Equivalence Transformations for Classes of Differential Equations

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy by

Ian G. Lisle

Department of Mathematics
University of British Columbia

Abstract

We consider classes $\mathcal{C}$ of differential equations characterized by the presence of *arbitrary elements,* that is, arbitrary functions or constants. Based on an idea of Ovsiannikov, we develop a systematic theory of *equivalence transformations,* that is, point changes of variables which map every equation in $\mathcal{C}$ to another equation in $\mathcal{C}$. Examples of nontrivial groups of equivalence transformations are found for some linear wave and nonlinear diffusion convection systems, and used to clarify some previously known results. We show how equivalence transformations may be inherited as symmetries of equations in $\mathcal{C}$, leading to a partial symmetry classification for the class $\mathcal{C}$. New symmetry results for a potential system form of the nonlinear diffusion convection equation are derived by this procedure.

Finally we show how to use equivalence group information to facilitate complete symmetry classification for a class of differential equations. The method relies on the geometric concept of a *moving frame,* that is, an arbitrary (possibly noncommuting) basis for differential operators on the space of independent and dependent variables. We show how to choose a frame which is invariant under the action of the equivalence group, and how to rewrite the determining equations for symmetries in terms of this frame. A symmetry classification algorithm due to Reid is modified to deal with the case of noncommuting operators. The result is an algorithm which combines features of Reid's classification algorithm and Cartan's equivalence method. The method is applied to the potential diffusion convection example, and yields a complete symmetry classification in a particularly elegant form.

Contents

1 Introduction **1**
- 1.1 Differential equations and their transformation . . . 1
- 1.2 Equivalence of differential equations: Examples . . . 3
- 1.3 Symmetries and differential equations . . . 9
- 1.4 Equivalence transformations . . . 10
- 1.5 Symmetry classification problem . . . 12

2 Transformation Groups and Differential Equations **15**
- 2.1 Transformation groups . . . 15
  - 2.1.1 Transformations, Lie groups . . . 15
  - 2.1.2 Infinitesimal operators . . . 18
  - 2.1.3 Invariant surface . . . 22
- 2.2 Extension . . . 23
  - 2.2.1 Notation for derivatives . . . 23
  - 2.2.2 Extension of transformation . . . 24
  - 2.2.3 Extension of group operator . . . 27
- 2.3 Differential equations and symmetry . . . 27
  - 2.3.1 Differential equations . . . 27
  - 2.3.2 Symmetries of differential equations . . . 28
  - 2.3.3 Algorithmic construction of symmetries . . . 31

3 The Equivalence Group **33**
- 3.1 Class of differential equations . . . 33
  - 3.1.1 Decoupled systems of d.e.'s . . . 33
  - 3.1.2 Class of d.e.'s . . . 34
- 3.2 Equivalence transformations . . . 36
- 3.3 Infinitesimal augmented transformations . . . 42
  - 3.3.1 Infinitesimal augmented transformations . . . 42
  - 3.3.2 Algebra of equivalence operators . . . 45
  - 3.3.3 Algorithm for construction of equivalence group . . . 49
  - 3.3.4 Proposition on form of infinitesimals . . . 50
  - 3.3.5 Structure of the equivalence group . . . 52
- 3.4 Examples of equivalence groups . . . 55
  - 3.4.1 Boltzmann's similarity solution for nonlinear diffusion . . . 55
  - 3.4.2 Nonlinear diffusion-convection equations . . . 61
  - 3.4.3 Wave equations . . . 68
  - 3.4.4 Hamilton's equations . . . 72

4 Symmetry Group Classification **75**
4.1 Symmetry classification problem . . . 75
4.1.1 Example: scalar diffusion convection . . . 76
4.2 Partial symmetry classification . . . 76
4.2.1 Symmetry inherited from equivalence group . . . 76
4.2.2 Optimal system of subalgebras . . . 80
4.2.3 Partial symmetry classification for nonlinear diffusion convection . . . 82
4.3 Modification of Reid algorithm . . . 90
4.3.1 Moving frame and determining equations . . . 92
4.3.2 Frame Reid method . . . 97
4.4 Invariant frame . . . 104
4.4.1 Augmented frame . . . 104
4.4.2 Invariant frame . . . 108
4.4.3 Differential invariants . . . 109
4.4.4 Tresse basis . . . 110
4.5 Symmetry classification . . . 112
4.5.1 Invariant form of group classification . . . 115
4.5.2 Potential diffusion convection system . . . 121

5 Conclusion **137**
5.1 Further work . . . 137
5.1.1 Isovector method for frame determining system . . . 138
5.2 Conclusions . . . 140

A Algorithms for Frame Systems **142**
A.1 Reduction to frame involutive form . . . 142
A.1.1 Orthonomic form . . . 142
A.1.2 Reduced orthonomic form . . . 143
A.1.3 Involutive form . . . 144
A.2 Group classification . . . 144

B Structure Constants **147**

C Similarity Solution for Nonlinear Diffusion **151**
C.1 Power law diffusivity . . . 151
C.1.1 Phase reduction . . . 153
C.1.2 Exact shooting . . . 154
C.1.3 Series solution . . . 154
C.2 Modified power law diffusivity . . . 155
C.2.1 Phase reduction . . . 155
C.2.2 Exact shooting . . . 156
C.2.3 Series solution . . . 156
C.3 Discussion . . . 157

List of Tables

3.1 Commutator table of equivalence algebra of nonlinear diffusion potential system. 49
3.2 Commutator table for equivalence operators of scalar diffusion convection equation 63
3.3 Commutation relations of equivalence algebra of scalar wave equation. 69

4.1 Symmetry classification for scalar nonlinear diffusion convection equation. 77
4.2 Optimal system of subalgebras for nonlinear diffusion convection potential system. 84
4.3 Partial symmetry classification for diffusion convection potential system: Case $K(u) = 0$ (diffusion equations). 86
4.4 Partial symmetry classification for diffusion convection potential system: Case with nonlinear convection. 87
4.5 Nonlocal symmetries inherited from equivalence group of diffusion convection potential system: Case $K(u) = 0$ (diffusion equations). 88
4.6 Nonlocal symmetries inherited from equivalence group of diffusion convection potential system: Case $K(u) \not\equiv 0$. 89
4.7 Commutator table of equivalence algebra of nonlinear diffusion equation. 116
4.8 Commutation relations for equivalence algebra of diffusion convection potential system 122

List of Figures

3.1 Relationship between linearizable diffusion convection potential systems 68

4.1 Classification tree for symmetries of nonlinear diffusion equation. . 120
4.2 Preliminary classification tree for potential diffusion convection system 131
4.3 Complete symmetry classification tree for diffusion convection potential system 136

C.1 Relation between concentration, flux, and spatial coordinate for Boltzmann's similarity solution with power law diffusivity. 152
C.2 Phase portrait for Boltzmann similarity solution with power law diffusivity 153

Acknowledgments

I wish to thank my supervisor Dr. G.W. Bluman for originally suggesting the study of equivalence transformations, and for his guidance, insight and unfailing sense of direction. Particular thanks are due for his careful and speedy reading of the drafts of my thesis.

My wholehearted appreciation also to Dr. Greg Reid for his constant encouragement and support, for his patient explanations of his symmetry classification algorithm, and for making available the MAPLE implementation of this algorithm. Alan Boulton was a constant source of enthusiasm over the last two years, participating in many lively discussions. The insightful lectures of Dr. K.Y. Lam on differential geometry provided me with the tools which made §4.3–§4.5 possible. The problem described in Appendix C resulted from discussions with Dr. J.Y. Parlange.

My fellow graduate students have maintained me with their friendship over the years, and have also enriched my knowledge of mathematics, computing, French, and softball. My most heartfelt appreciation to Maria Chiarolla for helping maintain my sanity when things did not go well. Honourable mention to the Stanley Cup playoffs, the few sunny days in winter, and Piper's Ale.

Funding for much of my stay in Canada was provided by the Canadian Commonwealth Scholarship and Fellowship Administration, and their support is gratefully acknowledged.

This thesis was typeset using LaTeX.

Chapter 1

Introduction

1.1 Differential equations and their transformation

In dealing with differential equations, a common situation is that one wishes to analyze simultaneously a whole class of equations of some given type. It is natural to consider 'the class of second order ordinary differential equations'

$$\frac{d^2y}{dx^2} = \omega(x, y, \frac{dy}{dx}) \tag{1.1}$$

or 'the nonlinear diffusion equation'

$$u_t = [D(u)u_x]_x$$

Here ω, D are arbitrary (smooth) functions of their arguments, at least in some suitable domain of definition. Thus the entire class of equations under consideration is specified by allowing these *arbitrary elements* to range over all possible functional forms.

In this dissertation, I will be concerned with the *transformation properties* of a given class $\mathcal{C}$ of differential equations. Attention will be restricted to *invertible* 'point' transformations, which act on a coordinate space of the independent and dependent variables. These are the usual 'changes of variables' in differential equations. For (1.1) for example, the most general such change of variables is

$$\begin{aligned} x' &= F(x, y) \\ y' &= G(x, y) \end{aligned} \tag{1.2}$$

(subject to the Jacobian $F_xG_y - G_xF_y$ being nonzero).

Any transformation applied to the variables in a differential equation (d.e.) yields another differential equation. Certain transformations are of particular interest:

symmetry A *symmetry* of a differential equation is a transformation which maps every solution of the differential equation to another solution of the *same equation*.

equivalence transformation An *equivalence transformation* for a differential equation in a given class is a change of variables which maps the equation to another equation in the *same class*.

We briefly discuss these types of transformations.

Knowledge of symmetries of a differential equation often assists in constructing (special or general) solutions of the d.e. In [13, 47, 9], symmetry methods for solving differential equations are described; [13] also discusses solutions of associated boundary value problems. Symmetry properties of a d.e. were also shown by Kumei and Bluman [13, 41, 14] to characterize whether a given differential equation can be mapped to a linear equation, and to give a means for constructing the linearizing map. We shall not be touching these applications (except briefly in §3.4.2). Rather, the methods we develop assist in constructing the symmetries themselves.

Equivalence transformations have been mainly used as a starting point for solving the *Cartan equivalence problem* (the problem is more properly due to Tresse [68], or even Lie [43]). Given a class of differential equations (for example all second order o.d.e.'s (1.1)), the Cartan equivalence problem is to find criteria for whether two d.e.'s are connected by a change of variables drawn from a transformation group $\mathcal{G}$ (for example all point changes of variable (1.2)). A method for constructing such criteria was given by Tresse [68], and subsequently used by him [69] to solve the equivalence problem for second order o.d.e.'s under point changes of variable. Cartan [19] radically reformulated the method, basing his solution method on the geometric theory of Pfaffian systems. The Cartan method (and Tresse's prior formulation) give equivalence criteria for the d.e.'s with respect to action of $\mathcal{G}$, but Cartan [19] showed that symmetry structure of the d.e.'s could also be found as a byproduct of his method.

Both Cartan and Tresse addressed the equivalence problem for classes of equations where some group $\mathcal{G}$ was already available. They were not concerned with the problem of *finding* a $\mathcal{G}$ 'suitable' to a given class of equations in the sense that each transformation in $\mathcal{G}$ maps an equation in the class to another equation in the class. Their examples were mainly concerned with finding equivalence criteria for 'geometrically natural' classes of objects, such as Riemannian metrics on a two dimensional space, or the set of second order o.d.e.'s. Following publication of Gardner's influential paper [25], such applications of the Cartan method have again become popular, with various authors treating ordinary and partial differential equations [35, 39, 34], Lagrangians [17, 61, 36, 37, 31], differential operators [38] and control problems [27]. In every case treated by these authors, the class of objects they analyze has associated with it a 'natural' group of transformations, usually the set of all point changes of variables or some subgroup thereof.

In contrast, one of our principal aims will be to show how to systematically *derive* a group $\mathcal{G}$ of transformations appropriate to a given class of d.e.'s. This line of reasoning was initiated by Ovsiannikov [52, §6.4], and has recently been applied by Ibragimov and coworkers [3, 4, 32] to various classes of partial differential equations. A theoretical foundation for their method of construction of this 'equivalence group' is not available, and we attempt to remedy this in Chapter 3. The advantage of dealing with the equivalence group is that it is often a 'small' (e.g., finite-parameter) group. The extensive geometric machinery of the Cartan equivalence method is geared to infinite transformation groups, and can often be dispensed with for finite groups. This permits us to obtain significant transformation information relatively easily.

With the equivalence group known, we may use it directly to map a solution of one d.e. in the class to a solution of another such d.e. However, just as the Cartan equivalence method incidentally yields *symmetry* information, so one of our principal uses of the equivalence group will be to assist in finding symmetries. In fact we shall devote an entire chapter §4 to this topic.

1.2 Equivalence of differential equations: Examples

Before developing any theory, we give a sequence of examples, illustrating various points about equivalence transformations.

Example 1.2.1. [Class closed under point transformations.] Consider the class (1.1) of second order ordinary differential equations (o.d.e.'s). Clearly *any* point transformation (1.2) maps a second order o.d.e. to another second order o.d.e. Substituting the change of variables (1.2) into an equation

$$\frac{d^2y'}{dx'^2} = \omega'(x', y', \frac{dy'}{dx'}) \tag{1.3}$$

shows that the undashed variables (x, y) satisfy

$$\begin{aligned}\frac{d^2y}{dx^2} = \Big\{(\Delta F)^3\,\omega'\Big(F, G, \tfrac{\Delta G}{\Delta F}\Big) - \Delta F\cdot(G_{xx} + 2pG_{xy} + G_{yy})\\ + \Delta G\cdot(F_{xx} + 2pF_{xy} + F_{yy})\Big\}\Big/(F_xG_y - F_yG_x)\end{aligned} \tag{1.4}$$

where the differential operator Δ is defined by

$$\Delta = \frac{\partial}{\partial x} + p\,\frac{\partial}{\partial y}$$

and $p \equiv \frac{dy}{dx}$.

Turning this around, it is seen that if two equations (1.3) and (1.1) are given, the dashed and undashed equations are connected by a change of variables (1.2) if and only if there exist functions $F(x, y)$, $G(x, y)$ such that

$$\begin{aligned}(\Delta F)^3\omega'\Big(F, G, \frac{\Delta G}{\Delta F}\Big) &= \Delta F(G_{xx} + 2pG_{xy} + p^2G_{yy})\\ &\quad - \Delta G(F_{xx} + 2pF_{xy} + p^2F_{yy}) + \omega\cdot(F_xG_y - F_yG_y).\end{aligned} \tag{1.5}$$

If such F, G can be found they can serve in the change of variables (1.2) to connect the two equations (1.1), (1.3). Such a criterion is useless in this form. For a given ω and ω' condition (1.5) represents a very complicated nonlinear partial d.e. in the unknowns F, G, and it is not apparent what to do with it. The equivalence problem, as treated by Tresse and Cartan, does not attempt to solve for F, G, but instead seeks *conditions on* ω *and* ω' *for this p.d.e. to have solutions.* The result is a complicated set of equations involving ω, ω' and their derivatives. The important point is that the functions F, G are not present. This means that whether equations are equivalent can be checked knowing only the equations: the equivalence problem (whether equations are equivalent) is thus separated from the more difficult problem of actually finding the transformation connecting the equations.

For this example, equivalence criteria were first found by Tresse [69], using his theory of equivalence [68] (see also [34] for a solution based on the Cartan equivalence method).

Example 1.2.2. Group of equivalence transformations.

Consider the class of nonlinear diffusion equations

$$u_t = [D(u)u_x]_x, \tag{1.6}$$

where $D(u) > 0$.

Under an arbitrary point transformation

$$\begin{aligned} x' &= \alpha(x,t,u) \\ t' &= \beta(x,t,u) \\ u' &= \gamma(x,t,u) \end{aligned}$$

equation (1.6) is certainly *not* mapped to another nonlinear diffusion equation. The most general point transformation which preserves the class of diffusion equations is the six-parameter equivalence group

$$\begin{aligned} x &= \lambda^2 c^{-1} x' + \varepsilon \\ t &= \lambda^2 t' + \delta \\ u &= au' + b, \qquad a, c, \lambda \neq 0 \end{aligned} \tag{1.7}$$

which maps (1.6) to a nonlinear diffusion equation with diffusivity

$$D'(u') = c^2 D(au' + b). \tag{1.8}$$

The simple transformations (1.7) reflect fundamental physical properties of the diffusion equation (arbitrary choice of units; arbitrary choice of origin for temperature), and are often used without comment for parameter elimination. They have the following significant properties:

Property (i) Correspondence (1.8) is established for *every* diffusivity D.

Property (ii) The *same* point transformation (1.7) establishes correspondence (1.8) for any diffusivity D.

Property (iii) Transformations (1.7) form a transformation *group* on (x, t, u) space.

In this case equivalence transformations (1.7) can be found by inspection.

Correspondence (1.8) for diffusivities is analogous to condition (1.5) for second order o.d.e.'s. The need for equivalence conditions on $D(u)$ (i.e., with parameters a, b, c eliminated) does not seem as pressing as in the o.d.e. case, but the only essential difference is finite versus infinite parameter groups (1.7) and (1.2) respectively.

Two diffusion equations with diffusivities $D(u)$, $D'(u')$ are connected by a transformation (1.7) if there exist constants a, b, c with $ac \neq 0$ such that D, D' are related by (1.8). This condition is analogous to (1.5) above: for given $D(u)$, $D'(u')$ it represents an equation to be satisfied by a, b, c. Whether this equation has solutions or not can be stated entirely in terms of D, D'. Denote derivatives of D with a dot. The criteria for equivalence with respect to (1.7) are

1. If both $\dot{D} = 0$ and $\dot{D}' = 0$ the equations are equivalent.
2. Suppose $\dot{D} \neq 0$, $\dot{D}' \neq 0$. Let
$$J := \left(\frac{D}{\dot{D}}\right)^{\cdot}$$
and let J' be the analogous quantity computed for the diffusivity D'. If J and J' are *constant* and *equal*, the equations are equivalent.
3. Suppose $J \neq \text{const}$, $J' \neq \text{const}$. Then the map $u \mapsto J$ is invertible, so J can serve as a coordinate instead of u. Let
$$K = \frac{D}{\dot{D}}\dot{J}$$
and let K' be the analogous quantity for the diffusivity D'. Express K as a function of J. If $K = f(J)$ and $K' = f(J')$ with the *same* function f, then the equations are equivalent.

In any other case the equations are not equivalent. □

A difference in emphasis is apparent between Examples 1.2.1 and 1.2.2. In the case of the ordinary d.e.'s, the class (1.1) comes equipped with a natural group of transformations (1.2) transforming equations to other equations. This is the kind of problem to which Cartan's equivalence method has usually been applied. In contrast, the diffusion equations (1.6) do not come equipped with a group, and we must somehow come up with transformations (1.7) as the appropriate ones. Once found, the group is sufficiently small that the equivalence criteria given above are superfluous: correspondence (1.8) intuitively seems more informative than "$K = f(J)$".

Example 1.2.3. Group of 'potential' equivalence transformations.

Consider the system of equations

$$\begin{aligned} v_x &= u \\ v_t &= D(u)u_x \end{aligned} \tag{1.9}$$

The scalar nonlinear diffusion equation (1.6) is *embedded* in system (1.9) in the following sense: if u, v satisfy (1.9), then u satisfies (1.6); conversely, if u satisfies (1.6) then there exists a function v such that u, v satisfy (1.9). Here v is a *potential* variable: we call (1.9) the *potential system* form of the nonlinear diffusion equation. As above, class (1.9) is not closed under arbitrary transformations of (x, t, u, v) space: the most general point transformation mapping a nonlinear diffusion potential system to another such system is given by the four-parameter family [3, 4]

$$\begin{aligned} v &= av' + bx' \\ x &= cv' + dx' \\ t &= t' \\ u &= \frac{au' + b}{cu' + d}, \qquad ad - bc \neq 0. \end{aligned} \tag{1.10}$$

The dashed variables satisfy a diffusion system with diffusivity D':

$$D'(u') = \frac{1}{(cu'+d)^2} D\Big(\frac{au'+b}{cu'+d}\Big). \tag{1.11}$$

Transformations (1.10) have the same properties (i), (ii), (iii) as in Example 1.2.2. When $c \neq 0$ in (1.10), the transformation is nontrivial, and cannot be found by inspection. For the potential system form, as for the scalar form of the diffusion equation, the main problem is to come up with the group (1.10).

This example is included to highlight the fact that transformation properties such as equivalence may vary depending on the form in which the original equation is written. We may think of the scalar equation (1.6) and potential system (1.9) as minor variants of the same equation, but from the viewpoint of point changes of variable, these forms differ radically, since the spaces (x, t, u) and (x, t, u, v) on which transformations act are different. In (1.10) if $c \neq 0$, then v' occurs explicitly in the transformation of x, and it is not possible to project (1.10) to a transformation acting on (x, t, u) space. Although transformations (1.10) establish a correspondence between scalar diffusion equations according to the schema

$$\begin{array}{ccc}
\text{potential system} & \xrightarrow{(1.10)} & \text{potential system} \\
D(u) & & \frac{1}{(cu+d)^2} D\Big(\frac{au+b}{cu+d}\Big) \\
\Big\uparrow & & \Big\downarrow \\
\text{scalar equation} & & \text{scalar equation} \\
D(u) & & \frac{1}{(cu+d)^2} D\Big(\frac{au+b}{cu+d}\Big),
\end{array}$$

the mapping is *nonlocal* because it involves v, which is an integral $\int u\,dx$. Transformations (1.10) are therefore *point* transformations for the potential system form (1.9), but *nonlocal* transformations of the scalar form (1.6) of the diffusion equation. This situation is analogous to symmetry properties of potential forms of equations. Where a symmetry transformation for a potential form of an equation is genuinely nonlocal, Bluman, et al. [15, 11] use the term 'potential symmetry'. Correspondingly, we may refer to (1.10) as 'potential equivalence transformations' of the scalar equation (1.6).

Knowledge of transformations (1.10) has immediate interesting consequences. For example, it follows that the system (1.9) with diffusivity $D(u) = (\alpha u + \beta)^{-2}$ can be mapped to the linear system

$$\begin{aligned} v_x &= u \\ v_t &= u_x \end{aligned}$$

for which u obeys the heat equation $u_t = u_{xx}$. This linearizing transformation was constructed in various ways in [10, 60, 15]. We discuss linearizing transformations for diffusion convection equations in §3.4.2.

Example 1.2.4. Classification for equivalence transformations.

Consider the class of ordinary differential equations

$$\frac{d^2y}{dx^2} = -\frac{k(x)}{2y} + m \exp\frac{dy}{dx} \tag{1.12}$$

with arbitrary function $k(x)$ and arbitrary constant m. The most general point transformation mapping each equation in the class to another such equation is the two-parameter family

$$\begin{aligned} x &= ax' + b \\ y &= ay', \qquad a \neq 0. \end{aligned} \tag{1.13}$$

A transformation of this form maps (1.12) to a similar 'dashed' equation, with new arbitrary elements

$$\begin{aligned} k'(x') &= k(ax' + b) \\ m' &= \tfrac{1}{a} m. \end{aligned}$$

Again these transformations share Properties (i), (ii), (iii) of Example 1.2.2: they act as a point group on the whole class of equations (1.12).

Now consider the *subclass* of equations

$$\frac{d^2y}{dx^2} = -\frac{k(x)}{2y} \tag{1.14}$$

obtained by setting $m = 0$ in (1.12). The most general transformation mapping any such equation to another of the same form is given by the *four*-parameter family

$$\begin{aligned} x &= \frac{ax' + b}{cx' + d} \\ y &= (ad - bc)\frac{y'}{cx' + d}, \qquad ad - bc \neq 0. \end{aligned} \tag{1.15}$$

Under the action of this transformation, equation (1.14) is mapped to a similar 'dashed' equation with arbitrary element

$$k'(x') = \frac{1}{(cx' + d)^2} k\Big(\frac{ax' + b}{cx' + d}\Big).$$

For this subclass (1.14) transformations (1.15) share Properties (i), (ii), (iii) above. However, if $c \neq 0$ in (1.15), the transformation does *not* map each equation in the class (1.12) to another equation of the same type. □

This example illustrates a general situation. For a given class of equations (such as (1.12)), we wish to find not only the equivalence transformations (like 1.13) which act on the class *as a whole*, but also to rationally *classify all subclasses* (like (1.14)) for which *additional* equivalence transformations appear. We do not address such classification questions here.

Example 1.2.5. Point equivalence transformations not forming a group.

Consider the class of nonlinear telegraph equations, written in potential system form:

$$\begin{aligned} v_t &= u_x \\ v_x &= c^2(u)u_t + b(u) \end{aligned} \tag{1.16}$$

In addition to some simple scaling and translation transformations (similar to Example 1.2.2), the one-parameter family of transformations

$$\begin{aligned} x' &= x - \varepsilon v \\ t' &= t - \varepsilon \int^u \frac{c^2(z)}{1-\varepsilon b(z)}\, dz \\ u' &= u \\ v' &= v \end{aligned} \tag{1.17}$$

maps each nonlinear telegraph system to another system with the same form, and new arbitrary elements b', c' given by [70]

$$\begin{aligned} b'(u') &= \frac{b(u')}{1-\varepsilon b(u')} \\ c'(u') &= \frac{c(u')}{1-\varepsilon b(u')}. \end{aligned} \tag{1.18}$$

This family of transformations shares Property (i) of Example 1.2.2: correspondence (1.18) is established for *any* arbitrary elements $b(u)$, $c(u)$. However Property (ii) is violated: (1.17) specifies a *different* point transformation for each choice of $b(u)$, $c(u)$. Moreover, Property (iii) is violated: transformations (1.17) do not form a point group acting on (x, t, u, v) space. For each nonlinear telegraph equation (1.16) (i.e., for each choice of $b(u)$, $c(u)$), mapping (1.17) is a point transformation. However, calculation of (x', t', u', v') depends not only on knowing *values* of the independent and dependent variables (x, t, u, v), but also on the *functions* $b(u)$, $c(u)$, so transformations (1.17) do *not* form a point group. However (1.17,1.18) may be regarded as a group of transformations acting on a function space with coordinates (x, t, u, v) and arbitrary functions $b(u)$, $c(u)$.

Example 1.2.6. Non-point equivalence.

Consider the linear hyperbolic equation

$$u_{xy} + A(x,y)u_x + B(x,y)u_y + C(x,y)u = 0, \tag{1.19}$$

which can be written in factored form:

$$(\partial_x + B)(\partial_y + A)u = hu$$

where $h = h(x, y)$ is the Laplace invariant [52, §9] $h = A_x + AB - C$. The factored equation may be written as the system

$$\begin{aligned} u_y + Au &= z \\ z_x + Bz &= hu \end{aligned} \tag{1.20}$$

Eliminating z yields (1.19) once again.

Suppose the Laplace invariant $h \neq 0$. Eliminating u from (1.20) yields another scalar equation

$$z_{xy} + A'(x,y)z_x + B'(x,y)z_y + C'(x,y)z = 0 \tag{1.21}$$

where

$$\begin{aligned} A' &= A - \frac{h_y}{h} \\ B' &= B \\ C' &= C + k - h - B\frac{h_y}{h} \end{aligned}$$

and $k = B_y + AB - C$ is the second Laplace invariant of (1.19). The two hyperbolic equations (1.19), (1.21) are put into correspondence by (1.20). The 'transformation' (1.20) is known as *Laplace's transformation* [52]. It is clearly *not* a point transformation (the map from $u(x, y)$ to $z(x, y)$ involves taking *derivatives* of u). Such non-point transformations are beyond the scope of the present investigation.

1.3 Symmetries and differential equations

Our main concern will be with construction and use of equivalence transformations for a class of equations. Before describing this, we review some ideas of Lie symmetry methods for differential equations. We give a more detailed account of this theory in §2, and for now limit ourselves to some comments on the general philosophy of Lie symmetry methods. The equivalence methods we describe are exactly parallel to these standard symmetry results.

A *point symmetry* of a differential equation (d.e.) is an invertible point transformation which maps every solution of the d.e. to another solution of the same d.e. The topic of symmetry for d.e.'s is by now well-studied. In the late nineteenth century, the Norwegian mathematician Sophus Lie developed the theory of continuous transformation groups (Lie groups) precisely to deal with such symmetries. He showed that the symmetries of a d.e. form a *group* (the *admitted group* of the equation). Knowledge of this group was shown by Lie to be of great assistance in understanding and constructing solutions of the d.e. The applications of symmetry groups to d.e.'s include [13, 47, 52]:

- mapping solutions to other solutions
- integration of ordinary d.e.'s in formula
- constructing invariant ('similarity') solutions, that is, solutions which are invariant under the action of a subgroup of the admitted group
- detection of linearizing transformations.

To execute any of these, a reliable method for finding symmetries of d.e.'s is required. In principle one could insert an arbitrary change of variables (e.g. (1.2)) into the equation (e.g. (1.1)) and then force the new variables (e.g. x', y') to satisfy the *same* differential equation. This yields a (usually large) number of (usually nonlinear) differential equations (the 'defining equations') to be satisfied by the transformation (e.g., the functions F, G of (1.2)). This direct approach is too cumbersome to be of much use: defining equations may be derived, but solving such a large system of nonlinear equations is usually out of the question.

The crucial insight of Lie was that this problem could be overcome by considering the 'infinitesimal' action of the group. An example is helpful. Consider the rotations of the (x, y) plane

$$\begin{aligned} x' &= x\cos\varepsilon + y\sin\varepsilon \\ y' &= -x\sin\varepsilon + y\cos\varepsilon, \end{aligned} \tag{1.22}$$

which form a group, whose transformations are parametrized by the angle ε of rotation. When $\varepsilon = 0$, the identity transformation results. Expanding in the neighbourhood of the identity $\varepsilon = 0$ gives

$$\begin{aligned} x' &= x + \varepsilon y + O(\varepsilon^2) \\ y' &= y - \varepsilon x + O(\varepsilon^2). \end{aligned} \tag{1.23}$$

The terms of order ε in (1.23) represent the derivative of (1.22) at the identity. Lie regards (1.23) as an 'infinitesimally small' rotation of the plane. His remarkable result (Lie's *first fundamental theorem*) is that the action of a group can be (essentially) *completely* recovered from the group's 'infinitesimal action', and involves only solution of an initial value problem for a finite system of *ordinary* differential equations.

Because of this the problem of finding symmetries reduces to the solution of a system of *linear* differential equations (the *determining equations*) for the infinitesimal group action. Once the infinitesimals are known, solving ordinary d.e. initial value problems suffices to recover the symmetry group. Most structural information about the group is available directly from the infinitesimals. Indeed, Reid [57, 58] shows how to extract structural information directly from the determining equations *without* knowing their solution. It cannot be overemphasized how important the 'infinitesimalizing' of symmetry calculations is: as Olver [47, p.43] observes, "... almost the entire range of applications of Lie groups to differential equations ultimately rests on this one construction".

When one is dealing with a *class* $\mathcal{C}$ of differential equations, the symmetry group admitted by the equations in the class will in general vary from equation to equation. The symmetry group *classification problem* for the class $\mathcal{C}$ is to rationally classify the equations in $\mathcal{C}$ into a hierarchy of cases according to the size and structure of their symmetry groups. This problem is considerably more difficult than finding the symmetries of a single differential equation. In fact we shall devote a great deal of our effort towards solving this problem.

Following the work of Ovsiannikov in the USSR in the late 1950's and 1960's [50] and of Bluman in the West in the late 1960's and 1970's [7, 9], there has been a major revival of interest in symmetry methods for differential equations. With the publication of the texts of Ovsiannikov [52], Olver [47], and Bluman and Kumei [13], there are now several comprehensive accounts of the basic theory, as well as more recent applications and generalizations. The central results of Lie's theory are outlined in Chapter 2; they allow the equivalence methods which follow to appear as a natural outgrowth, and in turn will provide a fruitful application of equivalence ideas.

1.4 Equivalence transformations

Just as symmetries of a differential equation transform solutions of the d.e. to other solutions of the same d.e., point equivalence transformations transform differential equations in some specified class $\mathcal{C}$ to other d.e.'s in the same class. Referring to the Examples in §1.1, it is apparent that several kinds of transformations map equations to equations in $\mathcal{C}$.

There may exist point transformations having the property that they map every equation in $\mathcal{C}$ to another equation in $\mathcal{C}$. In Examples 1.2.1–1.2.4, transformations (1.2, 1.7, 1.10, 1.13) respectively are of this kind. It is clearly of great interest to determine these transformations. They will certainly include those basic 'physical' transformations relating to choice of units etc. in the original equation, but as in Example 1.2.3, may also include less trivial transformations.

Ovsiannikov [52, §6] defined a suitable methodology and notation for dealing with such transformations, for which he used the term *equivalence transformations*. He derived some basic results about them, including the all-important property that they form a group. The defining properties of Ovsiannikov's equivalence transformations are (cf. Properties (i)–(iii), Example 1.2.2)

(i) The transformations act on *every* equation in the class $\mathcal{C}$. That is, they map every equation in $\mathcal{C}$ to another equation in $\mathcal{C}$.

(ii) The transformations are *fixed* point transformations, in the sense that they do not depend on the arbitrary elements, and are realized on the point space (independent and dependent variables) associated with the differential equations. In contrast, transformations (1.17) for the potential nonlinear telegraph equation (1.16) explicitly depend on the arbitrary functions $b(u)$, $c(u)$ occurring there.

(iii) The transformations act on the arbitrary elements as point transformations of an *augmented space* of independent and dependent variables and additional variables representing values taken by the arbitrary elements.

The collection of all such transformations constitute Ovsiannikov's *equivalence group*, which we denote by $\mathcal{Q}$. The action on the augmented space will be denoted by $\hat{\mathcal{Q}}$.

Although he intimates that determination of the equivalence group is possible using the Lie symmetry method, an explicit algorithm is not presented. The only example given is the equivalence group (1.7) for the scalar nonlinear diffusion equation (1.6), which unfortunately is available by inspection. Akhatov, Gazizov and Ibragimov [3] used Ovsiannikov's methodology in determining the infinitesimal form of transformations (1.10) for a potential form $v_t = D(v_x)v_{xx}$ of the nonlinear diffusion equation (compare 1.9). Subsequently, in the course of a heuristic investigation [4] of nonlocal symmetries, they further applied Ovsiannikov's ideas to several examples, giving sufficient detail for a general method to be discerned. They used the equivalence group to give a preliminary symmetry group classification for several examples, a technique which we describe below in §1.5, and more fully in §4.2. Ibragimov, Torrisi and Valenti [32] found the equivalence group $\mathcal{Q}$ for a large class of nonlinear hyperbolic equations and executed the preliminary classification for a finite-parameter subgroup of $\mathcal{Q}$. These are apparently the only significant uses of Ovsiannikov's equivalence ideas which have been made to date. It does not seem that a detailed theoretical exposition of the equivalence group is available, so a first goal in this dissertation will be to systematically develop a *theory* of equivalence transformations, and to show how to *algorithmically construct* them.

First, in §3.1 and §3.2, we develop the theory of equivalence transformations, filling in and extending the skeleton of theory provided by Ovsiannikov. We attempt to follow a course as closely parallel to Lie symmetry theory as possible. Calculation of equivalence transformations for a given class $\mathcal{C}$ of equations will be the subject of §3.3. We show how the problem can be formulated infinitesimally, and how this leads to a system of linear homogeneous determining equations for the infinitesimal equivalences. Calculating the equivalence group is often straightforward, because the method typically yields a large number of simple determining equations. We give examples of such equivalence calculations in §3.4, using nonlinear diffusion convection and linear wave equations as instances having nontrivial equivalence groups.

In §3.4 we make tangential reference to direct use of equivalence groups: mapping solutions of a 'simple' equation to solutions of related 'complicated' equations; or conversely, to simplify a complicated equation by mapping it to a simple (e.g. linear) equation. In Appendix C, using results derived in §3.4.1, we map some similarity solutions of the nonlinear diffusion equation with power law diffusivity $D(u) = u^m$ to solutions for the diffusivity $D(u) = u^m(1-u)^{-(m+2)}$. In §3.4.2, we clarify the process of linearizing some nonlinear diffusion convection equations. Finally, in §3.4.3, we give some nontrivial relationships between some 'potential symmetries' of linear wave equations.

However, our principal use of the equivalence group will be in classifying symmetries of a class of differential equations.

1.5 Symmetry classification problem

There are two broad approaches to classification of symmetries for a class of differential equations, which we characterize as *synthetic* and *analytic* methods. In *symmetry group analysis* of differential equations, one forms and analyzes determining equations for the infinitesimal symmetry transformations. A method due to Reid [57] for systematic group analysis is discussed below.

In contrast, synthetic methods bypass the construction of determining equations. Here we initially require a class of d.e.'s and a *given* group $\mathcal{G}$ of transformations mapping each d.e. in the class to another d.e. in the class. For example, given the class of second order o.d.e.'s (1.1), there is naturally available the group (1.2) of point transformations, which acts on the equations. With the algorithm of §3.3.3, we may provide such a group $\mathcal{G}$ for any class of d.e.'s, namely the equivalence group $\mathcal{Q}$. With $\mathcal{G}$ available one uses various algebraic and geometric processes to *construct* the d.e.'s admitting subgroups of $\mathcal{G}$ as symmetries. With this approach one can only extract those symmetries which are contained in the given group $\mathcal{G}$. For second order o.d.e.'s (1.1) this is no hindrance, since $\mathcal{G}$ (1.2) is the group of *all* point transformations. However, for a finite-parameter equivalence group such as (1.7) for nonlinear diffusion, it cannot be known how many symmetries lie outside $\mathcal{Q}$.

In §4.2 we describe a synthetic method which is appropriate for finite-parameter equivalence groups. The method is modelled on the classification of invariant solutions of d.e.'s, a theory which in turn relies on the classification of subgroups of the equivalence group using the adjoint group. This has the advantage of using very well known theory to derive the (necessarily partial) symmetry classification: recently, Ibragimov and others [4, 32] described this process, calling it the 'preliminary classification' method. Their examples include some quasilinear hyperbolic equations, potential forms of nonlinear diffusion equations, and fluid flow equations. We exemplify the partial classification method using nonlinear diffusion convection equations. The method provides 'quick and dirty' symmetry information: quick, because one often avoids dealing with infinite groups; dirty, because the information is not complete. For our example the equivalence group is small (a 10-parameter group), but it contains a surprising amount of symmetry information.

The premier synthetic classification methods are due to Tresse and Cartan (cf. §1.1). Here, given a class of d.e.'s and a group $\mathcal{G}$ acting on this class, one attempts to write the d.e.'s in a form which is *invariant* under the action of $\mathcal{G}$. Actually, their methods are geared to giving criteria for equivalence of two differential equations with respect to $\mathcal{G}$, with symmetry information appearing as a byproduct. Tresse described two variants of the process. In one, differential invariants of $\mathcal{G}$ are explicitly constructed, and the d.e.'s expressed in terms of them. In the other, one uses the action of $\mathcal{G}$ to reduce the class of d.e.'s to a small number of canonical forms. Cartan's reformulation of the equivalence method takes the second approach. Sophisticated geometric and algebraic machinery [26] uses the action of $\mathcal{G}$ to reduce a coframe to an invariant form where equivalence criteria and symmetry structure may be read off easily. Cartan [18] noted

> The general solution of this problem has already been given by the works of S. Lie and all those that they inspired. It is therefore only the *form* of the solution given here that is new.

(emphasis in original). On the other hand Cartan claimed as a genuinely new result his ability to extract structural properties of symmetry groups from his invariant coframes. Cartan's results take a new form in that they are expressed in the geometric language of differential forms. The extensive geometric machinery used in the Cartan equivalence method is required when finding equivalence criteria with respect to infinite groups $\mathcal{G}$, whose structure theory was described by Cartan [18] in terms of differential forms. However finite-parameter groups $\mathcal{G}$ have a structure theory which can be adequately described without forms, and less profound mathematical methods suffice. In a broad sense the methodology of the partial classification described in §4.2 is consistent with Cartan's method: the adjoint action of the group $\mathcal{G}$ is used to remove parameters and reduce the d.e.'s admitting symmetries from $\mathcal{G}$ to a finite number of canonical forms.

The advantage of synthetic methods is that they can use powerful geometric methods to uncover the symmetries which lie within the given group $\mathcal{G}$. The obvious deficiency is that one is tied completely to the group $\mathcal{G}$. It is not difficult to give a classification of the diffusion equations which admit some subgroup of (1.7) as symmetries, but what is really desired is a classification of *all* the point symmetries of the equations. There are two ways to approach this. Firstly, one can embed the given class of equations in a 'bigger' class, on which a suitably large group acts. For example, instead of analyzing diffusion equations, one might attempt to analyze all second order quasilinear p.d.e's. If the class is sufficiently enlarged, one can ensure that all the desired symmetry information is contained in the associated equivalence group, and then apply a synthetic procedure such as Cartan's. This has the obvious drawback of quickly leading to impracticably large classification problems. Indeed very few partial differential equation classifications have *ever* been found by Cartan's method.

One can instead attempt an *analytic* approach to symmetry classification. Here one accepts the given class of d.e.'s and attempts to sort it into subclasses on the basis of symmetry properties. The leading—and perhaps only—such method is that of Reid [57]. His approach is directly based on the Lie infinitesimal method for symmetries. First he calculates determining equations for the infinitesimal symmetries of equations in the class. By systematically appending compatibility conditions to a determining system, eventually a standard 'involutive' form is obtained, wherein the size and structure of the symmetry algebra can be read off by a simple process. When one does the same thing for a class of d.e.'s, Reid's method inevitably provides case splittings between the equations possessing symmetry groups of differing size and structure. Because his method involves only differentiation and algebraic processes, it is feasible to execute on a computer, and quite difficult classification problems can be solved by his procedure.

It is not clear how these two widely differing approaches can be combined. The Tresse and Cartan methods are geometric, while Reid's is analytic. However, in §4.5 we show how the two methods can be combined. This requires geometric machinery developed by Tresse (and much elaborated by Cartan) for dealing with *moving frames*, that is, non-commuting bases of differential operators. An outline of the necessary concepts is given in §4.3.1, where we show how to refer determining equations to a moving frame. Next, in §4.3.2, we develop a variant of Reid's method [55, 56] for reducing a system of d.e.'s to involutive form. Reid's original formulation is referred to a fixed *coordinate* basis, where the differential operators commute. We define a corresponding algorithm for reducing a frame system to a 'frame involutive' form.

The key idea, developed in §4.4, is to refer the determining equations to a moving frame which is *invariant* under the action of the equivalence group. Construction of such invariant frames was described by Tresse [68], and is at the heart of Cartan's method [25]. Referring the frame Reid algorithm to the invariant frame from Tresse's or Cartan's methods allows us to find a symmetry classification which is invariant under the action of the equivalence group. This process is described in §4.5. Our method may thus be regarded as either: a way to incorporate equivalence group information into Reid's method; or as a way to incorporate partial classification information (with respect to a 'small' group $\mathcal{Q}$) into a broader classification (with respect to the group of all point transformations).

Our new method fully utilizes equivalence group information in the construction of a complete point symmetry classification, thereby combining the best features of Cartan's and Reid's methods. There are some theoretical gaps in our treatment: in particular, a frame version of the Riquier integrability theorem remains to be proved. However, the method shows great promise. We believe it provides a powerful framework for dealing with p.d.e. symmetry classifications, including those which are computationally infeasible to either Reid or Cartan methods. In §4.5.2, we apply our method to a symmetry classification of nonlinear diffusion convection systems. For this example the order and simplicity in the classification which results is indeed remarkable.

Finally, in Chapter 5, we indicate some of the many directions in which the methods described here can be developed.

Chapter 2

Transformation Groups and Differential Equations

Before developing the theory of equivalence transformations, we first give the necessary background for dealing with (i) transformations and transformation groups (ii) differential equations (d.e.'s).

2.1 Transformation groups

We now establish the basic definitions and results on transformations and transformation groups, expecially as relating to differential equations. The results are standard, and no proof or motivation is offered. In the books of Bluman and Kumei [13], Olver [47] and Ovsiannikov [52] this material is developed in detail, and we refer to these sources for proofs of the theorems and illustrative examples.

2.1.1 Transformations, Lie groups

Spaces

We shall have use for various spaces representing independent variables, dependent variables, derivatives and so on. Without exception these are n-dimensional real spaces $\mathbb{R}^n$, or an open neighbourhood thereof. Rather than calling the spaces U, $\mathbb{R}^n$ etc., we shall mostly refer to spaces by their coordinates. Rather than "let $f\colon X \to U$ be a function", we shall say "let $u^i = f^i(x^1, x^2, \dots, x^n)$, $i = 1, 2, \dots, m$ be functions". The gain in readability should compensate for any loss of precision. We routinely make abbreviations such as $u = f(x)$ in preference to expressions using indices. In order to express calculational formulas, a debauch of indices is nevertheless necessary. We follow usual conventions for such indices: superscripts x^i, u^j to count coordinates of an ordinary space; subscripts for derivatives and similar objects. Thus u_i^j represents the i-th partial derivative of a component u^j. This conforms to geometric practice of placing covariant indices of a tensor as subscripts, and contravariant indices as superscripts; however not all of our indexed objects are tensors. We rigorously adhere to the summation convention: a repeated index occurring as a subscript and a superscript is to be summed over its range of values. Thus $u_i^j \dfrac{\partial f}{\partial u^j}$ is properly $\sum_{j=i}^{m} u_i^j \dfrac{\partial f}{\partial u^j}$. We often use Kronecker delta notation δ^i_j (1 for $i = j$, 0 otherwise). Changes of coordinates are denoted by $x' = f(x)$. For derivatives of a function K we use dot notation $\dot{K}$, $\ddot{K}$.

Transformations

Definition 2.1.1. A *transformation* of a space $x = (x^1, x^2, \ldots, x^n)$ is a smooth (C^∞) mapping $x' = \tau(x)$ such that $\tau = (\tau^1, \tau^2, \ldots, \tau^n)$ is one-one and onto.

The *inverse* transformation τ^{-1} of τ therefore always exists.

This definition is actually more stringent than required. If a mapping τ is defined only on some open subset of x-space we still use the term 'transformation of x'. Thus we regard the map $\tau\colon x \mapsto 1/x$ as a transformation 'of x', even though it is undefined at the point $x = 0$. That is, all our statements are *local* in nature. The local theory is congested with statements about 'neighbourhoods U of the point x_0', which can obscure the main thrust of the theory. We mostly omit reference to such neighbourhoods. Thus it must always be borne in mind that our results are not true as stated for 'a space x', but hold only on suitably small neighbourhoods of x. We occasionally underscore this point, but mostly let it pass without comment.

Lie groups

Our interest is in groups continuously parametrized by r real parameters:

Definition 2.1.2. Let r real parameters $\varepsilon = (\varepsilon^1, \varepsilon^2, \ldots, \varepsilon^r)$ lie in a space P. The space P is an *r-parameter Lie group* if there is defined a binary operation $*$ on P such that

- There is a unique *identity element* $e \in P$ such that $\varepsilon * e = e * \varepsilon = \varepsilon$ for all $\varepsilon \in P$.
- The operation $*$ is *associative*: $\varepsilon * (\delta * \gamma) = (\varepsilon * \delta) * \gamma$ for all $\varepsilon, \delta, \gamma \in P$.
- For every $\varepsilon \in P$, there exists an *inverse* element $\varepsilon^{-1} \in P$ such that $\varepsilon * \varepsilon^{-1} = \varepsilon^{-1} * \varepsilon = e$.
- Both the binary operation $*$ and the map $\varepsilon \mapsto \varepsilon^{-1}$ are *analytic*.

The identity element e can be taken as the origin 0, but there is no special necessity to do so.

Again this definition is more stringent than required. In general the parameters ε are local coordinates of an r-dimensional manifold. However, we only use the group elements near the identity, and locally we may treat the parameter space as $\mathbb{R}^r$. A local theory of such 'groups' is available [52, §12], [47, p19], wherein the binary operation and inverses are defined only in neighbourhoods of the identity element e. We understand that any reference to a 'Lie group' actually means 'elements sufficiently near the identity'. Any results based on global properties (connectedness, compactness, etc.) of the group are outside our domain of inquiry.

For example, the archetypal Lie group is the real numbers under addition: $P \equiv \mathbb{R}$, with the operation '$*$' being ordinary addition $+$. For adding angles of rotation the relevant group is the circle (addition modulo 2π). Although these two groups have differing global topologies (e.g., one is simply connected, the other is not), in the neighbourhood of the identity 0 they both represent simple addition, and from our viewpoint are identical.

Lie transformation group

The group P of parameters remains in the background; our interest is in *transformation groups*, i.e., collections of transformations *labelled* by the parameters ε of P.

Definition 2.1.3. A *Lie transformation group* on a space $x = (x^1, x^2, \dots, x^n)$ is a collection $\mathcal{G}$ of smooth transformations τ of x obtained as the homomorphic image of a Lie group of parameters. There is a map $\tau\colon P \to \mathcal{G}$ such that

- $\tau(e)$ is the identity map of x: $\tau(e)(x) = x$ for all x.
- $\tau(\varepsilon) \circ \tau(\delta) = \tau(\varepsilon * \delta)$ for all $\varepsilon, \delta \in P$.
- $\tau(\varepsilon^{-1}) = \tau(\varepsilon)^{-1}$
- The map $x' = F(x; \varepsilon) = \tau(\varepsilon)(x)$ is smooth (C^∞) in x and ε.

(see also [13, §2.1.3], [47, p.21], and [52, §16.1]). The binary operation on a transformation group is always composition $\circ$. Because transformation groups are our primary interest, the unqualified term 'Lie group' will always be taken to mean a Lie *transformation* group. If the underlying Lie group of parameters is used, we explicitly say so.

Example 2.1.4. The collection of transformations of (x, y)

$$\begin{aligned} x' &= \frac{x}{1 - \varepsilon x} \\ y' &= (1 - \varepsilon x)^2 y \end{aligned} \tag{2.1}$$

is a transformation group. Fixing ε specifies a (local) transformation $\tau(\varepsilon)$ of (x, y) space. Composing two such transformations $\tau(\varepsilon)$, $\tau(\delta)$ yields the transformation $\tau(\varepsilon + \delta)$, so the parameter ε lives on the additive group on $\mathbb{R}$. Note that none of the transformations is globally defined. Similarly, the image of a point (x, y) is not defined for every transformation in the group. However, the image of any point (x, y) is defined for every transformation sufficiently close to the identity $\varepsilon = 0$. This is sufficient for our purposes.

A transformation group $\mathcal{G}$ is specified by a map $x' = F(x; \varepsilon)$ with the properties

- for fixed ε, the map $\tau(\varepsilon)$ defined by $\tau(\varepsilon)(x) = F(x; \varepsilon)$ is a transformation of x.
- $F(x; e) = x$ for all x.
- $F(F(x; \varepsilon); \delta) = F(x; \delta * \varepsilon)$.
- F analytic in ε and C^∞ in x.

We also assume that if $F(x; \varepsilon) = x$ for all x, then $\varepsilon = e$, so that there are no 'unnecessary' parameters.

One-parameter transformation group

As a special case of a transformation group, let the single real parameter ε be additive.

Definition 2.1.5. A *one-parameter (ε) group* acting on a space x is a transformation group on x with the following properties:

- $\tau(0)$ is the identity transformation on x
- $\tau(\varepsilon) \circ \tau(\delta) = \tau(\varepsilon + \delta)$

See also [13, §2.1.4], [47, p28], [52, §1]. For example, (2.1) is a one-parameter (ε) Lie transformation group of (x, y) space. It seems a restriction to demand that the real parameter ε be additive in $\mathbb{R}$, but any other local group operation on $\mathbb{R}$ can be reparametrized to be addition [13, §2.2.1].

2.1.2 Infinitesimal operators

The key to practical construction of Lie transformation groups is an infinitesimal formulation of the problem, which replaces nonlinear conditions for a group with *linear* conditions.

Infinitesimal transformation

Consider a one-parameter (ε) group of transformations (Definition 2.1.5) acting on a space $x = (x^1, x^2, \ldots, x^n)$. In a neighbourhood of the identity $\varepsilon = 0$, the transformation

$$x' = F(x; \varepsilon)$$

can be expanded as

$$x' = x + \varepsilon\xi(x) + O(\varepsilon^2) \tag{2.2}$$

where $\xi = (\xi^1, \xi^2, \ldots, \xi^n)$ is given by

$$\xi^i(x) = \frac{\partial F^i}{\partial \varepsilon}(x; \varepsilon)\Big|_{\varepsilon=0}. \tag{2.3}$$

The quantities ξ^i are called *infinitesimals* of the one-parameter group: expansion (2.2) represents an 'infinitesimal transformation' from the group.

