains a full set of solutions to the differential equation $L = 0$. In other words, there exist $y_i, i = 1, 2, \dots, n$ where n is the order of L in V such that the Wronskian of y_1, y_2, \dots, y_n is not zero.

(Again the reader should have some inkling that this definition is similar to an extension of a polynomial ring up to the splitting field of some polynomial over that ring.)

With one more definition (that of a linear algebraic group), we can establish the Fundamental Theorem of Differential Galois Theory (see [Mag99] p. 1043).

Definition 32. A *linear algebraic group* over C is a subgroup of some general linear group $GL_n(C)$ that is the zero locus of a set of polynomials in the matrix entries.

Theorem 21. Fundamental Theorem of Differential Galois Theory: *Let $E \supset F$ be a Picard-Vessiot extension with common algebraically closed field C as the field of constants of both E and F . Let G be a group of differential automorphisms of E over F . Then G has structure of a linear algebraic group over C , and there is a one-to-one inclusion reversing correspondence between the intermediate differential field extensions $E \supset K \supset F$ and closed subgroups H of G , with K corresponding to*

$$\{g \in G \mid g(k) = k \text{ for all } k \in K\}$$

and H corresponding to

$$\{x \in E \mid h(x) = x \text{ for all } h \in H\}.$$

Moreover, H is a normal subgroup of G if and only if the corresponding extension K is a Picard-Vessiot extension of F , and if so, then the group of differential automorphisms of K over F is identified with G/H .

Of course, a reader familiar with the Fundamental Theorem of Galois Theory (c.f. Theorem 1) should see that these theorems are basically identical, with the group mentioned being the Galois Group in Theorem 1, and the *differential Galois group* in Theorem 21. The choice of this name is “appropriate and historical, Lie having pointed out the analogy in 1895.”³³ This connects back to differential equations for “the solvability of this group is then a necessary and sufficient condition for the [differential] equation to be solvable by quadratures.”³⁴

³³See [Mag99] p. 1043.

³⁴See [Hel94] p. 18.

9. CONCLUSION

We have seen that Sophus Lie, in his somewhat unique view of differential equations, created a means to examine the differential equations in a method that is not very familiar. Using this method of symmetries we are able to examine the differential equations and their solutions in a new light. It provides insight into aspects of differential equations that sometimes seem bizarre, or at least anti-intuitive such as the “Fundamental Source Solution” to the heat equation. It also enables us to be able to arrive very quickly at several key results, such as linearity of solutions and the principle of superposition. However, despite the seeming utility of this method, we have seen that the process of calculating the symmetries can become very cumbersome—perhaps this is why we are not as familiar with them as we are with the other methods that we have all seen in differential equations classes. However, exploring this method, has provided a new insight into differential equations and a better understanding of what is occurring in them. Furthermore, we have reached one of the major connections between differential equations and abstract algebra—one that is not commonly emphasized. This connection is related to the solvability of the equation. In the case of abstract algebra, we have the result that the polynomial equation is solvable by radicals if and only if the Galois group of the polynomial is a solvable group. The differential equation result is that if the equation admits a solvable group, then it is solvable by quadratures.

The field that Sophus Lie began as an attempt to better understand differential equations is one that has branched greatly. We have barely touched on all of the aspects here. Lie’s “group” (what today we would call a Lie Algebra), and the real group that we call a Lie Group have become fundamental subjects in their own right. There are many mathematicians studying these topics (and making new discoveries) today. (Several have been cited herein: Peter Olver at the University of Minnesota, Sigurdur Helgason at MIT, and Thomas Hawkins at Boston University.) In the interest of not having too many sources say the same thing, and to keep this from becoming a real book, we have, however, focused primarily on those found in Olver. Other sources include works by Peter Hydon from the University of Surrey and Hans Stephani at Friedrich-Schiller-Universität dealing with differential equations and their symmetries. The text mentioned by Bluman in the bibliography is also an excellent work on this topic (which is more readable than Olver). While we just mentioned it very briefly, the field of differential

Galois Theory is also a large field, with recent and authoritative works by Andy Magid at the University of Oklahoma and J.F. Pommaret, the Docteur ès-Sciences Ancien élève de l'École Polytechnique Ingénieur des Ponts et Chaussées. So, needless to say, we have seen the classical underpinnings of this theory, and there is definitely modern work continuing on many of the topics that are a result of Sophus Lie's *idée fixe*.

REFERENCES

- [BC96] David Bleecker and George Csordas, *Basic partial differential equations*, International Press, Cambridge, MA, 1996. MR MR1606155 (98m:35001)
- [BD65] William E. Boyce and Richard C. DiPrima, *Elementary differential equations and boundary value problems*, John Wiley & Sons Inc., New York, 1965. MR MR0179403 (31 #3651)
- [BK89] George W. Bluman and Sukeyuki Kumei, *Symmetries and differential equations*, Applied Mathematical Sciences, vol. 81, Springer-Verlag, New York, 1989. MR MR1006433 (91b:58284)
- [Haw00] Thomas Hawkins, *Emergence of the theory of Lie groups*, Sources and Studies in the History of Mathematics and Physical Sciences, Springer-Verlag, New York, 2000, An essay in the history of mathematics 1869–1926. MR MR1771134 (2001f:01031)
- [Hel90] Sigurdur Helgason, *Sophus Lie and the role of Lie Groups in Mathematics*, Report of the 14. Nordic LMFK-Conference (Reykjavik), 1990.
- [Hel94] ———, *Sophus Lie, the mathematician*, The Sophus Lie Memorial Conference (Oslo, 1992), Scand. Univ. Press, Oslo, 1994, pp. 3–21. MR MR1456457 (98g:01061)
- [Hun80] Thomas W. Hungerford, *Algebra*, Graduate Texts in Mathematics, vol. 73, Springer-Verlag, New York, 1980, Reprint of the 1974 original. MR MR600654 (82a:00006)
- [Hyd00] Peter E. Hydon, *Symmetry methods for differential equations*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2000, A beginner's guide. MR MR1741548 (2001c:58036)
- [Mag99] Andy R. Magid, *Differential Galois theory*, Notices Amer. Math. Soc. **46** (1999), no. 9, 1041–1049. MR MR1710665 (2001c:12007)
- [Olv93] Peter J. Olver, *Applications of Lie groups to differential equations*, second ed., Graduate Texts in Mathematics, vol. 107, Springer-Verlag, New York, 1993. MR MR1240056 (94g:58260)
- [OR] J. J. O'Connor and E.F. Robertson, *Marius Sophus Lie*, online <http://www.history.mcs.st-andrews.ac.uk/Biographies/Lie.html>, Accessed on 11/21/06.
- [San98] N. Sankaran, *Galois theory of differential fields—a survey*, Bull. Allahabad Math. Soc. **12/13** (1997/98), 57–83. MR MR1761034 (2001f:12010)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF HAWAI'I AT MĀNOA, HONOLULU, HI 96822

E-mail address: jburkman@math.hawaii.edu