Theorem 2.1.6 (First fundamental theorem of Lie). *The function F defining a one-parameter group of transformations (Definition 2.1.5) can be constructed from the infinitesimals ξ of the group as the solution $x' = F(x; \varepsilon)$ of the o.d.e. initial value problem*

$$\frac{dx'}{d\varepsilon} = \xi(x'), \qquad x'(0) = x. \tag{2.4}$$

For a proof see [13, §2.2.1], [47, §1.3], or [52, §2.3].

Example 2.1.7. Consider the one-parameter group (2.1) acting on (x, y) space. Differentiating with respect to ε (2.3) gives infinitesimals $(x^2, -2xy)$ corresponding to (x, y) respectively. The initial value problem (2.4) is here

$$\begin{aligned} \frac{dx'}{d\varepsilon} &= \quad x'^2 \qquad & x'(0) = x \\ \frac{dy'}{d\varepsilon} &= -2x'y' \qquad & y'(0) = y. \end{aligned}$$

Solving this indeed recovers the original one-parameter Lie group (2.1) (at least locally).

Thus the infinitesimals encode all information necessary to recover the action of a one-parameter group. Note that even when the infinitesimals ξ^i are smooth, existence and uniqueness results for o.d.e. initial value problems guarantee only *local* existence of a solution to the problem (2.4). Our local use of 'transformation' and 'transformation group' is thus natural for groups found by integration of (2.4).

Group operator

Definition 2.1.8. The *group operator* $\mathbf{X}$ of a one-parameter group with infinitesimals $\xi = (\xi^1, \xi^2, \dots, \xi^n)$ (2.2) is the first order differential operator

$$\mathbf{X} = \xi^i(x)\,\frac{\partial}{\partial x^i}\;. \tag{2.5}$$

For example, the group operator corresponding to (2.1) is $\mathbf{X} = x^2\,\partial_x - 2xy\,\partial_y$. Group operators will be denoted with boldface roman capitals $\mathbf{X}$, $\mathbf{Y}$. A group operator is a *vector field* on the space x: it attaches a vector $\xi(x)$ to each point x in the space. In §4.3 we use vector fields which do not naturally give rise to a transformation group, and we reserve this more geometric terminology for such circumstances. To save space we often write ∂_x instead of $\frac{\partial}{\partial x}$.

An operator $\mathbf{X}$ is a coordinate free object. It encodes information on the rate of change of a function f with respect to the group parameter ε as a point x is dragged along by the one-parameter group associated with $\mathbf{X}$:

$$\frac{d}{d\varepsilon} f(x'(\varepsilon))\Big|_{\varepsilon=0} = \mathbf{X}f(x).$$

Lie algebra of operators

An r-parameter Lie transformation group has associated r group operators $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_r$ which are linearly independent and form an r-dimensional vector space over $\mathbb{R}$. This vector space has the additional structure of being closed under commutation.

Definition 2.1.9. Let $\mathbf{X} = \xi^i\,\dfrac{\partial}{\partial x^i}$ and $\mathbf{Y} = \gamma^i\,\dfrac{\partial}{\partial x^i}$ be two group operators. Their *commutator* $[\,\mathbf{X}, \mathbf{Y}\,]$ is the first order operator

$$\mathbf{XY} - \mathbf{YX} = \Big(\xi^j\,\frac{\partial \gamma^i}{\partial x^j} - \gamma^j\,\frac{\partial \xi^i}{\partial x^j}\Big)\,\frac{\partial}{\partial x^i}\;. \tag{2.6}$$

The commutator bracket $[\ ,\ ]$ has the properties

- Bilinearity. $[\,\mathbf{X}, a\mathbf{Y} + b\mathbf{Z}\,] = a[\,\mathbf{X}, \mathbf{Y}\,] + b[\,\mathbf{X}, \mathbf{Z}\,]$ where a, b are real constants.

- Anticommutativity. $[\mathbf{X}, \mathbf{Y}] = -[\mathbf{Y}, \mathbf{X}]$
- Jacobi identity.

$$[\mathbf{X}, [\mathbf{Y}, \mathbf{Z}]] + [\mathbf{Y}, [\mathbf{Z}, \mathbf{X}]] + [\mathbf{Z}, [\mathbf{X}, \mathbf{Y}]] = 0 \tag{2.7}$$

Any vector space satisfying these three properties is called a *Lie algebra*, but our Lie algebras are always Lie algebras of operators, with commutator bracket defined by (2.6).

Correspondence between Lie group and Lie algebra

A Lie algebra of operators contains all the information necessary to reconstruct a Lie group.

Theorem 2.1.10. *To every r-parameter Lie transformation group $\mathcal{G}$ there corresponds an r-dimensional Lie algebra of operators L. An r-dimensional vector space L of operators derives from a Lie transformation group if and only if L is closed under commutation:*

$$[\mathbf{X}, \mathbf{Y}] \in L \qquad \text{for all } \mathbf{X}, \mathbf{Y} \in L\ .$$

Usually a finite-dimensional Lie algebra is resolved with respect to a basis $\mathbf{X}_i$, in which case this closure condition becomes

$$[\mathbf{X}_i, \mathbf{X}_j] = C_{ij}^k \mathbf{X}_k \tag{2.8}$$

for some constants C_{ij}^k which are called the *structure constants* of L. Antisymmetry of the commutator bracket shows $C_{ij}^k = -C_{ji}^k$, and the Jacobi identity (2.7) gives further relations. Practical construction of a Lie group proceeds from its Lie algebra of operators as follows.

Theorem 2.1.11. *Let $\mathcal{G}$ be an r-parameter Lie transformation group. Let $\mathbf{X}_1$, $\mathbf{X}_2$, ... , $\mathbf{X}_r$ be r linearly independent group operators corresponding to r independent one-parameter subgroups G_1, G_2, ... ,G_r of $\mathcal{G}$. Let $\tau_1(\varepsilon^1)$, $\tau_2(\varepsilon^2)$, ... , $\tau_r(\varepsilon^r)$ be transformations from the groups G_1, G_2, ... ,G_r respectively. Then every transformation $\tau(\varepsilon) \in \mathcal{G}$ sufficiently close to the identity can be realized by composing these:*

$$\tau(\varepsilon) = \tau_1(\varepsilon^1) \circ \tau_2(\varepsilon^2) \circ \ldots \circ \tau_r(\varepsilon^r). \tag{2.9}$$

See [52, §16.7].

The motivation for examining Lie algebras of operators is that they encode in linear form the local structure of a Lie group of transformations. That is, the structure constants C_{ij}^k determine a local Lie group up to isomorphism. To every structural feature of a local Lie group there is a corresponding structural feature of the Lie algebra, which can be discerned from its commutation relations. Details of this structural correspondence are given by Ovsiannikov [52, §15]: we state some of the more imporant ones here.

Normal subgroup, ideal

A subgroup $\mathcal{H}$ of a group $\mathcal{G}$ is *normal* if $\tau \circ \sigma \circ \tau^{-1} \in \mathcal{H}$ for all $\sigma \in \mathcal{H}$ and all $\tau \in \mathcal{G}$. A subalgebra I of L is an *ideal* if $[\mathbf{X}, \mathbf{Y}] \in I$ for all $\mathbf{X} \in I$ and all $\mathbf{Y} \in L$. The Lie algebra I of a normal subgroup $\mathcal{H}$ of $\mathcal{G}$ is an ideal in the Lie algebra L of $\mathcal{G}$. If $\{\mathbf{X}_i\}$ span an ideal I, and $\{\mathbf{X}_i, \mathbf{Y}_j\}$ span the whole Lie algebra L, the commutator table has the following form:

[,]	$\mathbf{X}_i$	$\mathbf{Y}_j$
$\mathbf{X}_i$	$\{\mathbf{X}_i\}$	$\{\mathbf{X}_i\}$
$\mathbf{Y}_j$	$\{\mathbf{X}_i\}$	$\{\mathbf{X}_i, \mathbf{Y}_j\}$

(2.10)

Direct and semidirect sum

A Lie algebra L is the *semidirect sum* of two subalgebras I, J, denoted by $L = I \oplus_s J$ if L is a vector space direct sum of I, J, and I is an ideal in L: $[I, J] \subseteq I$. If $\mathbf{X}_i$ are a basis for I and $\mathbf{Y}_j$ a basis for J, the commutator table of L takes the following form:

[,]	$\mathbf{X}_i$	$\mathbf{Y}_j$
$\mathbf{X}_i$	$\{\mathbf{X}_i\}$	$\{\mathbf{X}_i\}$
$\mathbf{Y}_j$	$\{\mathbf{X}_i\}$	$\{\mathbf{Y}_j\}$

This condition is stronger than requiring that I be an ideal in L, since we also require its vector space complement to be a subalgebra. Note that the semidirect sum operation is not commutative: $L \neq J \oplus_s I$. A semidirect sum becomes a *direct sum* of ideals when the off-diagonal blocks in the commutator table vanish. This is stronger than a semidirect sum: the two ideals I, J do not 'interact' at all.

Quotient group, quotient algebra

Let $\mathcal{H} \subseteq \mathcal{G}$ be normal. Define the equivalence relation $\sim$ on $\mathcal{G}$ by $\tau_1 \sim \tau_2$ iff $\tau_1 \circ \tau_2^{-1} \in \mathcal{H}$. Denote the equivalence class containing τ by $\bar{\tau}$. The collection of such equivalence classes is a group $\mathcal{G}/\mathcal{H}$ under an operation (which we also denote by $\circ$) induced by $\bar{\tau}_1 \circ \bar{\tau}_2 = \overline{\tau_1 \circ \tau_2}$. The group $\mathcal{G}/\mathcal{H}$ is called the *quotient group* (or factor group) of $\mathcal{G}$ over $\mathcal{H}$.

Let $I \subseteq L$ be an ideal. Define the equivalence relation $\sim$ on L by $\mathbf{X}_1 \sim \mathbf{X}_2$ iff $\mathbf{X}_1 - \mathbf{X}_2 \in I$. Denote the equivalence class containing $\mathbf{X}$ by $\bar{\mathbf{X}}$. The collection of such equivalence classes is a Lie algebra $\underline{L/I}$ under an operation (which we also denote by [,]) induced by $[\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2] = \overline{[\mathbf{X}_1, \mathbf{X}_2]}$. The algebra L/I is called the *quotient* or *factor algebra* of L over I.

Let $\mathcal{G}/\mathcal{H}$ be the quotient group of the Lie group $\mathcal{G}$ over the normal subgroup $\mathcal{H}$. Then the Lie algebra associated with this quotient group is the factor algebra L/I. The commutation relations of the factor algebra are obtained from those of L by taking the bottom right hand corner of the table (2.10) and dropping the X_i components. If L is a semidirect sum $I \oplus_s J$, the factor algebra L/I is isomorphic to J.

Isomorphism, homomorphism

A mapping ψ from a Lie group $\mathcal{G}$ onto the Lie group $\mathcal{M}$ is a *group homomorphism* if

$$\psi(\tau \circ \sigma) = \psi(\tau) \circ \psi(\sigma)$$

for all $\tau, \sigma \in \mathcal{G}$.
The *kernel* $\ker \psi$ of a homomorphism ψ is the set of transformations κ which are mapped to the identity e by ψ:

$$\ker \psi = \{\kappa \in \mathcal{G} \mid \psi(\kappa) = e\}.$$

It is a normal subgroup of $\mathcal{G}$. A homomorphism ψ is an *isomorphism* if it is one-to-one onto $\mathcal{M}$ (or equivalently, if $\ker \psi = \{e\}$). The image of a homomorphism ψ is isomorphic to the quotient group $\mathcal{G}/\ker \psi$.

A linear mapping A from a Lie algebra L onto the Lie algebra M is a *Lie algebra homomorphism* if

$$A([\mathbf{X}, \mathbf{Y}]) = [A(\mathbf{X}), A(\mathbf{Y})]$$

for all $\mathbf{X}, \mathbf{Y} \in L$. The *kernel* $\ker A$ of a homomorphism A is the set of operators which are mapped to zero by A:

$$\ker A = \{\mathbf{K} \in L \mid A(\mathbf{K}) = 0\}.$$

It is an ideal in L. The algebra homomorphism A is an *isomorphism* if it is one-to-one and onto M (equivalently, if $\ker A = \{0\}$). The image of a homomorphism A is isomorphic to the factor algebra $L/\ker A$. In particular, if $L = I \oplus_s J$, and $\ker A = I$, then the image of A is isomorphic to J.

Group and algebra homomorphisms and isomorphisms correspond, at least for local Lie groups and their associated Lie algebras.

2.1.3 Invariant surface

The construction of symmetries (or the equivalence transformations of Chapter 3) follows from an infinitesimal criterion for a transformation to leave invariant some surface. The criterion for the case involving derivatives relies on a corresponding result for algebraic equations, which we state first.

Definition 2.1.12. Let E be the set of points $x = (x^1, x^2, \ldots, x^n)$ satisfying the algebraic equations $f(x) = 0$, where $f = (f^1, f^2, \ldots, f^s)$ are s smooth functions. Let $\mathcal{G}$ be a Lie transformation group acting on the space x. The equation $f(x) = 0$ *admits* the group $\mathcal{G}$ (or is *invariant* under $\mathcal{G}$) if for every x satisfying $f(x) = 0$, and every transformation $\tau \in \mathcal{G}$, we have $f(\tau(x)) = 0$.

In short the group $\mathcal{G}$ transforms E to itself.

Theorem 2.1.13. *Assume the system $f(x) = 0$ is of maximal rank. That is, the Jacobian $\frac{\partial f^i}{\partial x^j}$ is of rank s at every point on $f(x) = 0$, where s is the number of equations. Then $f(x) = 0$ admits a Lie transformation group $\mathcal{G}$ if and only if*

$$\mathbf{X} f(x) = 0 \qquad \text{for all } x \text{ such that } f(x) = 0 \tag{2.11}$$

for every operator $\mathbf{X}$ of $\mathcal{G}$.

See [13, §2.2.7], [47, p83], or [52, §3.12] for a proof.

The rank condition on the Jacobian is essential in this theorem. If it is violated, either the system contains redundant equations which can be discarded, or the assignment of the surface $f(x) = 0$ is 'bad'. For instance the system $x^2 = 0$ does not satisfy the Jacobian condition. However the surface defined by this equation is clearly identical to that defined by $x = 0$, which does satisfy the rank condition. It is *always* possible to modify a system of equations so that it (locally) satisfies the Jacobian condition. Note that a system of equations written in solved form automatically satisfies the Jacobian condition, and hence we write equations in solved form whenever possible.

2.2 Extension

In the previous section we considered transformations of an arbitrary space x. When dealing with differential equations, we have spaces of independent and dependent variables, and properties of *derivatives* enter the picture. The process of taking an object defined on the base space of independent and dependent variables, and deriving the corresponding object on the space of derivatives is called *extension*. (The term 'prolongation' is also used.) For example, substituting $x' = x$, $y' = y/x$ in a scalar o.d.e. induces an action

$$\frac{dy'}{dx'} = \frac{1}{x}\frac{dy}{dx} - \frac{y}{x^2} \; .$$

on the first derivative dy/dx: this is the first extension of the transformation of (x, y) space.

We show in turn how to extend spaces, transformations, transformation groups, and group operators.

2.2.1 Notation for derivatives

We wish to develop results for differential equations in arbitrary numbers of independent and dependent variables. Let $x = (x^1, x^2, \ldots, x^n)$ be the n independent variables, and $u = (u^1, u^2, \ldots, u^m)$ be the m dependent variables. The transformations in which we are interested act on all such variables: the space (x, u) of independent and dependent variables will be called the *base space* of the d.e.

To deal with differential equations, we extend the base space by adjoining coordinates representing values taken by derivatives of u. Let $\theta(x)$ be a function $\theta = (\theta^1, \theta^2, \ldots, \theta^m)$ of the independent variables. The collection of all k-th order partial derivatives

$$\frac{\partial^k \theta^i(x)}{\partial x^{j_k} \ldots \partial x^{j_2} \partial x^{j_1}} \qquad \begin{array}{l} 1 \le i \le m \\ 1 \le j_1, j_2, \ldots, j_k \le n \end{array}$$

is denoted by $\underset{k}{\theta}(x)$. At each point x, the set of possible values of $\underset{k}{\theta}(x)$ forms a space of dimension $m\binom{n+k-1}{k}$. We assign coordinates

$$\underset{k}{u} = u^i_{j_1 j_2 \ldots j_k} \; , \qquad \begin{array}{c} 1 \le i \le m \\ 1 \le j_1, j_2, \ldots, j_k \le n \end{array}$$

to this space in the obvious way: these will collectively be denoted by $\underset{k}{u}$.

It is convenient to let $J = (j_1 j_2 \ldots j_k)$ denote a multi-index. The *order* of J is the number of elements in the multi-index (k in this case), and will be denoted by $|J|$. This allows convenient shorthand notations: the collection of k-th order derivatives of u can be concisely rendered as $\{u^j_J : |J| = k\}$. Concatenation of multi-indices is denoted in the obvious way, so that $Ji \equiv (j_1 j_2 \ldots j_k i)$. Equality of mixed partial derivatives implies that a multi-index is defined only up to permutation: if I is a rearrangement of the multi-index J, then $u_I \equiv u_J$.

We define the k-th *extension* space to be $(x, u, \underset{1}{u}, \ldots, \underset{k}{u})$, representing variables and derivatives up to order k.

2.2.2 Extension of transformation

A transformation

$$\begin{aligned} x' &= F(x, u) \\ u' &= G(x, u) \end{aligned} \tag{2.12}$$

of the base space induces an action on derivatives in a natural way. We now give the basis of this construction. We carefully distinguish when u is being treated as an independent *coordinate*, and when it is a function of x. Let u be assigned as $u = \theta(x)$, and correspondingly $\underset{1}{u} = \underset{1}{\theta}(x)$, ... , $\underset{k}{u} = \underset{k}{\theta}(x)$. A function $f(x, u, \underset{1}{u}, \ldots, \underset{k}{u})$ becomes

$$\theta^* f\,(x) = f(x, \theta(x), \underset{1}{\theta}(x), \ldots, \underset{k}{\theta}(x)), \tag{2.13}$$

which we call the *pullback* of f by θ. Although the name and notation are differential geometric, it need only be remembered that the notation $\theta^* f$ implies that we are treating u as a function of x in f.

Definition 2.2.1. The *graph* of a smooth function $\theta = (\theta^1, \theta^2, \ldots, \theta^m)$ is the set of points

$$\Gamma(\theta) = \{(x, u) \mid u = \theta(x)\}$$

The k-th extension of a graph $\Gamma(\theta)$ is the set of points

$$\Gamma(\underset{k}{\theta}) = \left\{(x, u, \underset{1}{u}, \ldots, \underset{k}{u}) \;\middle|\; (u, \underset{1}{u}, \ldots, \underset{k}{u}) = (\theta(x), \underset{1}{\theta}(x), \ldots, \underset{k}{\theta}(x)\right\} \tag{2.14}$$

A transformation τ on the base space (x, u) acts pointwise on a graph, mapping a point (x, u) on $\Gamma(\theta)$ to (x', u'). If the transformation τ is sufficiently close to the identity, these points (x', u') (at least locally) constitute the graph $\Gamma(\theta')$ of a function $u' = \theta'(x')$ [47, §2.2]. We say that τ transforms θ to θ'.

However, attempting the same argument on an extended graph $\Gamma(\underset{1}{\theta})$ by applying an arbitrary transformation of $(x, u, \underset{1}{u})$ will in general fail. Initially $\underset{1}{u} = \underset{1}{\theta}$ represents the slope of the plane tangent to the surface $u = \theta(x)$. After transformation, however, there is no guarantee that $\underset{1}{u'}$ agrees with the slope of the tangent plane to $u' = \theta'(x')$. The resulting locus is of no significance unless these derivatives 'match up'.

Definition 2.2.2. The k-th order *contact forms* $\underset{k}{C}$ on the space $(x, u, \underset{1}{u}, \ldots, \underset{k}{u})$ are the differential one forms

$$du^j_I - u^j_{Ii}\, dx^i, \qquad 0 \leq |I| \leq k-1 \tag{2.15}$$

The tangency conditions require that $\underset{k}{C}$ be preserved by a transformation $\underset{k}{\tau}$. They are expressed in terms of the following operators:

Definition 2.2.3. The (formally infinite) differential operator

$$D_{x^i} = \frac{\partial}{\partial x^i} + u^j_i \frac{\partial}{\partial u^j} + \cdots + u^j_{Ii} \frac{\partial}{\partial u^j_I} + \cdots . \tag{2.16}$$

is called the *total derivative* with respect to x^i.

Total derivative operators D_{x^i} are naturally dual to contact forms $\underset{k}{C}$ in the sense that they are annihilated by every such form. Although the sum defining D_{x^i} is formally infinite, we only apply total derivative operators to functions $f(x, u, \underset{1}{u}, \ldots, \underset{k}{u})$ defined on some finite order extension space, so only a finite number of terms is needed: the infinite sum is interpreted as "to whatever finite number of terms necessary". Functions $f(x, u, \underset{1}{u}, \ldots, \underset{k}{u})$ are sometimes called (k-th order) differential functions.

We wish to reserve the notation ∂_{x^i} for partial derivatives of a function $f(x, u, \underset{1}{u}, \ldots, \underset{k}{u})$, so that in $\partial_{x^i} f$, the coordinates $(u, u, \underset{1}{u}, \ldots, \underset{k}{u})$ are held constant. In contrast, D_{x^i} differentiates 'as though' u were a function of x. More precisely,

Proposition 2.2.4. *Let $f(x, u, \underset{1}{u}, \ldots, \underset{k}{u})$ be some function, and let $u = \theta(x)$. We have*

$$\theta^*(D_{x^i} f)(x) = \partial_{x^i}(\theta^* f)(x)$$

so that assigning $u = \theta(x)$ after *total* differentiation agrees with assigning $u = \theta(x)$ followed by *partial* differentiation. This justifies the name 'total derivative'. It is essential in the transformation theory to distinguish between ∂_{x^i} and D_{x^i}: the difference boils down to whether u is being treated as a coordinate (∂_{x^i}) or as a function of x (D_{x^i}). If we were *always* imagining u to be a function of x, the distinction would not be necessary, and in fact in §4.3 we allow the notation D_{x^i} to lapse and submit to the usual barbarism of confusing it with ∂_{x^i}. Until then, however, the distinction is carefully maintained.

Preserving contact conditions places strong restrictions on a transformation of the k-th extension space when $k \geq 1$.

Theorem 2.2.5. *Let $\underset{k}{\tau}$ be a transformation*

$$\begin{aligned} x'^i &= F^i(x, u, \underset{1}{u}, \ldots, \underset{k}{u}) \\ u'^j &= G^j(x, u, \underset{1}{u}, \ldots, \underset{k}{u}) \\ &\vdots \\ u'^j_I &= G^j_I(x, u, \underset{1}{u}, \ldots, \underset{k}{u}), \qquad 0 \leq |I| \leq k \end{aligned} \tag{2.17}$$

of k-th extension space. Define $A^i_j = D_{x^j}F^i$, along with the 'inverse matrix' B^j_l such that $B^j_l A^i_j = \delta^i_l$. (We may guarantee existence of this inverse by taking τ sufficiently close to the identity.) Then

$$D_{x'^i} = B^j_i D_{x^j}. \tag{2.18}$$

If $\underset{k}{\tau}$ preserves the k-th order contact conditions $\underset{k}{C}$ (Definition 2.2.2) then the functions $G^j_i, \dots, G^j_I$ for $1 \leq |I| \leq k$ are determined in terms of F, G by the recurrence (extension formula)

$$G^j_{Ii} = B^l_i D_{x^l} G^j_I, \qquad 0 \leq |I| \leq k-1. \tag{2.19}$$

See [13, §2.3.5] for further details and proof; also [47, Thm 2.36], [52, §4.5].

If F, G depend nontrivially on derivative components $(\underset{1}{u}, \dots, \underset{k}{u})$, the dependence cannot be arbitrary, since extension formula (2.19) apparently raises the order of derivatives each time it is applied. This results in strong restrictions on F, G:

Theorem 2.2.6 (Bäcklund).

(i) If the number of dependent variables $u = (u^1, u^2, \dots, u^m)$ is greater than one, the only transformations of $(x, u, \underset{1}{u}, \dots, \underset{k}{u})$ which preserve k-th order contact $\underset{k}{C}$, are extensions of transformations (2.12) of (x, u).

(ii) If the number of dependent variables u is one, the only transformations of $(x, u, \underset{1}{u}, \dots, \underset{k}{u})$ which preserve k-th order contact $\underset{k}{C}$ are extensions of transformations of $(x, u, \underset{1}{u})$.

Transformations obtained by extension of a base transformation (2.12) are called *extended point* transformations. Transformations obtained by extension of a transformation of $(x, u, \underset{1}{u})$ space are called *first order contact* transformations. We scarcely mention contact symmetries in this dissertation (see [13, §5.2.4] for details). Instead, from now on *we restrict ourselves without exception to extended point transformations.*

For later reference we mention *projectable* point transformations, of the form

$$\begin{aligned} x' &= F(x) \\ u' &= G(x, u). \end{aligned} \tag{2.20}$$

Sometimes the term 'fibre preserving' transformation is used.

Extension of a transformation *group* $\mathcal{G}$ on base space (x, u) to a group $\underset{k}{\mathcal{G}}$ on $(x, u, \underset{1}{u}, \dots, \underset{k}{u})$ is defined by extending each transformation in $\mathcal{G}$. The extended group $\underset{k}{\mathcal{G}}$ is isomorphic to $\mathcal{G}$.

Example 2.2.7. Consider the one-parameter group (2.1) of transformations. Suppose y is a dependent variable, and x the independent. Denoting the single first extension component by $\dot{y}$, so that the contact form is $dy - \dot{y}\,dx$, we compute from (2.2.5) the extended transformation

$$\dot{y}' = (1 - \varepsilon x)^3 \Big((1 - \varepsilon x)\dot{y} - 2\varepsilon y \Big)$$

2.2.3 Extension of group operator

The process (2.19) of extending a Lie transformation group $\mathcal{G}$ on the base space (x, u) to action on derivatives naturally induces an extension of the group operators (2.5) associated with $\mathcal{G}$. This is calculated by inserting the infinitesimal transformation (2.3) into extension formula (2.19).

Theorem 2.2.8. *Let*

$$\mathbf{X} = \xi^i(x,u)\,\frac{\partial}{\partial x^i} + \eta^j(x,u)\,\frac{\partial}{\partial u^j} \tag{2.21}$$

be an operator for a transformation group $\mathcal{G}$ acting on base space (x, u). Corresponding to the k-th extension group $\underset{k}{\mathcal{G}}$ is the operator

$$\begin{aligned}\underset{k}{\mathbf{X}} \;=\;& \xi^i(x,u)\,\frac{\partial}{\partial x^i} + \eta^j(x,u)\,\frac{\partial}{\partial u^j} + \eta^j_{(i)}(x,u,\underset{1}{u})\,\frac{\partial}{\partial u^j_i} + \cdots \\ &+ \eta^j_{(I)}(x,u,\underset{1}{u},\ldots,\underset{k}{u})\,\frac{\partial}{\partial u^j_I}, \qquad 0 \le |I| \le k\end{aligned} \tag{2.22}$$

where $\eta^j_{(I)}$, $1 \le |I| \le k$ are obtained from the recurrence

$$\eta^j_{(Ii)} = D_{x^i}\eta^j_{(I)} - u^j_{Iq}\,(D_{x^i}\xi^q) \tag{2.23}$$

where D_{x^i} is the total derivative operator (2.16).

See [13, §2.3.5], [47, p.108*ff*], [52, §4.8]. Our notation $\eta^j_{(I)}$ for extended infinitesimals is consistent with our placement downstairs of differentiation indices. The parentheses (I) are necessary to avoid confusion with partial derivatives η_I.

Example 2.2.9. Consider the group operator $\mathbf{X} = x^2\,\partial_x - 2xy\,\partial_y$ corresponding to the group (2.1). Extend using Theorem 2.2.8 to an operator $\underset{1}{\mathbf{X}} = \mathbf{X} + \eta_{(1)}\partial_{\dot{y}}$ on the space $(x, y, \dot{y})$. We find $\eta_{(1)} = -4x\dot{y} - 2y$. This agrees with the expression obtained by differentiation of the extended transformation noted in Example 2.2.7.

2.3 Differential equations and symmetry

2.3.1 Differential equations

From the outset, the term 'system of differential equations' will be taken to mean a general

system of	s	differential equations
of order	k	
in	n	independent variables
and	m	dependent variables.

Definition 2.3.1. A system E of s k-th order differential equations is defined by a function $f = (f^1, f^2, \ldots, f^s)$ on the k-th extension space $(x, u, \underset{1}{u}, \ldots, \underset{k}{u})$ as

$$f(x,u,\underset{1}{u},\ldots,\underset{k}{u}) = 0. \tag{2.24}$$

This is a system of algebraic equations with certain of the coordinates interpreted as coordinates of derivative spaces. The equation $f = 0$ (2.24) specifies a 'surface' E embedded in the space $(x, u, \underset{1}{u}, \dots, \underset{k}{u})$. Identifying a differential equation with this surface gives the theory a 'geometric' character. We do not make a distinction between this surface and the equations defining it, even though the same differential equations can be written in different forms (i.e. with various f's).

Definition 2.3.2. A (local) *solution* of equations E (2.24) through a point x_0 is a function $u = \theta(x)$ such that

$$f\Big(x, \theta(x), \underset{1}{\theta}(x), \dots, \underset{k}{\theta}(x)\Big) = 0 \tag{2.25}$$

for all x in some neighbourhood U of x_0.

In terms of the pullback (2.13), θ is a solution of $f = 0$ if $\theta^* f(x)$ vanishes identically for all $x \in U$. Alternatively, (2.25) states that the graph of θ lies in the surface E: $f = 0$ (2.24).

Example 2.3.3. To illustrate this notation, consider the scalar wave equation

$$u_{tt} = c^2(x) u_{xx}. \tag{2.26}$$

Here there are $n = 2$ independent variables (x, t); $m = 1$ dependent variable u; $s = 1$ equation; of order $k = 2$. The spaces involved are

$$\begin{aligned} x &= (x, t) \\ u &= (u) \\ \underset{1}{u} &= (u_x, u_t) \\ \underset{2}{u} &= (u_{xx}, u_{xt}, u_{tt}). \end{aligned}$$

The base space is (x, t, u); the twice-extended space is $(x, t, u, u_x, u_t, u_{xx}, u_{xt}, u_{tt})$. A *solution* of the wave equation (2.26) is a function $u = \theta(x, t)$ satisfying condition (2.25):

$$\frac{\partial^2 \theta}{\partial t^2}(x, t) = c^2(x) \frac{\partial^2 \theta}{\partial x^2}(x, t).$$

We use the abbreviated notation $x, u, \underset{1}{u}$ only in general theoretical statements: it is not particularly useful in concrete examples. Index notation $(x^1, x^2, \dots, x^n)$ for components of spaces will also be reserved for general theory: in examples it is preferable to give variables distinct names (x, t, etc.) reflecting their physical meaning.

2.3.2 Symmetries of differential equations

We now state the main results concerning symmetries of differential equations. Several results in Chapter 3 are established by parallel methods, so proofs of some of the theorems are given.

Definition 2.3.4. A transformation acting on the base space (x, u) of a system E of differential equations is a *point symmetry* of E if it maps every solution θ of E to another solution θ' of E.

('Mapping θ' means mapping the graph of θ.) As usual, the definition of symmetry is more stringent than really intended. Interpreted strictly, Definition 2.3.4 would disqualify a transformation from being a symmetry if there existed even one solution which was not mapped to another solution. The proper local statement [47, p.96] is to require that each solution $u = \theta(x)$ in a neighbourhood of a point x_0 be mapped to another local solution in a neighbourhood of x_0' by every symmetry transformation sufficiently close to the identity. This is sufficient for our purposes.

By Definition 2.3.4, symmetries act on *functions* θ representing solutions of the differential equations (2.24). The criterion for whether a point transformation is a symmetry would seem to demand we know all solutions of the differential equations. This is of no practical use: the 'function' criterion of Definition 2.3.4 must be replaced with a 'point-by-point' criterion.

Theorem 2.3.5. *Let* E *be a system of differential equations given by (2.24). Let* τ *be a transformation of the base space, whose extension* $\tau, \underset{1}{\tau}, \dots, \underset{k}{\tau}$ *leaves the surface* E *invariant. Then* τ *is a symmetry of* E.

Proof. Every solution $u = \theta(x)$ of E has its graph $\Gamma(\underset{k}{\theta})$ lying in the surface E. The extension $\underset{k}{\tau}$ of τ maps (extended) graphs to graphs, so there is a function $u' = \theta'(x')$ such that the $\underset{k}{\tau}$ maps $\Gamma(\underset{k}{\theta})$ to the graph $\Gamma(\underset{k}{\theta'})$ of θ'. But τ maps E to itself, so every point on the extended graph $\Gamma(\underset{k}{\theta'})$ lies in E. Hence θ' is a solution of the differential equations. □

This theorem is the basis for practical calculation of symmetries of differential equations. The 'point-by-point' nature of the criterion allows us to treat the differential equations as algebraic equations in extended space. Application of Theorem 2.1.13 yields an infinitesimal form of Theorem 2.3.5.

Definition 2.3.6. A system E of s differential equations $f = 0$ (2.24) satisfies the *Jacobian condition* if the Jacobian of f with respect to the variables $(u, \underset{1}{u}, \dots, \underset{k}{u})$ is of full rank s at all points on E.

The Jacobian condition guarantees that a system of d.e.'s can be (in principle) written in *solved form*, that is, s of the derivatives $(u, \underset{1}{u}, \dots, \underset{k}{u})$ can be isolated on the left hand side. Note that the variables x are omitted when considering the Jacobian rank condition: we do not allow the independent variables x to be bound by an algebraic relation. If such an algebraic relation is present, the system E has no solutions at all, and is *inconsistent*. The Jacobian condition is not sufficient to ensure consistency, since a relation among x could be implied as a compatibility condition of the equations in the original system.

Theorem 2.3.7. *Let* E *be a system of differential equations* $f = 0$ *(2.24) satisfying the Jacobian condition. Suppose* $\mathcal{G}$ *is a Lie transformation group such that*

$$\underset{k}{\mathbf{X}} f(x, u, \underset{1}{u}, \dots, \underset{k}{u}) = 0 \qquad \textit{whenever} \;\; f(x, u, \underset{1}{u}, \dots, \underset{k}{u}) = 0 \tag{2.27}$$

for every group operator $\mathbf{X}$ *of* $\mathcal{G}$*. Then* $\mathcal{G}$ *consists of symmetries of* E.

Proof. Applying Theorem 2.1.13 to the surface E shows that a point transformation τ leaves invariant the surface E if and only if infinitesimal condition (2.27) is satisfied. Theorem 2.3.5 then shows that τ is a symmetry of E. □

This theorem gives a constructive method for finding symmetries of a system of differential equations. To ensure that *all* symmetries are found, the differential equations E must satisfy additional hypotheses.

Definition 2.3.8. A system E (2.24) of differential equations $f = 0$ is *locally solvable* if through every point $(x, u, \underset{1}{u}, \ldots, \underset{k}{u})$ on E there passes the graph of a solution $u = \theta(x)$.

The importance of local solvability is motivated as follows. There are two surfaces of interest in the space $(x, u, \underset{1}{u}, \ldots, \underset{k}{u})$. First there is the surface E specifying the differential equations. Second there is the surface generated by the union of all graphs of *solutions* of the differential equations. Only for locally solvable systems may these two surfaces be identified. If a system is not locally solvable there are portions of surface E through which there are no solutions. In this case condition (2.27) would force a transformation τ to leave invariant a surface 'larger' than that generated by the solutions—in terms of which symmetry properties of E are defined. This leads to imposition of stronger conditions than necessary on τ.

For locally solvable systems, Theorem 2.3.5 admits a converse.

Theorem 2.3.9. *A locally solvable system E of differential equations (2.24) admits a symmetry τ if and only if $(\tau, \underset{1}{\tau}, \ldots, \underset{k}{\tau})$ leaves invariant the surface E (2.24) defining the equations.*

Proof. We have only to show the converse statement. Let $P = (x_0, u_0, \underset{1}{u}_0, \ldots, \underset{k}{u}_0)$ be a point on E. Local solvability guarantees existence of a solution $u = \theta(x)$ passing through P. Applying the symmetry $\underset{k}{\tau}$ to this solution, P is mapped to a point P' which lies on a solution $u' = \theta'(x')$ of equations E. Hence P' is on E. Thus every point on E is mapped to a point on E. □

Theorem 2.3.10. *Let E be a locally solvable system of differential equations $f = 0$ (2.24) satisfying the Jacobian condition. Then a Lie transformation group $\mathcal{G}$ is a point symmetry group of* E *if and only if infinitesimal condition (2.27) is satisfied for every operator* $\mathbf{X}$ *of $\mathcal{G}$.*

Proof. Just combine Theorems 2.3.7 and 2.3.9. □

This implies that the set of all operators $\mathbf{X}$ satisfying (2.27) generates the complete symmetry group of E. Further discussion of local solvability and related issues may be found in [47, §2.6]. For more detailed material on transformation of d.e.'s and their symmetries, see [13, §3,§4], [47, §2], [52, §5].

As it stands, the local solvability criterion could not be checked without knowing all the solutions of the d.e.'s. The following regularity conditions are more convenient.

Definition 2.3.11. We call a system E of differential equations

$$f(x, u, \underset{1}{u}, \ldots, \underset{k}{u}) = 0$$

regular if

(i) The function f is analytic in all its arguments.

(ii) f satisfies the Jacobian condition.

(iii) No further relations of order k or less can be derived from E by differentiation or taking compatibility conditions.

Theorem 2.3.12. *A regular system of differential equations is locally solvable.*

This is an immediate consequence of the Riquier-Janet theory [59, 33, 67, 56] on existence of solutions of involutive systems of d.e.'s.

For a given d.e. the conditions of Definition 2.3.11 can be checked by a finite algorithm [56], and this makes the criterion of regularity useful in practice. If condition (iii) above fails, the Janet theory [33], or Reid's variant thereof [56] gives an algorithmic procedure for appending compatibility conditions until the condition *is* satisfied. This is often of no concern. For example, scalar equations trivially can have no compatibility conditions. However, compatibility conditions are of utmost importance for dealing with determining equations for symmetries, because such systems are usually overdetermined, and imply many compatibility conditions. We shall return to this point at length in §4.3.

Example 2.3.13. Consider the potential system form of the nonlinear diffusion convection equation

$$\begin{aligned} v_x &= u \\ v_t &= D(u)u_x - K(u), \end{aligned} \tag{2.28}$$

where $D(u)$ and $K(u)$ are analytic functions. This system implies as compatibility condition

$$u_t = [D(u)u_x - K(u)]_x \tag{2.29}$$

i.e., the scalar diffusion convection equation. Provided $D(u) \neq 0$, this condition is of *second order*. Since no further compatibility conditions can be derived, we conclude that (2.28) is regular and therefore locally solvable. A separate treatment is required if $D(u) \equiv 0$, and in fact local solvability of (2.28) fails in that case.

2.3.3 Algorithmic construction of symmetries

Theorem 2.3.10 above leads to an algorithmic construction of the symmetry group $\mathcal{G}$ for a regular system E of differential equations. The key observation is that condition (2.27) contains extension variables $(\underset{1}{u}, \ldots, \underset{k}{u})$, which appear through the extension formula (2.23) and in the differential equations themselves. In both cases their occurrence is *explicitly* known. Hence condition (2.27) can be split up by powers of these extension variables, yielding a system of *determining equations* for the infinitesimals ξ, η. The details of this algorithm are as follows [13, §4.3.3], [47, §2.4], [52, §5.4].

Algorithm 2.3.14 (Lie).

1. Write the system E in solved form i.e., isolate derivatives on the left hand side of each equation. If necessary, append differential consequences of the system until it is regular.

2. Let ξ^i, $i = 1, \dots, n$; and η^j, $j = 1, \dots, m$ be arbitrary functions of (x, u). Write the formal operator

$$\mathbf{X} = \xi^i(x,u)\,\frac{\partial}{\partial x^i} + \eta^j(x,u)\,\frac{\partial}{\partial u^j} \tag{2.30}$$

3. Extend the operator $\mathbf{X}$ to $\underset{k}{\mathbf{X}}$ acting on $(\underset{1}{u}, \dots, \underset{k}{u})$, where k is the order of the differential equations E. This adds to $\mathbf{X}$ the terms

$$\cdots + \eta^j_{(i)}\,\frac{\partial}{\partial u^j_i} + \cdots + \eta^j_{(I)}\,\frac{\partial}{\partial u^j_I}, \qquad 1 \le |I| \le k.$$

where $\eta^j_{(I)}$, $1 \le |I| \le k$ are determined in terms of ξ, η and their derivatives by extension formulas (2.23, 2.16).

4. Apply the extended operator $\underset{k}{\mathbf{X}}$ to the function f which defines the system E of differential equations (2.24).

5. Restrict $\underset{k}{\mathbf{X}} f(x, u, \underset{1}{u}, \dots, \underset{k}{u})$ to the surface $f = 0$ by substituting for the derivatives which occur on the left hand side of E. Set the resulting expression to zero. This yields conditions (2.27) for an infinitesimal symmetry.

At this stage, conditions (2.27) are linear homogeneous partial differential equations for the infinitesimals ξ, η. The coefficients in these invariance conditions are (known) functions of x, u, and derivatives from the system E (i.e. some of $\underset{1}{u}, \dots, \underset{k}{u}$). Provided $\underset{1}{u}, \dots, \underset{k}{u}$ occur polynomially in the original system E (2.24), they occur polynomially in the invariance condition (Theorem 2.3.10), in an *explicitly* known manner. In this case, one is able to split up the invariance condition according to powers of these derivatives into a *finite* number of *determining equations* for the infinitesimals $\xi(x, u)$ and $\eta(x, u)$. Usually this system of determining equations is *overdetermined*, consisting of more equations than unknowns.

With this discussion we complete the algorithm:

6. Split up conditions (2.27) by powers of the variables $\underset{1}{u}, \dots, \underset{k}{u}$, to give determining equations for the infinitesimal symmetry group.

7. Solve the determining equations for the infinitesimals ξ, η.

8. For each infinitesimal operator in the algebra of symmetry operators, integrate the initial value problem (2.4) to yield a set of one-parameter subgroups of the symmetry group $\mathcal{G}$. Compose these subgroups (Theorem 2.1.11) to give (the connected component of) the symmetry group $\mathcal{G}$.

Steps 7 and 8 are not strictly algorithmic, since they involve integrations, which may not be able to be performed explicitly. However, in practice solution of the differential equations of 7 and 8 can often be accomplished, and the full symmetry group of E calculated. Even if they cannot be solved, an algorithm of Reid [56, 57] gives a standard form for the determining equations, from which size and structure of the symmetry group can be found without difficulty.

For a system of equations which is not locally solvable, there is nothing to prevent application of this algorithm, but there is no guarantee that the resulting list of symmetries is complete.

Chapter 3

The Equivalence Group

3.1 Class of differential equations

3.1.1 Decoupled systems of d.e.'s

We have frequent cause to deal with "systems of d.e.'s" which are *decoupled* into two or more subsystems, which are solved in sequence; the theory described in §2 must be modified to deal with this case.

For example, the first order o.d.e.'s

$$\frac{dv}{du} = f(u, v) \tag{3.1}$$

and

$$\frac{du}{dx} = g(x, u, v) \tag{3.2}$$

are a decoupled system. The two equations are solved sequentially: first (3.1) is solved for $v = \phi(u)$, which is then inserted into (3.2), turning it into a d.e. in (x, u):

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

The interpretation we wish to give decoupled systems is as specifying a *class of equations.*

Example 3.1.1. Consider the system A:

$$a_x = 0 \qquad a_t = 0 \tag{3.3}$$

for a as a function of (x, t, u), and the equation E for $u(x, t)$:

$$u_t = au_{xx} + a_u(u_x)^2. \tag{3.4}$$

The trivial system (3.3) has solution $a = D(u)$, where $D(u)$ is any function. Inserting this into (3.4) we obtain

$$u_t = (D(u)u_x)_x. \tag{3.5}$$

Equations (3.3), (3.4) therefore describe the class of nonlinear diffusion equations (3.5).

Note that in this latter example, equations (3.3), (3.4) cannot be collectively regarded as a system of d.e.'s, since there is no consistent assignment of independent and dependent variables. In particular, u is an independent variable in (3.3), but a dependent variable in (3.4).

Although the transformation theory of the remainder of this chapter is applicable to any decoupled system of d.e.'s, our motivation and terminology is for dealing with classes $\mathcal{C}$ of d.e.'s such as the diffusion equations above. Here the first half of the decoupled system specifies certain 'arbitrary elements' such as $a = D(u)$ above, which are then inserted into the second half of the system to give the class $\mathcal{C}$. Typically the arbitrary elements represent possible physical properties of media (e.g. 'the diffusivity $D(u)$', 'a fluid of uniform density ρ'); or mathematically natural collections of equations (e.g. second order o.d.e.'s $\ddot{y} = \omega(x, y, \dot{y})$).

3.1.2 Class of d.e.'s

Let A be a

system of	σ	differential equations
of order	κ	
in	ν	independent variables $w = (w^1, w^2, \dots, w^\nu)$
and	μ	dependent variables $a = (a^1, a^2, \dots, a^\mu)$.

described by the function $g = (g^1, g^2, \dots, g^\sigma)$ as

$$g(w, a, \underset{1}{a}, \dots, \underset{\kappa}{a}) = 0. \tag{3.6}$$

Let E be the system of s equations

$$f(x, u, \underset{1}{u}, \dots, \underset{k}{u}; a, \underset{1}{a}, \dots, \underset{\kappa}{a}) = 0 \tag{3.7}$$

defined by functions $f = (f^1, f^2, \dots, f^s)$, where $u = (u^1, u^2, \dots, u^m)$ are dependent variables, $x = (x^1, x^2, \dots, x^n)$ are independent variables, and $(x, u) \equiv w$. (Note that $m + n = \nu$). When $a = \phi(w) \equiv \phi(x, u)$ is assigned to lie on the graph of a solution of A (3.6) (and correspondingly $\underset{1}{a} = \underset{1}{\phi}$, etc.), we obtain the system E(ϕ)

$$f\Big(x, u, \underset{1}{u}, \dots, \underset{k}{u}, \phi(x, u), \underset{1}{\phi}(x, u), \dots, \underset{k}{\phi}(x, u)\Big) = 0, \tag{3.8}$$

which is a system of k-th order d.e.'s for u as a function of x. The decoupled system A, E (3.6), (3.7) represents a class $\mathcal{C}$ of differential equations, namely

$$\mathcal{C} = \{\mathrm{E}(\phi) \mid \phi \text{ solves A}\}. \tag{3.9}$$

We call the functions $\phi(x, u)$ solving A the *arbitrary elements* characterizing the class $\mathcal{C}$. The system A we call the *auxiliary system* of the class; E will be called the *primary system*.

Our convention is to use Roman indices (i, j, n, m, k, s) for the primary system E and its variables x^j, u^i; and Greek $(\beta, \gamma, \nu, \mu, \kappa, \sigma)$ for the auxiliary system A and its variables w^γ, a^β.

We use pullback notation (cf. (2.13)) to indicate that a has been assigned as a function of (x,u), so that (3.8) is written

$$\phi^{*} f\,(x,u,\underset{1}{u},\dots,\underset{k}{u}) = 0,$$

where ϕ is a solution of A, so that $\phi^{*} g\,(w) \equiv 0$. Thus a function $u=\theta(x)$ is a solution of an equation $\mathrm{E}(\phi) \in \mathcal{C}$ if ϕ solves A: $\phi^{*} g\,(w) \equiv 0$, and and θ solves $\mathrm{E}(\phi)$:

$$\theta^{*} \phi^{*} f\,(x) \equiv 0. \tag{3.10}$$

Writing this out in full, it says

$$f\Big(x,\theta(x),\underset{1}{\theta}(x),\dots,\underset{k}{\theta}(x),\phi(x,\theta(x)),\underset{1}{\phi}(x,\theta(x)),\dots,\underset{k}{\phi}(x,\theta(x))\Big) = 0,$$

which is as good an advertisement as any for pullback notation.

It is important to note that because assignment of independent and dependent variables is different in A (3.6) and E (3.7), there are two different extensions here; the 'underscripts' 1 in $\underset{1}{a}$ and in $\underset{1}{u}$ have distinct meanings. Components of $\underset{1}{u}$ are values of derivatives of $u=\theta(x)$ *with respect to* x, so $\underset{1}{u} = \{u^i_j\}$, for $i=1,\dots,m$ and $j=1,\dots,n$, where $u^i_j = \frac{\partial\theta^i}{\partial x^j}(x)$. In contrast, the components of $\underset{1}{a}$ are values of derivatives of $a=\phi(w)$ *with respect to* w, so $\underset{1}{a} = \{a^\beta_\gamma\}$, for $\beta = 1,\dots,\mu$ and $\gamma=1,\dots,\nu$, where $a^\beta_\gamma = \frac{\partial\phi^\beta}{\partial w^\gamma}(w)$. For example the first extensions $\underset{1}{a}$ for nonlinear diffusion Example 3.1.1 are $\underset{1}{a} = (a_x, a_t, a_u)$, whereas $\underset{1}{u} = (u_x, u_t)$.

Example 3.1.2. Consider the class of scalar wave equations

$$u_{tt} = c^2(x) u_{xx}, \tag{3.11}$$

where $c(x)$ is an arbitrary wavespeed function representing spatial inhomogeneity of the medium. This class of equations may be specified by a decoupled system with auxiliary system A:

$$a_u = 0 \qquad a_t = 0, \tag{3.12}$$

and primary system E:

$$u_{tt} = a^2 u_{xx}. \tag{3.13}$$

Equation (3.11) results when we assign $a = c(x)$ as the general solution of A.

Potential forms of wave equations can be constructed. The system [46]

$$\begin{aligned} v_x &= c^{-2}(x)\,[\,h(x,t)\,u_t - h_t(x,t)\,u\,] \\ v_t &= \qquad\quad h(x,t)\,u_x - h_x(x,t)\,u \end{aligned} \tag{3.14}$$

is a potential system for (3.11), if the function $h(x,t)$ satisfies

$$\frac{\partial^2 h}{\partial t^2}(x,t) = c^2(x)\frac{\partial^2 h}{\partial x^2}(x,t), \tag{3.15}$$

that is, if $h(x,t)$ is a solution of (3.11). In this case, the compatibility condition of (3.14) is the scalar wave equation (3.11). If $u=\theta(x,t)$, $v=\chi(x,t)$ solve the potential system (3.14) then $u = \theta(x,t)$ solves the scalar wave equation (3.11). Conversely, if $u = \theta(x,t)$ solves the scalar equation (3.11), then for each function h satisfying (3.15), there exists a $v = \chi(x,t)$ such that $u = \theta(x,t)$, $v = \chi(x,t)$ is a solution of the potential system form (3.14) of the equation.

This class of potential systems can be specified by the auxiliary system

$$\begin{aligned} a_t = a_u &= a_v = 0 \\ b_u &= b_v = 0 \\ b_{tt} &= a^2 b_{xx} \end{aligned} \tag{3.16}$$

with dependent variables a, b as functions of (x,t,u,v). A solution $a = c(x)$, $b = h(x,t)$ of this system is inserted into the system E:

$$\begin{aligned} v_x &= a^{-2}(bu_t - b_t u) \\ v_t &= \qquad bu_x - b_x u, \end{aligned} \tag{3.17}$$

turning it into the potential wave system (3.14).

Auxiliary systems frequently merely specify dependencies, as in (3.12), but as system (3.16) illustrates, this is by no means always so.

We have assumed that the arbitrary elements $a = \phi(w)$ depend only on $w = (x,u)$, but it is quite possible for dependence on derivatives of u to arise, so that we should take $w = (x, u, \underset{1}{u}, \ldots)$. For example, the class of second order o.d.e.'s $\ddot{y} = \omega(x,y,\dot{y})$ has an arbitrary element depending on $\dot{y}$. Dealing with derivative dependence requires only minor changes: it is largely a matter of notational inconvenience.

The point of specifying $\mathcal{C}$ as an auxiliary system A and primary system E is that it is described algebraically as a coordinate locus, instead of as a collection of equations parametrized by arbitrary elements. This is analogous to treating a d.e. as a surface in an extension space, rather than as a collection of solutions.

3.2 Equivalence transformations

With notation established for classes of differential equations, we now examine transformation properties of these equations.

Let $\mathcal{C}$ be a class of differential equations (§3.1.2). We seek transformations of (x,u) which map solutions of an equation $\mathrm{E}(\phi) \in \mathcal{C}$ to solutions of another equation $\mathrm{E}(\phi') \in \mathcal{C}$. There are several ways in which such transformations could be sought, each of them yielding different levels of generality: we consider only the most restrictive case.

Ovsiannikov [52, §6] considered *equivalence transformations* which act on solutions of equations as follows. An equivalence transformation is a point transformation τ on (x,u) space. Inserting this transformation τ into any equation $\mathrm{E}(\phi) \in \mathcal{C}$ maps it to another equation $\mathrm{E}(\phi') \in \mathcal{C}$. Most importantly, the relationship between the original arbitrary element $a = \phi(w)$ and its transform $a' = \phi'(w')$ is the result of a transformation $\hat{\tau}$ acting on (w,a) space as

$$\hat{\tau}: \qquad \begin{aligned} w' &= \tau(w) \\ a' &= \sigma(w,a). \end{aligned} \tag{3.18}$$

(Recall that $w \equiv (x,u)$). That is,

$$\phi'(w') = \sigma\left(\tau^{-1}(w'),\, \phi\circ\tau^{-1}(w')\right). \tag{3.19}$$

As in §2, our base space is (x,u), the space of independent and dependent variables of the primary system of d.e.'s. We call the space $(w,a) \equiv (x,u,a)$ the *augmented* space: it is the space of independent and dependent variables of the auxiliary system of d.e.'s. We continue to call a transformation τ of (x,u) space a 'point transformation'; a transformation $\hat\tau$ of the form (3.18) acting on (w,a) space will be called an *augmented transformation*. We follow the convention that an object (such as a transformation) on augmented space is 'hatted' ($\hat\tau$); its projection to a corresponding object on base space is unhatted (τ).

Example 3.2.1. A transformation τ of the form (1.7)

$$\begin{aligned} x &= \rho^{-1}x' \\ t &= t' \\ u &= \alpha u' + \beta, \qquad \alpha,\rho\neq 0 \end{aligned} \tag{3.20}$$

maps a nonlinear diffusion equation

$$u_t = (D(u)u_x)_x$$

to a diffusion equation

$$u'_{t'} = (D'(u')u'_{x'})_{x'},$$

where

$$D'(u') = \rho^2 D(\alpha u' + \beta).$$

In terms of the variable $a = D(u)$ introduced in Example 3.1.1, such a transformation of D results from the augmented transformation $\hat\tau$ consisting of (3.20) and

$$a = \rho^{-2}a'.$$

In this section we give a theory for equivalence transformations of the type described by Ovsiannikov [52]. In the next section we give the corresponding infinitesimal form of the theory.

Definition 3.2.2. Let $\mathcal{C}$ be a class of differential equations for $u = \theta(x)$, with arbitrary elements $a = \phi(w)$. An augmented transformation $\hat\tau$ is an *equivalence transformation* for $\mathcal{C}$ if every solution $u = \theta(x)$ of the equation $\mathrm{E}(\phi)\in\mathcal{C}$ is mapped to a solution $u' = \theta'(x')$ of the equation $\mathrm{E}(\phi')\in\mathcal{C}$, where ϕ' is the transform (3.19) of ϕ under the action of $\hat\tau$.

Since $\hat\tau$ transforms ϕ (solving the auxiliary system A) to ϕ' (also solving A), an equivalence transformation maps solutions to solutions of A, that is, it is a symmetry of A. Note that, as usual, 'solutions' on which equivalence transformations act are *local*, being defined only in a neighbourhood of some point x_0 (Definition 2.3.2). Similarly, the transformations themselves need not be defined globally.

Being an equivalence transformation does *not* depend on the particular equation $\mathrm{E}(\phi)$: the same augmented transformation may be applied to every equation in $\mathcal{C}$. This fact, and the fact that such transformations project to action on (x, u), are strong restrictions on the nature of mappings between equations. Firstly, transformations such as (1.15) in Example 1.2.4 which act only on a subclass of $\mathcal{C}$ are disqualified. Thus the very interesting problem of *classifying* mappings among equations is outside our ambit. Secondly, transformations such as (1.17) for non-linear telegraph systems (Example 1.2.5) are beyond our scope, since the action of (1.17) does not project to (x, t, u, v) space. Instead, (x', t', u', v') depend on the arbitrary elements $b(u)$, $c(u)$.

Before giving the main results, we clarify the extension process for an augmented transformation $\hat{\tau}$ (3.18). In extending a transformation $\hat{\tau}$ on (x, u, a) to one on $(x, u, \underset{1}{u}, \dots, \underset{k}{u}, a)$ we require invariance of contact forms $\underset{k}{C}$ (Definition 2.2.2)

$$du^j_I - u^j_{Ii}\, dx^i, \qquad 0 \le |I| \le k-1$$

where $I = (i_1 i_2 \dots i_l)$. This ensures that u^j_I transform 'like derivatives' of u with respect to x. Extending to action on $(w, a, \underset{1}{a}, \dots, \underset{\kappa}{a})$ requires invariance of contact forms $\underset{k}{\hat{C}}$

$$da^\beta_\Lambda - a^\beta_{\Lambda\gamma}\, dw^\gamma, \qquad 0 \le |\Lambda| \le \kappa - 1$$

where $\Lambda = (\lambda_1 \lambda_2 \dots \lambda_l)$ is a multi-index. Extension of $\hat{\tau}$ to action on $(x, u, \underset{1}{u}, \dots, \underset{k}{u}; a, \underset{1}{a}, \dots, \underset{\kappa}{a})$ will be denoted by $\underset{k\,\kappa}{\hat{\tau}}$. If one of the orders of extension is of no consequence we place a dot there, so $\underset{\cdot\,\kappa}{\hat{\tau}}$ means $\hat{\tau}$ has been extended to action on $(a, \underset{1}{a}, \dots, \underset{\kappa}{a})$ and may or may not have been extended to $(\underset{1}{u}, \underset{2}{u}, \dots)$.

To specify a solution θ of a d.e. $\mathrm{E}(\phi)$ in some class requires *two* functions: θ to specify which solution, ϕ to specify which equation $a = \phi(w)$. As in Definition 2.2.1, the graph $\Gamma(\theta)$ of this solution is the set of points

$$\Gamma(\theta) = \{(x, u) \mid u = \theta(x)\};$$

the extended graph $\Gamma(\underset{k}{\theta})$ is the set of points

$$\Gamma(\underset{k}{\theta}) = \Big\{ (x, u, \underset{1}{u}, \dots, \underset{k}{u}) \;\Big|\; (u, \underset{1}{u}, \dots, \underset{k}{u}) = \Big(\theta(x), \underset{1}{\theta}(x), \dots, \underset{k}{\theta}(x)\Big) \Big\}.$$

Similarly, we define the graph $\hat{\Gamma}(\phi)$ as the set of points in (x, u, a) space

$$\hat{\Gamma}(\phi) = \{(w, a) \mid a = \phi(w)\},$$

and its extensions to $\hat{\Gamma}(\underset{\kappa}{\phi})$ analogously. If both ϕ, θ are assigned we can define the locus

$$\hat{\Gamma}(\theta, \phi) = \{(x, u, a) \mid u = \theta(x), a = \phi(x, \theta(x))\},$$

which we call the *augmented graph* of θ, ϕ. The extension of this graph is, of course

$$\begin{aligned}\hat{\Gamma}(\underset{k}{\theta}, \underset{\kappa}{\phi}) = \Big\{ (x, u, \underset{1}{u}, \dots, \underset{k}{u}; a, \underset{1}{a}, \dots, \underset{\kappa}{a}) \;\Big|\; & (u, \underset{1}{u}, \dots, \underset{k}{u}) = (\theta(x), \underset{1}{\theta}(x), \dots, \underset{k}{\theta}(x)), \\ & (a, \underset{1}{a}, \dots, \underset{\kappa}{a}) = \Big(\phi(x, \theta(x)), \underset{1}{\phi}(x, \theta(x)), \dots, \underset{\kappa}{\phi}(x, \theta(x))\Big) \Big\}\end{aligned}$$

Criterion (3.10) that a function $u = \theta(x)$ is a solution of equation $\mathrm{E}(\phi)$ (3.8) in class $\mathcal{C}$, is that $\phi^* g(w) \equiv 0$, and $\theta^* \phi^* f\,(x) \equiv 0$. In terms of graphs, this says that θ solves $\mathrm{E}(\phi) \in \mathcal{C}$ if (i) the graph $\hat{\Gamma}(\underset{\kappa}{\phi})$ lies on the surface A: $g = 0$, (ii) the augmented graph $\hat{\Gamma}(\underset{k}{\theta}, \underset{\kappa}{\phi})$ lies on the surface E: $f = 0$.

We now establish the central results which lead to algorithmic determination of equivalence transformations.

Theorem 3.2.3. *Let $\mathcal{C}$ be a class of differential equations described by the surfaces*

$$\begin{aligned} \mathrm{A} &= \Big\{ (x, u, \underset{1}{u}, \ldots, \underset{k}{u}, a, \underset{1}{a}, \ldots, \underset{\kappa}{a}) \;\Big|\; g(w, a, \underset{1}{a}, \ldots, \underset{\kappa}{a}) = 0 \Big\} \\ \mathrm{E} &= \Big\{ (x, u, \underset{1}{u}, \ldots, \underset{k}{u}, a, \underset{1}{a}, \ldots, \underset{\kappa}{a}) \;\Big|\; f(x, u, \underset{1}{u}, \ldots, \underset{k}{u}, a, \underset{1}{a}, \ldots, \underset{\kappa}{a}) = 0 \Big\}. \end{aligned}$$

If $\hat{\tau}$ is an augmented transformation whose extension $\underset{k\,\kappa}{\hat{\tau}}$ leaves invariant the surfaces A: $g = 0$ *and* $\mathrm{E} \cap \mathrm{A}$: $f = 0, g = 0$, *then τ is an equivalence transformation for the class $\mathcal{C}$.*

Proof. Let $\hat{\tau}$ be an augmented transformation satisfying the conditions of the theorem. Let $u = \theta(x)$ be a solution of the d.e. $\mathrm{E}(\phi)$, where $a = \phi(w)$ is a solution of the auxiliary system A. Denote the transform (3.19) of ϕ under the action of $\hat{\tau}$ by ϕ'. Similarly, denote the transform of θ under the action of τ by θ'. We must show firstly, that ϕ' is a solution of A, and secondly, that θ' is a solution of $\mathrm{E}(\phi')$.

(i) The extension $\underset{\cdot\,\kappa}{\hat{\tau}}$ of $\hat{\tau}$ maps the surface A to itself, and Theorem 2.3.5 then shows $\hat{\tau}$ is a symmetry of the system A. That is, any solution ϕ of A maps to a solution ϕ' under the action (3.19) of $\hat{\tau}$.

(ii) Let $\hat{P}(x_0)$ be a point on the augmented graph $\hat{\Gamma}(\underset{k}{\theta}, \underset{\kappa}{\phi})$ in the space $(x, u, \underset{1}{u}, \ldots, \underset{k}{u}, a, \underset{1}{a}, \ldots, \underset{\kappa}{a})$, so that $\hat{P}(x)$ lies on the surface $\mathrm{E} \cap \mathrm{A}$ for all x in a neighbourhood of x_0. Transforming $\hat{\Gamma}(\underset{k}{\theta}, \underset{\kappa}{\phi})$ by $\underset{k\,\kappa}{\hat{\tau}}$ yields a set of points which lie on the graph $\hat{\Gamma}(\underset{k}{\theta'}, \underset{\kappa}{\phi'})$ (by definition of θ', ϕ'), and which lie on $\mathrm{E} \cap \mathrm{A}$ (by hypothesis). Since $\hat{\Gamma}(\underset{k}{\theta'}, \underset{\kappa}{\phi'})$ lies on $\mathrm{E} \cap \mathrm{A}$, $\underset{k}{\theta'}$ is a solution of $\mathrm{E}(\phi')$, with ϕ' a solution of A (from (i)). □

Theorem 3.2.3 replaces the 'function' criterion of Definition 3.2.2 (mapping solutions to solutions) with a pointwise criterion (mapping points to points): it is analogous to Theorem 2.3.5 for symmetries of differential equations.

Once again, to guarantee that *all* equivalence transformations are found, the class must satisfy additional hypotheses. These ensure that the surfaces E, A defining the class $\mathcal{C}$ of d.e.'s accurately reflect the collection of solutions of A and the solutions of equations $\mathrm{E}(\phi) \in \mathcal{C}$.

Definition 3.2.4. A class $\mathcal{C}$ of differential equations $f = 0$, $g = 0$ is *locally solvable* if

(i) Through every point $(w, a, \underset{1}{a}, \ldots, \underset{\kappa}{a})$ on the surface A: $g = 0$ there passes the graph of a solution $a = \phi(w)$ of the auxiliary system A. (i.e., the system A is locally solvable, Definition 2.3.8)

(ii) For every point

$$\hat{P} = (x, u, \underset{1}{u}, \dots, \underset{k}{u}, a, \underset{1}{a}, \dots, \underset{\kappa}{a}) \tag{3.21}$$

on the surface E ∩ A: $f = 0$, $g = 0$ there is a function ϕ solving A and a function θ solving E(ϕ) such that $\hat{P}$ lies on the augmented graph $\hat{\Gamma}(\theta, \phi)$. That is, $\hat{P}$ can be realized as

$$\hat{P} = \Big(x, \theta(x), \underset{1}{\theta}(x), \dots, \underset{k}{\theta}(x), \phi(x, \theta(x)), \underset{1}{\phi}(x, \theta(x)), \dots, \underset{\kappa}{\phi}(x, \theta(x))\Big).$$

For a locally solvable class, Theorem 3.2.3 admits a converse.

Theorem 3.2.5. *A locally solvable class $\mathcal{C}$ of differential equations admits an augmented equivalence transformation $\hat{\tau}$ if and only if $\hat{\tau}$ leaves invariant (i) the surface* A: $g = 0$ *specifying the auxiliary system, and (ii) the surface* E ∩ A: $f = 0$, $g = 0$.

Proof. We have only to show the converse statement. Let $\hat{\tau}$ be an augmented equivalence transformation. By Definition 3.2.2 of equivalence transformation, every solution $u = \theta(x)$ of the equation E(ϕ) is mapped by $\hat{\tau}$ to a solution $u' = \theta'(x')$ of the equation E(ϕ'), where ϕ' is given by (3.19). Since $\hat{\tau}$ acts on every solution ϕ of A by (3.19) to produce another solution ϕ' of A, it is a symmetry (Definition 2.3.4) of A. By hypothesis A is locally solvable, and Theorem 2.3.9 shows that $\hat{\tau}$ leaves the surface A invariant. This establishes (i).

Now let $\hat{P}$ (3.21) be a point on the surface E ∩ A: $f = 0$, $g = 0$. Local solvability (ii) of class $\mathcal{C}$ ensures there are functions $\theta(x)$, $\phi(w)$ such that $\hat{P}$ lies on the graph $\hat{\Gamma}(\underset{k}{\theta}, \underset{\kappa}{\phi})$ of a solution θ of equation E(ϕ) where ϕ solves A. Because $\hat{\tau}$ is an equivalence transformation, it maps θ to a solution θ' of equation E(ϕ') where ϕ' solves A. Hence $\underset{k\,\kappa}{\hat{\tau}}$ maps the graph $\hat{\Gamma}(\underset{k}{\theta}, \underset{\kappa}{\phi})$ to the graph $\hat{\Gamma}(\underset{k}{\theta'}, \underset{\kappa}{\phi'})$ of these transformed functions, and $\hat{\Gamma}(\underset{k}{\theta'}, \underset{\kappa}{\phi'})$ lies on the surface E ∩ A. In particular, it maps the point $\hat{P} \in \hat{\Gamma}(\underset{k}{\theta}, \underset{k}{\phi})$ to a point $\hat{P}' \in \hat{\Gamma}(\underset{k}{\theta'}, \underset{\kappa}{\phi'})$ on E ∩ A. Hence every point $\hat{P}$ on the surface E ∩ A is mapped to a point $\hat{P}'$ on E ∩ A. □

Ovsiannikov [52, §6.4] defines equivalence transformations slightly differently. His definition is by a pointwise property (mapping points on E to points on E). Since he does not ensure that class $\mathcal{C}$ is in a form where no integrability conditions occur, his surface E may not accurately represent the collection of solutions of the equations $\mathcal{C}$. Thus leaving surface E invariant may lead to stronger conditions than necessary, because one is leaving invariant portions of surface through which there pass no solutions. Assuming the class to be locally solvable obviates this possibility and allows us to establish the above correspondence between 'solution mapping' and 'point mapping' properties. Most importantly, Ovsiannikov did not take careful account of the auxiliary system constraining the arbitrary elements. In fact his only comment [52, p.66] is that it is "required to … know in advance possible special properties of the arbitrary element (for example, independence of some components of … $[(x, u)]$)". The correct way of dealing with the auxiliary system A may be discerned from the calculations of Akhatov, et al. [3, 4] and Ibragimov, et al. [32], although their auxiliary systems all serve merely to specify that the arbitrary elements are independent of certain components of (x, u). However, both these papers rely on Ovsiannikov's definition of equivalence transformation, so their formulation will fail for systems which are not locally solvable.

Theorem 3.2.3 requires that $\hat{\tau}$ (3.18) be a symmetry of the auxiliary system A. Note that this symmetry must be *projectable* i.e., the w component of $\hat{\tau}$ is independent of a. This makes construction of such symmetries simpler than finding the full point symmetry group of A.

The most important property of equivalence transformations is the following:

Theorem 3.2.6. *Point equivalence transformations for a locally solvable class $\mathcal{C}$ of differential equations form a group $\hat{\mathcal{Q}}$ acting on augmented space (x, u, a), and a group $\mathcal{Q}$ acting on the base space (x, u).*

Proof. The augmented transformations $\hat{\tau}$ leaving invariant the surfaces $\mathrm{A}\colon\ g = 0$ and $\mathrm{E}\cap\mathrm{A}\colon\ f = 0,\, g = 0$ form a group on (x, u, a) space. By Theorem 3.2.5 this is the augmented equivalence group $\hat{\mathcal{Q}}$. Projecting this group action onto the base space (x, u) (i.e., dropping the '$a' = \sigma(w, a)$' components of (3.18)), a group $\mathcal{Q}$ is obtained. This projection is a homomorphism of $\hat{\mathcal{Q}}$. □

We make extensive use of the augmented equivalence group $\hat{\mathcal{Q}}$, since it encodes information on how both the variables (x, u) and the arbitrary elements $a = \phi(w)$ transform. The base equivalence group $\mathcal{Q}$, giving the action on the independent and dependent variables (x, u) alone, will generally play a subsidiary role.

Example 3.2.7. To illustrate these group actions, consider the class $\mathcal{C}$ of scalar wave equations (3.11)

$$u_{tt} = c^2(x)u_{xx} \tag{3.22}$$

with wavespeed $a = c(x)$, specified by auxiliary system A (3.12):

$$a_u = 0 \qquad a_t = 0, \tag{3.23}$$

and primary system

$$u_{tt} = a^2 u_{xx}. \tag{3.24}$$

The augmented equivalence group consists of transformations

$$\begin{aligned} x &= \frac{\gamma_1 x' + \gamma_2}{\gamma_3 x' + \gamma_4} \\ t &= \rho t' + \kappa \\ u &= \frac{\lambda u' + \nu_0 + \nu_1 x' + \nu_2 t' + \nu_3 x' t'}{\gamma_3 x' + \gamma_4} \\ a &= \frac{1}{\rho} \cdot \frac{\pm a'}{(\gamma_3 x' + \gamma_4)^2} \end{aligned} \tag{3.25}$$

with ten independent parameters γ_i, κ, ρ, ν_i, λ satisfying $\lambda, \rho \neq 0$, $\gamma_1\gamma_4 - \gamma_2\gamma_3 = \pm 1$. It may be directly verified that every augmented transformation of this form leaves invariant the surfaces A (3.23) and $\mathrm{E} \cap \mathrm{A}$ (in this case just E (3.24)). Transformation (3.25) maps a wave equation (3.22) with wavespeed c to another such equation with wavespeed

$$c'(x') = \rho\,(\gamma_3 x' + \gamma_4)^2\; c\Big(\frac{\gamma_1 x' + \gamma_2}{\gamma_3 x' + \gamma_4}\Big)\,. \tag{3.26}$$

Action of the equivalence group on base space is obtained simply by dropping the a component of the augmented equivalence transformations.

3.3 Infinitesimal augmented transformations

Now that we have defined the equivalence group, we turn to methods for its calculation. As with symmetries (§1.2), the naive method would be to substitute an augmented transformation $\hat{\tau}$ (3.18) into equations E (3.7) and A (3.6) specifying the class $\mathcal{C}$, then to force the new variables $(x', u', \underset{1}{u}', \ldots, \underset{k}{u}',\ a', \underset{1}{a}', \ldots, \underset{\kappa}{a}')$ to satisfy identical equations. This yields a set of defining equations satisfied by the components of $\hat{\tau}$. In practice, the resulting enormous system of nonlinear equations is intractable: practical determination of the equivalence group requires an infinitesimal version of the process, parallelling that described in Section 2.3.2 for symmetries.

3.3.1 Infinitesimal augmented transformations

Augmented transformations, acting on (x, u, a), are of the form (3.18)

$$\left.\begin{array}{l} \left.\begin{array}{lcl} x' & = & F(x,u) \\ u' & = & G(x,u) \end{array}\right\}\tau \\ \begin{array}{lcl} a' & = & H(x,u,a) \end{array} \end{array}\right\}\hat{\tau} \tag{3.27}$$

A *one-parameter group of augmented transformations* (3.18) is a collection $\hat{\tau}(\varepsilon)$ of such transformations parametrized by an additive real parameter ε (cf. Definition 2.1.5). A transformation $\hat{\tau}(\varepsilon)$ is of the form

$$\begin{array}{lcl} x' & = & F(x,u;\varepsilon) \\ u' & = & G(x,u;\varepsilon) \\ a' & = & H(x,u,a;\varepsilon)\ . \end{array} \tag{3.28}$$

As in §2.1.2, transformations $\hat{\tau}(\varepsilon)$ near the identity $\varepsilon = 0$ may be expanded as

$$\begin{array}{lcl} x' & = & x + \varepsilon\xi(x,u) + O(\varepsilon^2) \\ u' & = & x + \varepsilon\eta(x,u) + O(\varepsilon^2) \\ a' & = & a + \varepsilon\alpha(x,u,a) + O(\varepsilon^2) \end{array} \tag{3.29}$$

where

$$\begin{array}{lcl} \xi(x,u) & = & \dfrac{d}{d\varepsilon}F(x,u;\varepsilon)\Big|_{\varepsilon=0} \\ \eta(x,u) & = & \dfrac{d}{d\varepsilon}G(x,u;\varepsilon)\Big|_{\varepsilon=0} \\ \alpha(x,u,a) & = & \dfrac{d}{d\varepsilon}H(x,u,a;\varepsilon)\Big|_{\varepsilon=0}. \end{array} \tag{3.30}$$

The quantities ξ, η, α defined by (3.30) are called the *infinitesimals* of the group (3.28). By Theorem 2.1.6, the group transformations (3.28) can be recovered from the infinitesimals (3.30) by solving the initial value problem

$$\begin{aligned} \frac{dx'}{d\varepsilon} &= \xi(x', u'), & x'(0) &= x \\ \frac{du'}{d\varepsilon} &= \eta(x', u'), & u'(0) &= u \\ \frac{da'}{d\varepsilon} &= \alpha(x', u', a'), & a'(0) &= a \end{aligned} \tag{3.31}$$

As in §2.2, infinitesimal information contained in ξ, η, α is conveniently stated in terms of a *group operator*

$$\hat{\mathbf{X}} = \xi^i(x,u)\,\frac{\partial}{\partial x^i} + \eta^j(x,u)\,\frac{\partial}{\partial u^j} + \alpha^\nu(x,u,a)\,\frac{\partial}{\partial a^\nu} \tag{3.32}$$

Extension of $\hat{\mathbf{X}}$ to action on derivatives $(\underset{1}{u}, \underset{2}{u}, \ldots, \underset{k}{u})$ of u and derivatives $(\underset{1}{a}, \underset{2}{a}, \ldots, \underset{\kappa}{a})$ of a is achieved by demanding that the infinitesimal transformation (3.29) leave invariant the contact forms $\underset{k}{C}$ (2.15)

$$du^j_I - u^j_{Ii}\, dx^i, \qquad 0 \le |I| \le k-1$$

and the contact forms $\underset{k}{\hat{C}}$

$$da^\beta_\Lambda - a^\beta_{\Lambda\gamma}\, dw^\gamma, \qquad 0 \le |\Lambda| \le \kappa - 1, \tag{3.33}$$

where I, Λ are multi-indices. The total derivative operators corresponding to these contact forms are

$$D_{x^i} = \frac{\partial}{\partial x^i} + u^j_i\,\frac{\partial}{\partial u^j} + \cdots + u^j_{Ii}\,\frac{\partial}{\partial u^j_I} + \cdots \tag{3.34}$$

and

$$\hat{D}_{w^\gamma} = \frac{\partial}{\partial w^\gamma} + a^\beta_\gamma\,\frac{\partial}{\partial a^\beta} + \cdots + a^\beta_{\Lambda\gamma}\,\frac{\partial}{\partial a^\beta_\Lambda} + \cdots \tag{3.35}$$

respectively. We adopt notation $\underset{k\kappa}{\hat{\mathbf{X}}}$ to indicate that $\hat{\mathbf{X}}$ has been extended to action on $\underset{k}{u}$ and $\underset{\kappa}{a}$. As in §3.2, if one of the orders of extension is immaterial, we place a dot there, so $\underset{\cdot\kappa}{\hat{\mathbf{X}}}$ means $\hat{\mathbf{X}}$ has been extended to action on $(a, \underset{1}{a}, \ldots, \underset{\kappa}{a})$, but may or may not have been extended to $(u, \underset{1}{u}, \ldots, \underset{k}{u})$.

The extension of $\hat{\mathbf{X}}$ is then

$$\underset{k\kappa}{\hat{\mathbf{X}}} = \underbrace{\xi^i \partial_{x^i} + \eta^j \partial_{u^j}}_{\zeta^\gamma \partial_{w^\gamma}} + \sum_{1 \le |I| \le k} \eta^j_{(I)} \partial_{u^j_I} + \alpha^\beta \partial_{a^\beta} + \sum_{1 \le |\Lambda| \le \kappa} \alpha^\beta_{(\Lambda)} \partial_{a^\beta_\Lambda} \tag{3.36}$$

where $\eta^j_{(I)}$ is a function of $(x, u, \underset{1}{u}, \ldots, \underset{|I|}{u})$ and $\alpha^\beta_{(\Lambda)}$ a function of $(w, a, \underset{1}{a}, \ldots, \underset{|\Lambda|}{a})$. The components $\eta^j_{(I)}$ follow from recurrence (2.23) as

$$\eta^j_{(Ii)} = D_{x^i}\eta^j_{(I)} - u^j_{Iq}\,(D_{x^i}\xi^q), \qquad 0 \le |I| \le k-1 \tag{3.37}$$

and similarly

$$\alpha^{\beta}_{(\Lambda\gamma)} = \hat{D}_{w^\gamma}\alpha^{\beta}_{(\Lambda)} - a^{\beta}_{\Lambda\rho}(\hat{D}_{w^\gamma}\zeta^\rho), \qquad 0 \le |\Lambda| \le \kappa - 1.$$

In fact, since ζ^ρ depends only on w, this last can be written

$$\alpha^{\beta}_{(\Lambda\gamma)} = D_{w^\gamma}\alpha^{\beta}_{(\Lambda)} - a^{\beta}_{\Lambda\rho}\,(\partial_{w^\gamma}\zeta^\rho), \qquad 0 \le |\Lambda| \le \kappa - 1. \tag{3.38}$$

Although this is all an immediate consequence of the general extension formula (2.23), great care is necessary. We have adopted the shorthand $w \equiv (x, u)$, so that a component w^γ of w could represent one of the x^i. We then have *three* operators representing 'differentiation with respect to x^i': ∂_{x^i}, $\hat{D}_{x^i}$ and D_{x^i}.

Example 3.3.1. Suppose there is one independent variable x, one dependent variable u and one arbitrary element a. There are three 'derivatives with respect to x':

$$\begin{aligned} &\partial_x \\ \hat{D}_x &= \partial_x + a_x\,\partial_a + a_{ux}\partial_{a_u} + a_{xx}\partial_{a_x} + \cdots \\ D_x &= \partial_x + u_x\,\partial_u + u_{xx}\partial_{u_x} + \cdots \end{aligned}$$

The operator ∂_x 'sees' u, u_x, a, a_x, a_u as coordinates, and accordingly $\partial_x u = 0$, $\partial_x a = 0$ etc. The operator $\hat{D}_x$ sees u, u_x as coordinates, but recognizes a as a function of (x, u), so $\hat{D}_x u = 0$, but $\hat{D}_x a = a_x$. The operator D_x recognizes u as a function of x so $D_x u = u_x$. One would imagine that $D_x a = a_x + u_x a_u$, but our operator D_x never operates on a, so there is no necessity to do so.

We compute an extension of the operator

$$\hat{\mathbf{X}} = \xi(x, u)\,\partial_x + \eta(x, u)\,\partial_u + \alpha(x, u, a)\,\partial_a$$

to the operator

$$\underset{11}{\hat{\mathbf{X}}} = \xi\,\partial_x + \eta\,\partial_u + \alpha\,\partial_a + \eta_{(x)}\partial_{u_x} + \alpha_{(x)}\partial_{a_x} + \alpha_{(u)}\partial_{a_u},$$

finding

$$\eta_{(x)} = D_x\eta - u_x D_x\xi$$

and

$$\begin{aligned} \alpha_{(x)} &= \hat{D}_x\alpha - a_x\,\partial_x\xi - a_u\,\partial_x\eta \\ \alpha_{(u)} &= \hat{D}_u\alpha - a_x\,\partial_u\xi - a_u\,\partial_u\eta \end{aligned}$$

with

$$D_x = \partial_x + u_x\,\partial_u + \cdots$$

and

$$\begin{aligned} \hat{D}_x &= \partial_x + a_x\,\partial_a + \cdots \\ \hat{D}_u &= \partial_u + a_u\,\partial_a + \cdots . \end{aligned}$$

The possibilities for confusion should be apparent.

In [4], Akhatov, et al. adopted a similar notation: their $\tilde{D}_x$ is our $\hat{D}_x$.

3.3.2 Algebra of equivalence operators

Now that infinitesimal augmented operators and their extensions have been defined, we examine properties of operators associated with the (augmented) equivalence group $\hat{\mathcal{Q}}$ of a class $\mathcal{C}$ of differential equations.

We first 'infinitesimalize' Theorem 3.2.3. As with symmetries, an additional hypothesis on the class of d.e.'s is required.

Definition 3.3.2. A class of differential equations specified by s equations $f = 0$, with σ auxiliary equations $g = 0$ satisfies the *Jacobian condition* if

(i) the Jacobian of g with respect to $a, \underset{1}{a}, \dots, \underset{\kappa}{a}$ is of full rank σ at all points satisfying $g = 0$.

(ii) the Jacobian of (f, g) with respect to $(u, \underset{1}{u}, \dots, \underset{k}{u}, a, \underset{1}{a}, \dots, \underset{\kappa}{a})$ is of full rank $s + \sigma$ at all points on $\mathrm{E} \cap \mathrm{A}\colon\ f = 0,\ g = 0$.

Theorem 3.3.3. *Let $\mathcal{C}$ be a class of differential equations satisfying the Jacobian condition. Suppose $\hat{\mathcal{Q}}$ is a Lie augmented transformation group such that*

$$\underset{\cdot\kappa}{\hat{\mathbf{X}}} g(w, a, \underset{1}{a}, \dots, \underset{\kappa}{a}) = 0 \qquad \text{when } g(w, a, \underset{1}{a}, \dots, \underset{\kappa}{a}) = 0 \tag{3.39}$$

and

$$\underset{k\kappa}{\hat{\mathbf{X}}} f(x, u, \underset{1}{u}, \dots, \underset{k}{u}, a, \underset{1}{a}, \dots, \underset{\kappa}{a}) = 0 \quad \text{when} \left\{ \begin{array}{r} g(w, a, \underset{1}{a}, \dots, \underset{\kappa}{a}) = 0 \\ f(x, u, \underset{1}{u}, \dots, \underset{k}{u}, a, \underset{1}{a}, \dots, \underset{\kappa}{a}) = 0 \end{array} \right. . \tag{3.40}$$

for every augmented operator $\hat{\mathbf{X}}$ of $\hat{\mathcal{Q}}$. Then $\hat{\mathcal{Q}}$ consists of equivalence transformations of $\mathcal{C}$.

Proof. Let the operators $\hat{\mathbf{X}}$ of the group $\hat{\mathcal{Q}}$ satisfy the conditions of the theorem. Applying Theorem 2.3.7 to condition (3.39) shows that $\hat{\mathcal{Q}}$ consists of symmetries of the auxiliary system A. Given (3.39) is satisfied, condition (3.40) is identical to

$$\left. \begin{array}{r} \underset{k\kappa}{\hat{\mathbf{X}}} f = 0 \\ \underset{\cdot\kappa}{\hat{\mathbf{X}}} g = 0 \end{array} \right\} \qquad \text{when } f = 0 \text{ and } g = 0.$$

Applying Theorem 2.1.13 to the surface $\mathrm{E} \cap \mathrm{A}\colon\ f = 0,\ g = 0$ shows that $\hat{\mathcal{Q}}$ consists of transformations leaving invariant $\mathrm{E} \cap \mathrm{A}$. Applying Theorem 3.2.3 shows that such transformations are equivalence transformations. □

We call the set of operators $\hat{\mathbf{X}}$ satisfying conditions (3.39, 3.40) the infinitesimal (augmented) equivalence group for the class $\mathcal{C}$ of equations.

Just as Theorem 2.3.10 does for symmetries, Theorem 3.3.3 gives a constructive method for finding equivalence transformations. For locally solvable classes of d.e.'s, Theorem 3.3.3 admits a converse., and here we can guarantee that *all* equivalence transformations can be found from the infinitesimal criteria.

Theorem 3.3.4. *Let $\mathcal{C}$ be a locally solvable class of differential equations $f = 0$, $g = 0$ satisfying the Jacobian condition. Then an augmented transformation group $\hat{\mathcal{Q}}$ is an augmented equivalence group of the class $\mathcal{C}$ if and only if infinitesimal conditions (3.39, 3.40) are satisfied for every operator $\hat{\mathbf{X}}$ of $\hat{\mathcal{Q}}$.*

Proof. We have only to prove the converse statement. Suppose $\hat{\mathcal{Q}}$ is the equivalence group of the class $\mathcal{C}$. Since $\mathcal{C}$ is locally solvable, Theorem 3.2.5 shows that every $\hat{\tau} \in \hat{\mathcal{Q}}$ leaves the surfaces $\mathrm{A}\colon\ g = 0$ and $\mathrm{E}\cap\mathrm{A}\colon\ f = 0,\, g = 0$ invariant. Theorem 2.1.13 then implies the infinitesimal conditions (3.39, 3.40). □

Thus the set of operators $\hat{\mathbf{X}}$ satisfying (3.39), (3.40) generates the complete point equivalence group of the class $\mathcal{C}$. The sequence of Theorems 3.2.3–3.3.4 exactly parallels the symmetry results of Theorems 2.3.5–2.3.10.

The algorithm for constructing the equivalence group from the infinitesimal criteria (3.39, 3.40) will be detailed in the next subsection. First we illustrate the result with an example.

Example 3.3.5. Consider the potential system form

$$\begin{aligned} v_x &= u \\ v_t &= D(u)u_x \end{aligned} \tag{3.41}$$

of the nonlinear diffusion equation

$$u_t = (D(u)u_x)_x$$

Letting $a = D(u)$ be the coordinate of 'diffusivity space', the primary system E is

$$\begin{aligned} v_x - u &= 0 \\ v_t - au_x &= 0 \end{aligned} \tag{3.42}$$

The auxiliary system A is

$$a_x = a_t = a_v = 0 \tag{3.43}$$

The most general transformation $\hat{\tau}$ in the equivalence group $\hat{\mathcal{Q}}$ is

$$\begin{aligned} v &= \frac{\lambda}{\rho}(\alpha v' + \beta x') + \kappa_0 \\ x &= \frac{\lambda}{\rho}(\gamma v' + \delta x') + \kappa_1 \\ t &= \frac{\lambda^2}{\rho}t' + \kappa_2 \\ u &= \frac{\alpha u' + \beta}{\gamma u' + \delta}, \qquad \alpha\delta - \beta\gamma = \pm 1 \\ a &= \rho^{-1}(\gamma u' + \delta)^2 a', \qquad \lambda \neq 0,\ \rho > 0. \end{aligned} \tag{3.44}$$

This group acts on diffusivity functions by

$$D'(u') = \frac{\rho}{(\gamma u' + \delta)^2} D\Big(\frac{\alpha u' + \beta}{\gamma u' + \delta}\Big), \qquad \alpha\delta - \beta\gamma = \pm 1. \tag{3.45}$$

Note that although $\hat{\mathcal{Q}}$ (3.44) has eight independent parameters, only four (independent) parameters affect the diffusivity $D(u)$. We return to this point in §3.3.5.

The infinitesimal operators corresponding to this eight-parameter group, obtained by differentiation of (3.44), are

$$
\begin{array}{ll}
\hat{\mathbf{X}}_1 = & \partial_x \\
\hat{\mathbf{X}}_2 = & \partial_t \\
\hat{\mathbf{X}}_3 = & x\,\partial_x + 2t\,\partial_t + v\,\partial_v \\
\hat{\mathbf{X}}_4 = & \partial_v \\
\hat{\mathbf{X}}_5 = & x\,\partial_v + \partial_u \\
\hat{\mathbf{X}}_6 = & -\tfrac{1}{2}x\,\partial_x + \tfrac{1}{2}v\,\partial_v + u\,\partial_u - a\,\partial_a \\
\hat{\mathbf{X}}_7 = & -v\,\partial_x + u^2\,\partial_u - 2ua\,\partial_a \\
\hat{\mathbf{X}}_8 = & x\,\partial_x + t\,\partial_t + v\,\partial_v + a\,\partial_a
\end{array}
\tag{3.46}
$$

These operators were found by Akhatov, et al. [3], although the form of potential equation they analyzed was slightly different. Only the operators $\mathbf{X}_5$, $\mathbf{X}_6$, $\mathbf{X}_7$, $\mathbf{X}_8$ affect the form of $D(u)$.

It may be directly verified that operators (3.46) satisfy conditions (3.39, 3.40) of Theorem 3.3.3. Consider for instance $\hat{\mathbf{X}}_7$. Computing the extension to $(x, t, u, v, u_x, v_x, v_t, a, a_x, a_t, a_v)$ space by the method of §3.3.1 gives

$$
\begin{aligned}
\underset{11}{\hat{\mathbf{X}}_7} \;=\; & -v\,\frac{\partial}{\partial x} + u^2\,\frac{\partial}{\partial u} + u_x(2u+v_x)\,\frac{\partial}{\partial u_x} + v_x^2\,\frac{\partial}{\partial v_x} + v_x v_t\,\frac{\partial}{\partial v_t} \\
& -2ua\,\frac{\partial}{\partial a} - 2ua_x\,\frac{\partial}{\partial a_x} - 2ua_t\,\frac{\partial}{\partial a_t} - (2ua_v - a_x)\,\frac{\partial}{\partial a_v}.
\end{aligned}
$$

Applying the extended operator $\underset{11}{\hat{\mathbf{X}}_7}$ to the left hand side of E (3.42) gives

$$v_x^2 - u^2, \qquad v_x(v_t - au_x)$$

which vanish identically on $\mathrm{E}\cap\mathrm{A}$. Applying it to A (3.43) gives

$$-2ua_x, \qquad -2ua_t, \qquad a_x - 2ua_v$$

which vanish identically on A, so $\hat{\mathbf{X}}_7$ indeed satisfies infinitesimal invariance conditions (3.39, 3.40).

The converse—that operators satisfying (3.39, 3.40) generate the equivalence group $\hat{\mathcal{Q}}$—will give the algorithm.

Example 3.3.5. **(cont.)** We show how to determine operators (3.46) for the potential diffusion system. First, an arbitrary operator $\hat{\mathbf{X}}$ on the augmented space (x, t, u, v, a) is of the form

$$\hat{\mathbf{X}} = \xi\,\partial_x + \tau\,\partial_t + \eta\,\partial_u + \sigma\,\partial_v + \alpha\,\partial_a$$

where ξ, τ, η, σ are functions of (x, t, u, v) and α is a function of (x, t, u, v, a). The necessary extension components are

$$\cdots + \eta_{(x)}\,\frac{\partial}{\partial u_x} + \sigma_{(x)}\,\frac{\partial}{\partial v_x} + \sigma_{(t)}\,\frac{\partial}{\partial v_t} + \alpha_{(x)}\,\frac{\partial}{\partial a_x} + \alpha_{(t)}\,\frac{\partial}{\partial a_t} + \alpha_{(v)}\,\frac{\partial}{\partial a_v}$$

where $\eta_{(x)}, \sigma_{(x)}, \sigma_{(t)}$ are functions of $(x, t, u, v, u_x, u_t, v_x, v_t)$ and $\alpha_{(x)}, \alpha_{(t)}, \alpha_{(v)}$ are functions of $(x, t, u, v, a, a_x, a_t, a_u, a_v)$. These components are computed by the method of §3.3.1. Enforcing conditions (3.39, 3.40) by applying the extended operator $\underset{11}{\hat{\mathbf{X}}}$ to equations A (3.43) and E (3.42) gives

$$\begin{array}{lll} \alpha_{(x)} = \alpha_{(t)} = \alpha_{(v)} = 0 & \text{on A} & \text{(a)} \\ \left.\begin{array}{rcl} \sigma_{(x)} & = & \eta \\ \sigma_{(t)} & = & a\eta_{(x)} + u_x\alpha \end{array}\right\} & \text{on E} \cap \text{A} & \text{(b)} \end{array} \tag{3.47}$$

Restriction to the surfaces A, E is achieved by substituting for a_x, a_t, a_v from (3.43) and for v_x, v_t from (3.42). This yields from (3.47a)

$$\begin{aligned} \alpha_x - a_u\eta_x &= 0 \\ \alpha_t - a_u\eta_t &= 0 \\ \alpha_v - a_u\eta_v &= 0. \end{aligned} \tag{3.48}$$

Since η, α do not depend on the coordinate a_u, (3.48) decomposes to give

$$\begin{aligned} \alpha_x = \alpha_t &= \alpha_v = 0 \\ \eta_x = \eta_t &= \eta_v = 0 \end{aligned} \tag{3.49}$$

Equation (3.47b) yields—after taking account of (3.49)—

$$\begin{aligned} \eta = (\sigma_x + u\sigma_v - u\xi_x - u^2\xi_v) + (\sigma_u - u\xi_u - a\tau_x - ua\tau_v) - a\tau_u u_x^2 \\ (\sigma_t - u\xi_t) + (\sigma_v - \tau_t)au_x + (\sigma_u - u\xi_u)u_t - a^2\tau_v u_x^2 = \\ ((\eta_u - \xi_x)a + \alpha)u_x - (\tau_x + u\tau_v)au_t - a\xi_u u_x^2 \end{aligned}$$

None of the infinitesimals ξ, τ, η, σ, α depend on derivatives u_x, u_t, so these equations can be split up by powers of u_x, u_t. Also, none of the ξ, τ, η, σ depend on the coordinate a, so equations not involving α may be split up by powers of a. Ultimately one arrives at a set of determining equations, which may be manipulated to involutive form:

$$\begin{gathered} \xi_t = \xi_u = 0 \\ \tau_x = \tau_u = \tau_v = 0 \\ \sigma_t = \sigma_u = 0 \\ \xi_{xx} = \xi_{xv} = \xi_{vv} = 0 \\ \sigma_{xx} = \sigma_{xv} = \sigma_{vv} = 0 \\ \tau_{tt} = 0 \\ \eta = \sigma_x + u(\sigma_v - \xi_x) - u^2\xi_v \\ \alpha = a(2\xi_x - \tau_t - 2u\xi_v) \end{gathered}$$

The general solution of these determining equations is easily found, and is given by an arbitrary linear combination of the operators $\hat{\mathbf{X}}_1, \ldots, \hat{\mathbf{X}}_8$ (3.46) (cf. [3]). Integrating the initial value problem (3.31), composing the one-parameter groups, and reparametrizing, yields the equivalence group $\hat{\mathcal{Q}}$ (3.44). (Or at least, its connected component: the discrete transformations $x \mapsto -x$, $v \mapsto -v$; and $u \mapsto -u$, $v \mapsto -v$ are not connected to the identity, and must be found somehow.)

In [3], Akhatov, Gazizov and Ibragimov used the equivalence algorithm just outlined to find what are essentially the above results.

[,]	$\hat{\mathbf{X}}_1$	$\hat{\mathbf{X}}_2$	$\hat{\mathbf{X}}_3$	$\hat{\mathbf{X}}_4$	$\hat{\mathbf{X}}_5$	$\hat{\mathbf{X}}_6$	$\hat{\mathbf{X}}_7$	$\hat{\mathbf{X}}_8$
$\hat{\mathbf{X}}_1$	0	0	$\hat{\mathbf{X}}_1$	0	$\hat{\mathbf{X}}_4$	$-\frac{1}{2}\hat{\mathbf{X}}_1$	0	$\hat{\mathbf{X}}_1$
$\hat{\mathbf{X}}_2$	0	0	$2\hat{\mathbf{X}}_2$	0	0	0	0	$\hat{\mathbf{X}}_2$
$\hat{\mathbf{X}}_3$	$-\hat{\mathbf{X}}_1$	$-2\hat{\mathbf{X}}_2$	0	$-\hat{\mathbf{X}}_4$	0	0	0	0
$\hat{\mathbf{X}}_4$	0	0	$\hat{\mathbf{X}}_4$	0	0	$-\frac{1}{2}\hat{\mathbf{X}}_4$	$-\hat{\mathbf{X}}_1$	$\hat{\mathbf{X}}_4$
$\hat{\mathbf{X}}_5$	$-\hat{\mathbf{X}}_4$	0	0	0	0	$\hat{\mathbf{X}}_5$	$2\hat{\mathbf{X}}_6$	0
$\hat{\mathbf{X}}_6$	$\frac{1}{2}\hat{\mathbf{X}}_1$	0	0	$\frac{1}{2}\hat{\mathbf{X}}_4$	$-\hat{\mathbf{X}}_5$	0	$\hat{\mathbf{X}}_7$	0
$\hat{\mathbf{X}}_7$	0	0	0	$\hat{\mathbf{X}}_1$	$-2\hat{\mathbf{X}}_6$	$-\hat{\mathbf{X}}_7$	0	0
$\hat{\mathbf{X}}_8$	$-\hat{\mathbf{X}}_1$	$-\hat{\mathbf{X}}_2$	0	$-\hat{\mathbf{X}}_4$	0	0	0	0

Table 3.1: Commutator table of equivalence algebra (3.46) for nonlinear diffusion potential system (3.41).

Since equivalence transformations form a group, the infinitesimal equivalence operators $\hat{\mathbf{X}}$ form a Lie algebra $\hat{L}$ of operators on the space (x, u, a). Hence, by analogy with the Lie symmetry algebra, we may call the operators satisfying (3.40, 3.40) the *Lie algebra of equivalence operators* for the class $\mathcal{C}$ of equations.

Example 3.3.5. **(cont.)** Consider the potential nonlinear diffusion system (3.41) discussed above. The Lie algebra structure of the equivalence operators (3.46) is given by the commutation relations in Table 3.1. The Lie algebra of operators $\hat{R} = \{\hat{\mathbf{X}}_5, \hat{\mathbf{X}}_6, \hat{\mathbf{X}}_7, \hat{\mathbf{X}}_8\}$ which actually affect $D(u)$ appears in the lower right hand corner. Note that the algebra is a semidirect sum $\hat{L} = \hat{K} \oplus_s \hat{R}$, where $\hat{K} = \{\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \hat{\mathbf{X}}_3, \hat{\mathbf{X}}_4\}$ is the algebra of operators which do not affect $D(u)$.

3.3.3 Algorithm for construction of equivalence group

Theorem 3.3.3 of the last subsection leads to an algorithmic construction of the equivalence group $\hat{\mathcal{Q}}$ for a class $\mathcal{C}$ of differential equations. Details of this construction have already been illustrated with the example of §3.3.2: no further theory is required.

Algorithm 3.3.6.

1. Let $\xi^i(x, u)$, $i = 1, \ldots, n$; $\eta^j(x, u)$, $j = 1, \ldots, m$ and $\alpha^\beta(x, u, a)$, $\beta = 1, \ldots, \mu$ be arbitrary functions of their arguments. Write the formal operator

$$\hat{\mathbf{X}} = \xi^i(x, u)\,\frac{\partial}{\partial x^i} + \eta^j(x, u)\,\frac{\partial}{\partial u^j} + \alpha^\beta(x, u, a)\,\frac{\partial}{\partial a^\beta}$$

2. Extend $\hat{\mathbf{X}}$ to action on derivatives $(\underset{1}{u}, \ldots, \underset{k}{u})$ of the dependent variables u (where k is the order of the differential equations in class $\mathcal{C}$), and to derivatives of the arbitrary elements $(\underset{1}{a}, \ldots, \underset{\kappa}{a})$, where κ is the order of the

auxiliary system. This gives

$$\underset{k\kappa}{\hat{\mathbf{X}}} = \hat{\mathbf{X}} + \sum_{1\le|I|\le k} \eta^j_{(I)} \frac{\partial}{\partial u^j_I} + \sum_{1\le|\Lambda|\le\kappa} \alpha^\beta_{(\Lambda)} \frac{\partial}{\partial a^\beta_\Lambda}$$

with $\eta^j_{(I)}$ determined from (3.37), and $\alpha^\beta_{(\Lambda)}$ from (3.38).

3. Apply the extended operator $\underset{k\kappa}{\hat{\mathbf{X}}}$ to the functions f, g which define the class $\mathcal{C}$ of differential equations (3.7, 3.6).

4. Enforce invariance of the auxiliary system A by restricting

$$\underset{\cdot\kappa}{\hat{\mathbf{X}}} g(w, a, \underset{1}{a}, \dots, \underset{\kappa}{a})$$

to the surface $g = 0$ and setting the resulting expression to zero.
Force $\underset{k\kappa}{\hat{\mathbf{X}}} f(x, u, \underset{1}{u}, \dots, \underset{k}{u}, a, \underset{1}{a}, \dots, \underset{\kappa}{a})$ to vanish on the surface $f = 0$, $g = 0$. This yields infinitesimal conditions (3.39, 3.40) for an infinitesimal equivalence transformation..

5. (Assuming the conditions are polynomial in the derivatives.) Split up conditions (3.39, 3.40) by powers of the variables $\underset{1}{u}, \dots, \underset{k}{u}$, $a, \underset{1}{a}, \dots, \underset{\kappa}{a}$ wherever possible. This gives the determining equations for the infinitesimal equivalence group. Manipulate these equations to involutive form by the algorithm of Reid [56].

6. Solve the determining equations for the infinitesimals ξ, η, α.

7. For each infinitesimal operator in the algebra of equivalence operators, integrate the initial value problem (3.31) to yield a set of one-parameter subgroups of the augmented equivalence group $\hat{\mathcal{Q}}$. Compose these subgroups to give the (connected component of) the equivalence group $\hat{\mathcal{Q}}$.

Comparing this equivalence algorithm with the symmetry method described in §2.3.3, the only essential difference is in the nature of the extensions needed. The above algorithm constructs 'symmetry transformations' of certain surfaces embedded in an augmented space. The equivalence group represents 'symmetries' of a *class* of d.e.'s, that is, transformations which leave the *class* invariant.

3.3.4 Proposition on form of infinitesimals

All necessary machinery is now in place to compute equivalence groups for specific examples. However, first we establish a proposition which gives the determining equations resulting from the commonest kind of auxiliary equations. Usually an arbitrary element $a^\beta = \phi^\beta(w)$ does not depend on all variables w, but only on a certain subset. In the wave equation (3.22), the wavespeed $a = c(x)$ is independent of the variables t, u, so the auxiliary system A includes the equations $a_t = a_u = 0$ (3.23). An equivalence operator $\hat{\mathbf{X}}$ leaves A invariant (Theorem 3.3.3), and enforcing invariance for equations of this simple type leads to simple determining equations (e.g. (3.49)). In the following, we assume the auxiliary system A is in solved form, i.e., certain 'leading derivatives' have been isolated on the left hand side. The derivatives not occurring on the left hand side are 'parametric derivatives'.

Proposition 3.3.7. *Let $\mathcal{C}$ be a class of d.e.'s characterized by arbitrary elements $a^\beta = \phi^\beta(w)$, $\beta = 1, 2, \dots, \mu$. Let $\mathcal{I}$ be the set of indices $\binom{\beta}{\gamma}$ such that equations $a^\beta_\gamma = 0$ are in the auxiliary system* A *of $\mathcal{C}$. Assume that no other first derivatives occur as leading derivatives in the auxiliary system. Then the infinitesimals ζ^γ, α^β of an equivalence operator*

$$\hat{\mathbf{X}} = \zeta^\gamma(w)\partial_{w^\gamma} + \alpha^\beta(w, a)\partial_{a^\beta}$$

satisfy the following equations: For each $\binom{\beta}{\gamma} \in \mathcal{I}$,

(i) $\dfrac{\partial \alpha^\beta}{\partial w^\gamma} = 0$

(ii) $\dfrac{\partial \zeta^\rho}{\partial w^\gamma} = 0$ *for ρ such that a^β_ρ is parametric in* A.

(iii) $\dfrac{\partial \alpha^\beta}{\partial a^\lambda} = 0$ *for λ such that a^λ_γ is parametric in* A.

Proof. Invariance of the equations

$$a^\beta_\gamma = 0, \qquad \tbinom{\beta}{\gamma} \in \mathcal{I} \tag{3.50}$$

of the auxiliary system A yields the conditions $a^\beta_{(\gamma)} = 0$ on A. That is,

$$\frac{\partial \alpha^\beta}{\partial w^\gamma} + a^\lambda_\gamma \frac{\partial \alpha^\beta}{\partial a^\lambda} - a^\beta_\rho \frac{\partial \zeta^\rho}{\partial w^\gamma} = 0 \quad \text{on A}, \qquad \tbinom{\beta}{\gamma} \in \mathcal{I}$$

Inserting (3.50) gives

$$\frac{\partial \alpha^\beta}{\partial w^\gamma} + \sum_{\lambda \in \mathcal{J}_\gamma} a^\lambda_\gamma \frac{\partial \alpha^\beta}{\partial a^\lambda} - \sum_{\rho \in \mathcal{J}^\beta} a^\beta_\rho \frac{\partial \zeta^\rho}{\partial w^\gamma} = 0, \qquad \tbinom{\beta}{\gamma} \in \mathcal{I} \tag{3.51}$$

where $\mathcal{J}_\gamma$ is the set of indices λ such that a^λ_γ is parametric in A.
$\mathcal{J}^\beta$ is the set of indices ρ such that a^β_ρ is parametric in A.

By hypothesis there are no first order leading derivatives other than (3.50), so condition (3.51) may be split up with respect to the remaining first derivatives. Note that a^λ_γ in the first sum of (3.51) could match a^β_ρ in the second only if $\lambda = \beta$, $\rho = \gamma$. But $\beta \notin \mathcal{J}_\gamma$, $\gamma \notin \mathcal{J}^\beta$, so this does not occur. Hence splitting (3.51) yields the proposition. □

The interpretation of this proposition is that an arbitrary element a^β which is independent of a certain variable w^γ must remain so after applying the infinitesimal transformation represented by $\hat{\mathbf{X}}$. Thus the infinitesimal α^β must be independent of w^γ and must also be independent of any components of a which depend on w^γ. Moreover, the transformation of the components w^ρ on which a^β may depend must *not* depend on w^γ (otherwise the transformed a^β could depend on w^γ indirectly through its arguments w^ρ).

This proposition is useful in generating determining equations without calculation. Note that the conclusions are unaffected if the auxiliary system A includes equations other than (3.50), provided they are of second or higher order. However, if A includes an equation with a first order leading derivative other than (3.50) (for example $a^\beta_\gamma = 2a^\beta$) then the conclusions of the proposition do not hold. This is because this leading derivative must be inserted into (3.51), thereby affecting its decomposition.

3.3.5 Structure of the equivalence group

We may now efficiently calculate equivalence groups, and it is probably worthwhile to skip to §3.4 to see some further examples. However in interpreting equivalence groups in these examples, we need the structural features to be described in this subsection.

In §3.2, §3.3, we obtained the main results leading to an algorithmic construction of the equivalence group. In Example 3.3.5 we noted that not all of the equivalence group affected the arbitrary element $D(u)$: there was present a four-parameter subgroup $\hat{\mathcal{K}} \prec \hat{\mathcal{Q}}$ with trivial action on $D(u)$. This subgroup $\hat{\mathcal{K}}$ therefore maps each equation $\mathrm{E}(\phi)$ in the class $\mathcal{C}$ to the *same* equation $\mathrm{E}(\phi)$, and therefore consists of *symmetries* common to every equation in the class. A point which did not arise in the diffusion example was the possible presence of augmented transformations which act *only* on the arbitrary elements, not transforming the base variables at all. Both these kinds of transformation are mostly of nuisance value, and our main interest is in how to factor them out.

Common symmetries

Definition 3.3.8. A *common symmetry* for a class $\mathcal{C}$ of differential equations is a transformation κ (of the base space (x, u)) which is a symmetry of every equation in the class.

Proposition 3.3.9. *The set of all common symmetries is a* group $\mathcal{K}$ *on base space* (x, u)*. Moreover* $\mathcal{K}$ *is a normal subgroup of the base equivalence group* $\mathcal{Q}$*.*

Proof. That $\mathcal{K}$ is a group is immediate. To show $\mathcal{K} \prec \mathcal{Q}$, note that any transformation $\kappa \in \mathcal{K}$ is a symmetry of every equation $\mathrm{E}(\phi) \in \mathcal{C}$, so it can be regarded as having action $\phi'(w') = \phi(w)$ on the arbitrary elements ϕ. Hence, by Definition 3.2.2, the augmented transformation $\hat{\kappa}$

$$\begin{aligned} (x', u') &= \kappa(x, u) \\ a' &= a. \end{aligned} \tag{3.52}$$

is an equivalence transformation. Disregarding the a component, we have $\mathcal{K}$ is a subgroup of $\mathcal{Q}$.

To show this subgroup is normal, let $\kappa \in \mathcal{K}$ be a common symmetry, and let $\hat{\tau} \in \hat{\mathcal{Q}}$ be any equivalence transformation. Suppose $\hat{\tau}$ maps $\mathrm{E}(\phi)$ to $\mathrm{E}(\phi')$, so that $\hat{\tau}^{-1}$ maps $\mathrm{E}(\phi')$ to $\mathrm{E}(\phi')$. Also, κ can be augmented to $\hat{\kappa}$ (3.52), which maps $\mathrm{E}(\phi')$ to itself for any ϕ'. Hence $\tau^{-1} \circ \kappa \circ \tau$ maps $\mathrm{E}(\phi)$ to itself for any ϕ, so by definition it is a common symmetry. Hence $\mathcal{K}$ is normal. □

The above properties of the common symmetry group $\mathcal{K}$ were noted by Ovsiannikov [52, 51].

Usually the common symmetries represent some very basic physical properties of the class of equations under consideration, such as homogeneity of a medium, or arbitrary choice of zero for a potential variable. In the potential diffusion system Example 3.3.5 considered above, the common symmetries are generated by operators $\hat{K} = \{\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \hat{\mathbf{X}}_3, \hat{\mathbf{X}}_4\}$, giving rise to the group $\hat{\mathcal{K}}$

$$\begin{aligned} v' &= \lambda v + \kappa_0 \\ x' &= \lambda x + \kappa_1 \\ t' &= \lambda^2 t + \kappa_2, \end{aligned} \tag{3.53}$$

whose transformations represent spatial (κ_1) and temporal (κ_2) homogeneity; arbitrary choice of zero for potential v (κ_0); and a dimensional scaling invariance (λ).

Trivial equivalence

The fussiness in the proof of Proposition 3.3.9 is due to the fact that augmentation of an equivalence transformation need not be unique: there can be many equivalence transformations $\hat{\tau}$ on (x, u, a) with the same component τ on the base space (x, u).

Definition 3.3.10. An augmented equivalence transformation $\hat{\rho}$ is *trivial* if it has trivial action on the base space, i.e., is of the form

$$\begin{aligned} x' &= x \\ u' &= u \\ a' &= \sigma(x, u, a) \end{aligned} \tag{3.54}$$

Trivial equivalence transformations reflect the extent to which augmentation is nonunique. Let $\hat{\tau}_1$, $\hat{\tau}_2$ be two equivalence transformations which both project to the same action τ on the base space (x, u). Then there is a trivial equivalence transformation $\hat{\rho}$ such that $\hat{\tau}_2 = \hat{\rho} \circ \hat{\tau}_1$, namely $\rho = \hat{\tau}_1^{-1} \circ \hat{\tau}_2$. Since $\hat{\tau}_1$ and $\hat{\tau}_2$ have the same action τ on the base space, ρ has trivial action on (x, u), so is a trivial equivalence transformation.

Proposition 3.3.11. *The set of trivial augmented equivalence transformations is a normal subgroup $\hat{\mathcal{M}}$ of the augmented equivalence group $\hat{\mathcal{Q}}$. The quotient group $\hat{\mathcal{Q}}/\hat{\mathcal{M}}$ is isomorphic to the base equivalence group $\mathcal{Q}$.*

Proof. That $\hat{\mathcal{M}}$ is a normal subgroup of $\hat{\mathcal{Q}}$ is immediate. The base equivalence group $\mathcal{Q}$ is the homomorphic image of $\hat{\mathcal{Q}}$, since it is obtained by projection (i.e. by dropping the a components of $\hat{\mathcal{Q}}$). The kernel of this homomorphism is the group $\hat{\mathcal{M}}$ of trivial equivalence transformations (3.54). Hence $\mathcal{Q} \simeq \hat{\mathcal{Q}}/\hat{\mathcal{M}}$. □

A trivial equivalence transformation projects to the identity transformation on base space (x, u), so two equations $\mathrm{E}(\phi)$, $\mathrm{E}(\phi')$ connected by such a transformation are in fact the *same equation*. Thus many differing arbitrary elements correspond to the *same* differential equation: in this case a 'differential equation' corresponds to an *equivalence class* of arbitrary elements ϕ. Our main interest is in how equations transform, rather than how arbitrary elements transform, so the quotient group $\hat{\mathcal{Q}}/\hat{\mathcal{M}}$ acting on these equivalence classes is of prime importance.

A familiar example arises in Hamilton's equations

$$\begin{aligned} \frac{dq^i}{dt} &= H_{p_i} \\ \frac{dp_i}{dt} &= -H_{q^i}, \end{aligned}$$

where the Hamiltonian $H(\mathbf{q}, \mathbf{p})$ is defined only to within an additive constant $H \mapsto H + \varepsilon$. Nevertheless, this trivial equivalence transformation will appear in the equivalence algebra.

Sometimes trivial equivalence transformations can be persuaded to disappear by redefining arbitrary elements. However, it is sometimes preferable to tolerate their presence. In Hamilton's equations, we could treat H_{q^i} and H_{p_i} as the arbitrary elements, but this is awkward, since it replaces a single arbitrary element with no auxiliary system by $2n$ arbitrary elements satisfying many compatibility conditions: this is a heavy price to pay to avoid some trifling indetermination in the Hamiltonian. In some cases trivial equivalences cannot be removed at all. For example, the class of linear partial differential equations

$$a^{ij}(x)u_{ij} + b^i(x)u_i + c(x)u = 0 \tag{3.55}$$

has trivial transformations

$$\begin{aligned} a^{ij\prime}(x) &= \lambda(x)\, a^{ij}(x) \\ b^{i\prime}(x) &= \lambda(x)\, b^i(x) \\ c'(x) &= \lambda(x)\, c(x) \end{aligned}$$

appearing in its augmented equivalence group: a linear p.d.e. determines its coefficients only to within a scaling. This cannot be avoided by redefining the coefficients a^{ij}, b^i, c.

Action on arbitrary elements

The interest in the equivalence group is to see how distinct equations can be transformed one to the other. In this context common symmetries are of no importance, since they do not affect the arbitrary elements at all. First we define the group $\hat{\mathcal{K}}$ to be the collection of equivalence transformations of the form (3.52), namely the common symmetries $\mathcal{K}$ augmented by appending a trivial action on a. We see that $\hat{\mathcal{K}}$ is a normal subgroup of $\hat{\mathcal{Q}}$.

Proposition 3.3.12. *The augmented equivalence group $\hat{\mathcal{Q}}$ homomorphically induces a group action on arbitrary elements ϕ. This group action is isomorphic to the quotient group $\hat{\mathcal{Q}}/\hat{\mathcal{K}}$.*

Proof. A transformation $\hat{\tau}$ from the augmented equivalence group acts on arbitrary elements $\phi(w)$ via (3.19). The group $\hat{\mathcal{Q}}$ of all such transformations thereby induces homomorphically a group action on functions $\phi(w)$. The kernel of the homomorphism consists of transformations $\hat{\tau}$ whose action on ϕ is trivial (i.e., such that $\phi' = \phi$ in (3.19)). But this is just the augmented common symmetry group $\hat{\mathcal{K}}$ by Proposition 3.3.9. □

For example, action (3.45) on diffusivity $D(u)$ in the potential diffusion system example has the structure of $GL_2(\mathbb{R})/\{I, -I\}$. This is indeed isomorphic to the quotient of $\hat{\mathcal{Q}}$ (3.44) over $\hat{\mathcal{K}}$ (3.53).

We frequently discard the group $\hat{\mathcal{K}}$ when considering an equivalence group $\hat{\mathcal{Q}}$. That is, we use the equivalence group 'modulo its common symmetries', so that we use a realization of $\hat{\mathcal{Q}}/\hat{\mathcal{K}}$. Interestingly, in the examples we consider, the equivalence algebra is a semidirect sum $\hat{K} \oplus_s \hat{R}$ for some subalgebra $\hat{R}$. For example, the equivalence algebra (Table 3.1) for the potential diffusion system (3.41) is a semidirect sum as marked. The basis must be carefully chosen to illustrate this feature: our choice for $\hat{\mathbf{X}}_6$ (3.46) was dictated by this consideration. The algebra $\hat{R} = \{\hat{\mathbf{X}}_5, \hat{\mathbf{X}}_6, \hat{\mathbf{X}}_7, \hat{\mathbf{X}}_8\}$ therefore generates the group of transformations which affect $D(u)$. We may speak of the equivalence group as being

$$\begin{aligned} v &= \rho^{-1}(\alpha v' + \beta x') \\ x &= \rho^{-1}(\gamma v' + \delta x') \\ t &= t' \\ u &= \frac{\alpha u' + \beta}{\gamma u' + \delta}, \qquad \alpha\delta - \beta\gamma = \pm 1 \\ a &= \rho^{-2}(\gamma u' + \delta)^2 a', \qquad \rho \neq 0, \end{aligned} \tag{3.56}$$

to within common symmetries. It is unclear whether there is always such a semidirect sum decomposition $\hat{L} = \hat{K} \oplus_s \hat{R}$ of the equivalence algebra.

Factoring out $\hat{\mathcal{K}}$ gives the action on arbitrary elements ϕ, but ultimately our interest is in how *distinct equations* in the class transform one to the other. That is, we wish to know how *equivalence classes* of arbitrary elements (with respect to $\hat{\mathcal{M}}$) transform. Hence we factor out both common symmetries *and* trivial equivalence transformations. In some sense we may include $\hat{\mathcal{M}}$ in the common symmetries: their base component is the identity transformation, so they map every equation in the class to itself. Factoring out both $\hat{\mathcal{K}}$ and $\hat{\mathcal{M}}$ is possible because they are commuting normal subgroups of the equivalence group $\hat{\mathcal{Q}}$. The group acting on distinct equations in the class $\mathcal{C}$ is isomorphic to the quotient group $\hat{\mathcal{Q}}/(\hat{\mathcal{M}}\hat{\mathcal{K}})$.

Structural features of a Lie algebra follow from the corresponding structure of the Lie group (§2.1.2). Hence the above discussion yields corresponding structural information about the Lie equivalence algebra $\hat{L}$. In particular, the common symmetry algebra and the trivial equivalence algebra are mutually commuting ideals in $\hat{L}$. We are basically interested in the quotient algebra which results when these are factored out.

3.4 Examples of equivalence groups

We now give some examples of nontrivial equivalence groups, which we use to illustrate some uses of equivalence transformations. The topics we touch on are: invariant solutions and their inherited equivalence group; potential equivalence transformations; and mapping nonlinear to linear p.d.e.'s.

3.4.1 Boltzmann's similarity solution for nonlinear diffusion

Derivation of similarity o.d.e.

In Example 3.3.5, we analyzed the potential system form

$$\begin{aligned} v_x &= u \\ v_t &= D(u) u_x \end{aligned} \tag{3.57}$$

of the nonlinear diffusion equation, and found an eight-parameter equivalence group (3.44). We concern ourselves with the scaling invariant similarity solution first discussed by Boltzmann: this is of physical importance, since the corresponding boundary conditions are easily realized in practice, and it is used as an approximation or asymptotic limit for many diffusion problems, as well as for measurement of the diffusivity $D(u)$ [21, 54].

We first give the necessary definitions.

Definition 3.4.1. Let E be a differential equation, admitting symmetry group $\mathcal{G}$. Let $\mathcal{H} \prec \mathcal{G}$ be a subgroup of $\mathcal{G}$. A solution $u = \theta(x)$ of E is *$\mathcal{H}$-invariant* if its graph $\Gamma_\theta = \{(x, u) \mid u = \theta(x)\}$ is invariant under the action of $\mathcal{H}$, that is,

$$h(\Gamma_\theta) = \Gamma_\theta, \qquad \forall h \in \mathcal{H}.$$

We write $\mathcal{H}(\Gamma_\theta) = \Gamma_\theta$.

The theory of invariant solutions is covered at length in [13, §4], [47, §3], [52, §19]; the concept is capable of generalization [8, 48].

Definition 3.4.2. An *invariant* of a group $\mathcal{H}$ acting on (x, u) is a function $F(x, u)$ such that

$$F \circ h\,(x, u) = F(x, u), \qquad \forall h \in \mathcal{H},$$

or more briefly, $F \circ \mathcal{H} = F$.

$\mathcal{H}$-invariant solutions satisfy a *reduced system* of d.e.'s (the '$\mathcal{H}$-reduced system' $\mathrm{E}/\mathcal{H}$) in a smaller number of variables, these being invariants of $\mathcal{H}$. There is a one-to-one correspondence between solutions of $\mathrm{E}/\mathcal{H}$ and $\mathcal{H}$-invariant solutions of E.

Boltzmann's similarity solution of (3.57) results from seeking solutions which are invariant under the scaling group $\mathcal{H}$ generated by the common symmetry operator

$$\hat{\mathbf{X}}_3 = v\,\partial_v + x\,\partial_x + 2t\,\partial_t.$$

Introduce invariants

$$\begin{aligned} z &= xt^{-1/2} \\ u & \\ y &= -\tfrac{1}{2}(v - xu)t^{-1/2} \end{aligned} \tag{3.58}$$

of $\mathcal{H}$: this gives the class of reduced o.d.e. systems

$$\begin{aligned} -2\frac{dy}{dz} + z\frac{du}{dz} &= 0 \\ y &= -D(u)\frac{du}{dz}. \end{aligned}$$

The variable y is related to the flux $q = -D(u)u_x$ by $y = qt^{1/2}$, which is why we prefer it over the obvious choice $vt^{-1/2}$. It is convenient to rewrite this using u as independent variable:

$$\begin{aligned} \frac{dy}{du} &= \frac{z}{2} \\ \frac{dz}{du} &= -\frac{D(u)}{y}. \end{aligned} \tag{3.59}$$

Equivalence group of similarity o.d.e.'s

We compute the equivalence group of the class of o.d.e.'s (3.59). Introducing the coordinate $a = D(u)$ and applying Proposition 3.3.7, we find the equivalence operator takes the form

$$\hat{\mathbf{Y}} = \xi(u)\,\partial_u + \zeta(u, z, y)\,\partial_z + \eta(u, z, y)\,\partial_y + \alpha(u, a)\,\partial_a.$$

Extending $\hat{\mathbf{Y}}$ by the method of §3.3.1, and enforcing invariance conditions (3.39, 3.40), gives determining equations for ξ, ζ, η, α, which may be completed by Reid's method [56], giving the involutive form

$$\begin{aligned}
\xi_z &= 0 & \eta_z &= 0 \\
\xi_y &= 0 & \eta_y &= \frac{1}{y}\,\eta \\
\xi_{uu} &= \frac{2}{y}\eta_u & \eta_{uu} &= 0
\end{aligned}$$
$$\zeta = 2\eta_u - z\xi_u + \frac{z}{y}\,\eta$$
$$\alpha = 2a(-\xi_u + \frac{1}{y}\,\eta).$$

This system is easily solved, giving an equivalence algebra generated by

$$\begin{array}{ll}
\hat{\mathbf{Y}}_1 = & \partial_u \\
\hat{\mathbf{Y}}_2 = & u\,\partial_u \quad - \tfrac{1}{2}z\,\partial_z + \tfrac{1}{2}y\,\partial_y - a\,\partial_a \\
\hat{\mathbf{Y}}_3 = & u^2\,\partial_u + (2y - uz)\,\partial_z + uy\,\partial_y - 2ua\,\partial_a \\
\hat{\mathbf{Y}}_4 = & z\,\partial_z + y\,\partial_y + 2a\,\partial_a.
\end{array} \tag{3.60}$$

This gives a four-parameter equivalence group $\hat{\mathcal{Q}}$:

$$\begin{aligned}
u &= \frac{\alpha u' + \beta}{\gamma u' + \delta} \\
z &= \rho^{-1}((\gamma u' + \delta)z' - 2\gamma y') \\
y &= \rho^{-1}\frac{y'}{\gamma u' + \delta}, \qquad \rho \neq 0 \\
a &= \rho^{-2}(\gamma u' + \delta)^2 a', \qquad \alpha\delta - \beta\gamma = \pm 1
\end{aligned} \tag{3.61}$$

relating (3.59) to the system with diffusivity

$$D'(u') = \frac{\rho^2}{(\gamma u' + \delta)^2}\; D'\Big(\frac{\alpha u' + \beta}{\gamma u' + \delta}\Big). \tag{3.62}$$

These transformations were noted by Lisle and Parlange [45]. Note that there are no common symmetry operators in the algebra (3.60): the only common symmetry in the equivalence group (3.61) is the discrete transformation $z \mapsto -z$, $y \mapsto -y$. Nevertheless, Proposition 3.3.12 still applies: the equivalence group $\hat{\mathcal{Q}}$ is a realization of the group $GL_2(\mathbb{R})$ of nonsingular 2×2 matrices, but in its action on $D(u)$, a matrix and its negative are identified, so that (3.62) has the structure $GL_2(\mathbb{R})/\{I, -I\}$.

An interesting point here is that we were able to explicitly construct the equivalence group for a system of first order o.d.e.'s. The corresponding problem of finding *symmetries* of a first order system of o.d.e.'s leads to a system of determining equations whose general solution requires solution of the original o.d.e.'s [13, §3.2.3], [52, §8]). Hence the symmetry problem is always indeterminate, but the equivalence group problem may be capable of solution.

Inherited equivalence

Note the resemblance between the equivalence group (3.61) for the similarity o.d.e. and the group (3.44) for the original p.d.e. system. It appears that the o.d.e.

inherits its equivalence group from the p.d.e. We give some general results in this direction. First we give a corresponding result about inherited symmetries [13, §7.2.7], [52, §20.4].

Definition 3.4.3. Let $\mathcal{H} \prec \mathcal{G}$ be a subgroup of a group $\mathcal{G}$. The *normalizer* of $\mathcal{H}$ in $\mathcal{G}$ is the group

$$N_{\mathcal{G}}(\mathcal{H}) = \{\tau \in \mathcal{G} \mid \tau \circ \mathcal{H} \circ \tau^{-1} = \mathcal{H}\}. \tag{3.63}$$

Theorem 3.4.4. *Let* E *be a system of d.e.'s admitting a point symmetry group* $\mathcal{G}$*. Let* $\mathcal{H} \prec \mathcal{G}$ *be a subgroup of* $\mathcal{G}$*. If* $\tau \in N_{\mathcal{G}}(\mathcal{H})$ *is a transformation in the normalizer of* $\mathcal{H}$*, then* τ *induces a point symmetry on the* $\mathcal{H}$*-reduced system* $\mathrm{E}/\mathcal{H}$*.*

Proof. The $\mathcal{H}$-reduced system is expressed in terms of invariants of $\mathcal{H}$. Let F be such an invariant. Action of τ on F gives a function $F' = F \circ \tau$. But, using Definition 3.4.2 of invariant,

$$F' \circ \mathcal{H} = F \circ \tau \circ \mathcal{H} = F \circ \mathcal{H} \circ \tau = F \circ \tau = F',$$

so F' is also an invariant of $\mathcal{H}$. Hence τ induces a mapping on the space of invariants of $\mathcal{H}$, i.e., on the variables in $\mathrm{E}/\mathcal{H}$.

Let Γ_θ be the graph of an $\mathcal{H}$-invariant solution $u = \theta(x)$ of E, so that (Definition 3.4.1) $\mathcal{H}(\Gamma_\theta) = \Gamma_\theta$. Let $\tau \in N_{\mathcal{G}}(\mathcal{H})$, so $\tau \circ \mathcal{H} = \mathcal{H} \circ \tau$. Let $\Gamma_{\theta'} = \tau(\Gamma_\theta)$. Then

$$\mathcal{H}(\Gamma_{\theta'}) = \mathcal{H} \circ \tau(\Gamma_\theta) = \tau \circ \mathcal{H}(\Gamma_\theta) = \tau(\Gamma_\theta) = \Gamma_{\theta'},$$

so θ' is an $\mathcal{H}$-invariant solution. Thus τ maps $\mathcal{H}$-invariant solutions of E to $\mathcal{H}$-invariant solutions. Since there is a one-to-one correspondence between $\mathcal{H}$-invariant solutions and solutions of $\mathrm{E}/\mathcal{H}$, τ induces a mapping from solutions of $\mathrm{E}/\mathcal{H}$ to solutions of $\mathrm{E}/\mathcal{H}$. From above this is a point transformation, and the Theorem is established. ☐

This result is given in a similar form by Ovsiannikov [52, Theorem 20.4]. In the above it is possible that the symmetry induced by τ on $\mathrm{E}/\mathcal{H}$ is trivial, i.e., is the identity map on the space of invariants of $\mathcal{H}$.

There is an infinitesimal version of the above result.

Definition 3.4.5. Let H be a Lie subalgebra of a Lie algebra L. The normalizer of H in L is the subalgebra

$$N_L(H) = \{\mathbf{X} \in L \mid [H, \mathbf{X}] \subseteq H\} \tag{3.64}$$

Theorem 3.4.6. *Let* E *be a system of d.e.'s with symmetry algebra* L*. Let* $H \subseteq L$ *be a subalgebra of* L*. If* $\mathbf{X} \in N_L(H)$ *is a transformation in the normalizer of* H*, then* $\mathbf{X}$ *induces a symmetry operator for* $\mathrm{E}/\mathcal{H}$*.*

We note that the theorem states that $\mathbf{X}$ may be written in terms of invariants of $\mathcal{H}$, but does not guarantee that the induced symmetry operator is nontrivial. Indeed it is quite possible for $\mathbf{X}$ to induce the zero operator (see below).

Bluman and Kumei [13, §7.2.7] used such a result to examine the inheritance of symmetries of the o.d.e. system (3.59) from the 'parent' p.d.e. system (3.57). Their version [13, Theorem 7.2.7-1] of Theorem 3.4.6 asserts that if E admits $\mathbf{X}$,

$\mathbf{Y}$ such that $[\mathbf{X},\mathbf{Y}] = \mu\mathbf{X}$ then $\mathbf{Y}$ induces a one-parameter symmetry group on the $\mathbf{X}$-reduced equation. The result is correct provided the word 'one-parameter' is deleted. For example, the equation $u_{xx} = u_x^2 u_t$ admits operators $\mathbf{X} = \partial_x$, $\mathbf{Y} = x\,\partial_x$ with the stipulated property. The invariants of $\mathbf{X}$ are t, u. It is true that $\mathbf{Y}$ acts on (t,u), but this action is trivial. Thus $\mathbf{Y}$ induces the identity transformation on (t,u), not a one-parameter group. (Their diffusion example is not affected by this observation.)

We now give corresponding results for inheritance of equivalence transformations of some class $\mathcal{C}$ of d.e.'s. We examine the group invariant solutions associated with a subgroup $\mathcal{H}$ of the common symmetries of $\mathcal{C}$. Reducing each equation $\mathrm{E}(\phi) \in \mathcal{C}$ gives a class $\mathcal{C}/\mathcal{H}$ of reduced systems $\mathrm{E}(\phi)/\mathcal{H}$. For example, reducing the diffusion p.d.e.'s (3.57) by the common symmetry subgroup generated by $\hat{\mathbf{X}}_3$ gives the class of reduced o.d.e.'s (3.59). In the following $\hat{\mathcal{K}}$, $\hat{\mathcal{H}}$ are essentially the same as $\mathcal{K}$, $\mathcal{H}$, since $\hat{\mathcal{K}}$ has trivial action on arbitrary element space a.

Theorem 3.4.7. *Let $\mathcal{C}$ be a class of equations with equivalence group $\hat{\mathcal{Q}}$ and common symmetries $\mathcal{K}$. Let $\mathcal{H} \prec \mathcal{K}$ be a subgroup of $\mathcal{K}$. Let $\hat{\tau} \in N_{\hat{\mathcal{Q}}}(\hat{\mathcal{H}})$ be an augmented equivalence transformation in the normalizer of $\hat{\mathcal{H}}$. Then $\hat{\tau}$ induces an equivalence transformation on the class of $\mathcal{H}$-reduced systems.*

Proof. By an identical argument to that used in the proof of Theorem 3.4.4, we find that $\hat{\tau} \in N_{\hat{\mathcal{Q}}}(\hat{\mathcal{H}})$ induces a transformation on the space of invariants of $\hat{\mathcal{H}}$. Moreover, since $\hat{\mathcal{H}}$ consists of common symmetries, its invariants are F, a, where F are the invariants of $\mathcal{H}$. As an equivalence transformation $\hat{\tau}$ is projectable to (x,u) space, so the transformation induced on invariants of $\hat{\mathcal{H}}$ is of the form

$$\begin{aligned} F' &= \psi(F) \\ a' &= \omega(F,a), \end{aligned}$$

that is, it is projectable to the space of variables F of $\mathrm{E}/\mathcal{H}$.

Let $u = \theta(x)$ be an $\mathcal{H}$-invariant solution of an equation $\mathrm{E}(\phi) \in \mathcal{C}$, and let $\Gamma_{\theta,\phi}$ be its augmented graph, so that $\hat{\mathcal{H}}(\Gamma_{\theta,\phi}) = \Gamma_{\theta,\phi}$. Let $\hat{\tau} \in N_{\hat{\mathcal{Q}}}(\mathcal{H})$, so that $\hat{\mathcal{H}} \circ \hat{\tau} = \hat{\tau} \circ \hat{\mathcal{H}}$. Suppose $\hat{\tau}$ maps solutions of $\mathrm{E}(\phi)$ to solutions of $\mathrm{E}(\phi')$; in particular, that it maps $u = \theta(x)$ solving $\mathrm{E}(\phi)$ to $u' = \theta'(x)$ solving $\mathrm{E}(\phi')$. Then

$$\hat{\mathcal{H}}(\Gamma_{\theta',\phi'}) = \hat{\mathcal{H}} \circ \hat{\tau}(\Gamma_{\theta,\phi}) = \hat{\tau} \circ \hat{\mathcal{H}}(\Gamma_{\theta,\phi}) = \hat{\tau}(\Gamma_{\theta,\phi}) = (\Gamma_{\theta',\phi'}),$$

so that $u' = \theta'(x')$ is an $\mathcal{H}$-invariant solution of $\mathrm{E}(\phi')$. Thus $\hat{\tau}$ maps $\mathcal{H}$-invariant solutions of $\mathrm{E}(\phi)$ to $\mathcal{H}$-invariant solutions of $\mathrm{E}(\phi')$, and hence induces a map from solutions of the reduced system $\mathrm{E}(\phi)/\mathcal{H}$ to solutions of the reduced system $\mathrm{E}(\phi')/\mathcal{H}$. From above, this is realized as an augmented transformation on the space of invariants of $\hat{\mathcal{H}}$ and is therefore an equivalence transformation of the class $\mathcal{C}/\mathcal{H}$ of reduced systems. □

Once again it is understood that the induced equivalence transformation can be trivial. Applying Theorem 3.4.4 to equations in $\mathcal{C}$ shows that the normalizer $N_{\mathcal{K}}(\mathcal{H})$ of $\mathcal{H}$ in $\mathcal{K}$ induces common symmetry transformations of the $\mathcal{H}$-reduced class $\mathcal{C}/\mathcal{H}$. Since $\hat{\mathcal{K}}$ is a normal subgroup of $\hat{\mathcal{Q}}$, the inherited common symmetries (from $N_{\hat{\mathcal{K}}}(\hat{\mathcal{H}})$) are a normal subgroup of the inherited equivalence transformations (from $N_{\hat{\mathcal{Q}}}(\hat{\mathcal{H}})$).

The infinitesimal form of Theorem 3.4.7 is:

Theorem 3.4.8. *Let $\mathcal{C}$ be a class of d.e.'s with equivalence algebra $\hat{Q}$ and common symmetry algebra K. Let $H \prec K$ be a subalgebra of common symmetries with associated symmetry group $\mathcal{H}$. Let $\mathbf{X} \in N_{\hat{Q}}(\hat{H})$ be an augmented equivalence operator in the normalizer of $\hat{H}$ in $\hat{Q}$. Then $\mathbf{X}$ induces an equivalence operator on the class $\mathcal{C}/\mathcal{H}$ of $\mathcal{H}$-reduced systems.*

Once again it is possible for the induced operator to be identically zero. The inherited common symmetry algebra is an ideal in the inherited equivalence algebra.

We now apply the above theory to the potential nonlinear diffusion example. From the commutator Table 3.1 we find the normalizer of $\{\hat{\mathbf{X}}_3\}$ is a five dimensional algebra with basis $\{\hat{\mathbf{X}}_3, \hat{\mathbf{X}}_5, \hat{\mathbf{X}}_6, \hat{\mathbf{X}}_7, \hat{\mathbf{X}}_8\}$. Writing these operators in terms of $\partial_u,\ \partial_z,\ \partial_y$ yields the four equivalence operators $\hat{\mathbf{Y}}$ (3.60) found above. Hence in this case the entire equivalence algebra is inherited from the 'parent' p.d.e. No common symmetry operators are inherited, but the discrete common symmetry $z \mapsto -z$, $y \mapsto -y$, is still inherited from the common symmetry group (3.53) for the p.d.e., and is indeed a normal subgroup of (3.61).

In Appendix C, we use symmetry and equivalence properties of the Boltzmann similarity o.d.e.'s (3.59) to assist in the solution of some boundary value problems for certain diffusivities.

Comparison with scalar diffusion

The scalar diffusion equation

$$u_t = (D(u)u_x)_x \tag{3.65}$$

admits only a six-parameter equivalence group, generated [52, §6.7] by

$$\begin{aligned}
\hat{\mathbf{X}}_1 &= \partial_x \\
\hat{\mathbf{X}}_2 &= \partial_t \\
\hat{\mathbf{X}}_3 &= x\,\partial_x + 2t\,\partial_t \\
\hat{\mathbf{X}}_5 &= \partial_u \\
\hat{\mathbf{X}}_6 &= -\tfrac{1}{2}x\,\partial_x + u\,\partial_u - a\,\partial_a \\
\hat{\mathbf{X}}_8 &= x\,\partial_x + t\,\partial_t + a\,\partial_a,
\end{aligned} \tag{3.66}$$

where the numbering is chosen to agree with the potential system case (3.46). These transformations consist of translations and scalings which are available by inspection. They establish the correspondence

$$D'(u') = \rho^2 D(\alpha u' + \beta). \tag{3.67}$$

The important point is that the scalar and potential system forms differ in the structure of their equivalence groups. Operator $\hat{\mathbf{X}}_7$ (3.46) cannot be expressed in terms of (x, t, u), and generates a nonlocal equivalence group of (3.65). These transformations were also found in [4, 3].

This behaviour is inherited by the Boltzmann similarity reduction, which for the scalar equation (3.65) gives

$$\frac{d}{dz}(-D(u)\frac{du}{dz}) = \tfrac{1}{2}z\frac{du}{dz}. \tag{3.68}$$

This o.d.e. has three-parameter equivalence group generated by

$$\begin{aligned}\hat{\mathbf{Y}}_1 &= \partial_u \\ \hat{\mathbf{Y}}_2 &= u\,\partial_u - \tfrac{1}{2}z\,\partial_z - a\,\partial_a \\ \hat{\mathbf{Y}}_4 &= z\,\partial_z + 2a\,\partial_a,\end{aligned} \tag{3.69}$$

all of which are inherited from (3.66). These simple scalings and translations are obvious by inspection. They act on $D(u)$ by (3.67). The operator $\hat{\mathbf{Y}}_3$ which was present in the equivalence algebra (3.60) of the o.d.e. system (3.59) is a nonlocal potential equivalence operator for the scalar o.d.e. (3.68). This situation is analogous to the situation for symmetry calculations [13, ch.7], [16].

3.4.2 Nonlinear diffusion-convection equations

Equivalence for scalar form

Consider the class of nonlinear diffusion convection equations

$$\begin{aligned} 0 &= u_t + q_x \\ q &= -D(u)u_x + K(u) \end{aligned} \tag{3.70}$$

governing the convection of a diffusing substance in one spatial dimension. Here u is concentration, q flux; the first equation expresses conservation of mass. Equation (3.70) governs for instance the flow of a liquid through a homogeneous porous medium [54], where the diffusive term $-D(u)u_x$ represents the effect of capillarity, and the convective term $K(u)$ the contribution of gravity. A particular equation is characterized by arbitrary functions $D(u)$ (diffusivity) and $K(u)$ (conductivity). Often q is eliminated from (3.70), giving the scalar equation

$$u_t = (D(u)u_x)_x - \dot{K}(u)u_x \tag{3.71}$$

where $\dot{K} \equiv \frac{dK}{du}$. Every point symmetry of system (3.70) is a contact symmetry of the scalar equation (3.71). It turns out that there are no genuine contact symmetries of (3.71), so the transformation properties of (3.71) and (3.70) are essentially identical. We analyze the system (3.70).

Introducing coordinates $a = D(u)$, $b = K(u)$ for arbitrary element space, class (3.70) has primary system

$$\begin{aligned} 0 &= u_t + q_x \\ q &= -au_x + b \end{aligned}$$

with a, b satisfying auxiliary system A:

$$\begin{aligned} a_x = a_t = a_q &= 0 \\ b_x = b_t = b_q &= 0. \end{aligned} \tag{3.72}$$

An augmented equivalence operator is sought in the form

$$\begin{aligned}\hat{\mathbf{X}} = \xi(x,t,u,q)\,\partial_x + \tau(x,t,u,q)\,\partial_t + \eta(u)\,\partial_u + \chi(x,t,u,q)\,\partial_q \\ + \alpha(u,a,b)\,\partial_a + \beta(u,a,b)\,\partial_b,\end{aligned}$$

where Proposition 3.3.7 has been used to simplify the forms of η, α, β. Enforcing conditions of Theorem 3.3.3 yields determining equations which are easily solved, giving an eight-dimensional basis of operators

$$\begin{array}{lcl}
\hat{\mathbf{X}}_1 & = & \partial_x \\
\hat{\mathbf{X}}_2 & = & \partial_t \\
\hat{\mathbf{X}}_3 & = & \partial_q + \partial_b \\
\hat{\mathbf{X}}_4 & = & t\,\partial_x + u\,\partial_q + u\,\partial_b \\
\hat{\mathbf{X}}_5 & = & \partial_u \\
\hat{\mathbf{X}}_6 & = & -\frac{1}{2}x\,\partial_x + u\,\partial_u + \frac{1}{2}q\,\partial_q + \frac{1}{2}b\,\partial_b - a\,\partial_a \\
\hat{\mathbf{X}}_7 & = & -x\,\partial_x - 2t\,\partial_t + q\,\partial_q + b\,\partial_b \\
\hat{\mathbf{X}}_8 & = & x\,\partial_x + t\,\partial_t + a\,\partial_a
\end{array} \tag{3.73}$$

Equivalence operators for the scalar equation (3.71) are obtained by dropping ∂_q components from (3.73). This algebra is a proper expansion of (3.66) for the scalar diffusion equation; the basis is chosen to reflect this, with only a renumbering of operators. Apart from scalings and translations, (3.73) includes the operator $\hat{\mathbf{X}}_4$ which represents a change to a uniformly moving coordinate frame,

$$\begin{array}{lll}
x \mapsto x + \nu t & u \mapsto u & b \mapsto b + \nu u \\
t \mapsto t & q \mapsto q + \nu u & a \mapsto a.
\end{array}$$

Altogether operators (3.73) generate a group $\hat{\mathcal{Q}}$

$$\begin{array}{rcl}
x & = & \rho\lambda\alpha^{-1/2}x' + \nu t' + \kappa_1 \\
t & = & \rho\lambda^2 t' + \kappa_2 \\
u & = & \alpha u' + \beta \\
q & = & \lambda^{-1}\alpha^{1/2}q' + \nu u' + \varepsilon \\
b & = & \lambda^{-1}\alpha^{1/2}b' + \nu u' + \varepsilon \\
a & = & \rho\alpha^{-1}a',
\end{array} \tag{3.74}$$

with the eight parameters κ_1, κ_2, ε, ν, α, β, λ, ρ, where $\alpha, \lambda, \rho > 0$. To these we may append three discrete equivalences

$$\begin{array}{lll}
R_1: \ x \mapsto -x & R_2: \ x \mapsto x & R_3: \ x \mapsto -x \\
\quad q \mapsto -q & \quad q \mapsto -q & \quad q \mapsto q \\
\quad u \mapsto u & \quad u \mapsto -u & \quad u \mapsto -u \\
\quad b \mapsto -b & \quad b \mapsto -b & \quad b \mapsto b,
\end{array} \tag{3.75}$$

so that the equivalence group consists of four disconnected sheets. Hence a diffusivity $D(u)$ and conductivity $K(u)$ are related by an equivalence transformation to any other $D'(u)$, $K'(u)$ of the form

$$\begin{array}{rcl}
K'(u) & = & \lambda K(\alpha u + \beta) + \nu u + \varepsilon \\
D'(u) & = & \rho D(\alpha u + \beta),
\end{array} \tag{3.76}$$

where the six parameters ε, ν, α, β, λ, ρ are not necessarily the same as in (3.74), and are subject to $\lambda, \alpha \neq 0$, $\rho > 0$. The common symmetries of (3.70) are the translations generated by $\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2$, represented in (3.74) by the parameters κ_1, κ_2.

[,]	$\hat{\mathbf{X}}_1$	$\hat{\mathbf{X}}_2$	$\hat{\mathbf{X}}_3$	$\hat{\mathbf{X}}_4$	$\hat{\mathbf{X}}_5$	$\hat{\mathbf{X}}_6$	$\hat{\mathbf{X}}_7$	$\hat{\mathbf{X}}_8$
$\hat{\mathbf{X}}_1$	0	0	0	0	0	$-\frac{1}{2}\hat{\mathbf{X}}_1$	$-\hat{\mathbf{X}}_1$	$\hat{\mathbf{X}}_1$
$\hat{\mathbf{X}}_2$	0	0	0	$\hat{\mathbf{X}}_1$	0	0	$-2\hat{\mathbf{X}}_2$	$\hat{\mathbf{X}}_2$
$\hat{\mathbf{X}}_3$	0	0	0	0	0	$\frac{1}{2}\hat{\mathbf{X}}_3$	$\hat{\mathbf{X}}_3$	0
$\hat{\mathbf{X}}_4$	0	$-\hat{\mathbf{X}}_1$	0	0	$-\hat{\mathbf{X}}_3$	$-\frac{1}{2}\hat{\mathbf{X}}_4$	$\hat{\mathbf{X}}_4$	0
$\hat{\mathbf{X}}_5$	0	0	0	$\hat{\mathbf{X}}_3$	0	$\hat{\mathbf{X}}_5$	0	0
$\hat{\mathbf{X}}_6$	$\frac{1}{2}\hat{\mathbf{X}}_1$	0	$-\frac{1}{2}\hat{\mathbf{X}}_3$	$\frac{1}{2}\hat{\mathbf{X}}_4$	$-\hat{\mathbf{X}}_5$	0	0	0
$\hat{\mathbf{X}}_7$	$\hat{\mathbf{X}}_1$	$2\hat{\mathbf{X}}_2$	$-\hat{\mathbf{X}}_3$	$-\hat{\mathbf{X}}_4$	0	0	0	0
$\hat{\mathbf{X}}_8$	$-\hat{\mathbf{X}}_1$	$-\hat{\mathbf{X}}_2$	0	0	0	0	0	0

Table 3.2: Commutator table for equivalence operators (3.73) of scalar diffusion convection equation (3.70). Bold outlines indicate the semidirect sum structure.

The structure of the equivalence algebra $\hat{L}$ is shown in the commutator Table 3.2. Note that the common symmetries $\{\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2\}$ are an ideal in $\hat{L}$. There are no trivial equivalences for the system form (3.70), but $\hat{\mathbf{X}}_3$ is trivial for the scalar form (3.71), and is a one-dimensional ideal in $\hat{L}$.

Although transformations (3.74) are useful, this result is disappointing since the Galileian transformation, scalings and translations which make up the equivalence group can be obtained by inspection.

Equivalence for potential form

The nonlinear diffusion convection equation (3.70) may also be written in various potential forms. Following the procedure of Bluman, Kumei and Reid [15], note that the first (continuity) equation of (3.70) is a divergence. Hence there exists a potential v such that

$$\begin{aligned} v_x &= u \\ v_t &= -q\,. \end{aligned} \tag{3.77}$$

The system (3.77), (3.70) for three dependent variables (u, q, v) is a potential system for the nonlinear diffusion convection equation. Eliminating q gives a system for u, v:

$$\begin{aligned} v_x &= u \\ v_t &= D(u)u_x - K(u) \end{aligned} \tag{3.78}$$

whose compatibility condition is the scalar equation (3.71). Elimination of u from this system (3.78) shows that v satisfies a scalar equation

$$v_t = D(v_x)v_{xx} - K(v_x). \tag{3.79}$$

Transformation properties of these three potential forms (3.70, 3.77), (3.78) and (3.79) are essentially identical. We calculate the equivalence group for the (u, v) potential system form (3.78).

Introducing coordinates $a = D(u)$, $b = K(u)$, the class (3.78) is specified by primary system

$$\begin{aligned} v_x &= u \\ v_t &= a u_x - b \end{aligned}$$

with auxiliary system

$$\begin{aligned} a_x &= a_t = a_v = 0 \\ b_x &= b_t = b_v = 0. \end{aligned}$$

Applying the method of §3.3.3, a system of determining equations is derived without difficulty.

The general solution of the determining equations is ten-dimensional. A basis for the equivalence operators is

$$\begin{aligned}
\hat{\mathbf{X}}_0 &= \partial_v \\
\hat{\mathbf{X}}_1 &= \partial_x \\
\hat{\mathbf{X}}_2 &= \partial_t \\
\hat{\mathbf{X}}_3 &= -t\,\partial_v + \partial_q + \partial_b \\
\hat{\mathbf{X}}_4 &= t\,\partial_x + u\,\partial_q + u\,\partial_b \\
\hat{\mathbf{X}}_5 &= x\,\partial_v + \partial_u \\
\hat{\mathbf{X}}_6 &= \tfrac{1}{2}v\,\partial_v - \tfrac{1}{2}x\,\partial_x + u\,\partial_u + \tfrac{1}{2}q\,\partial_q \quad \tfrac{1}{2}b\,\partial_b - a\,\partial_a \\
\hat{\mathbf{X}}_7 &= -v\,\partial_v - x\,\partial_x - 2t\,\partial_t + q\,\partial_q + b\,\partial_b \\
\hat{\mathbf{X}}_8 &= v\,\partial_v + x\,\partial_x + t\,\partial_t + a\,\partial_a \\
\hat{\mathbf{X}}_9 &= -v\,\partial_x + u^2\,\partial_u + uq\,\partial_q + ub\,\partial_b - 2ua\,\partial_a
\end{aligned} \tag{3.80}$$

Here we give action on not only (x, t, u, v, a, b) space, but also on the flux $q \equiv -v_t$ (3.77).

Operators $\hat{\mathbf{X}}_1, \ldots, \hat{\mathbf{X}}_8$ project onto corresponding operators (3.73) for the form (3.70) of the equation. However, operator $\hat{\mathbf{X}}_9$ is new. It generates the one-parameter group

$$\begin{aligned}
v &= v' & u &= \frac{u'}{1-\varepsilon u'} \\
x &= x' - \varepsilon v' & b &= \frac{b'}{1-\varepsilon u'} \\
t &= t' & a &= (1-\varepsilon u')^2 a' \\
q &= \frac{q'}{1-\varepsilon u'},
\end{aligned}$$

and maps a potential diffusion convection equation (3.78) to another such equation with

$$\begin{aligned}
D'(u') &= \frac{1}{(1-\varepsilon u')^2}\, D\Big(\frac{u'}{1-\varepsilon u'}\Big) \\
K'(u') &= (1-\varepsilon u')\, K\Big(\frac{u'}{1-\varepsilon u'}\Big).
\end{aligned}$$

It acts *nonlocally* on the base space (x, t, u, q) of (3.70), since transformation of x depends explicitly on v which is an *integral* $\int u\,dx$ (or $-\int q\,dt$) (3.77) of the local variables u, q . Following the terminology of Bluman, Kumei and Reid [15], we refer to such transformations as *potential equivalence transformations* (in [4], the terminology 'quasilocal' equivalence transformation is used.) The potential equivalence operator $\hat{\mathbf{X}}_9$ generalizes the transformations found by Akhatov, et al. [3, 4] to the case where a nonlinear convection term is present.

Altogether the operators $\hat{\mathbf{X}}_0, \ldots, \hat{\mathbf{X}}_9$ (3.80) generate a ten-parameter equivalence group $\hat{\mathcal{Q}}$

$$\begin{aligned}
v &= \frac{\lambda}{\rho}(\alpha v' + \beta x') + \theta t' + \kappa_0 \\
x &= \frac{\lambda}{\rho}(\gamma v' + \delta x') + \nu t' + \kappa_1 \\
t &= \frac{\lambda^2}{\rho} t' + \kappa_2 \\
u &= \frac{\alpha u' + \beta}{\gamma u' + \delta} \\
q &= \frac{\lambda q' + \rho(\alpha\nu - \gamma\theta)u' + \rho(\beta\nu - \delta\theta)}{\lambda^2(\gamma u' + \delta)} \\
b &= \frac{\lambda b' + \rho(\alpha\nu - \gamma\theta)u' + \rho(\beta\nu - \delta\theta)}{\lambda^2(\gamma u' + \delta)} \\
a &= \rho^{-1}(\gamma u' + \delta)^2 a'
\end{aligned} \tag{3.81}$$

with $\lambda, \rho > 0$ and $\alpha\delta - \beta\gamma = 1$. To this may be appended the discrete reflection equivalence R_2 (3.75) (with added component $v \mapsto -v$). The reflection R_1 (3.75) is connected to the identity in (3.81), while R_3 is connected to R_2. Hence the equivalence group of the potential system consists of *two* disconnected sheets. It is a realization of the matrix group

$$\begin{bmatrix} a & b & * & * \\ c & d & * & * \\ 0 & 0 & e & * \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{R}) \quad e > 0.$$

Action on diffusivity and conductivity functions is given by

$$\begin{aligned}
D'(u') &= \frac{\rho}{(\gamma u' + \delta)^2}\, D\Big(\frac{\alpha u' + \beta}{\gamma u' + \delta}\Big) \\
K'(u') &= \lambda(\gamma u' + \delta)\, K\Big(\frac{\alpha u' + \beta}{\gamma u' + \delta}\Big) + \mu u' + \varepsilon,
\end{aligned} \tag{3.82}$$

where

$$\begin{pmatrix} \mu \\ \varepsilon \end{pmatrix} := \frac{\rho}{\lambda} \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} -\nu \\ \theta \end{pmatrix} \tag{3.83}$$

The three parameters κ_i disappear here since these represent common translation symmetries. The common symmetry $\hat{\mathbf{X}}_0 = \partial_v$ which appears for the potential system is of trivial significance.

This result is apparently new. Note that the potential diffusion system (3.41) results by setting $K(u) \equiv 0$ in (3.78). The equivalence group (3.81) of the diffusion convection system (3.78) is therefore a proper generalization of the equivalence group (3.44) of the diffusion system, and the parametrization has been chosen to reflect this. Note that $\hat{\mathbf{X}}_3$ in (3.41) is a common symmetry for diffusion equations, but moves into the true equivalence group (operator $\hat{\mathbf{X}}_7$ in (3.80)) for (3.78)!

The equivalence group of the potential diffusion convection system is also a proper expansion of the equivalence group of the scalar diffusion convection equation (3.71). In fact, by adjoining the single hodograph-type transformation

$$\begin{aligned} x &= -v' & u &= 1/u' \\ v &= -x' & b &= b'/u' \\ t &= t' & a &= u'^2 a' \end{aligned} \tag{3.84}$$

to (3.74), we can obtain the equivalence group (3.81) of the potential system. Unfortunately, addition of this one transformation has the consequence that symmetry group classification is significantly more difficult in the potential form than in the scalar form of the diffusion convection system, a topic we discuss in §4.2, §4.3.

Mapping nonlinear to linear equations

One of the most important mapping problems for differential equations is to determine whether a nonlinear d.e. can be mapped to a linear d.e., and if so, to determine the mapping. Where a class of equations includes a linear equation, equivalence transformations acting on this equation can give rise to a nonlinear equation. Therefore the equivalence group detects and constructs certain linearizing mappings.

By setting $D(u) \equiv 1$, $K(u) \equiv 0$ in (3.70), the linear heat system

$$\begin{aligned} v_x &= u \\ v_t &= u_x \end{aligned} \tag{3.85}$$

results. Applying transformations (3.81) from the equivalence group of (3.78) to this heat system gives the equations linearizable by a change of variables of this type. We take the view that the scalings, translations etc. of (3.74) are obvious, and may freely be used to remove parameters. With this understood, the heat equation maps to the equation

$$\begin{aligned} v'_{x'} &= u' \\ v'_{t'} &= u'^{-2} u'_{x'} \end{aligned} \tag{3.86}$$

by the hodograph-type transformation (3.84). Linearization of equation (3.86) was discovered in a different form by Bluman and Kumei [10] (see also [66, 60, 15, 14]). Equivalence transformations (3.81) are a proper generalization of the Bluman-Kumei mapping to the case of arbitrary diffusivity and conductivity. Only for (3.86) does this mapping linearize the equation; for all other cases it maps a nonlinear equation to another nonlinear equation.

Another well known diffusion convection equation is Burgers' equation,

$$U_T = U_{XX} - 2UU_X, \tag{3.87}$$

which results from assigning $D(U) = 1$, $K(U) = U^2$. In potential system form, Burgers' equation is

$$\begin{aligned} V_X &= U \\ V_T &= U_X - U^2 \ . \end{aligned} \tag{3.88}$$

This is mapped to the linear heat system (3.85) by the Cole-Hopf transformation

$$\begin{aligned} v &= e^{-V} \qquad & x &= X \\ u &= -Ue^{-V} \qquad & t &= T \end{aligned} \ . \tag{3.89}$$

Because this transformation is not contained in the equivalence group (3.81), it is *not* detected by the present method as applied to class (3.78). In order to find linearizing transformations, the general method of Kumei and Bluman [41] must be applied: the equivalence group can give interesting results, but they are incomplete in nature.

Given that Burgers' equation can be mapped to the heat equation, the equivalence group (3.81) of the *potential* system allows Burgers' system to be mapped to the system of Fokas and Yortsos [24]

$$\begin{aligned} V'_{X'} &= U' \\ V'_{T'} &= \frac{1}{U'^2}U'_{X'} + \frac{1}{U'} \end{aligned} \tag{3.90}$$

by the hodograph-type transformation (3.84). By composing this with the Cole-Hopf transformation (3.89), we find that the nonlinear Fokas-Yortsos system (3.90) is mapped to the linear heat system (3.85) by

$$\begin{aligned} X' &= -\log v \qquad & V' &= x \\ T' &= t \qquad & U' &= -vu^{-1}. \end{aligned} \tag{3.91}$$

With the equivalence group of class (3.78) available, the result of Fokas and Yortsos [24] therefore is 'predictable': it results from the Cole-Hopf transformation combined with equivalence properties common to the whole class. The relationship between their case (3.90) and Burgers' system (3.88) is identical to that between the Bluman-Kumei system (3.86) and the heat system (3.85).

Note that all of the transformations (3.84, 3.89, 3.91) explicitly involve the v coordinate in the transformation of (x, t, u). Thus they are all *nonlocal* transformations of (x, t, u) space: they are not point transformations for the scalar form (3.71) of the diffusion convection equation. The discovery of these transformations in [10, 24] was by means of generalized (Lie-Bäcklund) symmetries of the *scalar* form of the equation. This gives the results only after difficult calculations, and necessitates an awkward statement of the transformations. For the potential system (3.78) the transformations take their most transparent form. The idea that analysis of a potential system can lead to significant nonlocal results for a scalar equation is due to Bluman, Kumei and Reid [15] (see also [11]). The potential system approach is much simpler, since one deals only with *point* transformations acting on a *different space* (e.g. transformations of (x, t, u, v) as opposed to (x, t, u)).

The linearization described above for the potential Fokas-Yortsos system (3.90) results by composing two nonlocal transformations—the hodograph-type transformation (3.84), and the Cole-Hopf transformation (3.89). Further composing this

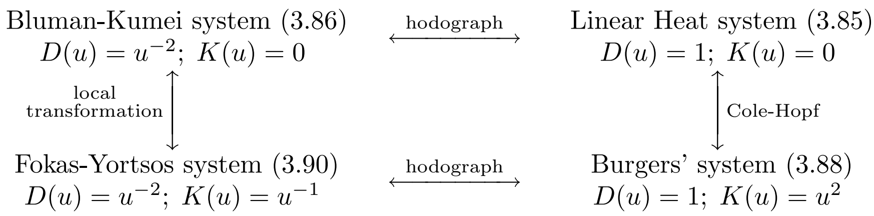

Figure 3.1: Relationship between linearizable diffusion convection potential systems. Only Fokas-Yortsos and Bluman-Kumei equations are connected by a local transformation.

linearizing map with transformation (3.84) maps the Fokas-Yortsos system (3.90) to the Bluman-Kumei system (3.86) (see Figure 3.1). This leads to the following

Proposition 3.4.9. *The scalar form*

$$U'_{T'} = (U'^{-2}U'_{X'} - U'^{-1})_{X'}$$

of the Fokas-Yortsos equation is point *equivalent to the scalar Bluman-Kumei equation*

$$u'_{t'} = (u'^{-2}u'_{x'})_{x'}$$

by the transformation

$$\begin{aligned} X' &= -\log x' \\ T' &= t' \\ U' &= -x'u'. \end{aligned} \tag{3.92}$$

This is remarkable since the transformations between any other pair drawn from the four systems (3.85, 3.86, 3.88, 3.90) are nonlocal, explicitly involving the v-coordinate in the transformation of (x, t, u). Proposition 3.4.9 reduces the linearization results of Fokas and Yortsos [24] to those of Bluman and Kumei [10]. The scalar linear heat equation admits infinitely many point symmetries; Burgers' equation admits five; and Bluman-Kumei and Fokas-Yortsos four each. Hence it is certain that Bluman-Kumei and Fokas-Yortsos are the only equations in Figure 3.1 connected by a local transformation on (x, t, u) space.

3.4.3 Wave equations

Consider the class of linear wave equations (3.22)

$$u_{tt} = c^2(x)u_{xx} \tag{3.93}$$

characterized by wavespeed function $c(x)$. This system was examined in Example 3.2.7; the equivalence group was described in Example 3.2.7. It is clear on physical grounds that (3.93) is equivalent to a wave equation with wavespeed

$$c'(x') = \gamma\, c(\alpha x' + \beta) \tag{3.94}$$

[,]	$\hat{\mathbf{X}}_1$	$\hat{\mathbf{X}}_2$	$\hat{\mathbf{X}}_3$	$\hat{\mathbf{X}}_4$	$\hat{\mathbf{X}}_5$	$\hat{\mathbf{X}}_6$	$\hat{\mathbf{X}}_7$	$\hat{\mathbf{X}}_8$	$\hat{\mathbf{X}}_9$	$\hat{\mathbf{X}}_{10}$
$\hat{\mathbf{X}}_1$	0	$-\hat{\mathbf{X}}_2$	$-\hat{\mathbf{X}}_3$	$-\hat{\mathbf{X}}_4$	$-\hat{\mathbf{X}}_5$	0	0	0	0	0
$\hat{\mathbf{X}}_2$	$\hat{\mathbf{X}}_2$	0	0	0	0	0	0	0	0	$\hat{\mathbf{X}}_3$
$\hat{\mathbf{X}}_3$	$\hat{\mathbf{X}}_3$	0	0	0	0	0	0	$-\hat{\mathbf{X}}_2$	$-\frac{1}{2}\hat{\mathbf{X}}_3$	0
$\hat{\mathbf{X}}_4$	$\hat{\mathbf{X}}_4$	0	0	0	0	$-\hat{\mathbf{X}}_2$	$-\hat{\mathbf{X}}_4$	0	$\frac{3}{2}\hat{\mathbf{X}}_4$	$\hat{\mathbf{X}}_5$
$\hat{\mathbf{X}}_5$	$\hat{\mathbf{X}}_5$	0	0	0	0	$-\hat{\mathbf{X}}_3$	$-\hat{\mathbf{X}}_5$	$-\hat{\mathbf{X}}_4$	$-\frac{1}{2}\hat{\mathbf{X}}_5$	0
$\hat{\mathbf{X}}_6$	0	0	0	$\hat{\mathbf{X}}_2$	$\hat{\mathbf{X}}_3$	0	$\hat{\mathbf{X}}_6$	0	0	0
$\hat{\mathbf{X}}_7$	0	0	0	$\hat{\mathbf{X}}_4$	$\hat{\mathbf{X}}_5$	$-\hat{\mathbf{X}}_6$	0	0	0	0
$\hat{\mathbf{X}}_8$	0	0	$\hat{\mathbf{X}}_2$	0	$\hat{\mathbf{X}}_4$	0	0	0	$\hat{\mathbf{X}}_8$	$2\hat{\mathbf{X}}_9$
$\hat{\mathbf{X}}_9$	0	$\frac{1}{2}\hat{\mathbf{X}}_2$	$\frac{1}{2}\hat{\mathbf{X}}_3$	$-\frac{3}{2}\hat{\mathbf{X}}_4$	$\frac{1}{2}\hat{\mathbf{X}}_5$	0	0	$-\hat{\mathbf{X}}_8$	0	$\hat{\mathbf{X}}_{10}$
$\hat{\mathbf{X}}_{10}$	0	$-\hat{\mathbf{X}}_3$	0	$-\hat{\mathbf{X}}_5$	0	0	0	$-2\hat{\mathbf{X}}_9$	$-\hat{\mathbf{X}}_{10}$	0

Table 3.3: Commutation relations of equivalence algebra (3.95) of the scalar wave equation (3.93). The algebra is a semidirect sum of the subalgebras of dimensions 6 and 4 shown.

with α, β, γ arbitrary constants satisfying $\alpha, \gamma \neq 0$.

Execution of Algorithm 3.3.6 on the class (3.93) (that is, (3.24, 3.23)), gives equivalence operators

$$\begin{aligned}
\hat{\mathbf{X}}_1 &= u\,\partial_u \\
\hat{\mathbf{X}}_2 &= \partial_u \\
\hat{\mathbf{X}}_3 &= x\,\partial_u \\
\hat{\mathbf{X}}_4 &= t\,\partial_u \\
\hat{\mathbf{X}}_5 &= xt\,\partial_u \\
\hat{\mathbf{X}}_6 &= \partial_t \\
\hat{\mathbf{X}}_7 &= t\,\partial_t - a\,\partial_a \\
\hat{\mathbf{X}}_8 &= \partial_x \\
\hat{\mathbf{X}}_9 &= x\,\partial_x + \tfrac{1}{2}u\,\partial_u + a\,\partial_a \\
\hat{\mathbf{X}}_{10} &= x^2\,\partial_x + xu\,\partial_u + 2xa\,\partial_a.
\end{aligned} \tag{3.95}$$

Commutation relations for this algebra are shown in Table 3.3.

The algebra is a semidirect sum $\hat{L} = \hat{K} \oplus_s \hat{R}$. The common symmetry algebra $\hat{K}$ is spanned by $\hat{\mathbf{X}}_1, \ldots, \hat{\mathbf{X}}_6$; it generates the six-parameter common symmetry group

$$\begin{aligned}
x &= x' \\
t &= t' + \kappa \\
u &= \lambda u' + \nu_0 + \nu_1 x' + \nu_2 t' + \nu_3 x' t' \\
a &= a', \qquad \lambda \neq 0
\end{aligned} \tag{3.96}$$

to which can be appended the discrete time-reversal symmetry $t \mapsto -t$. The parameters ν_i represent superposition symmetries. The functions 1, x, t, xt are solutions common to every wave equation and can be added to any solution of *any* wave equation, yielding another solution of the same equation. The parameter λ gives the scaling associated with any linear equation, while κ represents invariance under time translation.

The complement $\hat{R} = \{\hat{\mathbf{X}}_7, \hat{\mathbf{X}}_8, \hat{\mathbf{X}}_9, \hat{\mathbf{X}}_{10}\}$ of $\hat{K}$ generates the 'true' equivalence transformations

$$\begin{aligned} x &= \frac{\gamma_1 x' + \gamma_2}{\gamma_3 x' + \gamma_4} \\ t &= \rho t' + \kappa \\ u &= \frac{u'}{\gamma_3 x' + \gamma_4}, \qquad \gamma_1\gamma_4 - \gamma_2\gamma_3 = \pm 1 \\ a &= \frac{1}{\rho}\frac{\pm a'}{(\gamma_3 x' + \gamma_4)^2}, \qquad \rho > 0 \end{aligned} \tag{3.97}$$

with four independent parameters γ_i, ρ. Groups (3.96) and (3.97) generate the whole equivalence group (3.25) of the scalar wave equation. Transformation (3.97) maps the wave equation (3.93) with wavespeed c to another such equation with wavespeed

$$c'(x') = \rho\,(\gamma_3 x' + \gamma_4)^2\; c\Big(\frac{\gamma_1 x' + \gamma_2}{\gamma_3 x' + \gamma_4}\Big). \tag{3.98}$$

Three of the parameters here are associated with the obvious equivalences (3.94). However the interesting case is $\gamma_3 \neq 0$, which leads to projective transformations not available by inspection, and apparently not previously known. Action (3.98) on the wavespeed c reflects the factoring out of the six common symmetries (3.96). Disregarding the indetermination of sign of c, the structure of this action is $GL_2(\mathbb{R})/\{I, -I\}$ (compare with (3.62)).

As noted in Example 3.1.2, a wave equation (3.93) may also be written in the very general potential form (3.14):

$$\begin{aligned} v_x &= c^{-2}(x)\,[\,h(x,t)\,u_t - h_t(x,t)\,u\,] \\ v_t &= h(x,t)\,u_x - h_x(x,t)\,u \end{aligned} \tag{3.99}$$

where $h(x,t)$ is any nonzero solution of the scalar wave equation (3.93). This class is specified by systems (3.16), (3.17). Execution of Algorithm 3.3.6 gives equivalence operators

$$\begin{aligned} \hat{\mathbf{X}}_0 &= \partial_v \\ \hat{\mathbf{X}}_1 &= u\,\partial_u + v\,\partial_v \\ \hat{\mathbf{X}}_6 &= \partial_t \\ \hat{\mathbf{X}}_7 &= t\,\partial_t + v\,\partial_v - a\,\partial_a \\ \hat{\mathbf{X}}_8 &= \partial_x \\ \hat{\mathbf{X}}_9 &= x\,\partial_x + \tfrac{1}{2}u\,\partial_u + a\,\partial_a + \tfrac{1}{2}b\,\partial_b \\ \hat{\mathbf{X}}_{10} &= x^2\,\partial_x + xu\,\partial_u + 2xa\,\partial_a + xb\,\partial_b \\ \hat{\mathbf{X}}_{11} &= v\,\partial_v + b\,\partial_b \end{aligned} \tag{3.100}$$

The operators $\hat{\mathbf{X}}_1$ and $\hat{\mathbf{X}}_6$, … , $\hat{\mathbf{X}}_{10}$ correspond to operators in the Lie algebra (3.95) of the scalar equation. The new operators $\hat{\mathbf{X}}_0$, $\hat{\mathbf{X}}_{11}$ have trivial action on

(x, t, u) and are of no significance. Hence no potential equivalence transformations arise in this case. Altogether the operators (3.100) generate an eight-parameter group. Once again the Lie algebra is a semidirect sum of common symmetries $\hat{K} = \{\hat{\mathbf{X}}_0, \hat{\mathbf{X}}_1\}$ and their complement $\hat{R} = \{\hat{\mathbf{X}}_6, \ldots, \hat{\mathbf{X}}_{11}\}$. The common symmetries represent a scaling due to linearity of system (3.99); and the fact that v is a potential variable, and only determined to within a constant. Note that the operator $\hat{\mathbf{X}}_6$, representing time translation, moves from the common symmetry group for the scalar form (3.93) to the true equivalences for the system (3.99). The common superposition symmetries which were present for the scalar equation disappear completely.

The group of 'true' equivalences generated by $\hat{R}$ is

$$\begin{aligned} x &= \frac{\gamma_1 x' + \gamma_2}{\gamma_3 x' + \gamma_4} \\ t &= \rho t' + \kappa \\ u &= \frac{u'}{(\gamma_3 x' + \gamma_4)} \\ v &= \frac{1}{\rho\mu} v' \\ a &= \frac{1}{\rho} \frac{\pm a'}{(\gamma_3 x' + \gamma_4)^2} \qquad \rho, \mu \neq 0 \\ b &= \frac{1}{\rho\mu} \frac{b'}{\gamma_3 x' + \gamma_4}, \qquad \gamma_1\gamma_4 - \gamma_2\gamma_3 = \pm 1. \end{aligned} \tag{3.101}$$

The effect of these transformations on the arbitrary elements c, h is to map them to new elements

$$\begin{aligned} c'(x') &= \rho(\gamma_3 x' + \gamma_4)^2 \; c\Big(\frac{\gamma_1 x' + \gamma_2}{\gamma_3 x' + \gamma_4}\Big) \\ h'(x', t') &= \rho\mu(\gamma_3 x' + \gamma_4) \; h\Big(\frac{\gamma_1 x' + \gamma_2}{\gamma_3 x' + \gamma_4}, \; \rho t' + \kappa\Big). \end{aligned} \tag{3.102}$$

The availability of transformations (3.101) has some interesting consequences. There is particular interest in the potential forms (3.99) where the function $h(x, t)$ is

$$h(x, t) = \nu_0 + \nu_1 x + \nu_2 t + \nu_3 x t\,. \tag{3.103}$$

An h of this form is a solution of *every* scalar wave equation (3.93), and therefore the corresponding potential form (3.99) is valid for every wavespeed. Use of obvious scaling and translation equivalences shows that any h of the form (3.103) can be reduced to the six canonical cases

$$\begin{aligned} h(x,t) = \;& 1, \quad t, \quad x + t, \\ & x, \quad xt, \quad xt + 1\,. \end{aligned} \tag{3.104}$$

However, application of the transformation

$$x = \frac{1}{x'} \qquad u = \frac{u'}{x'} \tag{3.105}$$

maps each member in the second row of (3.104) to the corresponding member in the first row. Hence the canonical list can be shortened to

$$h(x,t) = 1,\ t\ , x+t\ .$$

Thus knowing the symmetry groups of the potential system (3.99) for the above three cases allows the symmetries for every case (3.103) to be recovered. In particular, the new group classification found by Ma [46] for (3.99) in the case $h(x,t) = x$ can be obtained by transformation (3.105) from the group classification found by Bluman and Kumei [11] for the case $h(x,t) = 1$. In [11], more-or-less explicit formulas for the wavespeeds $c(x)$ admitting a nontrivial symmetry group when $h(x,t) = 1$ were found. Their information on qualitative behaviour of $c(x)$ may be transformed through (3.105) to give qualitative behaviour of wavespeeds admitting a nontrivial symmetry group for $h(x,t) = x$. For instance, the wavespeed considered in [12] is bounded by two constants c_1, c_2: $0 < c_1 < c(x) < c_2$. Application of transformation (3.105) (see also (3.102)) takes any such wavespeed to one which vanishes at $x = 0$ and is unbounded at $x' \to \pm\infty$. Thus the physically interesting behaviour of $c(x)$ is lost in the transformation process.

3.4.4 Hamilton's equations

We apply the equivalence group Algorithm 3.3.6 to Hamilton's equations in the $2n$ dependent variables $\mathbf{q} = (q^1, q^2, \ldots, q^n)$ ('coordinates') and $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ ('momenta'):

$$\begin{aligned} \frac{dq^i}{dt} &= \frac{\partial H}{\partial p_i}(\mathbf{q}, \mathbf{p}, t) \\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial q^i}(\mathbf{q}, \mathbf{p}, t) \end{aligned} \tag{3.106}$$

characterized by the arbitrary element $H(\mathbf{q}, \mathbf{p}, t)$ (the 'Hamiltonian').

The 'canonical formalism' of classical mechanics [28, ch.9], [5, ch.9] is concerned with finding transformations of $(\mathbf{q}, \mathbf{p})$ space ('phase space') which leave the form of Hamilton's equations invariant. Our goal here is to show how our equivalence group construction for the class (3.106) leads to these canonical transformations.

Introducing a coordinate $h = H(\mathbf{q}, \mathbf{p}, t)$ for 'Hamiltonian space', and coordinates h_{q^i}, h_{p_i} for derivatives of H, the class (3.106) is written

$$\begin{aligned} \dot{q}^i &= h_{p_i} \\ \dot{p}_i &= -h_{q^i}. \end{aligned}$$

where the dot represents $\frac{d}{dt}$. The function H is arbitrary, so there is no auxiliary system. The equivalence group operator is sought in the form

$$\begin{aligned} \hat{\underset{1\,1}{\mathbf{X}}} &= \tau \frac{\partial}{\partial t} + \kappa^i \frac{\partial}{\partial q^i} + \pi_i \frac{\partial}{\partial p_i} + \kappa^i_{(t)} \frac{\partial}{\partial \dot{q}^i} + \pi_{i(t)} \frac{\partial}{\partial \dot{p}_i} \\ &\quad + \gamma \frac{\partial}{\partial h} + \gamma_{(q^i)} \frac{\partial}{\partial h_{q^i}} + \gamma_{(p_i)} \frac{\partial}{\partial h_{p_i}} \end{aligned}$$

with τ, κ^i, π_i functions of $(\mathbf{q}, \mathbf{p}, t)$, and $\gamma(\mathbf{q}, \mathbf{p}, t, h)$. Enforcing invariance conditions (Theorem 3.3.3) yields the determining equations

$$\begin{array}{ll} \dfrac{\partial \tau}{\partial q^i} = 0 & \dfrac{\partial \tau}{\partial p_i} = 0 \\ \dfrac{\partial \kappa^i}{\partial p_j} - \dfrac{\partial \kappa^j}{\partial p_i} = 0 & \dfrac{\partial \kappa^i}{\partial t} - \dfrac{\partial \gamma}{\partial p_i} = 0 \\ \dfrac{\partial \pi_i}{\partial q^j} - \dfrac{\partial \pi_j}{\partial q^i} = 0 & \dfrac{\partial \pi_i}{\partial t} + \dfrac{\partial \gamma}{\partial q^i} = 0 \\ \multicolumn{2}{c}{\dfrac{\partial \kappa^i}{\partial q^j} + \dfrac{\partial \pi_j}{\partial p_i} = \delta^i_j(\dfrac{\partial \tau}{\partial t} + \dfrac{\partial \gamma}{\partial h})} \end{array} \tag{3.107}$$

along with the dependency conditions

$$\frac{\partial \tau}{\partial h} = \frac{\partial \kappa^i}{\partial h} = \frac{\partial \pi_i}{\partial h} = 0 \tag{3.108}$$

Computing compatibility conditions of (3.107) shows that

$$\frac{\partial \tau}{\partial t} + \frac{\partial \gamma}{\partial h} = k \tag{3.109}$$

where k is a constant.

The determining equations can be solved by writing them as the integrability conditions of certain equations. This integration can be concisely stated in terms of differential forms as follows. Define the differential one form ϕ on augmented space $(\mathbf{q}, \mathbf{p}, t, h)$ by

$$\phi = \pi_i \, dq^i - \kappa^i \, dp_i + \tau \, dh - \gamma \, dt \, .$$

Let θ be the one form

$$\theta = p_i \, dq^i - h \, dt. \tag{3.110}$$

In terms of these forms, determining equations (3.107, 3.109) are just

$$d\phi = k \, d\theta \, . \tag{3.111}$$

The general solution of this equation, after taking account of dependency conditions (3.108), is

$$\phi = k\theta - d\Big(W(\mathbf{q}, \mathbf{p}, t) + \tau(t)h\Big) \tag{3.112}$$

where W, τ are arbitrary functions of their arguments. Thus the general equivalence transformation of Hamilton's equations is characterized by one arbitrary constant k, one arbitrary function τ of one variable, and one arbitrary function W of $2n + 1$ variables.

Writing the solution (3.112) componentwise, we see that the equivalence group of Hamilton's equations (3.106) is generated by operators

$$\Lambda = p_i \partial_{p_i} + h \partial_h \tag{3.113a}$$

$$T(\tau) = \tau \, \partial_t - \tau_t h \partial_h \tag{3.113b}$$

$$\Omega(W) = -W_{p_i} \partial_{q^i} + W_{q^i} \partial_{p_i} - W_t \partial_h \tag{3.113c}$$

The scaling operator Λ is associated with arbitrary choice of units for h. The operators $T(\tau)$ permit arbitrary variation in the time metric, which induces a corresponding variation in the Hamiltonian. These operators are usually ignored in mechanics texts, since their significance is trivial. The important generators are $\Omega(W)$, which represent *infinitesimal canonical transformations* [28]. The commutator of two such transformations is

$$[\,\Omega(W_1), \Omega(W_2)\,] = \Omega(\{W_1, W_2\})$$

where $\{\ ,\ \}$ is the Poisson bracket:

$$\{W_1, W_2\} = \frac{\partial W_1}{\partial q^i}\frac{\partial W_2}{\partial p_i} - \frac{\partial W_1}{\partial p_i}\frac{\partial W_2}{\partial q^i}\ . \tag{3.114}$$

Among equivalence operators (3.113c) are some trivial equivalences, generated by setting W to be a function of t only: $\hat{\mathbf{X}} = W_t(t)\partial_h$. These reflect the fact that addition of any function of t to the Hamiltonian H

$$H'(\mathbf{q}, \mathbf{p}, t) = H(\mathbf{q}, \mathbf{p}, t) + E(t)$$

does not affect Hamilton's equations. It is understood that 'a Hamiltonian' actually refers to an equivalence class of Hamiltonians connected by such transformations.

In older classical mechanics books (e.g. [28, 42]), the derivation of canonical transformations is essentially that given here, except that a fixed time parametrization $\tau(t) = 0$ is usually enforced *a priori*. Sometimes the scalings Λ are also suppressed.

More abstract treatments of classical mechanics take a geometric viewpoint, with canonical transformations being defined as transformations of phase space which leave invariant a 'symplectic two form' ω [5]. Our derivation of the determining equations (3.111) can be rephrased as follows. Let $\omega = d\theta$ (with θ given by (3.110)), so that

$$\omega = dp_i \wedge dq^i - dh \wedge dt.$$

The condition that a transformation map a Hamiltonian system (3.106) to another such system is equivalent to demanding that ω be transformed to a multiple of itself. In infinitesimal form, this requires that

$$\mathcal{L}_{\hat{\mathbf{X}}}\omega = \alpha\omega \tag{3.115}$$

where α is a scalar function and $\mathcal{L}_{\hat{\mathbf{X}}}$ denotes Lie derivative with respect to $\hat{\mathbf{X}}$. In fact, taking exterior derivative d shows that α is a constant k; condition (3.115) can be written

$$d(\mathcal{L}_{\hat{\mathbf{X}}}\theta) = k\, d\theta$$

which is exactly (3.111).

Our derivation of canonical transformations differs from the usual one mostly in notation and terminology. The important point is that our method is part of an overall theoretical and algorithmic machinery applicable to very general classes of equations. The example of Hamilton's equations is simplified by the fact that the auxiliary system is null. When a nontrivial auxiliary system is present, the machinery of §3.2, §3.3 becomes essential.

Chapter 4

Symmetry Group Classification

4.1 Symmetry classification problem

The symmetry group classification problem for a class $\mathcal{C}$ of differential equations is to find and construct the symmetry group of each equation in $\mathcal{C}$. One attempts to find conditions on the arbitrary elements so that symmetries are present. One approach to this problem is to derive and attempt to solve the determining equations for the infinitesimals of the symmetry operators. In the course of this process one hopes to find 'classifying conditions', that is, conditions on the arbitrary elements of $\mathcal{C}$ which split the calculation into two branches, depending whether the arbitrary elements take this or that form. Consider the standard example of the scalar nonlinear diffusion equation (3.65). Ovsiannikov [52, §6.7] obtains a determining equation of the form

$$(D/\dot{D})^{\cdot\cdot}\,(2\xi_x - \eta_t) = 0.$$

The next step depends on whether $(D/\dot{D})^{\cdot\cdot} = 0$ or not, so this condition is classifying. Reid [55, 57] showed that classification can be performed *without* solving the determining equations. His method algorithmically finds classification conditions by appending compatibility conditions to the determining system.

However classified, the determining equations must ultimately be solved to find the symmetry operators. This results in a list of functional forms for the arbitrary elements, each with its associated symmetry operators. For the scalar diffusion equation (3.65), if $D(u) = (au + b)^m$, $(a \neq 0)$ then the equation admits a symmetry

$$a(m+1)x\,\partial_x + a(m+2)t\,\partial_t + (au+b)\,\partial_u\ .$$

The parameters a, b here may 'without loss of generality' be set to $a = 1$, $b = 0$, and this kind of parameter elimination is customary in presenting results of symmetry classification. However, even the 'w.l.o.g.' parameter removal above relies on knowledge *extrinsic* to the symmetry classification procedure, namely availability of equivalence transformations (3.74). A method such as Reid's, based on analysis of determining equations will *find* symmetries, but it cannot specify which equations are related by a change of variables, and hence parameter elimination is not achieved. Nevertheless, use of equivalence transformations to remove parameters from classifying conditions is an integral (albeit implicit) part of the symmetry group classification process. The parameter reduction effected using whatever equivalence transformations are available by inspection is *ad hoc*, although it may suffice for simple examples. A more complete and systematic parameter elimination is possible using the full equivalence group calculated by the method of §3.3.3.

4.1.1 Example: scalar diffusion convection

For later reference we now give a symmetry group classification for the scalar nonlinear diffusion convection equation (3.70):

$$u_t = (D(u)u_x - K(u))_x \, . \tag{4.1}$$

This will permit a comparison with the potential system form (3.78), for which classification is more difficult. Calculations for the scalar form can be completed by hand, and although lengthy, *ad hoc* methods suffice to sort out all the cases: the method is not reproduced here. A symmetry classification for (4.1) is shown in Table 4.1, where cases which are distinct under equivalence transformations (3.74) are shown. The general forms of $D(u)$, $K(u)$ and their associated symmetry operators may be obtained by use of the equivalence group (3.74). For instance Case **2a** below is $D(u) = u^m$, $K(u) = u^n$, $n \neq 0, 1$, representing the general family

$$\begin{aligned} D'(u') &= (\alpha u' + \beta)^m \\ K'(u') &= \lambda(\alpha u' + \beta)^n + \nu u' + \varepsilon, \qquad \alpha, \lambda \neq 0. \end{aligned} \tag{4.2}$$

The symmetry $\mathbf{Y}_3$ for this family becomes

$$\mathbf{Y}_3 = \alpha((m-n+1)x' - \nu(n-1)t')\,\partial_{x'} + \alpha(m-2n+2)t'\,\partial_{t'} + (\alpha u' + \beta)\,\partial_{u'} \, .$$

If $K(u) = \nu u + \varepsilon$ for some ν, ε then the convection term can be removed, and we are left (Case **1.**) with the nonlinear diffusion equation (3.65). Symmetry classification for this case is well known [52, §6.7]. Burgers' equation $D(u) = 1$, $K(u) = u^2$ (Case **2a***i*) also has well known symmetry properties [13, p.266]. The remainder of the symmetry classification is adapted from Lisle [44] (with corrections). In [49], Oron and Rosenau claim to give a classification for the scalar diffusion convection equation, but the results in their Table 3(a) are seriously in error. In particular Cases **2a***i*, **2a***ii*, **2c**, and **2e** of our Table 4.1 are not detected at all, and Cases **2a** and **2d** are only partially detected (they impose spurious restrictions $n = m + 1$ and $m = -1$ respectively). Case 5 of their Table 1(a) of symmetries of the scalar diffusion equation also appears to be spurious.

4.2 Partial symmetry classification

4.2.1 Symmetry inherited from equivalence group

Reid's method [55, 57] in principle solves the symmetry group classification problem for a class of differential equations, but the calculations involved are lengthy even when a computer algebra package is used. Instead of attempting a full point symmetry group classification, we may seek those point symmetries belonging to

1. $K = 0$ (diffusion) $\quad \mathbf{Y}_3 = v\,\partial_v + x\,\partial_x + 2t\,\partial_t$

a.	$D(u) = 1$	$\mathbf{Y}_4^\dagger = -2t\,\partial_x + xu\,\partial_u$ $\mathbf{Y}_5^\dagger = -4xt\,\partial_x - 4t^2\,\partial_t + (x^2+2t)u\,\partial_u$ $\mathbf{Y}_6 = u\,\partial_u$ $\mathbf{Y}_\infty^\dagger = \theta(x,t)\,\partial_u$ $u = \theta(x,t)$ any solution of $u_{xx} = u_t$
b.	$D(u) = u^m, m \neq 0$	$\mathbf{Y}_4 = mx\,\partial_x + 2u\,\partial_u$
b*i*.	$m = -\frac{4}{3}$	$\mathbf{Y}_5^\dagger = x^2\,\partial_x - 3xu\,\partial_u$
c.	$D(u) = e^u$	$\mathbf{Y}_4 = x\,\partial_x + 2\,\partial_u$

2. $K \not\equiv 0$ (nonlinear convection)

a.	$D(u) = u^m$, $K(u) = u^n,\ n \neq 0, 1$	$\mathbf{Y}_3 = (m-n+1)x\,\partial_x + (m-2n+2)t\,\partial_t + u\,\partial_u$
a*i*.	$m = 0$ $n = 2$	$\mathbf{Y}_4 = 2t\,\partial_x + \partial_u$ $\mathbf{Y}_5^\dagger = 2xt\,\partial_x + 2t^2\,\partial_t + (x - 2tu)\,\partial_u$
a*ii*.	$m = -2$ $n = -1$	$\mathbf{Y}_4^\dagger = e^{-x}\,\partial_x + e^{-x}u\,\partial_u$
b.	$D(u) = u^m$ $K(u) = \log u$	$\mathbf{Y}_3 = (m+1)x\,\partial_x + (m+2)t\,\partial_t + u\,\partial_u$
c.	$D(u) = u^m$ $K(u) = u\log u$	$\mathbf{Y}_3 = (mx+t)\,\partial_x + mt\,\partial_t + u\,\partial_u$
d.	$D(u) = e^{mu}$ $K(u) = e^u$,	$\mathbf{Y}_3 = (m-1)x\,\partial_x + (m-2)t\,\partial_t + \partial_u$
e.	$D(u) = e^u$ $K(u) = u^2$	$\mathbf{Y}_3 = (x+2t)\,\partial_x + t\,\partial_t + \partial_u$

Table 4.1: Full symmetry classification for scalar form of nonlinear diffusion convection equation (4.1). Operators shown are in addition to common symmetries $\mathbf{Y}_1 = \partial_x$, $\mathbf{Y}_2 = \partial_t$. All symmetry operators are inherited from the equivalence algebra (3.73) except those marked with a dagger $\mathbf{Y}_i^\dagger$.

the equivalence group of the class. This gives only a partial symmetry classification for the class, but is often much easier to obtain than the full classification. The results we state are easy adaptations of methods for group invariant solutions (§3.4.1).

We make the following observations. The equivalence group $\hat{\mathcal{Q}}$ is a group of symmetries for the auxiliary system A satisfied by the arbitrary elements. Let $\hat{\mathcal{H}} \prec \hat{\mathcal{Q}}$ be some subgroup of $\hat{\mathcal{Q}}$, and let $a = \phi(w)$ be an $\mathcal{H}$-invariant solution of A, that is,

$$\hat{\mathcal{H}}(\Gamma_\phi) = \Gamma_\phi,$$

where $\Gamma_\phi = \{(w, a) \mid a = \phi(w)\}$ is the graph of ϕ. Thus $\mathcal{H}$ maps every solution $u = \theta(x)$ of $\mathrm{E}(\phi)$ to a solution of the *same* equation $\mathrm{E}(\phi)$, and is therefore a *symmetry* group of $\mathrm{E}(\phi)$. Therefore finding and classifying all subgroups $\hat{\mathcal{H}} \prec \hat{\mathcal{Q}}$ of the equivalence group of $\mathcal{C}$ leads to a classification of all $\hat{\mathcal{H}}$-invariant solutions of A, and hence to a classification of symmetry groups of equations $\mathrm{E}(\phi) \in \mathcal{C}$. This symmetry classification is *partial*, since the symmetries found will all lie in $\mathcal{Q}$; there may be additional symmetries outside $\mathcal{Q}$. All of the machinery described in [13, §4], [47, §3], [52, §19] for classification of invariant solutions may now be brought to bear, and no new theory is required.

This simple insight appears to be new. Recently it was discovered independently by Akhatov, Gazizov and Ibragimov [4], who used these ideas to assist in several symmetry classification calculations. Subsequently Ibragimov, Torrisi and Valenti [32] executed the method on a more difficult example. Note that common symmetries $\mathcal{K}$ have a special, trivial role here: *every* solution $a = \phi(w)$ of A is $\hat{\mathcal{K}}$-invariant. Hence we can neglect $\hat{\mathcal{K}}$ and concern ourselves only with classifying subgroups of $\hat{\mathcal{Q}}/\hat{\mathcal{K}}$.

We use an infinitesimal formulation of the above.

Proposition 4.2.1. *Let* $\hat{\mathbf{X}}$ *be an equivalence operator for a class* $\mathcal{C}$ *of differential equations:*

$$\hat{\mathbf{X}} = \zeta^\gamma(w)\partial_{w^\gamma} + \alpha^\beta(w, a)\partial_{a^\beta}. \tag{4.3}$$

Let $\mathrm{E}(\phi) \in \mathcal{C}$ *be a differential equation in* $\mathcal{C}$*, and suppose* $\hat{\mathbf{X}}$ *is such that*

$$\alpha^\beta(w, \phi(w)) = \zeta^\gamma(w)\frac{\partial \phi^\beta}{\partial w^\gamma}(w), \qquad \beta = 1, 2, \ldots, \mu. \tag{4.4}$$

Then $\mathbf{X}$ *is a symmetry operator for* $\mathrm{E}(\phi)$*.*

Proof. Property (4.4) asserts that the vector field $\hat{\mathbf{X}}$ is everywhere tangent to the surface $a = \phi(w)$, i.e., that $\hat{\mathbf{X}}(a - \phi(w)) = 0$ whenever $a = \phi(w)$. By Theorem 2.1.13, the one-parameter group $\hat{\mathcal{H}}$ associated with $\hat{\mathbf{X}}$ leaves invariant the surface $a = \phi(w)$. From the notes above, $\mathcal{H}$ consists of symmetries of $\mathrm{E}(\phi)$. Hence the operator $\mathbf{X}$ associated with $\mathcal{H}$ is an infinitesimal symmetry of $\mathrm{E}(\phi)$. □

For each equivalence operator $\hat{\mathbf{X}}$, infinitesimal condition (4.4) is a system of first order differential equations

$$\zeta^\gamma(w)a^\beta_\gamma = \alpha^\beta(w, a) \tag{4.5}$$

for the arbitrary elements $a = \phi(w)$, with ζ^γ and α^β being known functions. For a differential equation $\mathrm{E}(\phi)$ to admit $\mathbf{X}$, the function ϕ must be a solution of both the auxiliary system and equations (4.5). Note that condition (4.4) is quite distinct from requiring $a' = a$: it is not the *coordinate* a which must be preserved, but the *function* $a = \phi(w)$.

Example 4.2.2. Consider the scalar diffusion convection equation (4.1) with diffusivity $D(u)$ and conductivity $K(u)$. The most general operator from the augmented equivalence group is a linear combination $\sum_{i=1}^{8} c_i \hat{\mathbf{X}}_i$ of operators (3.73). Condition (4.4) that the equation $a = D(u)$, $b = K(u)$ admit this operator is

$$\begin{aligned} (c_6 u + c_5)\dot{D}(u) &= (c_8 - c_6)D(u) \\ (c_6 u + c_5)\dot{K}(u) &= (c_7 + \tfrac{1}{2}c_6)K(u) + (c_4 u + c_3). \end{aligned} \tag{4.6}$$

The system (4.5) is

$$\begin{aligned} (c_6 u + c_5)\dot{a} &= (c_8 - c_6)a \\ (c_6 u + c_5)\dot{b} &= (c_7 + \tfrac{1}{2}c_6)b + (c_4 u + c_3); \end{aligned} \tag{4.7}$$

if $a = D(u)$, $b = K(u)$ are solutions of this, the diffusion convection equation (4.1) with diffusivity $D(u)$ and conductivity $K(u)$ admits the symmetry $\mathbf{X} = \sum_{i=1}^{8} c_i \mathbf{X}_i$. The constants c_1, c_2 associated with common translation symmetries $\mathbf{Y}_1 = \partial_x$, $\mathbf{Y}_2 = \partial_t$ may be assigned arbitrarily. The solution of o.d.e.'s (4.7) is easily found for the various values of $c_3, \ldots, c_8$. The result is a collection of functions $D(u)$ and $K(u)$ each with their associated symmetries (of equivalence type). Equivalence transformations (3.74) can then be used to remove parameters $c_3, \ldots, c_8$ from $D(u)$, $K(u)$.

The completeness of neither the resulting collection of $D(u), K(u)$ nor their associated symmetries can be guaranteed, since only those symmetries *inherited* from the equivalence algebra are found in this way. What is remarkable in this case is how much of the symmetry classification of Table 4.1 is recovered. *Every* functional form of $D(u)$, $K(u)$ with symmetry beyond the common translations is found, although subcases **1b***i* ($m = -4/3$ diffusion) and **2a***ii* (Fokas-Yortsos) are not singled out as exceptional. Not only are the forms of $D(u)$, $K(u)$ found, but almost all the symmetries for these $D(u)$, $K(u)$ are found in this way. Partial classification for this example yields Table 4.1 apart from the cases marked with daggers:

- Linear heat equation Case **1a.**, operators $\mathbf{Y}_4$, $\mathbf{Y}_5$, $\mathbf{Y}_\infty$.
- Burgers' equation Case **2a***i*, operator $\mathbf{Y}_5$.
- Fokas-Yortsos equation Case **2a***ii*, operator $\mathbf{Y}_4$.
- $u^{-4/3}$ diffusion equation Case **1b***i*., operator $\mathbf{Y}_5$.

Note that most of the classification of the scalar diffusion equation [52, §6.7] is inherited from the simple equivalence group (3.66). This *partial* symmetry classification detects several symmetries which were missed by Oron and Rosenau [49] in their alleged classification of the scalar diffusion convection equation. Thus even when it is feasible to calculate a complete symmetry classification, partial results obtained from the equivalence group can offer a valuable check.

Equivalence transformations (3.74) all reflect physical properties of diffusion convection processes (rescaling of units, Galileian invariance etc.). The symmetries found by the partial classification procedure are hence predictable on physical grounds. Only the daggered operators in Table 4.1 appear 'out of the blue'. In the usual method for symmetry classification one has to wait until the end of a long calculation to find even very simple symmetries, and there is no ready criterion for distinguishing 'predictable' symmetries from 'exceptional' ones.

The usual method for symmetry classification is 'analytic': for a given class of equations, we seek symmetries by analyzing determining equations. The partial classification is *synthetic*, that is, given a group operator we *construct* the equations which admit it as a symmetry. Determining equations for symmetries are never formed. It is this which permits partial classification results to be obtained with small computational labour, at least for finite-parameter equivalence groups. The process is the same as finding invariant solutions of a d.e. There we suppose that a symmetry group $\mathcal{G}$ of the d.e. is known, and seek solutions $u = \theta(x)$ which are invariant under the action of a given subgroup of $\mathcal{G}$. Constructing group invariant solutions begins with a *known* group of (symmetry) transformations, just as partial classification begins with a known group of (equivalence) transformations.

4.2.2 Optimal system of subalgebras

When executing a partial classification one obtains d.e.'s (such as (4.7)) for the arbitrary elements. Typically these d.e.'s contain many parameters—one for each operator from the equivalence algebra. These d.e.'s could be integrated with parameters in place; the actual result depends on the parameter values, giving a classification of the arbitrary elements. Following integration, equivalence group action can be used to remove parameters, giving a short list of the essentially different cases. Instead, we describe how to use the equivalence group action to simplify the d.e.'s *before* integrating them. The method is based on the following considerations.

Let $\hat{\mathcal{Q}}$ be the augmented equivalence group of a class of d.e.'s and let $\hat{\mathcal{H}} \subseteq \hat{\mathcal{Q}}$ be a subgroup. Let ϕ be an $\hat{\mathcal{H}}$-invariant solution of the auxiliary system A (so $\hat{\mathcal{H}}(\Gamma_\phi) = \Gamma_\phi$ where Γ_ϕ is the graph $a = \phi(w)$). Thus $\mathrm{E}(\phi)$ admits $\mathcal{H}$ as symmetries. Let $\hat{\tau} \in \hat{\mathcal{Q}}$ be an equivalence transformation mapping $\mathrm{E}(\phi)$ to $\mathrm{E}(\phi') \in \mathcal{C}$ Then ϕ' is an invariant solution of A with respect to the conjugate subgroup $\hat{\tau} \circ \hat{\mathcal{H}} \circ \hat{\tau}^{-1}$, and hence $\mathrm{E}(\phi')$ admits $\tau \circ \mathcal{H} \circ \tau^{-1}$ as symmetries. Hence we need only consider reduction of A with respect to subgroups $\hat{\mathcal{H}} \prec \hat{\mathcal{Q}}$ which are distinct under conjugation by $\hat{\tau}$ This is the usual process of classification of group invariant solutions [47, §3.3], [52, §20.5].

An infinitesimal form of this result is given in terms of the action of 'conjugation by $\hat{\tau}$' on equivalence operators. Let $\mathcal{G}$ be a Lie group, with associated Lie algebra L. Let $\mathbf{X} \in L$ be a group operator, generating a one-parameter group $\mathcal{H}$ of transformations $\sigma(\varepsilon)$. The one-parameter subgroup obtained by conjugation by some $\tau \in \mathcal{G}$ is $\mathcal{H}'_\tau = \tau \circ \mathcal{H} \circ \tau^{-1}$. It consists of transformations $\sigma'(\varepsilon)$, where $\sigma'(\varepsilon) = \tau \circ \sigma(\varepsilon) \circ \tau^{-1}$. The group operator $\mathbf{X}'_\tau$ associated with $\mathcal{H}'_\tau$ is found by differentiation as $\mathbf{X}'_\tau = \frac{d}{d\varepsilon}\sigma'(\varepsilon)\big|_{\varepsilon=0}$. The linear mapping from $\mathbf{X}$ to $\mathbf{X}'_\tau$ is denoted by $\mathrm{Ad}\,\tau$. The map $\tau \mapsto \mathrm{Ad}\,\tau$ is a homomorphism of $\mathcal{G}$ onto a group of linear transformations of the Lie algebra L. The group $\mathrm{Ad}\,\mathcal{G}$ of matrices is called the *adjoint group* of $\mathcal{G}$. It gives the action of a transformation group on its Lie algebra.

Calculation and use of the adjoint group is described in [47, §3.3], [52, §14].

Proposition 4.2.3. *Let $\hat{H} \subseteq \hat{L}$ be a subalgebra of the equivalence algebra $\hat{L}$ for a class $\mathcal{C}$ of d.e.'s with equivalence group $\hat{\mathcal{Q}}$. Let ϕ be an $\hat{H}$-invariant solution of the auxiliary system* A *of $\mathcal{C}$, so that* $\mathrm{E}(\phi) \in \mathcal{C}$ *admits H as a symmetry algebra. Let $\hat{\tau} \in \hat{\mathcal{Q}}$ be an equivalence transformation, mapping* $\mathrm{E}(\phi)$ *to* $\mathrm{E}(\phi') \in \mathcal{C}$*. Then* $\mathrm{E}(\phi')$ *admits symmetry algebra* $\mathrm{Ad}\,\tau(H) \subseteq L$*.*

Proof. Let $\hat{H}$ have associated group $\hat{\mathcal{H}} \prec \hat{\mathcal{Q}}$. From the comments above, $\mathrm{E}(\phi')$ admits the symmetry group $\tau \circ \mathcal{H} \circ \tau^{-1}$. The algebra associated with this group is $\mathrm{Ad}\,\tau(H)$. □

Hence we need only find $\hat{H}$-invariant solutions of A with respect to subalgebras of $\hat{L}$, distinct under the adjoint action of $\hat{\mathcal{Q}}$.

The adjoint action of the equivalence group on the equivalence algebra can be used to construct an *optimal system* [52, §14], [47, §3.3] of one dimensional subalgebras. An optimal system is a collection of equivalence operators which are *essentially different* (no two operators in the optimal system are connected by an equivalence transformation) and *complete* (every operator in the equivalence algebra is equivalent to an operator in the optimal system). Each operator in the optimal system gives rise to a classifying d.e. (4.5) for the arbitrary elements. No two such classifying d.e.'s are connected by an equivalence transformation. Also every possible classifying d.e. is connected to a classifying d.e. associated with an operator in the optimal system. Hence the optimal system of one dimensional subalgebras gives a minimally short list of classifying d.e.'s, which are then integrated one by one.

Integration of (4.5) gives rise to additional parameters—the constants of integration. These constants of integration may be removable using the action of any remaining equivalence transformations. Let $\hat{\mathcal{H}} \prec \hat{\mathcal{Q}}$ be a subgroup of the equivalence group $\hat{\mathcal{Q}}$ of a class $\mathcal{C}$ of d.e.'s with auxiliary system A, and let $\mathcal{C}_{\hat{\mathcal{H}}} \subseteq \mathcal{C}$ be the *subclass* of d.e.'s

$$\mathcal{C}_{\hat{\mathcal{H}}} = \{\mathrm{E}(\phi) \in \mathcal{C} \mid \phi \text{ is an } \hat{\mathcal{H}}\text{-invariant solution of A}\}.$$

Thus $\mathcal{C}_{\hat{\mathcal{H}}}$ is the set of equations in $\mathcal{C}$ admitting $\mathcal{H}$ as a symmetry group.

Proposition 4.2.4. *The subclass $\mathcal{C}_{\hat{\mathcal{H}}}$ of d.e.'s inherits the normalizer*

$$N_{\hat{\mathcal{Q}}}(\hat{\mathcal{H}}) = \{\hat{\tau} \in \hat{\mathcal{Q}} \mid \hat{\tau} \circ \hat{\mathcal{H}} \hat{\tau}^{-1} = \hat{\mathcal{H}}\}$$

of $\hat{\mathcal{H}}$ in its equivalence group. In particular, $\mathcal{C}_{\hat{\mathcal{H}}}$ has common symmetries $\hat{\mathcal{K}}$ and $\hat{\mathcal{H}}$.

Proof. From Theorem 3.4.4, $N_{\hat{\mathcal{Q}}}(\hat{\mathcal{H}})$ maps $\hat{\mathcal{H}}$-invariant solutions of A to $\hat{\mathcal{H}}$-invariant solutions of A. Hence $N_{\hat{\mathcal{Q}}}(\hat{\mathcal{H}})$ consists of transformations mapping solutions of $\mathrm{E}(\phi) \in \mathcal{C}_{\hat{\mathcal{H}}}$ to solutions of $\mathrm{E}(\phi') \in \mathcal{C}_{\hat{\mathcal{H}}}$. Hence $N_{\hat{\mathcal{Q}}}(\hat{\mathcal{H}})$ consists of equivalence transformations of $\mathcal{C}_{\hat{\mathcal{H}}}$. The statement about common symmetries is trivial. □

This result is related to, but distinct from, Theorem 3.4.7, which was concerned with inheritance of equivalence transformations when the d.e.'s $\mathrm{E}(\phi) \in \mathcal{C}$ were reduced with respect to the action of a subgroup of the common symmetries $\hat{\mathcal{K}}$. In Proposition 4.2.4, no group reduction of $\mathrm{E}(\phi)$ is being effected: instead, by group reduction of the auxiliary system A, one picks out a subclass of equations which share some symmetry group $\mathcal{H}$.

4.2.3 Partial symmetry classification for nonlinear diffusion convection

We now perform a partial classification of the potential system form (3.78)

$$\begin{aligned} v_x &= u \\ v_t &= D(u)u_x - K(u) \end{aligned} \tag{4.8}$$

of the nonlinear diffusion convection equation. The equivalence algebra $\hat{L}$ (3.80) for this class is ten-dimensional, with seven operators having a nontrivial action on the diffusivity and conductivity functions $D(u), K(u)$. To write condition (4.5) it is convenient to project these operators to action on (u, a, b) space:

$$\begin{aligned}
\tilde{\mathbf{X}}_3 &= \partial_b \\
\tilde{\mathbf{X}}_4 &= u\,\partial_b \\
\tilde{\mathbf{X}}_5 &= \partial_u \\
\tilde{\mathbf{X}}_6 &= u\,\partial_u + \tfrac{1}{2} b\,\partial_b - a\,\partial_a \\
\tilde{\mathbf{X}}_9 &= u^2\,\partial_u + ub\,\partial_b - 2ua\,\partial_a \\
\tilde{\mathbf{X}}_7 &= b\,\partial_b \\
\tilde{\mathbf{X}}_8 &= a\,\partial_a
\end{aligned} \tag{4.9}$$

This neatly removes the common symmetries $\hat{\mathbf{X}}_0$, $\hat{\mathbf{X}}_1$, $\hat{\mathbf{X}}_2$. We denote this projected Lie algebra by $\tilde{L}$. The operators $\tilde{\mathbf{X}}_3, \ldots, \tilde{\mathbf{X}}_9$ generate the Lie group (see 3.82)

$$\begin{aligned}
u &= \frac{\alpha u' + \beta}{\gamma u' + \delta} \\
b &= \lambda \frac{b' - \mu u' - \varepsilon}{\gamma u' + \delta} \\
a &= \rho(\gamma u' + \delta)^2\, a'
\end{aligned} \tag{4.10}$$

where $\rho > 0$, $\lambda \neq 0$, $\alpha\delta - \beta\gamma = 1$. This projected group action is denoted by $\tilde{\mathcal{Q}}$ (compare (3.82)).

Applying a general linear combination

$$\tilde{\mathbf{Y}} = \sum_{i=3}^{9} c_i \tilde{\mathbf{X}}_i \tag{4.11}$$

of equivalence operators (4.9) to the equations $a = D(u)$, $b = K(u)$ yields a classifying system (4.5) in the form

$$\begin{aligned}
(c_9 u^2 + c_6 u + c_5)\dot{a} &= (-2c_9 u + c_8 - c_6)a \\
(c_9 u^2 + c_6 u + c_5)\dot{b} &= (c_9 u + \tfrac{1}{2}c_6 + c_7)b + (c_4 u + c_3).
\end{aligned} \tag{4.12}$$

A solution $a = D(u)$, $b = K(u)$ of this classifying system of d.e.'s gives a diffusion convection system admitting the symmetry $\mathbf{Y} = \sum_{i=3}^{9} c_i \mathbf{X}_i$, in addition to the common translation symmetries

$$\mathbf{Y}_0 = \partial_v, \qquad \mathbf{Y}_1 = \partial_x, \qquad \mathbf{Y}_2 = \partial_t\,. \tag{4.13}$$

Instead of a frontal assault on (4.12), we use Proposition 4.2.3 to remove parameters c_i before integrating. Direct substitution of the group (4.10) into $\tilde{\mathbf{Y}}$ (4.11) gives the adjoint action on the constants c_i specifying $\tilde{\mathbf{Y}}$. Suppose that $\tilde{\mathbf{Y}} = \sum c_i \tilde{\mathbf{X}}_i$ maps to $\tilde{\mathbf{Y}}' = \sum c_i' \tilde{\mathbf{X}}_i'$, where in an obvious notation, $\tilde{\mathbf{X}}_4' = u' \frac{\partial}{\partial b'}$ and so forth. The adjoint action is a linear map, with matrix form

$$
\begin{bmatrix} c_8 \\ \hline c_7 \\ \hline c_9 \\ c_6 \\ c_5 \\ \hline c_4 \\ c_3 \end{bmatrix}' = \left[\begin{array}{c|c|ccc|cc} 1 & & & & & & \\ \hline & 1 & & & & & \\ \hline & & & & & & \\ & & & C & & & \\ & & & & & & \\ \hline & \mu & & B & & A & \\ & \varepsilon & & & & & \end{array}\right] \begin{bmatrix} c_8 \\ \hline c_7 \\ \hline c_9 \\ c_6 \\ c_5 \\ \hline c_4 \\ c_3 \end{bmatrix} \tag{4.14}
$$

with submatrices A, B, C

$$
\begin{aligned}
A &= \lambda \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \qquad \alpha\delta - \beta\gamma = \pm 1,\ \lambda \neq 0 \\
B &= \begin{pmatrix} \alpha(\beta\mu - \alpha\varepsilon) & \alpha(\delta\mu - \gamma\varepsilon) - \frac{1}{2}\mu & \gamma(\delta\mu - \gamma\varepsilon) \\ \beta(\beta\mu - \alpha\varepsilon) & \beta(\delta\mu - \gamma\varepsilon) - \frac{1}{2}\varepsilon & \delta(\delta\mu - \gamma\varepsilon) \end{pmatrix} \\
C &= \begin{pmatrix} \alpha^2 & \alpha\gamma & \gamma^2 \\ 2\alpha\beta & \alpha\delta + \beta\gamma & 2\gamma\delta \\ \beta^2 & \beta\delta & \delta^2 \end{pmatrix}
\end{aligned} \tag{4.15}
$$

The adjoint action (4.14) is used to simplify the operator $\tilde{\mathbf{Y}}$. For this it is helpful to know the three invariants

$$
J = c_4^2 - 4c_3c_9, \qquad c_7, \qquad c_8 \tag{4.16}
$$

of the adjoint group action. For example, knowing that J is invariant under adjoint actions, it is immediately possible to assert that diffusivities $D(u)$ satisfying equation (4.12) are sorted into three distinct classes, namely $J < 0$, $J = 0$ and $J > 0$. The invariants may be obtained by the methods described in [52, §17.4]. First the Lie algebra $\tilde{L}$ is broken into a direct sum of the centre $\{\tilde{\mathbf{X}}_8\}$ and its complement $\tilde{L}^6 = \{\tilde{\mathbf{X}}_3, \tilde{\mathbf{X}}_4, \tilde{\mathbf{X}}_5, \tilde{\mathbf{X}}_6, \tilde{\mathbf{X}}_9, \tilde{\mathbf{X}}_7\}$. The coefficient c_8 of $\tilde{\mathbf{X}}_8$ in this decomposition is an invariant of the adjoint action. The invariant c_7 is obvious by inspection, while J can be found from the Killing polynomial of $\tilde{L}^6$. Finally an optimal system of one-parameter subalgebras of $\tilde{L}$ is constructed: such a system is shown in Table 4.2. For each subalgebra we also give its normalizer.

With the optimal system of one dimensional subalgebras known, it remains to integrate classifying equations (4.12) for each operator in the optimal system. We illustrate the procedure for Case **1** from Table 4.2, for which $c_6 = 1$, $c_7 = n - \frac{1}{2}$, $c_8 = m + 1$, and $n \neq 0, 1$. Classifying system (4.12) becomes

$$
u\dot{a} = m\,a, \qquad u\dot{b} = n\,b,
$$

which is easily integrated, yielding

$$
a = D(u) = D_0\,|u|^m, \qquad b = K(u) = K_0 |u|^n. \tag{4.17}
$$

We require the diffusivity $D(u)$ to be positive, so $D_0 > 0$; the constant K_0 is unrestricted. By Proposition 4.2.4, this subclass (4.17) of diffusion convection

Case	Operator	Normalizer	Diffusivity $D(u)$	Conductivity $K(u)$
***1.**	$\mathbf{X}_6 + (n - \frac{1}{2})\mathbf{X}_7 + (m+1)\mathbf{X}_8,\ n \neq 0, 1$	$\mathbf{X}_7$	$\lvert u\rvert^m$	$K_0\lvert u\rvert^n$
2.	$\mathbf{X}_4 + \mathbf{X}_6 + \frac{1}{2}\mathbf{X}_7 + (m+1)\mathbf{X}_8$	$\mathbf{X}_4$	$\lvert u\rvert^m$	$u \log \lvert u\rvert + K_0 u$
3.	$\mathbf{X}_5 + \mathbf{X}_7 + m\mathbf{X}_8$	$\mathbf{X}_7$	e^{mu}	$K_0 e^u$
4.	$2\mathbf{X}_4 + \mathbf{X}_5 + \mathbf{X}_8$	$\mathbf{X}_3$	e^u	$u^2 + K_0$
5.	$\mathbf{X}_5 + \mathbf{X}_8$	$\mathbf{X}_3, \mathbf{X}_7$	e^u	K_0
6.	$2\mathbf{X}_4 + \mathbf{X}_5$	$\mathbf{X}_3, \mathbf{X}_6 + \frac{3}{2}\mathbf{X}_7$	1	$u^2 + K_0$
7.	$\mathbf{X}_5$	$\mathbf{X}_3, \mathbf{X}_6, \mathbf{X}_7$	1	K_0
8.	$\mathbf{X}_7 + m\mathbf{X}_8$	$\mathbf{X}_5, \mathbf{X}_6, \mathbf{X}_9$	$\begin{cases} \text{arbitrary} & \text{if } m = 0 \\ \text{inconsistent} & \text{otherwise} \end{cases}$	0
9.	$\mathbf{X}_8$	(Whole algebra)	inconsistent	
10.	$\mathbf{X}_3 + \mathbf{X}_8$	$\mathbf{X}_4, \mathbf{X}_5, \mathbf{X}_6 - \frac{1}{2}\mathbf{X}_7$	inconsistent	
11.	$\mathbf{X}_3$	$\mathbf{X}_4, \mathbf{X}_5, \mathbf{X}_6, \mathbf{X}_7$	inconsistent	
†**12.**	$\mathbf{X}_5 + \mathbf{X}_9 + n\mathbf{X}_7 + m\mathbf{X}_8$	$\mathbf{X}_7$	$\dfrac{1}{1+u^2}\exp(m \tan^{-1} u)$	$K_0\sqrt{1+u^2}\exp n \tan^{-1} u$

$*$ Parameters (m, n) can be further restricted to lie in the region $\Omega = \{m > -1, n \neq 0, 1\} \cup \{m = -1, n \geq \frac{1}{2}, \neq 1\}$

† Parameters (m, n) may be taken to lie in the region $\Lambda = \{n > 0\} \cup \{n = 0, m \geq 0\}$

Table 4.2: Optimal system of one-dimensional subalgebras of $\tilde{L}$ (4.9) for nonlinear diffusion convection potential system (4.8). The normalizer of the algebra spanned by $\mathbf{Y}$ consists of $\mathbf{Y}$ itself, $\mathbf{X}_8$, and the operators in the 'Normalizer' column. $D(u)$ and $K(u)$ are found by integration of (4.12). A multiplicative constant D_0 has been removed from $D(u)$, using action of $\mathbf{X}_8$.

systems takes the normalizer $\mathbf{X}_7$, $\mathbf{X}_8$ as equivalence operators; by construction $\mathbf{X}_6 + (n - \frac{1}{2})\mathbf{X}_7 + (m+1)\mathbf{X}_8$ is a common symmetry of the subclass. Action of $\mathbf{X}_8$ can scale D_0 to unity; similarly $\mathbf{X}_7$ can scale K_0 to 1 unless $K_0 = 0$. If $K_0 = 0$, $\hat{\mathbf{X}}_7$ has trivial action on the equation, and is another symmetry (in fact, the Boltzmann scaling group for diffusion equations). There remains a certain amount of bookkeeping to ensure cases are not repeated. For example, $D(u) = |u|^m$, $K(u) = 0$ occurs again in Case **8**. This equation admits two symmetry operators, which occur as distinct cases in the optimal system.

Several operators in the optimal system of Table 4.2 have no associated solutions of (4.12), or else give $D(u) = 0$, which we ruled out *a priori*. These cases are marked 'inconsistent'.

Case **8** with $m = 0$ ($K(u) = 0$, $D(u)$ arbitrary) gives the class of potential diffusion systems (3.57), and shows that this class inherits as common symmetries $\hat{\mathbf{X}}_0$, $\hat{\mathbf{X}}_1$, $\hat{\mathbf{X}}_2$ from the diffusion convection equation, plus the additional common symmetry $\hat{\mathbf{X}}_7$. The four remaining operators $\mathbf{X}_5$, $\mathbf{X}_6$, $\mathbf{X}_9$, $\mathbf{X}_8$ in the normalizer of $\hat{\mathbf{X}}_7$ give the four 'true' equivalences (3.56) of the diffusion system. This nicely illustrates Proposition 4.2.4 on inheritance of equivalences by subclasses with symmetry.

Once all the repeats are weeded out, we finally obtain the partial classification of diffusion convection potential systems shown in Tables 4.3, 4.4. Constants of integration have been removed where possible. In addition to the symmetry operators shown, certain discrete symmetries are inherited from the equivalence group: these are noted in the table. Mostly these are obvious reflections, but we draw attention to the cases $D(u) = u^{-1}$, $K(u) = 0$ and $D(u) = u^{-1}$, $K(u) = u^{1/2}$, which have the hodograph-type transformation (3.84) as a discrete symmetry.

The symmetries listed in Tables 4.3, 4.4 for the diffusion convection potential system (4.8) are local symmetries of the scalar diffusion convection equation (3.71) unless $\mathbf{X}_9$ has a nonzero component. Thus the only listed equations with nonlocal inherited symmetries are the '$e^{m \tan^{-1} u}$' cases. (Because of the nonlocality, this case does not appear for the scalar equation (Table 4.1)). However, many nonlocal inherited symmetries are *hidden* by the parameter removal we have effected. The coefficients listed in Table 4.4 are representatives of families $D(u)$, $K(u)$ obtained by applying equivalence transformations (3.81) to the listed $D(u)$, $K(u)$. Since (3.81) includes nonlocal 'potential equivalences', the local symmetries listed in Tables 4.3 4.4 may correspond to *nonlocal* symmetries of the related $D(u)$, $K(u)$.

To exhibit the nonlocal inherited symmetries, we insert parameters using the potential equivalence group (3.81), then remove as many as possible using just the local equivalence group (3.74). The resulting nonlocal symmetries are shown in Table 4.6. We emphasize that all $D(u)$, $K(u)$ in Table 4.6 correspond to cases in Tables 4.3, 4.4. The close resemblance between Cases **1b** and **1d**; and between Cases **2a** and **2f**, reflects the fact that they are connected by a complex-valued transformation $u \mapsto iu$, $m \mapsto im$.

The partial classification gives significant insight into the symmetry structure of the diffusion convection potential system (4.8), and several interesting potential symmetries [15] are uncovered by the method. On the basis of the analysis we can make no statement about the possible completeness of the results in Tables 4.3, 4.4. A direct approach to symmetry classification of the diffusion convection system (4.8) is significantly more difficult than for the case where convection is absent, which is analyzed in [3, 15, 4]. This is due to the presence of *two* arbitrary func-

Case	Equation	Symmetry Operators
1.	$K(u) = 0$ (Diffusion case)	$\mathbf{Y}_3 = \mathbf{X}_7 = v\,\partial_v + x\,\partial_x + 2t\,\partial_t$
a.	$D(u) = 1$ (linear heat)	$\mathbf{Y}_4 = \mathbf{X}_6 - \frac{1}{2}\mathbf{X}_7 + \mathbf{X}_8 = v\,\partial_v + u\,\partial_u$ $\mathbf{Y}_5 = \mathbf{X}_5 = x\,\partial_v + \partial_u$
***b.**	$D(u) = \lvert u\rvert^m, \quad m \geq -1, \neq 0$	$\mathbf{Y}_4 = \mathbf{X}_6 - \frac{1}{2}\mathbf{X}_7 + (m+1)\mathbf{X}_8$ $= (m+2)v\,\partial_v + (m+1)x\,\partial_x + (m+2)t\,\partial_t + u\,\partial_u$
c.	$D(u) = e^u$	$\mathbf{Y}_4 = \mathbf{X}_5 + \mathbf{X}_8 = (v+x)\,\partial_v + x\,\partial_x + t\,\partial_t + \partial_u$
†**d.**	$D(u) = \dfrac{1}{1+u^2}\exp(m\tan^{-1}u),$ $m \geq 0$	$\mathbf{Y}_4 = \mathbf{X}_5 + \mathbf{X}_9 + m\mathbf{X}_8$ $= (mv+x)\,\partial_v + (mx-v)\,\partial_x + mt\,\partial_t + (1+u^2)\,\partial_u$

$*$ Reflection symmetry $v \mapsto -v$ $u \mapsto -u$.

The case $m = -1$ also admits the hodograph $x \mapsto v$, $v \mapsto x$, $u \mapsto 1/u$.

† Case $m = 0$ admits reflection symmetry $v \mapsto -v$, $u \mapsto -u$.

Table 4.3: Partial symmetry classification for diffusion convection potential system (4.8): Case $K(u) = 0$ (diffusion equations). Operators shown are in addition to common translations $\mathbf{Y}_0 = \partial_v$, $\mathbf{Y}_1 = \partial_x$, $\mathbf{Y}_2 = \partial_t$. Operator $\mathbf{Y}_3$ and reflection symmetry $x \mapsto -x$, $v \mapsto -v$ are common to all cases.

Case **2.**	Equation	Symmetry Operators
***a.**	$D(u) = \|u\|^m$	$\mathbf{Y}_3 = \mathbf{X}_6 + (n - \frac{1}{2})\mathbf{X}_7 + (m+1)\mathbf{X}_8$
	$K(u) = \|u\|^n, \qquad (m,n) \in \Omega$	$= (m - n + 2)v\,\partial_v + (m - n + 1)x\,\partial_x + (m - 2n + 2)t\,\partial_t + u\,\partial_u$
b.	$D(u) = 1$	$\mathbf{Y}_3 = \mathbf{X}_6 + \frac{3}{2}\mathbf{X}_7 + \mathbf{X}_8 = -x\,\partial_x - 2t\,\partial_t + u\,\partial_u$
	$K(u) = u^2$	$\mathbf{Y}_4 = 2\mathbf{X}_4 + \mathbf{X}_5 \qquad = x\,\partial_v + 2t\,\partial_x + \partial_u$
†c.	$D(u) = \|u\|^m$	$\mathbf{Y}_3 = \mathbf{X}_4 + \mathbf{X}_6 + \frac{1}{2}\mathbf{X}_7 + (m+1)\mathbf{X}_8$
	$K(u) = u \log \|u\|$	$= (m+1)v\,\partial_v + (t + mx)\,\partial_x + mt\,\partial_t + u\,\partial_u$
d.	$D(u) = e^{mu}$	$\mathbf{Y}_3 = \mathbf{X}_5 + \mathbf{X}_7 + m\mathbf{X}_8$
	$K(u) = e^u$	$= (x + mv - v)\,\partial_v + (m-1)x\,\partial_x + (m-2)t\,\partial_t + \partial_u$
e.	$D(u) = e^u$	$\mathbf{Y}_3 = \mathbf{X}_4 + \mathbf{X}_5 + \mathbf{X}_8$
	$K(u) = u^2$	$= (v + x)\,\partial_v + (x + t)\,\partial_x + t\,\partial_t + \partial_u$
‡f.	$D(u) = \dfrac{1}{1+u^2}\exp(m \tan^{-1} u)$	$\mathbf{Y}_3 = \mathbf{X}_5 + \mathbf{X}_9 + n\mathbf{X}_7 + m\mathbf{X}_8$
	$K(u) = \sqrt{1+u^2}\exp(n \tan^{-1} u), \quad (m,n) \in \Lambda$	$= (mv - nv + x)\,\partial_v + (-v + mx - nx)\,\partial_x + (m - 2n)t\,\partial_t + (1 + u^2)\,\partial_u$

∗ Admits reflection symmetry $x \mapsto -x$ $u \mapsto -u$.
Parameter region $\Omega = \{n > \frac{1}{2}, \neq 1\} \cup \{n = \frac{1}{2}, m \geq -1\}$
Case $(m,n) = (-1, 1/2)$ also admits hodograph $x \mapsto v$, $v \mapsto x$, $u \mapsto 1/u$.
† Admits reflection symmetry $v \mapsto -v$, $u \mapsto -u$.
‡ Parameter region $\Lambda = \{n > 0\} \cup \{n = 0, m \geq 0\}$.
Case $(m,n) = (0,0)$ admits reflection symmetry $x \mapsto -x$, $u \mapsto -u$.

Table 4.4: Partial symmetry classification for diffusion convection potential system (4.8): Case with nonlinear convection present. Operators shown are in addition to common translations $\mathbf{Y}_0 = \partial_v$, $\mathbf{Y}_1 = \partial_x$, $\mathbf{Y}_2 = \partial_t$.

Case	Diffusivity (Conductivity $K(u) = 0$)	Nonlocal symmetry operator
1a.	$D(u) = u^{-2}$	$\mathbf{Y} = v\,\partial_x - u^2\,\partial_u$
b.	$D(u) = \dfrac{1}{\lvert 1-u^2\rvert}\left\lvert\dfrac{1+u}{1-u}\right\rvert^{m/2}, \quad m \neq \pm 2$	$\mathbf{Y} = (mv + x)\,\partial_v + (v + mx)\,\partial_x + mt\,\partial_t + (1 - u^2)\,\partial_u$
c.	$D(u) = u^{-2}e^{1/u}$	$\mathbf{Y} = v\,\partial_v + (v + x)\,\partial_x + t\,\partial_t - u^2\,\partial_u$
d.	$D(u) = \dfrac{1}{1+u^2}\exp\left(m\tan^{-1}u\right)$	$\mathbf{Y} = (mv + x)\,\partial_v + (-v + mx)\,\partial_x + mt\,\partial_t + (1 + u^2)\,\partial_u$

Table 4.5: Nonlocal symmetries inherited from equivalence group (3.81) of diffusion convection potential system (4.8): Case $K(u) = 0$ (diffusion equations). Case numbering matches Table 4.3, but in **1b** the parameter m does not correspond. Parameters have been removed using local equivalences (3.74) only. Note that $\left\lvert\dfrac{1+u}{1-u}\right\rvert^{m/2} \equiv \exp(m\tanh^{-1}u)$.

Case	Diffusivity, Conductivity	Nonlocal symmetry operator
2a.	$D(u) = \dfrac{1}{\lvert 1-u^2\rvert}\left\lvert\dfrac{1+u}{1-u}\right\rvert^{m/2}$ $K(u) = \lvert 1-u^2\rvert^{1/2}\left\lvert\dfrac{1+u}{1-u}\right\rvert^{n/2},\ n \neq \pm 1$	$\mathbf{Y} = ((m-n)v + x)\,\partial_v + (v + (m-n)x)\,\partial_x + (m-2n)t\,\partial_t + (1-u^2)\,\partial_u$
b.	$D(u) = u^{-2}$ $K(u) = u^{-1}$	$\mathbf{Y} = 2t\,\partial_v + v\,\partial_x - u^2\,\partial_u$
c.	$D(u) = \dfrac{1}{\lvert 1-u^2\rvert}\left\lvert\dfrac{1+u}{1-u}\right\rvert^{m/2}$ $K(u) = (1+u)\log\left\lvert\dfrac{1+u}{1-u}\right\rvert$	$\mathbf{Y} = ((m-1)v + x - 2t)\,\partial_v + (v + (m-1)x + 2t)\,\partial_x + (m-2)t\,\partial_t + (1-u^2)\,\partial_u$
d.	$D(u) = u^{-2}e^{m/u}$ $K(u) = ue^{1/u}$	$\mathbf{Y} = (m-1)v\,\partial_v + (v + (m-1)x)\,\partial_x + (m-2)t\,\partial_t - u^2\,\partial_u$
e.	$D(u) = u^{-2}e^{1/u}$ $K(u) = u^{-1}$	$\mathbf{Y} = (v-2t)\,\partial_v + (v+x)\,\partial_x + t\,\partial_t - u^2\,\partial_u$
f.	$D(u) = \dfrac{1}{1+u^2}\exp(m\tan^{-1}u)$ $K(u) = \sqrt{1+u^2}\,\exp(n\tan^{-1}u)$	$\mathbf{Y} = ((m-n)v + x)\,\partial_v + (-v + (m-n)x)\,\partial_x + (m-2n)t\,\partial_t + (1+u^2)\,\partial_u$

Table 4.6: Nonlocal symmetries inherited from equivalence group (3.81) of diffusion convection potential system (4.8). Case numbering matches Table 4.4, but in **2a** and **2c**, the parameter m does not correspond. Parameters have been removed using local equivalences (3.74) only. Note that $\left\lvert\dfrac{1+u}{1-u}\right\rvert^{m/2} \equiv \exp(m\tanh^{-1}u)$.

tions $D(u)$, $K(u)$ instead of one. In fact it has not been possible to complete the calculations by hand. Reid's method [55, 57] as implemented in the symbolic language MAPLE can complete the calculations only after great labour. Due to the complexity of the output classifying equations (a typical case is the fourth order nonlinear system reproduced as (4.107)), it is difficult to interpret the results. We shall take up this point again in §4.5, where the complete symmetry classification is calculated by incorporating equivalence group information into a modified version of Reid's method. The result is that the partial classification above misses symmetries in only two cases, namely the linearizable equations discussed in §4.2.3: linear heat/Bluman-Kumei; and Burgers'/Fokas-Yortsos' systems. Apart from this the partial classification is complete. This is remarkable considering the relatively small labour involved in deriving these symmetry results.

Oron and Rosenau [49, Table 3(b)] give a symmetry classification for the scalar potential form (3.79) of the diffusion convection equation, whose symmetry and equivalence properties are identical to the potential system form (4.8). Like their classification for the scalar form (3.71), there are many errors and omissions in their results. In particular, *none* of the potential symmetries of Table 4.6 are found. Moreover they do not detect Case **2e** at all, and of the cases $D(u) = u^m$, $K(u) = \log|u|$; $D(u) = u^m$, $K(u) = u \log|u|$ related to **2c** they find only the special case $D(u) = 1$, $K(u) = \log u$. Partial classification offers a valuable check on symmetry calculations using other methods: just knowing classifying system (4.12) alerts one to major omissions in stated results.

Akhatov, Gazizov and Ibragimov [4] independently discovered the partial classification method (their 'preliminary classification'). They use the adjoint group in the same way as we do to construct an optimal system of subalgebras (see also [32]) and give a short list of essentially different cases. Our Proposition 4.2.4 is new. Combined with the results in §3.4.1, it gives a simple but powerful way of predicting how much symmetry and equivalence is inherited by a subclass $\mathcal{C}_{\hat{\mathcal{H}}}$ of d.e.'s from its parent class $\mathcal{C}$, or by a class of group invariant reduced equations $\mathcal{C}/\mathcal{H}$ from $\mathcal{C}$. These results may be chained together. This process is appealing, since it uses *one* calculation of an equivalence group to the maximum possible extent, making symmetry results available without calculating or solving determining equations.

The price paid for these easily obtained symmetry and equivalence results is that there is no guarantee of completeness: there may be symmetries and equivalences other than the inherited ones.

In [4], Akhatov, et al. performed a partial classification for various potential forms of one-dimensional gas dynamics equations. For one of their examples, a nonlocal equivalence transformation appears. They were also able to perform a complete symmetry classification, and as in our example, almost all the symmetries are inherited from the equivalence group. Another partial classification example is given in [32]. This case is less interesting, since there are no nonlocal equivalence transformations present.

4.3 Modification of Reid algorithm

We now consider symmetry group classification for a class $\mathcal{C}$ of d.e.'s, where now we ask for the full group, not just symmetries inherited from the equivalence group of $\mathcal{C}$. The goal is to discover 'classifying conditions' for the arbitrary elements which discriminate between cases with different symmetry properties (see the discussion at the beginning of §4.2.1).

The systematic classification of symmetry groups for a class of differential equations is due to Reid [55, 57], whose method is based on an algorithm for reducing an arbitrary system of p.d.e.'s to involutive form. His method does not utilize the transformation information contained in the equivalence group. This is an advantage in the sense that the input to the method is an easily derived and standard object, namely the determining equations for the symmetry operators. However, there are less desirable consequences. For example, two d.e.'s which are connected by an equivalence transformation may appear on different branches of the classification. Moreover, the algorithm can produce 'classifying' conditions which are completely spurious—the symmetry results on both branches can be identical. An analogous situation arises with linear algebraic systems with symbolic entries: the solution structure of

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

depends only on the determinant $ad - bc$, whereas Gaussian elimination also splits the calculation on two different paths depending whether $a = 0$ or not. Reid's method is a 'Gaussian elimination' method for linear *differential* equations, and can give the same kind of spurious case splittings.

As we have seen, the equivalence group $\hat{\mathcal{Q}}$ includes nontrivial symmetry information. This information can be made available through partial classification (§4.2) or the Tresse-Cartan equivalence method. Both methods proceed synthetically, constructing equations with symmetries without forming determining equations. However, these methods can only yield symmetries which lie in the equivalence group. One could circumvent this by sufficiently enlarging the class of equations until its equivalence group includes every possible symmetry transformation. However, this quickly leads to computationally intractable problems. One feels that there ought to be a way to use the partial symmetry information from the equivalence group to one's advantage when finding the complete classification. Until now there has been no way to do this. In [4], Akhatov, et al. compute both partial and complete symmetry classifications for one example in gas dynamics, but the two calculations are disjoint. No advantage accrues from knowing the equivalence group until the end, when it is used merely to remove parameters.

We seek to combine the best features of Reid's approach with those of the Tresse-Cartan method. Rather than avoiding determining equations altogether, we use the equivalence group $\hat{\mathcal{Q}}$ to *rewrite* the determining equations in a radically different form. Distinct equations $\mathrm{E}(\phi)$, $\mathrm{E}(\phi')$ in a class $\mathcal{C}$ in general give rise to distinct determining systems. This is true even when they are connected by a transformation from $\mathcal{Q}$. We rewrite the determining system in a form which is *invariant* under the action of $\hat{\mathcal{Q}}$, so that if $\mathrm{E}(\phi)$ and $\mathrm{E}(\phi')$ are connected by an equivalence transformation, they have *identical* determining systems. The equivalence information is then 'built in'. Provided subsequent manipulations are in terms of $\hat{\mathcal{Q}}$-invariant operations, the invariance property of the determining system under action of $\hat{\mathcal{Q}}$ is preserved. Symmetry classification is then performed by reformulating Reid's algorithm in terms of invariant operations.

Our method takes the results of the Tresse-Cartan equivalence method, and uses them as *input* to the invariant Reid method. If no equivalence information is available, or if the equivalence group is trivial, our method is just Reid's. Conversely, if the equivalence group is known to contain every symmetry, our method becomes the Tresse-Cartan method. Usually we are in the intermediate situation, where a nontrivial equivalence group is available, but is not guaranteed to contain all symmetries.

We believe this major reformulation of infinitesimal symmetry methods to be new. Our method should permit resolution of group classification problems which are computationally infeasible to Cartan or Reid used alone. The simplification and structure this process gives to the Reid classification method can be very great indeed. As an example, we give the complete symmetry group classification for the nonlinear diffusion convection equation in potential form.

4.3.1 Moving frame and determining equations

Initially we concentrate our attention on Reid's algorithm for reducing a system of determining equations to involutive form. His calculations are carried out in a fixed *coordinate system*. In particular, the differential operators ∂_{w^j} in the system represent derivatives with respect to a coordinate system. Moreover the dependent variables ζ^j are components of a vector field $\mathbf{Y} = \zeta^j \partial_{w^j}$ referred to this coordinate basis ∂_{w^j}. Subsequent manipulations, in particular computation of compatibility conditions, are performed by taking derivatives ∂_{w^j} of equations in the determining system. Our first goal is to modify Reid's involution algorithm so that all these steps are referred to an *arbitrary moving frame*. We begin by introducing the necessary theoretical machinery.

Definition 4.3.1. Let W be a ν-dimensional space. A *moving frame* on W is a set of ν smooth vector fields $\Delta_1, \Delta_2, \ldots, \Delta_\nu$ which are linearly independent at each point in W.

Ultimately any moving frame can be referred to the coordinate frame

$$\partial_{w^1}, \partial_{w^2}, \ldots, \partial_{w^\nu}$$

associated with the coordinate system w, as

$$\Delta_i = A_i^j(w) \partial_{w^j}.$$

The $\nu \times \nu$ matrix $A_i^j(w)$ of smooth functions is to be nonsingular at every point w. More generally, given a frame $\{\Delta_i\}_{i=1}^{\nu}$, we may change frame to

$$\Delta_i' = A_i^j(w) \Delta_j$$

where the smooth nonsingular matrix $A_i^j(w)$ is the *change of frame matrix*. It is convenient to write this explicitly in vector-matrix form

$$\begin{pmatrix} \Delta_1' \\ \Delta_2' \\ \vdots \\ \Delta_\nu' \end{pmatrix} = \begin{pmatrix} A_1^1(w) & A_1^2(w) & \ldots & A_1^\nu(w) \\ A_2^1(w) & A_2^2(w) & \ldots & A_2^\nu(w) \\ \vdots & \vdots & \ddots & \vdots \\ A_\nu^1(w) & A_\nu^2(w) & \ldots & A_\nu^\nu(w) \end{pmatrix} \begin{pmatrix} \Delta_1 \\ \Delta_2 \\ \vdots \\ \Delta_\nu \end{pmatrix}$$

or, more briefly,

$$\Delta' = A\Delta \tag{4.18}$$

If a *change of coordinates* $w \mapsto w'$ is executed, the change of coordinate frame from $\partial_w = (\partial_{w^1}\ \partial_{w^2}\ \ldots\ \partial_{w^\nu})$ to $\partial_{w'} = (\partial_{w'^1}\ \partial_{w'^2}\ \ldots\ \partial_{w'^n})$ is given by (4.18) with A being the Jacobian matrix, $A_i^j = \dfrac{\partial w^j}{\partial w'^i}$. In general, however, the matrix A is arbitrary, and not derived from a change of coordinates.

We shall always denote frame operators by upper case Greek letters Δ, Λ, etc. Moving frames are widely used in geometry, and we refer to [65, Vol. II, §7] or any other modern differential geometry text for further information. Our orientation is more computational than most treatments.

Vector field referred to frame

A vector field $\mathbf{Y}$ may be referred to a coordinate frame ∂_w as

$$\mathbf{Y} = \zeta^i(w)\partial_{w^i}.$$

A vector field may also be referred to a moving frame Δ, since the basis vector fields Δ_i are linearly independent at every point w:

$$\mathbf{Y} = \theta^i \Delta_i.$$

It is convenient to write $\theta = (\theta^1\ \theta^2\ \ldots\ \theta^\nu)^T$, so that

$$\mathbf{Y} = \theta^T \Delta.$$

We used θ earlier to name functions $u = \theta(x)$, but no confusion should arise since we have no further need to assign u as a function of x.

Suppose a change of frame (4.18) is executed. Then

$$\mathbf{Y} = \theta'^T \Delta'$$

where

$$\theta' = (A^T)^{-1}\theta, \tag{4.19}$$

showing how components of a vector field transform under change of frame.

Example 4.3.2. The following frame arises in the analysis of the nonlinear diffusion convection potential system. We draw our examples in this section from this equation, its equivalence group and its symmetry analysis. A complete symmetry analysis based on the methods of this section will be presented in §4.5.2.

Let $\partial = (\partial_v, \partial_x, \partial_t, \partial_u)^T$ be a coordinate frame on a space (v, x, t, u). Introduce the moving frame Δ given by

$$\begin{aligned}
\Delta_1 &= \partial_v \\
\Delta_2 &= \partial_x \\
\Delta_3 &= \partial_t + \dot{K}(u)\partial_x + (u\dot{K}(u) - K(u))\partial_v \\
\Delta_4 &= \partial_u
\end{aligned} \tag{4.20}$$

where $K(u)$ is some smooth function. The determinant of the change of frame matrix is 1, in particular it is nonzero, so Δ is indeed a moving frame.

Let

$$\mathbf{Y} = \chi\partial_v + \xi\partial_x + \tau\partial_t + \eta\partial_u$$

be an arbitrary vector field, with χ, ξ, τ, η functions of (v, x, t, u). Resolving this with respect to the moving frame Δ, i.e. $\mathbf{Y} = \theta^i \Delta_i$, we find

$$\begin{aligned} \theta^1 &= \chi - (u\dot{K}(u) - K(u))\tau \\ \theta^2 &= \xi - \dot{K}(u)\tau \\ \theta^3 &= \tau \\ \theta^4 &= \eta. \end{aligned} \tag{4.21}$$

Structure relations

Since the commutator $[\,\Delta_i, \Delta_j\,]$ of two vector fields Δ_i, Δ_j from a moving frame is a vector field, it must be expressible as a linear combination of Δ_k at each point:

$$[\,\Delta_i, \Delta_j\,] = \gamma_{ij}^k \Delta_k, \tag{4.22}$$

where γ_{ij}^k are functions of w, which we call the *structure functions*. Relations (4.22) will be called the *structure relations* for the moving frame Δ. Clearly γ_{ij}^k is antisymmetric in the lower indices $\gamma_{ij}^k = -\gamma_{ji}^k$.

Example 4.3.2. **(cont.)** For the moving frame Δ (4.20), we compute

$$\begin{array}{lll} [\Delta_1, \Delta_2] = 0 & [\Delta_1, \Delta_3] = 0 & [\Delta_1, \Delta_4] = 0 \\ & [\Delta_2, \Delta_3] = 0 & [\Delta_2, \Delta_4] = 0 \\ & & [\Delta_3, \Delta_4] = -\ddot{K}(u)(\Delta_2 + u\Delta_1). \end{array} \tag{4.23}$$

In general a moving frame does *not* derive from a coordinate system: it is a standard result [65, Vol. I, Theorem 5.14] that a frame represents a coordinate system if and only if $[\Delta_i, \Delta_j] = 0$ for each i, j. In general, frame operators Δ_i may be algebraically manipulated in much the same way as partial derivatives—in particular they are linear differential operators obeying Leibniz' rule $\Delta_i(fg) = f\Delta_i g + g\Delta_i f$—except that they may not be freely permuted. Instead the structure relations (4.22) must be consulted to execute any changes in order of application of the frame operators Δ_i. We examine this process in more detail.

Let $J = (j^1, j^2, \dots, j^p)$ be an ordered multi-index. Denote the number of indices by $|J| = p$. We shall often write

$$\Delta_J \equiv \Delta_{j_p} \dots \Delta_{j_2} \Delta_{j_1}$$

for brevity, so that $\Delta_{ij} = \Delta_j \Delta_i$. The ordering of indices should be carefully noted: this is consistent with the convention that $u_{xy} \equiv \partial_y \partial_x u$.

Suppose I is a permutation of J. If Δ_i represented derivatives with respect to a coordinate system we would have $\Delta_I = \Delta_J$. In this case, the only feature of importance in J is the number of 1's, 2's, ..., ν's. For general frames this is not so, and order in a multi-index is important. Nevertheless Δ_I and Δ_J are "essentially" the same in the sense that they differ only by "lower order terms".

Proposition 4.3.3. *Let I be a permutation of a p-th order multi-index J, and let Δ be a frame. Then*

$$\Delta_I = \Delta_J + \sum_{|K| \le p-1} C^K \Delta_K$$

where C^K are certain coefficient functions expressible in terms of the frame's structure functions γ_{ij}^k and their frame derivatives.

Proof. For $p = 1$ the proposition is trivial; structure relations (4.22) in the form $\Delta_{ij} = \Delta_{ji} + \gamma^k_{ij}\Delta_k$ express its truth for $p = 2$. For arbitrary p the structure relations (4.22) show how to effect pairwise interchanges of neighbouring elements. Let J' be obtained from J by such a pairwise interchange of indices: $J = (j_1 j_2 \dots j_k l m j_{k+3} \dots j_p)$, and $J' = (j_1 j_2 \dots j_k m l j_{k+3} \dots j_p)$. Then

$$\Delta_{J'} = \Delta_J + \Delta_{j_{k+3}\dots j_p}(\gamma^q_{lm}\Delta_q)\Delta_{j_1 j_2 \dots j_k}.$$

The second term is of order $p - 1$ and by Leibniz' rule may ultimately be written as $\sum_{|K| \le p-1} C^K \Delta_K$, for some coefficients C^K. Since any permutation can be effected by a sequence of pairwise interchanges of neighbouring elements, repeated application of this argument yields the proposition. □

In future we refer to the lower order correction terms as 'permutation terms': they are expressible solely in terms of the structure functions γ^k_{ij} and their frame derivatives.

Change of frame

Suppose a frame Δ has structure relations (4.22). Executing a change of frame (4.18) to Λ, a direct calculation gives

$$[\Lambda_i, \Lambda_j] = \beta^k_{ij}\Lambda_k$$

with new structure functions

$$\beta^k_{ij} = B^k_l \Big(A^p_i A^q_j \gamma^l_{pq} + A^p_i \Delta_p(A^l_j) - A^p_j \Delta_p(A^l_i) \Big) \tag{4.24}$$

where the matrix $B = [B^i_j]$ is the inverse $B = A^{-1}$ of the change of frame matrix, so $A^i_k B^k_j = \delta^i_j$. Thus structure functions for the new frame are available in terms of the old structure functions, the change of frame matrix, its inverse, and its frame derivatives. Note that unlike the structure constants C^k_{ij} of a Lie algebra, the structure functions do not generally constitute a tensor.

Example 4.3.2. **(cont.)** Suppose we change frame from Δ (4.20) to Λ given by

$$\begin{aligned} \Lambda_1 &= D^{1/2}(u)\,(\Delta_2 + u\Delta_1) \\ \Lambda_2 &= \dot{D}(u)D^{-3/2}(u)\,\Delta_2 + D^{-3/2}(u)(u\dot{D}(u) + 2D(u))\Delta_1 \\ \Lambda_3 &= \Delta_3 \\ \Lambda_4 &= 1/D(u)\,\Delta_4 \end{aligned} \tag{4.25}$$

where $D(u) > 0$ is a strictly positive smooth function. We compute, for example

$$\begin{aligned} [\Lambda_2, \Lambda_4] &= D^{-5/2}\dot{D}\,[\Delta_2, \Delta_4] + D^{-5/2}(u\dot{D} + 2D)\,[\Delta_1, \Delta_4] \\ &\quad + \Big(D^{-3/2}\dot{D}\Delta_2(D^{-1}) + D^{-3/2}(u\dot{D} + 2D)\Delta_1(D^{-1})\Big)\Delta_4 \\ &\quad + D^{-1}\Delta_4(\dot{D}D^{-3/2})\Delta_2 - D^{-1}\Delta_4(D^{-3/2}(u\dot{D} + 2D))\Delta_1. \end{aligned}$$

The first two terms vanish by virtue of structure relations (4.23) for Δ_i. Frame derivatives of u, $D(u)$ and $\dot{D}(u)$ are required. However, the original definition (4.20) of Δ shows $\Delta_1 D(u) = \Delta_2 D(u) = 0$, while $\Delta_4 D(u) = \dot{D}(u)$. Hence

$$[\Lambda_2, \Lambda_4] = -D^{-7/2}(D\ddot{D} - 3/2\dot{D}^2)(\Delta_2 + u\Delta_1)$$

We must now express this in terms of Λ_i:

$$[\Lambda_2, \Lambda_4] = -L(u)\Lambda_1$$

where the function $L(u)$ is given by

$$L(u) = \frac{D(u)\ddot{D}(u) - \frac{3}{2}\dot{D}(u)^2}{D(u)^4}. \tag{4.26}$$

Carrying out similar manipulations gives structure relations for Λ:

$$\begin{array}{lll} [\Lambda_1, \Lambda_2] = 0 & [\Lambda_1, \Lambda_3] = 0 & [\Lambda_1, \Lambda_4] = -\frac{1}{2}\Lambda_2 \\ & [\Lambda_2, \Lambda_3] = 0 & [\Lambda_2, \Lambda_4] = -L(u)\Lambda_1 \\ & & [\Lambda_3, \Lambda_4] = -I(u)\Lambda_1 \end{array} \tag{4.27}$$

where $I(u) = \ddot{K}(u)D^{-3/2}(u)$.

Infinitesimal determining equations

The determining equations are linear homogeneous p.d.e.'s for the components ξ^i, η^j of a symmetry vector field

$$\mathbf{Y} = \xi^i \partial_{x^i} + \eta^j \partial_{u^j}.$$

In the determining equations there is no distinction between the independent and dependent variables x, u of the original d.e.'s: all x, u are *independent* variables of the determining system. As in §3.1 we use the notation $w = (x, u)$, and as in §3.3.1 denote the corresponding infinitesimals by $\zeta = (\xi, \eta)$, so that $\mathbf{Y} = \zeta^i \partial_{w^i}$. Thus the determining equations are expressed in terms of

(i) Differential operators ∂_{w^i}, which operate on the dependent variables ζ^i of the determining system.

(ii) Components ζ^i of a vector field, referred to the coordinate frame ∂_{w^i}.

(iii) Coefficient functions, which are functions of w.

Instead of referring determining equations to the w coordinate system, we refer them to an arbitrary moving frame. By the process described above, we change frame to Δ. The determining equations will be expressed in terms of

(i′) Differential operators Δ_i, operating on the 'dependent variables' θ^i of the system, and given by (4.18).

(ii′) Components θ^i of a vector field, referred to the moving frame Δ_i, and given by (4.19).

(iii′) Coefficient functions, which are functions of w. The new coefficient functions are expressed in terms of the old coefficients, and the change of frame matrix, its inverse, and its frame derivatives of various orders.

Clearly the determining equations are linear and homogeneous in θ^i. However, when written with respect to the frame they are *not* (as it stands) differential equations. A *frame system* is a system of d.e.'s referred to a moving frame.

Example 4.3.4. The diffusion convection potential system (4.8)

$$\begin{aligned} v_x &= u \\ v_t &= D(u)u_x - K(u) \end{aligned}$$

leads to determining equations

$$\begin{aligned}
&\partial_v \tau = 0 \\
&\partial_x \tau = 0 \\
&\partial_u \tau = 0 \qquad \partial_u \xi = 0 \qquad \partial_u \chi = 0 \\
&\eta - \partial_x \chi - u\,\partial_x \xi - u^2\,\partial_v \xi = 0 \\
&\dot{D}\eta + D(-\partial_x \xi - \partial_v \chi + \partial_t \tau + \partial_u \eta) = 0 \\
&\dot{K}\eta + K(-\partial_v \chi + \partial_t \tau + u\,\partial_v \xi) - D(\partial_x \eta + u\,\partial_v \eta) + \partial_t \chi - u\,\partial_t \xi = 0
\end{aligned} \tag{4.28}$$

for the components of a symmetry vector field $\mathbf{Y} = \xi\,\partial_x + \tau\,\partial_t + \chi\,\partial_v + \eta\,\partial_u$. Introducing the moving frame Δ (4.20) in place of the operators ∂_x, ∂_t, ∂_v, ∂_u; and the field components θ^i (4.21) in place of ξ, τ, χ, η, these equations become

$$\begin{aligned}
&\Delta_1\theta^3 = 0 \qquad \Delta_4\theta^2 + \ddot{K}\theta^3 = 0 \qquad \Delta_4\theta^1 + u\ddot{K}\theta^3 = 0 \\
&\Delta_2\theta^3 = 0 \\
&\Delta_4\theta^3 = 0 \\
&\theta^4 - \Delta_2\theta^1 - u\Delta_1\theta^1 - u\Delta_2\theta^2 - u^2\Delta_1\theta^2 = 0 \\
&\dot{D}\theta^4 + D(-\Delta_1\theta^1 - \Delta_2\theta^2 + \Delta_3\theta^3 + \Delta_4\theta^4) = 0 \\
&-D(\Delta_2\theta^4 + u\Delta_1\theta^4) + \Delta_3\theta^1 - u\Delta_3\theta^2 = 0
\end{aligned} \tag{4.29}$$

4.3.2 Frame Reid method

Reid [55, 56] described an algorithm for bringing a linear homogeneous system of partial differential equations to an *involutive form*, whose compatibility conditions yield no new relations. Now suppose we have a frame system, i.e., a linear homogeneous system for frame derivatives of certain dependent variables θ^i. We seek to generalize these ideas to construct a 'frame involutive' system. We successively define orthonomic, reduced orthonomic, and involutive systems with respect to a frame. These concepts are straight adaptations from the Riquier-Janet-Reid theory [33, 67, 55] for systems of d.e.'s. We attempt to stay as close as possible to Reid's methods. We are not aware of other attempts at a 'frame Riquier-Janet' theory in the literature: frames are usually used in conjunction with geometric integrability theorems (Frobenius theorem, Cartan-Kähler theorem), which obscure the relationship with Reid's method.

Frame derivatives

Let $\{\Delta_i\}_{i=1}^{\nu}$ be a moving frame on a space $(w^1, w^2, \ldots, w^\nu)$, with structure relations $[\Delta_i, \Delta_j] = \gamma_{ij}^k \Delta_k$. Let $\{\theta^i\}_{i=1}^{\mu}$ be certain dependent variables. In our application θ^i are dependent variables in determining equations for symmetries, and are components of a vector field $\mathbf{Y} = \theta^i \Delta_i$. However this fact is not used until §4.5, so we let θ^i represent any variables. We first establish notation for frame derivatives. Proposition 4.3.3 shows that Δ_I and Δ_J are equivalent to within lower order terms if I is a permutation of J. In certain circumstances the ordering of a multi-index J is of no importance. In this case we denote the multi-index J by $[J]$, which represents the equivalence class of J under arbitrary permutation. Thus $[I] = [J]$ if and only if I is a permutation of J; the characterizing feature of $[J]$ is the number of 1's, 2's, ... , ν's contained in J. We define $N_i(J)$ to be the number of occurrences of i in the multi-index J.

Let I and J be two multi-indices of orders p_1, p_2 respectively, with $p_1 \leq p_2$. We say $I \subseteq J$ if there exists a (p_2-p_1)-th order multi-index L such that $[J] = [IL]$. Thus $(133) \subseteq (3131)$, since $[3131] = [(133)(1)]$. Alternatively, $I \subseteq J$ if $N_i(I) \leq N_i(J)$ for all i.

Definition 4.3.5. Let $\Delta_I\theta$ and $\Delta_J\theta$ be two frame derivatives of θ. We say $\Delta_J\theta$ is a $((|J|-|I|)$-th order) *frame derivative* of $\Delta_I\theta$ if $I \subseteq J$.

Clearly $\Delta_{133}\theta$ is a derivative of $\Delta_1\theta$, since $\Delta_{133}\theta = \Delta_{33}(\Delta_1\theta)$. Note that $\Delta_{133}\theta$ is also a (first order) derivative of $\Delta_{31}\theta$, since $[133] = [(31)(3)]$. If $I \subseteq J$ then $\Delta_J\theta$ may be obtained from $\Delta_I\theta$ by application of a frame operator Δ_L plus permutation terms. Thus

$$\Delta_{133}\theta = \Delta_{313}\theta + \Delta_3(\gamma^k_{31}\Delta_k\theta).$$

Note that if I is a permutation of J, then $\Delta_J\theta$ is a 'zeroth-order derivative' of $\Delta_I\theta$.

Order relation on frame derivatives

The Reid [56] and Janet [33] methods for rendering a p.d.e. system involutive rely crucially on *ordering* the partial derivatives $\partial_J u^j$ which occur there (here J is a multi-index). From our point of view, their method assigns an order relation not on derivatives $\partial_J u^j$, but on equivalence classes of derivatives $\partial_{[J]}u^j$. If I is a permutation of J, $[I] = [J]$, $\partial_I u^j$ and $\partial_J u^j$ are regarded as the same derivative. In a frame system, the frame derivatives $\Delta_I\theta^j$ and $\Delta_J\theta^j$ are distinct objects, even when I is a permutation of J, . However, for purposes of ordering we regard them as identical. We denote the set of frame derivatives equivalent to $\Delta_J\theta^j$ under permutation by $\Delta_{[J]}\theta^j$.

Definition 4.3.6. A *Janet ordering* of frame derivatives is a total order relation $\prec$ on equivalence classes $\Delta_{[J]}\theta^j$ of derivatives with the properties:

1. (transitivity) If $\Delta_{[I]}\theta^i \prec \Delta_{[J]}\theta^j$ and $\Delta_{[J]}\theta^j \prec \Delta_{[K]}\theta^k$, then $\Delta_{[I]}\theta^i \prec \Delta_{[K]}\theta^k$.
2. (trichotomy) If $\Delta_{[I]}\theta^i$, $\Delta_{[J]}\theta^j$ are two derivatives, exactly one of (a) $\Delta_{[I]}\theta^i \prec \Delta_{[J]}\theta^j$, (b) $\Delta_{[J]}\theta^j \prec \Delta_{[I]}\theta^i$, (c) $\Delta_{[J]}\theta^j = \Delta_{[I]}\theta^i$, is true.
3. (preservation under differentiation) If $\Delta_{[I]}\theta^i \prec \Delta_{[J]}\theta^j$, then $\Delta_{[IL]}\theta^i \prec \Delta_{[JL]}\theta^j$ for all arbitrary order multi-indices L.
4. (respects differentiation) $\Delta_{[J]}\theta^j \prec \Delta_{[JL]}\theta^j$ for all nonempty multi-indices L.

These properties are essential in determining which derivative should be isolated on the left hand side of an equation. By using Proposition 4.3.3, we can pass freely between $\Delta_I\theta^j$ and $\Delta_J\theta^j$, and we do not distinguish them in the ordering. If desired, frame derivatives could be further ordered within the equivalence classes $[J]$, to give a total order relation on the set of all frame derivatives.

As an example of a Janet ordering, consider the lexicographic ordering $\Delta_{[I]}\theta^i \prec \Delta_{[J]}\theta^j$ if

1. $|I| < |J|$
2. $|I| = |J|$, but $i < j$
3. $|I| = |J|$, $i = j$ but the first nonzero member of the sequence

$$N_1(I) - N_1(J), \quad N_2(I) - N_2(J), \quad \dots, \quad N_\nu(I) - N_\nu(J)$$

   is negative.

(Recall $N_i(I)$ is the number of i's in the sequence I). For example, suppose there are two dependent variables θ^1, θ^2, and two frame operators Δ_1, Δ_2. Lexicographic order is

$$\begin{aligned} \theta^1 \prec \theta^2 \prec \Delta_1\theta^1 \prec \Delta_2\theta^1 \prec \Delta_1\theta^2 \prec \Delta_2\theta^2 \prec \Delta_{11}\theta^1 \prec \begin{Bmatrix} \Delta_{12}\theta^1 \\ \Delta_{21}\theta^1 \end{Bmatrix} \\ \prec \Delta_{22}\theta^1 \prec \Delta_{11}\theta^2 \prec \begin{Bmatrix} \Delta_{12}\theta^2 \\ \Delta_{21}\theta^2 \end{Bmatrix} \prec \Delta_{22}\theta^2 \end{aligned}$$

Any other convenient ordering satisfying (i)—(iv) of Definition 4.3.6 may be chosen: Janet orders do not have to be lexicographic. In practice we choose the ordering during the course of a hand calculation. Our failure to distinguish between $\Delta_{12}\theta$ and $\Delta_{21}\theta$ has the consequence that the standard 'involutive' form eventually attained by our system is not unique: An equation $\Delta_{12}\theta = \text{rhs}$ could be replaced by $\Delta_{21}\theta = \text{rhs}+\text{permutation terms}$, without upsetting our ordering. This could be resolved by ordering frame derivatives within each permutation class.

Assuming that an ordering has been chosen for a frame system, we seek to append all compatibility conditions to the system. Algorithms for this are shown in Appendix A.1; we illustrate the ideas involved by example.

Orthonomic system

We adapt the concept of orthonomic system [55, 33] to frame systems.

Definition 4.3.7. A linear homogeneous frame system is in *orthonomic form* if

(i) Each equation is resolved in the form

$$\Delta_I\theta^i = \sum_{j,J} C_j^J \Delta_J\theta^j$$

(ii) $\Delta_I\theta^i$ is strictly higher in the ordering than any terms $\Delta_J\theta^j$ on the right hand side.

(iii) A given derivative $\Delta_J\theta^j$ cannot appear in both the left and right hand sides of the system.

Achieving orthonomic form is basically a linear algebra problem, which is solved by Gauss-Jordan elimination (see Appendix A.1.1). Requirement (ii) adds the complication that certain ordering conditions must be respected in the process.

The highest order derivative occurring in an equation will be called the *leading* derivative.

Example 4.3.8. Consider a frame system with dependent variable θ, referred to the frame Λ considered above (4.25):

$$\begin{aligned} -\dot{D}\Lambda_1\theta + D^2\Lambda_2\theta &= 0 \\ (2D + u\dot{D})\Lambda_1\theta - uD^2\Lambda_2\theta &= 0 \\ \Lambda_4\theta &= 0 \\ \Lambda_4\Lambda_3\theta - \Lambda_1\theta &= 0 \end{aligned} \tag{4.30}$$

The frame has structure relations (4.27). Here $D = D(u)$ is some *nonzero* function of u.

First we lexicographically order the derivatives occurring in the system,

$$\theta \prec \Lambda_1\theta \prec \Lambda_2\theta \prec \Lambda_3\theta \prec \Lambda_4\theta \prec \Lambda_4\Lambda_3\theta.$$

The highest ordered derivative occurring is $\Lambda_4\Lambda_3\theta$, and we isolate it on the left hand side $\Lambda_4\Lambda_3\theta = \Lambda_1\theta$. The next highest is $\Lambda_4\theta$, which is already isolated. Neither of these leading derivatives occur elsewhere in the system, so no substitutions are required yet. The next highest derivative is $\Lambda_2\theta$, which is the leading derivative in both the first two equations. Choosing the first one arbitrarily, and isolating $\Lambda_2\theta$ gives $\Lambda_2\theta = \frac{\dot{D}}{D^2}\Lambda_1\theta$. Substituting this into the second equation yields $\Lambda_1\theta = 0$. *This* must now be substituted throughout, giving $\Lambda_2\theta = 0$, and $\Lambda_4\Lambda_3\theta = 0$. Finally we achieve orthonomic form:

$$\Lambda_1\theta = 0, \qquad \Lambda_2\theta = 0, \qquad \Lambda_4\theta = 0, \qquad \Lambda_4\Lambda_3\theta = 0 \tag{4.31}$$

Note that choice of ordering helps matters here. If we had $\Lambda_2\theta \prec \Lambda_1\theta$ in the order, division by either $\dot{D}$ or $2D + u\dot{D}$ would have been necessary. This requires one of these coefficients to be nonvanishing, which imposes restrictions on $\dot{D}$ which were not needed in the ordering originally chosen. In a hand calculation one can vary the ordering of derivatives during the procedure in order to avoid such divisions for as long as possible.

Reduced Orthonomic Form

A frame system in orthonomic form separates the derivatives unambiguously into two classes (those which occur on the left hand side and those which do not). However, the resolution of the system is unsatisfactory in that one may have derivatives which are derivatives of leading derivatives.

Definition 4.3.9. A *reduced* orthonomic system is a frame system in orthonomic form (satisfying (i), (ii), (iii) of Definition 4.3.7) and also

(iv) No derivative in the system is the derivative of any derivative on the left hand side.

Note that, since we regard $\Delta_i\Delta_j\theta$ as a derivative of $\Delta_i\theta$, a system with both $\Delta_i\theta$ and $\Delta_i\Delta_j\theta$ on the left hand side would not be in reduced orthonomic form.

Suppose we are given an orthonomic system, in which $\Delta_J\theta^j$ is a derivative of some leading derivative $\Delta_I\theta^j$. Thus $[J] = [IL]$ for some multi-index L, and $\Delta_J\theta^j = \Delta_L\Delta_I\theta^j$ + permutation terms. The system includes an equation $\Delta_I\theta^j$ = rhs (since $\Delta_I\theta^j$ is leading). Execute the substitution of this into $\Delta_J\theta^j$, replacing $\Delta_J\theta^j$ by $\Delta_L(\text{rhs})+$permutation terms. Reid calls this an *implicit substitution*, and we retain this terminology. The only additional feature in our process is the presence of permutation terms resulting from noncommuting frame operators. Reduced orthonomic form is achieved by executing all possible implicit substitutions throughout a system (see Appendix A.1.2).

Example 4.3.10. We bring system (4.30) to reduced orthonomic form. First the system is brought to orthonomic form (4.31). Now we note that $\Lambda_4\Lambda_3\theta$ is a derivative of a leading derivative $\Lambda_4\theta$. From the structure relation $[\Lambda_3, \Lambda_4] = -\ddot{K}D^{-3/2}\Lambda_1$ (4.27) we find

$$\Lambda_4\Lambda_3\theta = \Lambda_3(\Lambda_4\theta) + \ddot{K}D^{-3/2}\Lambda_1\theta.$$

while the system states $\Lambda_4\Lambda_3\theta = 0$. Hence inserting $\Lambda_4\theta = 0$, we find $\ddot{K}D^{-3/2}\Lambda_1\theta = 0$. The system is now no longer in orthonomic form. Bringing it to orthonomic form just eliminates this last equation, and we obtain the reduced orthonomic system

$$\Lambda_1\theta = 0, \qquad \Lambda_2\theta = 0, \qquad \Lambda_4\theta = 0 \tag{4.32}$$

Compatibility Conditions

Let a frame system be given in reduced orthonomic form. Let

$$\Delta_I\theta^i = \text{rhs}_1 \tag{4.33}$$
$$\Delta_J\theta^i = \text{rhs}_2 \tag{4.34}$$

be two equations in the system. Define the 'union' $[I \cup J]$ of multi-indices I, J by

$$N_j[I \cup J] = \max\{N_j(I), N_j(J)\}, \qquad j = 1, 2, \dots, \nu$$

The 'union' is only defined to within a permutation of its indices. Let $U \in [I \cup J]$ be some multi-index in the union of I and J. Thus $\Delta_U\theta^i$ is a derivative of both $\Delta_I\theta^i$ and $\Delta_J\theta^i$. Also U is the "smallest" such multi-index, in the sense that any multi-index K of order $|K| \le |U|$ with this property is a permutation of U, $[K] = [U] = [I \cup J]$. For example, let $I = (13312)$, and $J = (212)$. The 'union' of I, J is $[I \cup J] = [112233]$, to within permutation. Suppose $[U] = [IL] = [JM]$ for some multi-indices L, M. Then

$$\begin{aligned}\Delta_U\theta^i &= \Delta_L\Delta_I\theta^i + \text{permutation terms}\\ &= \Delta_L(\text{rhs}_1) + \text{permutation terms}\end{aligned}$$

and

$$\begin{aligned}\Delta_U\theta^i &= \Delta_M\Delta_J\theta^i + \text{permutation terms}\\ &= \Delta_M(\text{rhs}_2) + \text{permutation terms.}\end{aligned}$$

Equating these two expressions yields the *compatibility condition*

$$\Delta_L(\text{rhs}_1) - \Delta_M(\text{rhs}_2) + \text{permutation terms} = 0$$

of (4.33). Substitutions and implicit substitutions from the original reduced orthonomic system are applied to the resulting expression, which then involves only *non*leading derivatives. (In many cases this simplification leads to triviality $0 = 0$.)

Example 4.3.11. We compute compatibility conditions of the reduced orthonomic system (4.32). Consider the two equations $\Lambda_1\theta = 0$ and $\Lambda_4\theta = 0$. The 'union' of (1) and (4) is just (14). The compatibility condition is $\Lambda_1(0)-\Lambda_4(0)-[\Lambda_1,\Lambda_4]\theta = 0$. Structure relations (4.27)) show $[\Lambda_1,\Lambda_4] = -\frac{1}{2}\Lambda_2$. Hence the compatibility condition is $\Lambda_2\theta = 0$. Simplification of this by the system yields a triviality. In fact *all* the compatibility conditions of this system are trivial.

Suppose now the structure relations had been different, and that we had $[\Lambda_1,\Lambda_4] = \Lambda_3$. Compatibility of the same equations would have given us $\Lambda_3\theta = 0$, a nontrivial equation. Even the simplest equations can generate nontrivial compatibility conditions through structure relations.

Frame involutive form

If we adjoin compatibility conditions to a reduced orthonomic frame system, the composite system is no longer in solved form, and must be brought once again to reduced orthonomic form. We distinguish systems where this process does not lead to addition of further relations.

Definition 4.3.12. A reduced orthonomic frame system R is *involutive* (or 'passive') if the compatibility conditions of R become trivial after carrying out implicit substitutions from R.

A frame system may be brought to involutive form by putting it into reduced orthonomic form, appending compatibility conditions, then repeating the process (see Appendix A.1.3). By an argument originally due to Tresse [68] (see also [55]), this process must be finite.

Example 4.3.13. Consider the reduced orthonomic system (4.32). Its compatibility conditions—partly computed above—are $\Lambda_2\theta = 0$, $L(u)\Lambda_1\theta = 0$, with $L(u)$ defined by (4.26). Reducing these using the original system gives trivialities $0 = 0$. Hence system (4.32) is involutive.

Associated with an involutive system are two sets of (equivalence classes of) frame derivatives. The derivatives which occur on the left hand side of the system have values which are specified in terms of those on the right hand side. By differentiation, any derivative of these is also expressed in terms of the derivatives on the right hand side.

Definition 4.3.14. If $\Delta_J\theta^j$ occurs on the left hand side, or is a derivative of $\Delta_I\theta^j$ occurring on the left hand side of a frame involutive system, then it is called a *leading* or *principal* derivative of the system. If $\Delta_J\theta^j$ is not a principal derivative it is called a *parametric* derivative.

Note that the criterion for whether a frame derivative is principal (and hence also for parametric) respects permutation of the multi-index J defining it. Thus we could not have $\Delta_{12}\theta$ being principal and $\Delta_{21}\theta$ being parametric.

Consider for example, the involutive system (4.32). The principal derivatives are $\Lambda_1\theta$, $\Lambda_2\theta$, $\Lambda_4\theta$ (which occur on the left hand side) and their derivatives $\Lambda_{11}\theta$ $\Lambda_{12}\theta$, $\Lambda_{21}\theta$, etc. The parametric derivatives are θ, $\Lambda_3\theta$, $\Lambda_{33}\theta$, ...

Frame Riquier theory

The importance of involutive systems of p.d.e.'s is that there is available a theory for existence of a unique solution in the neighbourhood of initial data obtained by specifying each parametric derivative at a point. The theory is due to Riquier [67] (see also Reid [55, 56]). We state the principal existence theorem in a restricted form, sufficient for our analysis of determining equations.

Theorem 4.3.15. *Let a linear homogeneous involutive system DQ in independent variables w and dependent variables θ be given in a* coordinate *frame. Let w_0 be a point at which the coefficient functions are analytic. If the parametric derivatives of DQ are of finite number r, and values of these parametric derivatives are specified at w_0, there exists a unique analytic solution of DQ in a neighbourhood of w_0 satisfying these initial conditions. In particular the system has an r-dimensional solution space. If the parametric derivatives are not finite in number the system has an infinite-dimensional solution space.*

Briefly put, the solution space dimension of an involutive system is equal to the number of parametric derivatives in the system.

Example 4.3.16. Consider a linear homogeneous system in two dependent variables (ξ, τ), and three independent variables (x, t, u):

$$\begin{aligned} \xi_{tt} &= 0 & \qquad \tau_{tt} &= 0 \\ \xi_x &= \tfrac{1}{x}\xi + \tau_t & \qquad \tau_x &= 0 \\ \xi_u &= 0 & \qquad \tau_u &= 0. \end{aligned}$$

This system is involutive, and has as parametric derivatives τ, τ_t, ξ, ξ_t. The Riquier theorem therefore asserts that at a point (x_0, t_0, u_0) with $x \neq 0$, we may specify $\tau(x_0, t_0, u_0) = c_1$, $\tau_t(x_0, t_0, u_0) = c_2$, $\xi(x_0, t_0, u_0) = c_3$ and $\xi_t(x_0, t_0, u_0) = c_4$ arbitrarily: associated with each choice of $c_1, \ldots, c_4$ is a unique solution of the system. The solution space is therefore four-dimensional. The general solution is in fact $\tau(x, t, u) = c_1 + c_2 t$, $\xi(x, t, u) = c_2 x \log x + c_3 x + c_4 xt$, where we have taken $x_0 = 1$, $t_0 = 0$, $u_0 = 0$ as a suitable initial data point. The main point is that this explicit solution is not needed to count the solution space dimension.

In its full generality, the theory is not restricted to linear systems, and a careful enumeration of initial data sufficient to guarantee existence and uniqueness for the infinite dimensional case is also performed. Reid [56] gives details, examples and computational algorithms for his variant of this process. Involutivity is essential for establishing uniqueness in the Riquier theorem. The criterion of involutivity depends explicitly on the ordering of derivatives chosen. In turn, the objects ordered, namely $\frac{\partial^K \theta^j}{\partial (w^1)^{i_1} \partial (w^2)^{i_2} \cdots \partial (w^\nu)^{i_\nu}}$ explicitly depend on the coordinate system w employed. The order relation makes no sense if a change of variables is executed: a new ordering must be devised, written in terms of the new variables.

Now suppose that a system is referred to a moving frame. An ordering of frame derivatives is devised, and the system brought to frame involutive form. Because it is referred to a coordinate system, the Riquier theory does not apply to this frame involutive system. We could attempt to circumvent this by (notionally) 'translating' the frame involutive system back into a system of p.d.e.'s and then applying the Riquier theory. However, a frame involutive system will not be in involutive form when thus translated, since involutivity of p.d.e.'s is defined in terms not of frame derivatives but partial derivatives with respect to the coordinate system. In particular, the system will not be in solved form, and restoration of solved form requires an ordering of partial derivatives distinct from the ordering used in the original frame system. Hence a proof of the following 'frame Riquier' theorem is essential if we are to extract information on solution space dimension directly from the frame involutive system.

Conjecture 4.3.17. *Let a linear homogeneous frame involutive system DQ in the dependent variables θ be given, referred to a moving frame Δ. Let w_0 be a point at which the coefficient functions are analytic. Partition the parametric frame derivatives of DQ into equivalence classes under permutation. If such equivalence classes are of finite number r, and values of one parametric derivative in each class are specified at w_0, there exists a unique solution of DQ in a neighbourhood of w_0 satisfying these initial conditions. In particular the system has an r-dimensional solution space. If the parametric derivatives are not finite in number the system has an infinite-dimensional solution space.*

We shall critically rely on this presumed result in the following material. Note that the result is stated in terms of equivalence classes of derivatives. Suppose we have $\Delta_{12}\theta$ and $\Delta_{21}\theta$ as parametric derivatives. These are not independent, since $\Delta_{21}\theta = \Delta_{12}\theta + \gamma^k_{12}\Delta_k\theta$. Hence we can prescribe only one of them as initial data, the other being determined in terms of it.

For systems DQ of determining equations for a Lie symmetry algebra, Reid [57] showed how to find structure constants of the algebra by Taylor expansion. This idea was subsequently improved in [58]. *Assuming* the frame Riquier existence conjecture, Appendix B gives an elegant and algorithmic way to find the structure constants of the symmetry algebra from the frame involutive form, without solving the determining system. We use this in our examples.

4.4 Invariant frame

The frame Reid method described above applies to any moving frame. We now show how to choose a frame in which calculations become particularly simple. This is achieved by requiring the frame to be invariant under the action of the (augmented) equivalence group. Despite the obvious geometric flavour of all the following material, we refrain from overtly using geometric concepts such as pullbacks, induced maps, sections of bundles, etc. Instead we state results from Ovsiannikov [52, §24], who uses analytic methods and terminology. Note that Ovsiannikov's Lemma 24.2 incorrectly asserts that invariant operators constitute a Lie algebra over the 'field' of invariant functions because they form a vector space which is closed under commutation. This is false, since commutation does not distribute linearly over scalar multiplication by an element of this field.

4.4.1 Augmented frame

As in Example 4.3.2 the frames we use depend upon arbitrary elements $\phi(w)$. The following discussion applies to any set of independent and dependent variables, with extension to derivatives. Our independent variables are $w = (x, u)$ (the ν independent variables in the determining equations or the constraining system); our dependent variables are the μ coordinates $a = \phi(w)$ of arbitrary element space. First we introduce some terminology and notation.

Definition 4.4.1. We define

$$\underset{k}{z} = (w, a, \underset{1}{a}, \ldots, \underset{k}{a}) \tag{4.35}$$

to be the collection of independent and dependent variables and derivatives up to order k.

This notation is convenient because $(w, a, \underset{1}{a}, \ldots, \underset{k}{a})$ occurs so frequently.

Definition 4.4.2. A real-valued function $f(\underset{k}{z})$ of independent and dependent variables, and derivatives up to order k will be called a (k-th order) *differential function.*

If the order k is not specified, we understand f to be of arbitrary finite order.

Definition 4.4.3. Let $g(w, a, \underset{1}{a}, \ldots, \underset{k}{a})$ be a k-th order differential function. If the dependent variables $a = \phi(w)$ are assigned as a function of the independent variables we define the function $\phi^* g$ by

$$\phi^* g\,(w) = g(w, \phi(w), \underset{1}{\phi}(w), \ldots, \underset{k}{\phi}(w)) \tag{4.36}$$

This notation was used earlier in §3.1.2.

In Example 4.3.2 we used a moving frame with vector fields such as $\Delta_4 = \frac{1}{D(u)}\,\partial_u$. It is natural to introduce a coordinate $a = D(u)$ for diffusivity space and to write this as $\Delta_4 = \frac{1}{a}\,\partial_u$. This is misleading notation for the following reason. Action of ∂_u on a function $D(u)$ gives $\dot{D}(u)$. However, action of ∂_u on the differential function a gives 0. Clearly this is because the *total* derivative operator $\hat{D}_u$ is appropriate here, and we should define $\Delta_4 = \frac{1}{a}\hat{D}_u$, so that Δ_4 'sees' a as a function of u.

Proposition 4.4.4. *Let Δ be the differential operator*

$$\Delta = g^i(\underset{k}{z})\hat{D}_{w^i}$$

where $\hat{D}_{w^i}$ is the total derivative operator

$$\hat{D}_{w^i} = \partial_{w^i} + a^j_i \partial_{a^j} + \cdots + a^j_{Ji}\partial_{a^j_J} + \cdots .$$

Let $\phi^*\Delta$ be the vector field on w space

$$\phi^*\Delta = \phi^* g\,(w)\partial_{w^i} \tag{4.37}$$

Then Δ and $\phi^*\Delta$ agree in their actions on differential functions in the sense that

$$(\phi^*\Delta)(\phi^* f)(w) = \phi^*(\Delta f)(w) \tag{4.38}$$

for any differential function f.

First we clarify the nature of the various terms in (4.38): f and Δf are both differential functions; after inserting $a = \phi(w)$, $\phi^*(\Delta f)$ is a function of w alone. On the left hand side, $\phi^* f$ is a function of w; $\phi^*\Delta$ is a vector field on w space, so $(\phi^*\Delta)(\phi^* f)$ is a function of w.

Proof.

$$\Delta f = g^i(\underset{k}{z})\,(\hat{D}_{w^i} f)(\underset{l}{z})$$

so

$$\phi^*(\Delta f)(w) = (\phi^* g^i)(w)(\phi^*(\hat{D}_{w^i} f))(w)$$

and by the fundamental property (Proposition 2.2.4) of total derivatives,

$$= (\phi^* g^i)(w)\partial_{w^i}(\phi^* f)(w).$$

By definition (4.38) of $\phi^*\Delta$, this equals $(\phi^*\Delta)(\phi^* f)(w)$. □

We are primarily concerned with moving frames of the following form.

Definition 4.4.5. An *augmented moving frame* Δ with respect to independent variables $(w^1, w^2, \ldots, w^\nu)$ and dependent variables a is an ordered set of ν differential operators of the form

$$\Delta_i = g_i^j(\underset{k}{z})\hat{D}_{w^j} \tag{4.39}$$

such that the matrix $G(\underset{k}{z}) = [\, g_i^j(\underset{k}{z})\,]$ is nonsingular for all values of $\underset{k}{z}$.

Once arbitrary elements $a = \phi(w)$ are assigned as functions of the independent variables, we obtain the moving frame $\phi^*\Delta$:

$$\phi^*\Delta_i = \phi^* g_i^j(w)\partial_{w^j}. \tag{4.40}$$

Although the frames before and after assignment of arbitrary elements are conceptually distinct entities, in examples we wantonly confuse the two. Such notational abuse is possible because of Proposition 4.4.4, and is true to the Leibnizian tradition of confusing a function with its value. In a calculation we write ϕ, $\underset{1}{\phi}$ and so on, manipulating them as coordinates, but 'imagining' that they are functions of w. Comment on this situation would not be necessary if it were not that in other calculations (e.g. §3.3) it is essential that $a = \phi(w)$ *not* be imagined as functions of w. Since we are now always imagining a to be a function of w, the only relevant 'derivative with respect to u' is $\hat{D}_u$, (i.e., ∂_u plays no role). It is perverse to continue using total derivative notation in this case, and from now on we write ∂_u when $\hat{D}_u$ is meant.

Example 4.4.6. Consider the augmented frame (4.25)

$$\begin{aligned}
\Lambda_1 &= a^{1/2}(\hat{D}_x + u\hat{D}_v)\\
\Lambda_2 &= \dot{a}a^{-3/2}\hat{D}_x + a^{-3/2}(u\dot{a} + 2a)\hat{D}_v\\
\Lambda_3 &= \hat{D}_t + \dot{b}\hat{D}_x + (u\dot{b} - b)\hat{D}_v\\
\Lambda_4 &= \tfrac{1}{a}\hat{D}_u
\end{aligned} \tag{4.41}$$

on (x, t, u, v) space, with dependent variables (a, b), and $\dot{a} \equiv a_u$ etc. Assign $a = D(u)$, $b = K(u)$, and let $\phi = (D, K)$: this yields the frame

$$\begin{aligned}
\phi^*\Lambda_1 &= D(u)^{1/2}(\,\partial_x + u\,\partial_v)\\
\phi^*\Lambda_2 &= \dot{D}(u)D(u)^{-3/2}\,\partial_x + D(u)^{-3/2}(u\dot{D}(u) + 2D(u))\,\partial_v\\
\phi^*\Lambda_3 &= \partial_t + \dot{K}(u)\,\partial_x + (u\dot{K}(u) - b)\,\partial_v\\
\phi^*\Lambda_4 &= \tfrac{1}{D(u)}\,\partial_u
\end{aligned}$$

In fact our usual notation will be to write $\Lambda_1 = D^{1/2}(\,\partial_x + u\,\partial_v)$ etc., leaving it ambiguous whether the arbitrary elements have yet been assigned.

Mapping of vector field

A transformation $\tau\colon w' = \tau(w)$ of base space naturally induces an action on total derivative operators by

$$\hat{D}_{w^i} = \frac{\partial \tau^j}{\partial w^i}(w)\hat{D}_{w'^j} = \frac{\partial \tau^j}{\partial w^i} \circ \tau^{-1}(w')\hat{D}_{w'^j}.$$

Hence if $\hat{\tau}$ represents an action on $\underset{k}{z}$ space,

$$\begin{aligned} w' &= \tau(w) \\ a' &= \sigma(w,a) \end{aligned}$$

this transformation naturally induces an action on an augmented vector field $\Delta = g^i \hat{D}_{w^i}$ by

$$\hat{\tau}_* \Delta = g'^j(\underset{k}{z})\hat{D}_{w^j}$$

where

$$g'^j \circ \hat{\tau}(\underset{k}{z}) = g^i(\underset{k}{z})\frac{\partial \tau^j}{\partial w^i}(w). \tag{4.42}$$

The notation τ_* follows differential geometric conventions on mapping tangent vectors.

Example 4.4.7. Consider the augmented vector field

$$\mathbf{Y} = a^{1/2}(\hat{D}_x + u\hat{D}_v) \tag{4.43}$$

under the action of a transformation $\hat{\tau}\colon (v,x,t,u,a,b) \mapsto (v',x',t',u',a',b')$ given by

$$\begin{aligned} v &= \alpha v' + \beta x' \\ x &= \gamma v' + \delta x' \\ t &= t' \\ u &= \frac{\alpha u' + \beta}{\gamma u' + \delta} \\ a &= (\gamma u' + \delta)^2 a' \\ b &= \frac{b'}{\gamma u' + \delta}, \qquad \alpha\delta - \beta\gamma = 1. \end{aligned} \tag{4.44}$$

(These transformations constitute a three-parameter group $\hat{\mathcal{G}}^3$.) Then

$$\begin{aligned} \hat{D}_x &= \alpha \hat{D}_{x'} - \beta \hat{D}_{v'} \\ \hat{D}_v &= -\gamma \hat{D}_{x'} + \delta \hat{D}_{v'}, \end{aligned}$$

so

$$\begin{aligned} \mathbf{Y} &= (\gamma u' + \delta)a'^{1/2}\Big((\alpha\hat{D}_{x'} - \beta\hat{D}_{v'}) + \Big(\frac{\alpha u' + \beta}{\gamma u' + \delta}\Big)(-\gamma\hat{D}_{x'} + \delta\hat{D}_{v'})\Big) \\ &= a'^{1/2}(\hat{D}_{x'} + u'\hat{D}_{v'}) \end{aligned} \tag{4.45}$$

Thus the transformed vector field is

$$\hat{\tau}_* \mathbf{Y} = a^{1/2}(\hat{D}_x + u\hat{D}_v)$$

4.4.2 Invariant frame

The example above illustrates the following concept.

Definition 4.4.8. An augmented vector field $\mathbf{Y}$ is *invariant* under the action of a transformation $\hat{\tau}$ if

$$\hat{\tau}_* \mathbf{Y} = \mathbf{Y}$$

The definition asserts that the functions g'^j (4.42) are identical to g^j in the original vector field. That is

$$g^j \circ \hat{\tau}(\underset{k}{z}) = \partial_{w^j} \tau^i(w)\, g^j(\underset{k}{z}) \tag{4.46}$$

for all points $\underset{k}{z}$. In practical terms, we take a vector field $\mathbf{Y} = g^j(\underset{k}{z}) D_{w^j}$ and express it in terms of new variables $(w', a', \underset{1}{a}', \ldots, \underset{k}{a}')$ as

$$\mathbf{Y} = g'^j(\underset{k}{z}') D_{w'^j}.$$

The vector field is invariant if g'^j are the same functions of $\underset{k}{z}'$ as g^j are of $\underset{k}{z}$.

Example 4.4.7. **(cont.)** The vector field $\mathbf{Y}$ (4.43) is invariant under the action of transformations (4.44). Its expression (4.45) in dashed variables is identical to its expression in terms of the original variables, so that $\hat{\tau}_* \mathbf{Y} = \mathbf{Y}$.

Now that invariance of a vector field under *one* transformation has been defined, we define invariance of a frame under a transformation group by requiring that each vector field in the frame is invariant under each transformation in the group.

Definition 4.4.9. Let $\hat{\mathcal{Q}}$ be a group of augmented transformations $\hat{\tau}(\varepsilon)$:

$$\begin{aligned} w' &= \tau(w; \varepsilon) \\ a' &= \sigma(w, a; \varepsilon) \end{aligned}$$

An augmented frame Δ is invariant under the action of $\hat{\mathcal{Q}}$ if

$$\hat{\tau}_*(\varepsilon)\Delta_i = \Delta_i$$

for each Δ_i, $i = 1, 2, \ldots, \nu$, and all transformations $\hat{\tau}(\varepsilon) \in \hat{\mathcal{Q}}$.

Example 4.4.10. Consider the augmented frame (4.20) under the action of the two-parameter group $\hat{\mathcal{G}}^2$

$$\begin{aligned} v &= v' + \mu t' & u &= u' \\ x &= x' + \varepsilon t' & D &= D' \\ t &= t' & K &= K' + \varepsilon u' - \mu \end{aligned} \tag{4.47}$$

(We are here freely confusing notations for frames before and after assigning arbitrary elements $a = D(u)$, $b = K(u)$.) We find

$$\begin{aligned} \Delta_1 &= \partial_{v'} \\ \Delta_2 &= \partial_{x'} \\ \Delta_3 &= \partial_{t'} + \dot{K}' \partial_{x'} + (u' \dot{K}' - K')\partial_{v'} \\ \Delta_4 &= \partial_{u'} \end{aligned}$$

so that Δ takes the same form in the new variables, whatever values are taken by the group parameters ε, μ. Hence Δ is an invariant frame with respect to this group action.

Similarly, we note that the frame Λ (4.25) is invariant under the action of both $\hat{\mathcal{G}}^2$ and the three-parameter group $\hat{\mathcal{G}}^3$ (4.44).

4.4.3 Differential invariants

In addition to invariant frame differential operators, we require the following more familiar concept:

Definition 4.4.11. Let $f(\underset{k}{z})$ be a (k-th order) differential function. Let $\hat{\mathcal{Q}}$ be a transformation group acting on $z = (w, a)$. If

$$f(\hat{\tau}(\underset{k}{z})) = f(\underset{k}{z})$$

for all $\underset{k}{z}$, and all transformations $\hat{\tau} \in \hat{\mathcal{Q}}$, then f is a (k-th order) *differential invariant* of $\hat{\mathcal{Q}}$.

This concept was used earlier in §3.4.1. We may write more briefly $f \circ \hat{\mathcal{Q}} = f$.

Example 4.4.12. Consider the two-parameter group $\hat{\mathcal{G}}^2$ (4.47) acting on $(x, t, u, v; D, K)$ space. Clearly u is a differential invariant of $\hat{\mathcal{G}}^2$. In addition, D and its derivatives $\dot{D}$, $\ddot{D}$, ... are differential invariants, as are $\ddot{K}$, $\dddot{K}$. We define $J := \ddot{K}$.

A less trivial calculation of invariants is obtained from the action of the three-parameter group $\hat{\mathcal{G}}^3$ (4.44). Suppose we seek invariants of the five-parameter group $\hat{\mathcal{G}}^5$ obtained by composing $\hat{\mathcal{G}}^2$ and $\hat{\mathcal{G}}^3$. Action of $\hat{\mathcal{G}}^3$ on (v, x, t, u, D, K) induces action on derivatives $\dot{D}$, $\dot{K}$ etc. and hence on the differential invariants of $\hat{\mathcal{G}}^2$ just noted. We find

$$\begin{aligned}
u &= \frac{\alpha u' + \beta}{\gamma u' + \delta} \\
D &= (\gamma u' + \delta)^2 D' \\
\dot{D} &= (\gamma u' + \delta)^3 \big((\gamma u' + \delta)\dot{D}' + 2\gamma D'\big) \\
\ddot{D} &= (\gamma u' + \delta)^4 \big((\gamma u' + \delta)^2 \ddot{D}' + 6\gamma(\gamma u' + \delta)\dot{D}' + 6\gamma^2 D'\big) \\
J &= (\gamma u' + \delta)^3 J'.
\end{aligned}$$

With this, we note for instance that

$$\begin{aligned}
I &:= |J| D^{-3/2} = |\ddot{K}| D^{-3/2} \\
L &:= \frac{D\ddot{D} - 3/2\dot{D}^2}{D^4}
\end{aligned} \tag{4.48}$$

are differential invariants—not only of $\hat{\mathcal{G}}^3$ (4.44), but also of $\hat{\mathcal{G}}^2$ (4.47). Hence I and L are differential invariants of $\hat{\mathcal{G}}^5$.

If an invariant frame for a group $\hat{\mathcal{Q}}$ is known, certain differential invariants of $\hat{\mathcal{Q}}$ are immediately available:

Proposition 4.4.13. *Let Δ be an augmented frame, invariant under the action of a group $\hat{\mathcal{Q}}$. The structure functions γ_{ij}^k*

$$[\Delta_i, \Delta_j] = \gamma_{ij}^k \Delta_k$$

are differential invariants of $\hat{\mathcal{Q}}$.

Tresse [68] first noted this property. A more modern discussion of differential invariants and invariant operators is given in [52, §24].

Example 4.4.14. As noted above, the frame Δ (4.20) is invariant under the action of the group $\hat{\mathcal{G}}^2$ (4.47). Its structure relations (4.23) have all commutators vanishing except $[\Delta_3, \Delta_4] = -J(\Delta_2 + u\Delta_1)$. The coefficients are expressed in terms of u and $J \equiv \ddot{K}$, which are invariants of $\hat{\mathcal{G}}^2$ (see above).

Now consider the frame Λ (4.25), which is invariant under the five-parameter group $\hat{\mathcal{G}}^5$ obtained by composition of $\hat{\mathcal{G}}^2$ and $\hat{\mathcal{G}}^3$ (4.44). The coefficients in the structure relations (4.27) for Λ are constants except $[\Lambda_2, \Lambda_4] = -L\Lambda_1$, and $[\Lambda_3, \Lambda_4] = -I\Lambda_1$. Both L and I are invariants (4.48) of the group $\hat{\mathcal{G}}^5$.

Once one has found a differential invariant, the invariant frame provides a means for generating *additional* invariants:

Proposition 4.4.15. *If J is a differential invariant of a group $\hat{\mathcal{Q}}$, and Δ is an invariant augmented vector field, then ΔJ is also a differential invariant of $\hat{\mathcal{Q}}$.*

Generally if J is a k-th order invariant, ΔJ is of order $k+1$, although it can happen that ΔJ vanishes or is constant.

Example 4.4.16. Consider the action of the operator Λ_4 (4.41) on the invariant L (4.48). We are assured that $\Lambda_4 L$ is an invariant of the group $\hat{\mathcal{G}}^5$. Tracing the definition of Λ_4, we find $\Lambda_4 = \frac{1}{D}\hat{D}_u$, and we compute

$$\Lambda_4 L = \frac{D^2\dddot{D} - 6D\dot{D}\ddot{D} + 6\dot{D}^3}{D^6}.$$

It may be directly verified that this is a third order differential invariant of $\hat{\mathcal{G}}^5$.

In practice there is no gain in expanding an expression such as $\Lambda_4 L$: the point is to manipulate the invariants of $\hat{\mathcal{G}}^5$ as painlessly as possible, and this is achieved by treating $\Lambda_4 L$ as an entity in its own right.

4.4.4 Tresse basis

Generally differential invariants and invariant frames may not be defined at all points. The following material may be found in Eisenhart [22] or Ovsiannikov [52, §24].

Definition 4.4.17. Let an r-parameter group on an N-dimensional space y be generated by a Lie algebra with basis $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_r$, with $\mathbf{X}_i = \xi_i^j(y)\partial_{y^j}$. Define the $r \times N$ matrix $\Xi(y) = [\,\xi_i^j(y)\,]$ of infinitesimals. The *rank* of the system of operators at a point y is defined to be $\rho(y) = \operatorname{rank}\Xi(y)$. It is a property of the group action and is independent of the basis chosen for the Lie algebra of operators.

The rank $\rho(y)$ is in fact the dimension of the group orbit passing through y. Clearly we have $\rho(y) \le r$ at all points y. Hence $\rho(y)$ attains a maximum value $\bar{\rho}$. It is easily shown that if $\rho(y_0) = \bar{\rho}$ then $\rho(y) = \bar{\rho}$ for all y in some neighbourhood of y_0, since we always assume ξ_i^j smooth. Indeed if ξ_i^j are analytic, the maximum rank $\bar{\rho}$ is attained 'generically', i.e. $\rho(y) < \bar{\rho}$ on sets of dimension strictly less that n.

Definition 4.4.18. A point y at which $\rho(y) = \bar{\rho}$ is called *regular*. If $\rho(y) < \bar{\rho}$, we call the point y *singular*.

The generic rank $\bar{\rho}$ gives a count of the invariants of the group action:

Proposition 4.4.19. *Let a group $\mathcal{G}$ act on a space y with dimension N. In the neighbourhood of a regular point y there exist exactly $t = N - \bar{\rho}$ invariants of $\mathcal{G}$.*

Our interest is in group action on a base space (w, a) and its extensions to $(w, a, \underset{1}{a}, \dots, \underset{k}{a})$. Extension of the group appends columns to the matrix $\Xi(w, a)$, leading to a sequence of generic ranks

$$\bar{\rho}_0, \bar{\rho}_1, \dots, \bar{\rho}_k$$

for each order of extension. If the k-times extended space has dimension N_k, Proposition 4.4.19 shows there are $t_k = N_k - \bar{\rho}_k$ invariants of the k times extended space. The dimension of the extension spaces is unbounded as $k \to \infty$, while the ranks $\bar{\rho}_k$ are bounded above since $\bar{\rho}_k \le r$, so it follows that the number of differential invariants t_k is unbounded as the order of extension $k \to \infty$. However Proposition 4.4.15 can be used to generate a sequence of differential invariants ΔJ, $\Delta^2 J$, ... from one invariant J. We may reasonably hope to generate all differential invariants of a group by application of invariant frame operators to a finite number of such invariants. This is indeed the case.

Theorem 4.4.20 (Tresse basis). [52, §24]

(i) *For every r-parameter group $\hat{\mathcal{Q}}$ acting on (w, a) space, there is a finite order χ such that the generic rank $\bar{\rho}_\chi$ of the χ times extended group exactly equals r.*

(ii) *Let $\underset{\chi}{z}$ be a regular point of $\hat{\mathcal{Q}}$ in the χ times extended space $\underset{\chi}{z} = (w, a, \underset{1}{a}, \dots, \underset{\chi}{a})$. Then in the neighbourhood of $\underset{\chi}{z}$ there exists an augmented frame Δ,*

$$\Delta_i = g_i^j(\underset{\chi}{z}) D_{w^j}$$

invariant under the action of $\hat{\mathcal{Q}}$.

(iii) *(Tresse basis) Every differential invariant of a group $\hat{\mathcal{Q}}$ may be obtained by application of invariant frame operators Δ_i to the differential invariants of order $\le \chi + 1$.*

The bound $\chi + 1$ in (iii) is not sharp, i.e., in some cases differential invariants of order lower than $\chi + 1$ may suffice.

The structure functions γ_{ij}^k of a $\hat{\mathcal{Q}}$-invariant frame Δ yield certain differential invariants (by Proposition 4.4.13). In many cases these are complete in the sense that every differential invariant of $\hat{\mathcal{Q}}$ is obtainable from γ_{ij}^k by application of the frame operators Δ_i. If this is so, the invariant frame encodes *all* of the invariant information of the group.

The results above were stated for *finite*-parameter Lie transformation groups. Most have analogous statements for infinite-parameter Lie groups, but it is beyond our scope to describe this theory: the Cartan equivalence method is specifically tailored for dealing with this case. Methods for calculation of differential invariants and invariant frames are canvassed in Ovsiannikov [52, §17,§24] and by Tresse [68]. Ovsiannikov covers only methods for constructing invariants from the infinitesimal operators of the group $\hat{\mathcal{Q}}$, which involves too many integrations to be useful in practice.

Instead, it is preferable first to construct the group action. A naive elimination of group parameters is then often practical to find the differential invariants and invariant frame. Alternatively the Cartan equivalence procedure may be used to find the frame (actually the dual coframe), and ultimately its invariants. We content ourselves with an example of the naive procedure.

Example 4.4.21. Consider the action of the group $\hat{\mathcal{G}}^2$ (4.47) on the coordinates u, D, K, along with extensions to $\dot{K}$, $\dot{D}$, etc.:

$$\begin{aligned} u &= u' \\ D &= D' \\ K &= K' + \varepsilon u' - \mu \\ \dot{K} &= \dot{K}' + \varepsilon. \end{aligned}$$

We may solve for the group parameters ε and μ as

$$\begin{aligned} \varepsilon &= \dot{K}' - \dot{K} \\ \mu &= K' - K + u'(\dot{K}' - \dot{K}). \end{aligned}$$

The action of $\hat{\mathcal{G}}^2$ on the coordinate frame $(\partial_v, \partial_x, \partial_t, \partial_u)$ is

$$\begin{aligned} \partial_{v'} &= \partial_v \\ \partial_{x'} &= \partial_x \\ \partial_{t'} &= \partial_t + \mu\,\partial_v + \varepsilon\,\partial_x \\ \partial_{u'} &= \partial_u. \end{aligned}$$

Hence ∂_v, ∂_x, ∂_u are invariant operators. Substituting for ε, μ from above gives, after some rearrangement

$$\partial_{t'} - K'\,\partial_v + \dot{K}'(\partial_x + u'\,\partial_v) = \partial_t - K\,\partial_v + \dot{K}(\partial_x + u'\,\partial_v).$$

Noting that $\partial_v = \partial_{v'}$, $\partial_x = \partial_{x'}$, and $u = u'$, this becomes

$$\partial_{t'} - K'\partial_{v'} + \dot{K}'(\partial_{x'} + u'\partial_{v'}) = \partial_t - K\,\partial_v + \dot{K}(\partial_x + u\,\partial_v),$$

which gives us the remaining invariant operator. Altogether we have derived a frame Δ (4.20), which is invariant under $\hat{\mathcal{G}}^2$.

This procedure is less elegant than the Cartan method (or Tresse's 'reduced forms'). Its principal disadvantage is that much calculation is duplicated, since everything is found in both primed and unprimed coordinates. Nevertheless it is surprisingly effective for finite-parameter groups which are not too large.

4.5 Symmetry classification

Reid's algorithm for bringing determining equations DQ to involutive form is effective when DQ contains arbitrary elements, i.e., when $DQ(\phi)$ is derived from a d.e. $E(\phi)$ drawn from some class. Bringing the determining system to orthonomic form requires division by coefficients of the leading derivatives, which may now depend on the arbitrary elements a, $\underset{1}{a}$, ... For example, we might have

$$\dot{a}\,\partial_u \xi - \xi = 0.$$

Whether $\partial_u \xi$ can be isolated depends on whether the arbitrary elements are such as to make this coefficient vanish. For example, if $\dot{a} \neq 0$ we find $\partial_u \xi = 1/\dot{a}\,\xi$, whereas if the 'pivot' $\dot{a}$ vanishes identically, the equation becomes $\xi = 0$. To effect a complete classification, every such branch must be pursued until involutive form is attained. The Riquier theory then yields the dimension of the Lie symmetry algebra, and the method of Appendix B gives its commutation relations.

This process requires minor modification when we refer the Reid method to a moving frame Δ. The auxiliary system A must be restated in terms of the frame Δ: it is originally written in terms of w, a and derivatives with respect to w, namely $\hat{D}_{w^j}$. When a frame Δ is introduced, we replace the operators $\hat{D}_{w^j}$ in A by their expressions in terms of Δ, so that A becomes a frame system in Δ. Note that in the determining equations DQ, the dependent variables ζ^j are affected by a change of frame, since they are components $\zeta^j \partial_{w^j}$ of a vector field. However, the dependent variables a in A are *scalars*, and are unaffected by the change to Δ. We denote the collection of κ-th order frame derivatives of a by $\Delta^\kappa a$.

The vital classification step occurs when isolating a frame derivative on the left hand side of an equation in DQ. Suppose we attempt to isolate a derivative $\Delta_J \theta^j$, and to do so requires division by a coefficient $H(w, a, \Delta a, \dots, \Delta^\kappa a)$, which we follow Reid [55, 57] in calling a *pivot*. To effect division requires knowledge of whether or not the pivot vanishes. At the beginning of the classification, we have some information about a, namely that it satisfies the auxiliary frame system A. Substitution from this system may reveal definitively that a pivot vanishes. However, if the *classifying equation*

$$H(w, a, \Delta a, \dots, \Delta^\kappa a) = 0$$

is not an implication of A, a branching appears: we must separately attempt involutive form for the cases

(i) The arbitrary elements satisfy system A and the inequality

$$H(w, \Delta a, \dots, \Delta^\kappa a) \neq 0$$

(ii) The arbitrary elements satisfy the system A and

$$H(w, \Delta a, \dots, \Delta^\kappa a) = 0$$

obtained by adjoining the classifying equation to the original auxiliary system.

Thus we build up a *tree* of possibilities, accumulating a *classifying system* CQ—consisting of the original auxiliary frame system A along with additional classifying *frame equations* which have arisen—and a set CI of classifying *frame inequalities* which result from demanding that various pivots *not* vanish.

With appropriate modifications to the procedures presented in §A.1, the classification algorithm is capable of concise recursive definition. These modifications are given in Appendix A.2. Firstly, the classifying equations CQ and classifying inequalities CI must be made available to all procedures. Secondly, each process now has two possible returns: a 'successful' one (e.g., involutive form was achieved) and an 'indecisive' one (division by a pivot could not be resolved, and the process halted in an incomplete state). Assuming we have available a procedure *involutive* which reduces a frame system to involutive form and is modified in this way, we define a function *classify* recursively as

Algorithm 4.5.1. (classify)

function *classify*(DQ,CQ,CI)
INPUT: DQ ... frame determining system
CQ ... frame classifying system
CI ... classifying frame inequalities
OUTPUT: Nothing
SIDE EFFECT: Involutive form and corresponding classifying systems and inequalities for each leaf of the tree are printed out.

$DQ :=$ *involutive*(DQ,CQ,CI,*pivot*)
if *pivot* = (null) **then**
print(DQ,CQ,CI)
else
classify(DQ,CQ,$\{CI, pivot \neq 0\}$)
classify(DQ,$\{CQ, pivot = 0\}$,CI)
fi
end

This procedure concisely describes the generation of a classification tree, and mirrors the process used in hand calculation. Initially we invoke *classify* with DQ being the original frame determining system, CQ being the auxiliary frame system A, and CI being empty.

Our recursive generation of the tree is both more natural and more efficient than Reid's original statement [57] of the classification procedure. Reid [57] originally advocated division by coefficients as though they were nonzero, but retaining them in a pivot list. His procedure then restarts "from scratch but subject to one of the pivots being identically zero." Our calculation is restarted at the point where an unresolved division occurred, so repetition of calculations is avoided.

Unlike Reid's procedure, ours has not been implemented on a computer algebra system, although Appendix A may be regarded as an outline for such an implementation. When performing hand calculations, there is considerable scope for modifying the methods just described. As long as care is taken to respect an ordering of the derivatives, and not to execute circular chains of reasoning (substituting an equation into itself), the steps can be executed in almost any order desired. Typically one works with a simple subsystem of the determining system DQ, simplifying it as much as possible, computing its compatibility conditions and so forth. Later the remaining equations in the system are adjoined one by one. By doing this one can defer dealing with complicated equations until many simple equations are available. Typically also, we vary the ordering of derivatives used during the course of the calculation, attempting to defer for as long as possible division by troublesome coefficients.

4.5.1 Invariant form of group classification

We can now execute symmetry classification of a class of differential equations in an invariant manner, i.e., with each step being invariant with respect to the action of the equivalence group of the class. We simply execute the above classification algorithm referred to a frame invariant under the action of the equivalence group.

1. Derive the equivalence group $\hat{\mathcal{Q}}$ of the class.

2. Derive determining equations $DQ(\phi)$ for symmetries of an equation $\mathrm{E}(\phi)$.

3. Construct invariants and invariant augmented frame(s) of $\hat{\mathcal{Q}}$, along with their structure relations. (Different frames may be necessary for different arbitrary elements ϕ.)

4. Rewrite DQ in terms of the invariant frame, with invariant coefficients.

5. Rewrite the auxiliary system A in terms of invariants and frame operators.

6. Invoke the frame Reid classification procedure *classify* with the classifying system CQ initially equal to A, and CI initially null.

7. For each leaf of the resulting tree there is a frame involutive form of DQ: find the size and structure of the Lie symmetry algebra associated with these involutive DQ's.

This new method for symmetry classification is therefore a generalization of Reid's [57] to a case where an equivalence group is available. Equivalence information is built into the method through the invariant frame. Once the invariant frame is calculated, most of the hard work is over. In many cases completion to frame involutive form can be achieved by hand, even for systems requiring large amounts of computer time and memory in Reid's MAPLE implementation of his algorithm. This is presumably because a great deal of the symmetry information is in the equivalence group, so that factoring this out reduces the computational complexity. An especially useful feature of our new method is that, since it is expressed in terms of invariants of the equivalence group, the case splittings involved are likewise invariant. This means that two equations connected by an equivalence transformation must end up on the same branch of the classification tree. This drastically reduces the number of spurious case splittings generated by Reid's method, with consequent gains in interpretability of the tree.

Example 4.5.2. Before giving the results of a major classification calculation, we demonstrate the method on a very simple example. Consider the nonlinear diffusion equation

$$u_t = (D(u)u_x)_x. \tag{4.49}$$

The determining equations for a symmetry operator

$$\mathbf{Y} = \xi\,\partial_x + \tau\,\partial_t + \eta\,\partial_u$$

[,]	$\hat{\mathbf{X}}_1$	$\hat{\mathbf{X}}_2$	$\hat{\mathbf{X}}_3$	$\hat{\mathbf{X}}_4$	$\hat{\mathbf{X}}_5$	$\hat{\mathbf{X}}_6$
$\hat{\mathbf{X}}_1$	0	0	$\hat{\mathbf{X}}_1$	$\hat{\mathbf{X}}_1$	0	0
$\hat{\mathbf{X}}_2$	0	0	$2\hat{\mathbf{X}}_2$	0	0	0
$\hat{\mathbf{X}}_3$	$-\hat{\mathbf{X}}_1$	$-2\hat{\mathbf{X}}_2$	0	0	0	0
$\hat{\mathbf{X}}_4$	$-\hat{\mathbf{X}}_1$	0	0	0	0	0
$\hat{\mathbf{X}}_5$	0	0	0	0	0	$\hat{\mathbf{X}}_5$
$\hat{\mathbf{X}}_6$	0	0	0	0	$-\hat{\mathbf{X}}_5$	0

Table 4.7: Commutator table of equivalence algebra (4.51) of nonlinear diffusion equation (4.49).

are [52, eq.6.7.3]

$$\begin{aligned} \partial_x \tau = 0, \qquad & \partial_u \tau = 0 \\ \partial_u \xi = 0, \qquad & \partial_u^2 \eta = 0 \\ D(2\,\partial_x \xi - \partial_t \tau) - \dot{D}\eta &= 0 \\ D(2\,\partial_x\,\partial_u \eta - \partial_x^2 \xi) + 2\dot{D}\,\partial_x \eta + \partial_t \xi &= 0 \\ D\,\partial_x^2 \eta - \partial_t \eta &= 0. \end{aligned} \tag{4.50}$$

Here and throughout, it is understood that ∂_x, ∂_u etc. 'see' $D(u)$ as a function of u (i.e., they are really $\hat{D}_x$, $\hat{D}_u$ etc.). We leave ambiguous whether D refers to the coordinate a of diffusivity space or to the function $D(u)$.

We attempt to construct invariants and invariant frames of the six-parameter equivalence group generated by (3.66), which we rewrite here as

$$\begin{aligned} \hat{\mathbf{X}}_1 &= \partial_x \\ \hat{\mathbf{X}}_2 &= \partial_t \\ \hat{\mathbf{X}}_3 &= x\,\partial_x + 2t\,\partial_t \\ \hat{\mathbf{X}}_4 &= x\,\partial_x + 2a\,\partial_a \\ \hat{\mathbf{X}}_5 &= \partial_u \\ \hat{\mathbf{X}}_6 &= u\,\partial_u, \end{aligned} \tag{4.51}$$

where a renumbering and change of basis has been executed. The equivalence algebra structure is shown in Table 4.7. Note that the algebra is solvable, with the chain of normal subgroups

$$\{\hat{\mathbf{X}}_1\} \prec \{\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2\} \prec \cdots \prec \{\hat{\mathbf{X}}_1, \ldots, \hat{\mathbf{X}}_6\}.$$

Instead of attacking the whole equivalence group at once, we proceed in steps through this normal subgroup chain.

As we enlarge from a subgroup $\hat{\mathcal{H}}$ to the next largest group $\hat{\mathcal{G}}$, we require expressions for

1. A change of frame from an $\hat{\mathcal{H}}$-invariant frame to a $\hat{\mathcal{G}}$-invariant frame Δ.
2. The invariant infinitesimals for Δ, i.e., quantities θ such that a vector field $\mathbf{Y} = \theta^i \Delta_i$.
3. Invariants of $\hat{\mathcal{G}}$ in terms of those of $\hat{\mathcal{H}}$.
4. Structure relations of the $\hat{\mathcal{G}}$-invariant frame—expressed in terms of invariants of $\hat{\mathcal{G}}$.
5. The auxiliary system A in terms of invariants of $\hat{\mathcal{G}}$ and the frame Δ.
6. The determining system DQ written in terms of Δ, θ and the invariants of $\hat{\mathcal{G}}$.

Common translation symmetries

We treat the common symmetry operators $\hat{\mathbf{X}}_1 = \partial_x$, $\hat{\mathbf{X}}_2 = \partial_t$ together. They generate a group $\hat{\mathcal{G}}^2$

$$\begin{aligned} x' &= x + \kappa_1 \\ t' &= t + \kappa_2 \\ u' &= u \\ a' &= a. \end{aligned} \tag{4.52}$$

The coordinate frame ∂_x, ∂_t, ∂_u is invariant, and ξ, τ, η are invariant infinitesimals. The invariants are u, D, subject to the auxiliary system

$$\partial_x D = 0, \qquad \partial_t D = 0. \tag{4.53}$$

Note that the determining system (4.50) is already expressed in terms of these invariant quantities: in particular, x and t do not appear explicitly.

Boltzmann scaling

We now adjoin the scaling symmetry operator $\hat{\mathbf{X}}_3$, giving the three-parameter common symmetry group $\hat{\mathcal{G}}^3$. This has action

$$\begin{aligned} \partial_{x'} &= \lambda^{-1}\,\partial_x & & \\ \partial_{t'} &= \lambda^{-2}\,\partial_t & u' &= u \\ \partial_{u'} &= \partial_u & D' &= D \end{aligned}$$

on the invariants of $\hat{\mathcal{G}}^2$. Invariants (u, D), and even an invariant operator ∂_u are available, but the parameter λ cannot be completely eliminated: *there is no invariant frame for* $\hat{\mathcal{G}}^3$. The frame Reid method works regardless of the frame to which it is referred. Hence our calculations do not rely critically on the frame being invariant, and we just ignore $\hat{\mathbf{X}}_3$.

Scaling group—diffusivity

We adjoin the scaling operator $\hat{\mathbf{X}}_4$, giving a four-parameter subgroup $\hat{\mathcal{G}}^4$. This has action

$$\begin{aligned} \partial_{x'} &= \rho^{-1}\,\partial_x & u' &= u \\ \partial_{t'} &= \partial_t & D' &= \rho^2 D \\ \partial_{u'} &= \partial_u & \dot{D}' &= \rho^2 \dot{D} \end{aligned}$$

on the invariants of $\hat{\mathcal{G}}^2$. We easily find an invariant frame Δ, with dual infinitesimals θ, and invariants u, I:

$$\begin{array}{lll} \Delta_1 = D^{1/2}\,\partial_x & \theta^1 = D^{-1/2}\xi & \\ \Delta_2 = \partial_t & \theta^2 = \tau & \\ \Delta_3 = \partial_u & \theta^3 = \eta & I := \dfrac{\dot{D}}{D} \end{array} \tag{4.54}$$

The invariants u, I are subject to constraints

$$\begin{array}{ll} \Delta_1 u = 0 & \Delta_1 I = 0 \\ \Delta_2 u = 0 & \Delta_2 I = 0 \\ \Delta_3 u = 1, & \end{array} \tag{4.55}$$

with $\Delta_3 I = \dot{I}$ being free. The structure relations of the frame Δ are

$$[\Delta_1, \Delta_2] = 0, \qquad [\Delta_1, \Delta_3] = -\tfrac{1}{2} I \Delta_1, \qquad [\Delta_2, \Delta_3] = 0. \tag{4.56}$$

We do not show the determining system (4.50) in this frame.

Translation in u

We adjoin the translation operator $\hat{\mathbf{X}}_5 = \partial_u$, which has trivial action on Δ and I: hence Δ, θ, I are invariants. The only change from above is that u is removed from the list of invariants.

Scaling of u

Finally we adjoin the scaling operator $\hat{\mathbf{X}}_6 = u\,\partial_u$, which has action

$$\begin{array}{ll} \Delta_1' = \Delta_1 & \\ \Delta_2' = \Delta_2 & \\ \Delta_3' = \alpha^{-1}\Delta_3 & I' = \alpha^{-1} I \end{array}$$

on the invariants Δ, I. The calculation now splits into two cases.
Case a. $I \neq 0$.

Here we may divide by I to eliminate the parameter α. An invariant frame Γ, invariant infinitesimals ζ, and invariant J are easily found:

$$\begin{array}{lll} \Gamma_1 := \Delta_1 & \zeta^1 := \theta^1 & \\ \Gamma_2 := \Delta_2 & \zeta^2 := \theta^2 & \\ \Gamma_3 := I^{-1}\Delta_3 & \zeta^3 := I\theta^3 & J := \dot{I}/I^2, \end{array} \tag{4.57}$$

where we retain dot notation $\dot{I} \equiv \Delta_3 I$, since $\Delta_3 = \partial_u$. The invariant J is constrained by

$$\Gamma_1 J = 0, \qquad \Gamma_2 J = 0. \tag{4.58}$$

The structure relations for Γ are

$$[\Gamma_1, \Gamma_2] = 0, \qquad [\Gamma_1, \Gamma_3] = -\tfrac{1}{2}\Gamma_1, \qquad [\Gamma_2, \Gamma_3] = 0. \tag{4.59}$$

Determining system (4.50) becomes

$$\begin{array}{ll}\Gamma_1\zeta^2 = 0 & (\Gamma_1)^2\zeta^3 = \Gamma_2\zeta^3 \\ \Gamma_3\zeta^2 = 0 & \Gamma_1\Gamma_3\zeta^3 = \frac{1}{2}(\Gamma_1)^2\zeta^1 + (J-1)\Gamma_1\zeta^3 - \frac{1}{2}\Gamma_2\zeta^1 \\ \Gamma_3\zeta^1 = -\frac{1}{2}\zeta^1 & (\Gamma_3)^2\zeta^3 = J\Gamma_3\zeta^3 + (\Gamma_3 J)\zeta^3 \\ \Gamma_2\zeta^2 = 2\Gamma_1\zeta^1 - \zeta^3. & \end{array} \tag{4.60}$$

Case b. $I = 0$.

This is the linear equation case $\dot{D} = 0$. Here the parameter α cannot be eliminated, and Δ is as close to an invariant frame as we can manage. There are no invariants. Reduction to involutive form shows there are infinitely many symmetries. We leave aside this case and pursue Case **a**.

Reduction to involutive form

Now that the frame Γ has been introduced, we apply the frame Reid method described in §4.3.2 to bring (4.60) to involutive form. The system is already in reduced orthonomic form. We compute compatibility conditions, for example

$$\Gamma_2(\Gamma_3\zeta^2) - \Gamma_3(\Gamma_2\zeta^2) = \Gamma_2(0) - \Gamma_3(2\Gamma_1\zeta^1 - \zeta^3).$$

Structure relations (4.59) simplify the left hand side, giving

$$0 = 2\Gamma_3\Gamma_1\zeta^1 - \Gamma_3\zeta^3.$$

Note that the first term is a derivative of the leading derivative $\Gamma_3\zeta^1$. Implicit substitution for $\Gamma_3\zeta^1$ from (4.60) gives $\Gamma_3\zeta^3 = 0$, which may be appended to the system. Inserting this into the other equations in (4.60), we find

$$(\Gamma_3 J)\zeta^3 = 0,$$

so that $\Gamma_3 J$ is a pivot.

If $\Gamma_3 J \neq 0$, we have $\zeta^3 = 0$, and the system quickly collapses to an involutive form with three-dimensional solution space. The three symmetries are, of course, the common symmetries $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$. We therefore do not present this case.

If $\Gamma_3 J = 0$ (so that J is a constant), we continue computing compatibility conditions, ultimately bringing the system to the form

$$\begin{array}{lll}\Gamma_1\zeta^2 = 0 & (\Gamma_1)^2\zeta^1 = 2(1-J)\Gamma_1\zeta^3 & (\Gamma_1)^2\zeta^3 = 0 \\ \Gamma_2\zeta^2 = 2\Gamma_1\zeta^1 - \zeta^3 & \Gamma_2\zeta^1 = 0 & \Gamma_2\zeta^3 = 0 \\ \Gamma_3\zeta^2 = 0 & \Gamma_3\zeta^1 = -\frac{1}{2}\zeta^1 & \Gamma_3\zeta^3 = 0 \end{array} \tag{4.61}$$

along with

$$(3-4J)\Gamma_1\zeta^3 = 0,$$

so that $(3-4J)$ is a pivot.

If $J \neq 3/4$, we have $\Gamma_1\zeta^3 = 0$, and the system collapses to the involutive form

$$\begin{array}{lll}\Gamma_1\zeta^2 = 0 & (\Gamma_1)^2\zeta^1 = 0 & \Gamma_1\zeta^3 = 0 \\ \Gamma_2\zeta^2 = 2\Gamma_1\zeta^1 - \zeta^3 & \Gamma_2\zeta^1 = 0 & \Gamma_2\zeta^3 = 0 \\ \Gamma_3\zeta^2 = 0 & \Gamma_3\zeta^1 = -\frac{1}{2}\zeta^1 & \Gamma_3\zeta^3 = 0. \end{array}$$

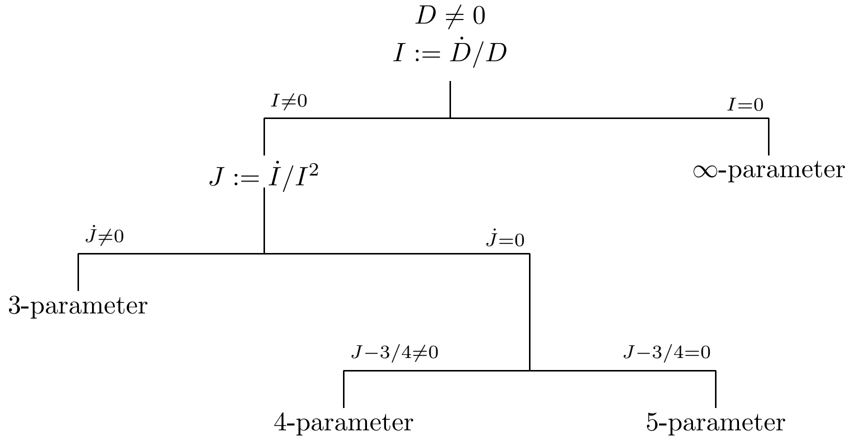

Figure 4.1: Classification tree for symmetries of nonlinear diffusion equation.

Here there are four parametric derivatives ζ^1, $\Gamma_1\zeta^1$, ζ^2, ζ^3, so there is a four parameter symmetry group. Applying the method of Appendix B, we find the symmetry algebra has structure

[,]	$\mathbf{Y}_1$	$\mathbf{Y}_2$	$\mathbf{Y}_3$	$\mathbf{Y}_4$
$\mathbf{Y}_1$	0	0	$\mathbf{Y}_1$	0
$\mathbf{Y}_2$	0	0	$2\mathbf{Y}_2$	$-\mathbf{Y}_2$
$\mathbf{Y}_3$	$-\mathbf{Y}_1$	$-2\mathbf{Y}_2$	0	0
$\mathbf{Y}_4$	0	$\mathbf{Y}_2$	0	0

If $J = 3/4$, system (4.61) is involutive. There are five parametric derivatives ζ^1, $\Gamma_1\zeta^1$, ζ^2, ζ^3, $\Gamma_1\zeta^3$, so there is a five-parameter symmetry group, whose commutation relations are easily found.

All in all we have generated the classification tree shown in Figure 4.1.

We make some remarks about this classification, comparing it with the results in [52, §6.7]. Firstly, the classification tree Figure 4.1 results from two kinds of splittings. The top branch is due to our method of constructing frames and occurs even before we consider determining equations. Subsequent branches are generated from the determining system by the frame Reid method.

Secondly, the case splittings revealed by the frame method agree with those of the usual classification. The invariant J has the expression $J = -(D/\dot{D})^{\cdot}$, so that the classifying equation $\Gamma_3 J = 0$ is $(D/\dot{D})^{\cdot\cdot} = 0$, in agreement with [52, eq.6.7.12]. The diffusivities satisfying this are $D(u) = (au + b)^m$, for which $J = m^{-1}$, and $D(u) = ae^{bu}$, for which $J = 0$. That e^u and u^m are not split in our classification tree reflects the fact that e^u is merely a limiting case of the power law diffusivities: $e^u = \lim_{m\to\infty}(1 + \frac{u}{m})^m$. The commutation relations we computed above are identical for all values of J. This fact is obscured in [52], where the inessential parameter m appears, and algebras for the u^m and e^u cases appear to be different.

Some remarks are in order on why the operator $\hat{\mathbf{X}}_3$ gave difficulty above. First we note that our failure in this case does not contradict Theorem 4.4.20(ii), which guarantees existence of an invariant frame only at *regular* points of the equivalence group action. However auxiliary system (4.53) specifies a locus of *singular* points.

The problem seems to be due to there being 'too much' symmetry. There is no difficulty in finding an invariant frame for the symmetry operators $\hat{\mathbf{X}}_1$, $\hat{\mathbf{X}}_2$. When the additional symmetry $\hat{\mathbf{X}}_3$ arises, its action on all the invariants u, D, $\dot{D}$, ... , is trivial, so there is no way to eliminate the group parameter λ. We expect difficulty whenever the rank of the system of common symmetry operators is less than the number of such operators, i.e., the symmetry group acts multiply transitively on its orbits. In this case we will not be able to find a frame which is invariant under the action of the equivalence group $\hat{\mathcal{Q}}$. Instead we may find an 'almost invariant' frame, on which $\hat{\mathcal{Q}}$ has nontrivial action. The residuum of group action presumably reflects the structure of the isotropy (stabilizer) subgroup of the common symmetry group. This phenomenon affects only our ability to find an invariant frame, and does *not* affect operation of the frame Reid algorithm.

Finally we note that for this simple example the overhead of computing and substituting for the invariant frames is scarcely worth the effort. Despite the cleaner appearance of the classification tree and commutation relations, the diffusion example is not difficult enough to justify use of such 'heavy machinery'.

4.5.2 Potential diffusion convection system

We now give a substantially more difficult computational example, applying the frame method to the diffusion convection potential system

$$\begin{aligned} v_x &= u \\ v_t &= Du_x - K \end{aligned} \tag{4.62}$$

with auxiliary system

$$\begin{aligned} \partial_v D &= 0 \qquad & \partial_v K &= 0 \\ \partial_x D &= 0 \qquad & \partial_x K &= 0 \\ \partial_t D &= 0 \qquad & \partial_t K &= 0, \end{aligned} \tag{4.63}$$

specifying permissible diffusivity and conductivity functions $D(u)$, $K(u)$. We make no notational distinction between D, K as coordinates and D, K as functions. Dot notation is used for derivatives $\partial_u D = \dot{D}$, $\partial_u K = \dot{K}$, etc. From the outset we impose the inequality $D \neq 0$: if $D = 0$ the equation ceases to be parabolic, and is not locally solvable.

Seeking a symmetry operator in the form

$$\mathbf{Y} = \chi\,\partial_v + \xi\,\partial_x + \tau\,\partial_t + \eta\,\partial_u$$

yields determining equations (4.28) which can be written in the suggestive form

$$
\begin{aligned}
&\partial_v \tau = 0 \\
&\partial_x \tau = 0 \\
&\partial_u \tau = 0 \qquad \partial_u \xi = 0 \qquad \partial_u \chi = 0 \\
&\dot{D}(\partial_x + u\,\partial_v)(\chi - u\xi) - 2D(\partial_x + u\,\partial_v)\xi + D\,\partial_t \tau = 0 \\
&(\partial_t + \dot{K}(\partial_x + u\,\partial_v) - K\,\partial_v)(\chi - u\xi) + K\,\partial_t\tau - D(\partial_x + u\,\partial_v)^2(\chi - u\xi) = 0 \\
&\eta = (\partial_x + u\,\partial_v)(\chi - u\xi).
\end{aligned} \tag{4.64}
$$

A 10-parameter equivalence group $\hat{\mathcal{Q}}$ (3.81) for the diffusion convection system (4.62) was found in §3.4. A basis for the Lie algebra $\hat{L}$ of $\hat{\mathcal{Q}}$ is given in (3.80). The commutation relations are shown in Table 4.8.

	$\hat{\mathbf{X}}_0$	$\hat{\mathbf{X}}_1$	$\hat{\mathbf{X}}_2$	$\hat{\mathbf{X}}_3$	$\hat{\mathbf{X}}_4$	$\hat{\mathbf{X}}_5$	$\hat{\mathbf{X}}_6$	$\hat{\mathbf{X}}_9$	$\hat{\mathbf{X}}_7$	$\hat{\mathbf{X}}_8$
$\hat{\mathbf{X}}_0$	0	0	0	0	0	0	$\frac{1}{2}\hat{\mathbf{X}}_0$	$-\hat{\mathbf{X}}_1$	$-\hat{\mathbf{X}}_0$	$\hat{\mathbf{X}}_0$
$\hat{\mathbf{X}}_1$	0	0	0	0	0	$\hat{\mathbf{X}}_0$	$-\frac{1}{2}\hat{\mathbf{X}}_1$	0	$-\hat{\mathbf{X}}_1$	$\hat{\mathbf{X}}_1$
$\hat{\mathbf{X}}_2$	0	0	0	$-\hat{\mathbf{X}}_0$	$\hat{\mathbf{X}}_1$	0	0	0	$-2\hat{\mathbf{X}}_2$	$\hat{\mathbf{X}}_2$
$\hat{\mathbf{X}}_3$	0	0	$\hat{\mathbf{X}}_0$	0	0	0	$\frac{1}{2}\hat{\mathbf{X}}_3$	$\hat{\mathbf{X}}_4$	$\hat{\mathbf{X}}_3$	0
$\hat{\mathbf{X}}_4$	0	0	$-\hat{\mathbf{X}}_1$	0	0	$-\hat{\mathbf{X}}_3$	$-\frac{1}{2}\hat{\mathbf{X}}_4$	0	$\hat{\mathbf{X}}_4$	0
$\hat{\mathbf{X}}_5$	0	$-\hat{\mathbf{X}}_0$	0	0	$\hat{\mathbf{X}}_3$	0	$\hat{\mathbf{X}}_5$	$2\hat{\mathbf{X}}_6$	0	0
$\hat{\mathbf{X}}_6$	$-\frac{1}{2}\hat{\mathbf{X}}_0$	$\frac{1}{2}\hat{\mathbf{X}}_1$	0	$-\frac{1}{2}\hat{\mathbf{X}}_3$	$\frac{1}{2}\hat{\mathbf{X}}_4$	$-\hat{\mathbf{X}}_5$	0	$\hat{\mathbf{X}}_9$	0	0
$\hat{\mathbf{X}}_9$	$\hat{\mathbf{X}}_1$	0	0	$-\hat{\mathbf{X}}_4$	0	$-2\hat{\mathbf{X}}_6$	$-\hat{\mathbf{X}}_9$	0	0	0
$\hat{\mathbf{X}}_7$	$\hat{\mathbf{X}}_0$	$\hat{\mathbf{X}}_1$	$2\hat{\mathbf{X}}_2$	$-\hat{\mathbf{X}}_3$	$-\hat{\mathbf{X}}_4$	0	0	0	0	0
$\hat{\mathbf{X}}_8$	$-\hat{\mathbf{X}}_0$	$-\hat{\mathbf{X}}_1$	$-\hat{\mathbf{X}}_2$	0	0	0	0	0	0	0

Table 4.8: Commutation relations for equivalence algebra (3.80) of diffusion convection potential system (4.62). A chain of normal subalgebras is outlined. The algebra is a semidirect sum of the diagonal blocks.

We seek to construct invariants and invariant frames for $\hat{\mathcal{Q}}$, writing the determining system (4.64) in terms of these. As above, we compute invariants of $\hat{\mathcal{Q}}$ in steps. We use a chain of normal subgroups

$$\hat{\mathcal{G}}^3 \prec \hat{\mathcal{G}}^5 \prec \hat{\mathcal{G}}^8 \prec \hat{\mathcal{G}}^9 \prec \hat{\mathcal{G}}^{10} = \hat{\mathcal{Q}}$$

of the equivalence group, corresponding to algebras of dimension 3, 5, 8, 9, 10 starting at the top left of Table 4.8. In this case the operators appended at each stage themselves form a subalgebra, i.e., $\hat{L}$ is a semidirect sum of the subalgebras

$$L^3\{\hat{\mathbf{X}}_0, \hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2\} \oplus_s L^2\{\hat{\mathbf{X}}_3, \hat{\mathbf{X}}_4\} \oplus_s L^3\{\hat{\mathbf{X}}_5, \hat{\mathbf{X}}_6, \hat{\mathbf{X}}_9\} \oplus_s L^1\{\hat{\mathbf{X}}_7\} \oplus_s L^1\{\hat{\mathbf{X}}_8\}.$$

After using the connected component of $\hat{\mathcal{Q}}$, we adjoin the discrete transformation R_2 (3.75).

As in Example 4.5.2, in enlarging from one subgroup to the next we must find: an invariant frame; invariant infinitesimals; differential invariants; structure relations; invariant auxiliary system; and frame determining system.

Common translation symmetries

The algebra $L^3\{\hat{\mathbf{X}}_0, \hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2\}$ generates the translation symmetries

$$\begin{aligned} v' &= v + \kappa_2 & D' &= D \\ x' &= x + \kappa_0 & K' &= K \\ t' &= t + \kappa_1 & & \\ u' &= u. & & \end{aligned} \tag{4.65}$$

The coordinate frame $\partial_v,\ \partial_x,\ \partial_t,\ \partial_u$ is invariant, as are the infinitesimals are χ, ξ, τ, η. The structure functions all vanish. The invariants u, D, K are subject to the constraining system (4.63), plus obvious properties such as $\partial_x u = 0,\ \partial_u u = 1$.

Galilean transformation

Now consider the group generated by $L^2\{\hat{\mathbf{X}}_3, \hat{\mathbf{X}}_4\}$:

$$\begin{aligned} v' &= v - \varepsilon t & D' &= D \\ x' &= x + \delta t & K' &= K + \mu u + \varepsilon \\ t' &= t & & \\ u' &= u. & & \end{aligned} \tag{4.66}$$

An invariant frame Δ and corresponding infinitesimals θ are given by

$$\begin{aligned} \Delta_1 &= \partial_v & \theta^1 &= \chi - (u\dot{K} - K)\tau \\ \Delta_2 &= \partial_x & \theta^2 &= \xi - \dot{K}\tau \\ \Delta_3 &= \partial_t - K\,\partial_v + \dot{K}(\,\partial_x + u\,\partial_v) & \theta^3 &= \tau \\ \Delta_4 &= \partial_u & \theta^4 &= \eta. \end{aligned} \tag{4.67}$$

The invariants of the group action are

$$u, \quad D, \quad J := \ddot{K} \tag{4.68}$$

The structure relations for Δ are

$$\begin{aligned} &[\Delta_1, \Delta_2] = 0 \qquad [\Delta_1, \Delta_3] = 0 \qquad [\Delta_1, \Delta_4] = 0 \\ & [\Delta_2, \Delta_3] = 0 \qquad [\Delta_2, \Delta_4] = 0 \\ & [\Delta_3, \Delta_4] = -uJ\Delta_1 \end{aligned} \tag{4.69}$$

The invariants u, D, J are subject to (from (4.63))

$$\begin{aligned} \Delta_1 u &= 0 & \Delta_1 D &= 0 & \Delta_1 J &= 0 \\ \Delta_2 u &= 0 & \Delta_2 D &= 0 & \Delta_2 J &= 0 \\ \Delta_3 u &= 0 & \Delta_3 D &= 0 & \Delta_3 J &= 0. \\ \Delta_4 u &= 1 & & & & \end{aligned} \tag{4.70}$$

The frame derivatives $\Delta_4 D$ and $\Delta_4 J$ are free. Since Δ_4 is ∂_u, we retain dot notation and write $\Delta_4 J = \dot{J}$ etc. Finally, determining equations (4.64) become

$$\begin{aligned}
&\Delta_1\theta^3 = 0 \\
&\Delta_2\theta^3 = 0 \\
&\Delta_4\theta^3 = 0 \qquad \Delta_4\theta^2 = -J\theta^3 \qquad \Delta_4\theta^1 = -uJ\theta^3 \\
&\dot{D}(\Delta_2 + u\Delta_1)(\theta^1 - u\theta^2) - 2D(\Delta_2 + u\Delta_1)\theta^2 + D\Delta_3\theta^3 = 0 \\
&\Delta_3(\theta^1 - u\theta^2) - D(\Delta_2 + u\Delta_1)^2(\theta^1 - u\theta^2) = 0 \\
&\theta^4 = (\Delta_2 + u\Delta_1)(\theta^1 - u\theta^2)
\end{aligned} \tag{4.71}$$

Group isomorphic to $SL_2(\mathbb{R})$

Now consider the group generated by $L^3\{\hat{\mathbf{X}}_5, \hat{\mathbf{X}}_6, \hat{\mathbf{X}}_9\}$:

$$\begin{aligned}
v' &= \alpha v + \beta x & D' &= (\gamma u + \delta)^2 D \\
x' &= \gamma v + \delta x & K' &= \frac{K}{\gamma u + \delta} \\
t' &= t \\
u' &= \frac{\alpha u + \beta}{\gamma u + \delta} & & \alpha\delta - \beta\gamma = 1,
\end{aligned} \tag{4.72}$$

acting on the coordinate frame by

$$\begin{aligned}
\partial_{v'} &= \delta\,\partial_v - \gamma\,\partial_x \\
\partial_{x'} &= -\beta\,\partial_v + \alpha\,\partial_x \\
\partial_{t'} &= \partial_t \\
\partial_{u'} &= (\gamma u + \delta)^2\,\partial_u.
\end{aligned}$$

The action on the quantities Δ, u, D, J is

$$\begin{aligned}
u' &= \frac{\alpha u + \beta}{\gamma u + \delta} & \Delta_1' &= \delta\Delta_1 - \gamma\Delta_2 \\
a' &= (\gamma u + \delta)^2 a & \Delta_2' &= -\beta\Delta_1 + \alpha\Delta_2 \\
J' &= (\gamma u + \delta)^3 J & \Delta_3' &= \Delta_3 \\
& & \Delta_4' &= (\gamma u + \delta)^2\Delta_4.
\end{aligned}$$

An invariant frame Λ is given by

$$\begin{aligned}
\Lambda_1 &= \pi D^{1/2}(\Delta_2 + u\Delta_1) \\
\Lambda_2 &= \pi D^{-3/2}\big(2D\Delta_1 + \dot{D}(\Delta_2 + u\Delta_1)\big) \\
\Lambda_3 &= \Delta_3 \\
\Lambda_4 &= \tfrac{1}{D}\Delta_4,
\end{aligned} \tag{4.73}$$

with corresponding infinitesimals λ, defined by

$$\begin{aligned}
\lambda^1 &= -\tfrac{1}{2}\pi D^{-3/2}\big(\dot{D}(\theta^1 - u\theta^2) - 2D\theta^2\big) \\
\lambda^2 &= \tfrac{1}{2}\pi D^{1/2}(\theta^1 - u\theta^2) \\
\lambda^3 &= \theta^3 \\
\lambda^4 &= D\theta^4.
\end{aligned} \tag{4.74}$$

The quantity π appearing throughout is a sign $\pm$. If $J \neq 0$ then we take $\pi = \operatorname{sgn} J$. If $J = 0$ this choice is impermissible, and we take $\pi = 1$. This is discussed below.

The invariants of the group action are

$$L := \frac{D\ddot{D} - 3/2\dot{D}^2}{D^4} \qquad I := |J|D^{-3/2} \tag{4.75}$$

The structure relations of the frame Λ are

$$\begin{array}{lll} [\Lambda_1, \Lambda_2] = 0 & [\Lambda_1, \Lambda_3] = 0 & [\Lambda_1, \Lambda_4] = -\frac{1}{2}\Lambda_2 \\ & [\Lambda_2, \Lambda_3] = 0 & [\Lambda_2, \Lambda_4] = -L\Lambda_1 \\ & & [\Lambda_3, \Lambda_4] = -I\Lambda_1 \end{array} \tag{4.76}$$

From (4.70) the invariants L, I are subject to

$$\begin{array}{ll} \Lambda_1 L = 0 & \Lambda_1 I = 0 \\ \Lambda_2 L = 0 & \Lambda_2 I = 0 \\ \Lambda_3 L = 0 & \Lambda_3 I = 0, \end{array} \tag{4.77}$$

with $\Lambda_4 L$ and $\Lambda_4 I$ unconstrained. Finally, determining equations (4.71) become

$$\begin{array}{lll} \Lambda_1\lambda^3 = 0 & (\Lambda_1)^2\lambda^2 = \Lambda_3\lambda^2 & \Lambda_1\lambda^1 = \frac{1}{2}\Lambda_3\lambda^3 \\ \Lambda_2\lambda^3 = 0 & \Lambda_4\lambda^2 = -\frac{1}{2}\lambda^1 & \Lambda_4\lambda^1 = -L\lambda^2 - I\lambda^3 \\ \Lambda_4\lambda^3 = 0 & & \end{array}$$
$$\lambda^4 = 2\Lambda_1\lambda^2 \tag{4.78}$$

In this beautiful form only two terms have nonconstant coefficients, and the simplicity of structure of the determining system (4.64) is revealed.

Consider the sign π which appeared above in (4.73), (4.74). If $J \neq 0$, we may choose $\pi = \operatorname{sgn} J$, and the frame Λ is then invariant under the action of the whole SL_2 subgroup (4.72). However, if $J = 0$, Λ_1, Λ_2 (4.73) change sign under action of (4.72), so that (4.73) is not invariant. This is because when $I = 0$ (pure diffusion) the transformation $x \mapsto -x$, $v \mapsto -v$ becomes a reflection *symmetry*. In this case we set $\pi = 1$ and continue, with Λ not invariant. We should give this as a case splitting now, but it appears immediately below, so we don't bother.

Scaling group—convection

Most of the hard work is now over, and things begin to become interesting. Consider the group generated by $L^1\{\hat{\mathbf{X}}_7\}$:

$$\begin{array}{ll} v' = \mu^{-1}v & D' = D \\ x' = \mu^{-1}x & K' = \mu K \\ t' = \mu^{-2}t & \mu > 0. \\ u' = u & \end{array} \tag{4.79}$$

This has action on Λ, L, I given by

$$\begin{array}{ll} \Lambda_1' = \mu\Lambda_1 & L' = L \\ \Lambda_2' = \mu\Lambda_2 & I' = \mu I \\ \Lambda_3' = \mu^2\Lambda_3 & \\ \Lambda_4' = \Lambda_4. & \end{array}$$

We must now consider the possibilities **a.** $I \neq 0$ **b.** $I = 0$.

Case a. $I \neq 0$.

Here we can effect division by I and eliminate the group parameter μ. An invariant frame Γ and corresponding infinitesimals ζ are given by

$$\begin{aligned} \Gamma_1 &= I^{-1}\Lambda_1 & \zeta^1 &= I\lambda^1 \\ \Gamma_2 &= I^{-1}\Lambda_2 & \zeta^2 &= I\lambda^2 \\ \Gamma_3 &= I^{-2}\Lambda_3 & \zeta^3 &= I^2\lambda^3 \\ \Gamma_4 &= \Lambda_4 & \zeta^4 &= \lambda^4. \end{aligned} \tag{4.80}$$

The invariants of the group action are

$$L, \qquad M := I^{-1}\Lambda_4 I \tag{4.81}$$

The structure relations of Γ are

$$\begin{array}{lll} [\Gamma_1, \Gamma_2] = 0 & [\Gamma_1, \Gamma_3] = 0 & [\Gamma_1, \Gamma_4] = -\frac{1}{2}\Gamma_2 + M\Gamma_1 \\ & [\Gamma_2, \Gamma_3] = 0 & [\Gamma_2, \Gamma_4] = -L\Gamma_1 + M\Gamma_2 \\ & & [\Gamma_3, \Gamma_4] = -\Gamma_1 + 2M\Gamma_3 \end{array} \tag{4.82}$$

From (4.77) the invariants L, M are subject to

$$\begin{aligned} \Gamma_1 L &= 0 & \Gamma_1 M &= 0 \\ \Gamma_2 L &= 0 & \Gamma_2 M &= 0 \\ \Gamma_3 L &= 0 & \Gamma_3 M &= 0, \end{aligned} \tag{4.83}$$

with $\Gamma_4 L$ and $\Gamma_4 M$ free. Finally, determining system (4.78) becomes

$$\begin{array}{lll} \Gamma_1\zeta^3 = 0 & (\Gamma_1)^2\zeta^2 = \Gamma_3\zeta^2 & \Gamma_1\zeta^1 = \frac{1}{2}\Gamma_3\zeta^3 \\ \Gamma_2\zeta^3 = 0 & \Gamma_4\zeta^2 = -\frac{1}{2}\zeta^1 + M\zeta^2 & \Gamma_4\zeta^1 = M\zeta^1 - L\zeta^2 - \zeta^3 \\ \Gamma_4\zeta^3 = 2M\zeta^3 & & \\ & \zeta^4 = 2\Gamma_1\zeta^2 & \end{array} \tag{4.84}$$

Case b. $I = 0$.

The condition $I = 0$ is equivalent to $J = 0$, that is, $\ddot{K} = 0$. This case is equivalent to a pure diffusion equation $K = 0$. Here division by I cannot be effected, so the group parameter μ cannot be eliminated. This is the case encountered in Example 4.5.2: the Boltzmann scaling has become a symmetry. An invariant L is available, but there is no invariant frame. Note that this case also inherits the symmetry $x \mapsto -x$, $v \mapsto -v$ from the SL_2 group (4.72): this is due to our failure to eliminate the sign π.

Scaling group—diffusion

Finally we account for the group generated by $L^1\{\hat{\mathbf{X}}_8\}$, namely

$$\begin{aligned} v' &= \rho v & D' &= \rho D \\ x' &= \rho x & K' &= K \\ t' &= \rho t & & \rho > 0. \\ u' &= u & & \end{aligned} \tag{4.85}$$

Computing action on the quantities Λ, L, I yields

$$\begin{aligned} \Lambda_1' &= \rho^{-1/2}\Lambda_1 & L' &= \rho^{-2}L \\ \Lambda_2' &= \rho^{-3/2}\Lambda_2 & I' &= \rho^{-3/2}I \\ \Lambda_3' &= \rho^{-1}\Lambda_3 \\ \Lambda_4' &= \rho^{-1}\Lambda_4. \end{aligned} \tag{4.86}$$

Now consider the branches $I \neq 0$ and $I = 0$ from above.

Case a. $I \neq 0$.

We compute the action on Γ, L, M:

$$\begin{aligned} \Gamma_1' &= \rho\Gamma_1 & L' &= \rho^{-2}L \\ \Gamma_2' &= \Gamma_2 & M' &= \rho^{-1}M \\ \Gamma_3' &= \rho^2\Gamma_3 \\ \Gamma_4' &= \rho^{-1}\Gamma_4. \end{aligned}$$

The calculation splits into two subcases, depending on whether L vanishes.

Case aa. $I \neq 0$, $L \neq 0$.

Here we can effect division by L and thereby eliminate the group parameter ρ. An invariant frame Σ and corresponding infinitesimals β are

$$\begin{aligned} \Sigma_1 &= |L|^{1/2}\Gamma_1 & \beta^1 &= |L|^{-1/2}\zeta^1 \\ \Sigma_2 &= \Gamma_2 & \beta^2 &= \zeta^2 \\ \Sigma_3 &= L\Gamma_3 & \beta^3 &= L^{-1}\zeta^3 \\ \Sigma_4 &= |L|^{-1/2}\Gamma_4 & \beta^4 &= |L|^{1/2}\zeta^4. \end{aligned} \tag{4.87}$$

The invariants of the group action are

$$P := |L|^{-3/2}\Gamma_4 L, \qquad Q := M|L|^{-1/2}, \qquad \sigma := \operatorname{sgn} L. \tag{4.88}$$

The sign σ is truly invariant, and cannot be removed through stealth or art. The absolute value signs throughout must be carefully respected, e.g., $\Delta|L| = \sigma\Delta L$. The structure relations of the frame Σ are

$$\begin{aligned} &[\Sigma_1,\Sigma_2] = 0 \quad &&[\Sigma_1,\Sigma_3] = 0 \quad &&[\Sigma_1,\Sigma_4] = \tfrac{1}{2}(2Q-\sigma P)\Sigma_1 - \tfrac{1}{2}\Sigma_2 \\ & &&[\Sigma_2,\Sigma_3] = 0 &&[\Sigma_2,\Sigma_4] = -\sigma\Sigma_1 + Q\Sigma_2 \\ & && &&[\Sigma_3,\Sigma_4] = -\sigma\Sigma_1 + (2Q-\sigma P)\Sigma_3 \end{aligned} \tag{4.89}$$

From (4.83), the invariants P, Q are subject to

$$\begin{aligned} \Sigma_1 P &= 0 & \Sigma_1 Q &= 0 \\ \Sigma_2 P &= 0 & \Sigma_2 Q &= 0 \\ \Sigma_3 P &= 0 & \Sigma_3 Q &= 0, \end{aligned} \tag{4.90}$$

with $\Sigma_4 P$ and $\Sigma_4 Q$ unconstrained. Finally, determining system (4.84) becomes

$$\begin{aligned} \Sigma_1\beta^3 &= 0 & (\Sigma_1)^2\beta^2 &= \sigma\Sigma_3\beta^2 \\ \Sigma_2\beta^3 &= 0 & \Sigma_4\beta^2 &= -\tfrac{1}{2}\beta^1 + Q\beta^2 \\ \Sigma_4\beta^3 &= (2Q-\sigma P)\beta^3 \\ & & \Sigma_1\beta^1 &= \tfrac{1}{2}\Sigma_3\beta^3 \\ \beta^4 &= 2\Sigma_1\beta^2 & \Sigma_4\beta^1 &= \tfrac{1}{2}(2Q-\sigma P)\beta^1 - \sigma\beta^2 - \sigma\beta^3 \end{aligned} \tag{4.91}$$

Case ab. $I \neq 0$, $L = 0$.

Here division by L cannot be effected. Another splitting appears, depending on whether M vanishes.

Case aba. $I \neq 0$, $L = 0$, $M \neq 0$.

Here we can use M to eliminate the group parameter ρ. An invariant frame Ξ and corresponding infinitesimals ψ are:

$$\begin{aligned} \Xi_1 &= M\Gamma_1 & \psi^1 &= M^{-1}\zeta^1 \\ \Xi_2 &= \Gamma_2 & \psi^2 &= \zeta^2 \\ \Xi_3 &= M^2\Gamma_3 & \psi^3 &= M^{-2}\zeta^3 \\ \Xi_4 &= M^{-1}\Gamma_4 & \psi^4 &= M\zeta^4. \end{aligned} \tag{4.92}$$

The invariants of the group action are

$$R := M^{-2}\Gamma_4 M, \qquad S := LM^{-2}. \tag{4.93}$$

The quantities R, S, Ξ are well defined whenever $M \neq 0$ (regardless of L). Here we have $L = 0$, so $S = 0$. The structure relations of Ξ are

$$\begin{aligned} &[\Xi_1, \Xi_2] = 0 \qquad [\Xi_1, \Xi_3] = 0 \qquad [\Xi_1, \Xi_4] = -\tfrac{1}{2}\Xi_2 + (1-R)\Xi_1 \\ &\qquad\qquad\qquad\;\; [\Xi_2, \Xi_3] = 0 \qquad [\Xi_2, \Xi_4] = \Xi_2 \\ &\qquad\qquad\qquad\qquad\qquad\qquad\quad\; [\Xi_3, \Xi_4] = -\Xi_1 + 2(1-R)\Xi_3 \end{aligned} \tag{4.94}$$

From (4.83) the invariant R is subject to

$$\Xi_1 R = 0 \qquad \Xi_2 R = 0 \qquad \Xi_3 R = 0, \tag{4.95}$$

with $\Xi_4 R$ unconstrained. Finally, determining system (4.84) becomes

$$\begin{aligned} &\Xi_1\psi^3 = 0 & &(\Xi_1)^2\psi^2 = \Xi_3\psi^2 & &\Xi_1\psi^1 = \tfrac{1}{2}\Xi_3\psi^3 \\ &\Xi_2\psi^3 = 0 & &\Xi_4\psi^2 = -\tfrac{1}{2}\psi^1 + \psi^2 & &\Xi_4\psi^1 = (1-R)\psi^1 - \psi^3 \\ &\Xi_4\psi^3 = 2(1-R)\psi^3 & & & & \\ & & &\psi^4 = 2\Xi_1\psi^2 & & \end{aligned} \tag{4.96}$$

Case abb. $I \neq 0$, $L = 0$, $M = 0$.

Here there is no way to eliminate the group parameter ρ. The best we can do for a frame is Γ. This singular case is again associated with equivalence transformations moving into the symmetry group. Conditions $L = 0$, $M = 0$, $I \neq 0$ are

$$D\ddot{D} - \tfrac{3}{2}\dot{D}^2 = 0, \qquad (\ddot{K}D^{-3/2})^{\cdot} = 0, \qquad \ddot{K} \neq 0$$

which lead to

$$\begin{aligned} D(u) &= (eu+f)^{-2} \\ K(u) &= \frac{au^2+bu+c}{eu+f} \end{aligned}$$

where at least one of e, f is nonvanishing, and $eu+f$ does not divide au^2+bu+c. These are the equations (including the Fokas-Yortsos system (3.90)) which are equivalent to Burgers' system (3.88). It is interesting that the frame calculations pick this out as a singular case even though the linearizing transformation taking Burgers' to the heat equation is not detected.

All branches with $I \neq 0$ have now been exhausted, so we pass on to

Case b. $I = 0$.

The action of the scaling group on our 'almost invariant' frame Λ was given above (4.86). We still have $L' = \rho^{-2}L$, so we get another splitting.

Case ba. $I = 0$, $L \neq 0$.

Here we can use division by L to eliminate the group parameter ρ. An invariant frame Υ, and corresponding infinitesimals ω are:

$$\begin{aligned} \Upsilon_1 &= |L|^{-1/4}\Lambda_1 & \omega^1 &= |L|^{1/4}\lambda^1 \\ \Upsilon_2 &= |L|^{-3/4}\Lambda_2 & \omega^2 &= |L|^{3/4}\lambda^2 \\ \Upsilon_3 &= |L|^{-1/2}\Lambda_3 & \omega^3 &= |L|^{1/2}\lambda^3 \\ \Upsilon_4 &= |L|^{-1/2}\Lambda_4 & \omega^4 &= |L|^{1/2}\lambda^4. \end{aligned} \tag{4.97}$$

The invariants of the group action are

$$P := |L|^{-3/2}\Lambda_4 L, \qquad \sigma := \operatorname{sgn} L. \tag{4.98}$$

This P is the same as (4.88), merely rewritten in new notation. The structure relations of the frame Υ are

$$\begin{aligned} &[\Upsilon_1, \Upsilon_2] = 0 \quad & &[\Upsilon_1, \Upsilon_3] = 0 \quad & &[\Upsilon_1, \Upsilon_4] = \tfrac{1}{4}\sigma P \Upsilon_1 - \tfrac{1}{2}\Upsilon_2 \\ & & &[\Upsilon_2, \Upsilon_3] = 0 & &[\Upsilon_2, \Upsilon_4] = -\sigma\Upsilon_1 + \tfrac{3}{4}\sigma P \Upsilon_2 \\ & & & & &[\Upsilon_3, \Upsilon_4] = -\tfrac{1}{2}\sigma P \Upsilon_3. \end{aligned} \tag{4.99}$$

From (4.77) the invariant P is subject to

$$\Upsilon_1 P = 0 \qquad \Upsilon_2 P = 0 \qquad \Upsilon_3 P = 0, \tag{4.100}$$

with $\Upsilon_4 P$ unconstrained. Finally, determining system (4.78) becomes

$$\begin{aligned} \Upsilon_1\omega^3 &= 0 & (\Upsilon_1)^2\omega^2 &= \Upsilon_3\omega^2 & \Upsilon_1\omega^1 &= \tfrac{1}{2}\Upsilon_3\omega^3 \\ \Upsilon_2\omega^3 &= 0 & \Upsilon_4\omega^2 &= -\tfrac{1}{2}\omega^1 + \tfrac{3}{4}\sigma P\omega^2 & \Upsilon_4\omega^1 &= \tfrac{1}{4}\sigma P\omega^1 - \sigma\omega^2 \\ \Upsilon_4\omega^3 &= \tfrac{1}{2}\sigma P\omega^3 & & & & \\ & & \omega^4 &= 2\Upsilon_1\omega^2 & & \end{aligned} \tag{4.101}$$

Case bb. $I = 0$, $L = 0$.

Here there is no way to eliminate the group parameter ρ, and we are stuck with the frame Λ. This singular case corresponds to D, K satisfying

$$D\ddot{D} - \tfrac{3}{2}\dot{D}^2 = 0, \qquad \ddot{K} = 0$$

which lead to

$$\begin{aligned} D(u) &= (eu + f)^{-2} \\ K(u) &= au + b \end{aligned}$$

where at least one of e, f is nonvanishing. These are the equations (including the Bluman-Kumei system (3.86)) which are equivalent to the linear heat system. In this case the linearizing transformation is in the equivalence group, so it is expected that this case should be picked out as singular.

Reflection equivalence

The calculation of invariant and 'almost invariant' frames for the connected component of the equivalence group is now finished. For completeness, we should adjoin the reflection equivalence R_2 (3.75) $v \mapsto -v$, $u \mapsto -u$. This has little effect on the branches above. It causes the quantities M, P, Q to change sign, as well as various operators Λ, Σ, Ξ, Γ. Two cases are affected by these sign changes:

Case **aa.** $I \neq 0$, $L \neq 0$. Here if $P \neq 0$, we may use $\operatorname{sgn} P$ to eliminate sign anomalies. Thus $\tilde{P} := \operatorname{sgn} P \cdot P$ and $\tilde{Q} := \operatorname{sgn} P \cdot Q$ are invariant, as are the operators $\tilde{\Sigma}_4 := \operatorname{sgn} P \cdot \Sigma_4$, etc. If $P = 0$ but $Q \neq 0$, we use $\operatorname{sgn} Q$ in the same way. The structure relations (4.89), determining system (4.91) etc. are identical, except that P, Q, Σ_4 are replaced by their 'sign corrected' relatives $\tilde{P}$, $\tilde{Q}$, $\tilde{\Sigma}_4$. The only interesting case is $P = Q = 0$, where the sign cannot be compensated. Here a discrete symmetry is inherited from the equivalence group. This is the case $D(u) = |\alpha u^2 + \beta u + \gamma|^{-1}$ (Fujita diffusivity [23]) and $K(u) = K_0|\alpha u^2 + \beta u + \gamma|^{1/2}$, where $\beta^2 - 4\alpha\gamma \neq 0$, $K_0 \neq 0$. Up to equivalence there are two distinct cases:

- $D(u) = |u|^{-1}$, $K(u) = |u|^{1/2}$, admitting the hodograph-type transformation (3.84) as a symmetry.
- $D(u) = (1+u^2)^{-1}$, $K(u) = \sqrt{1+u^2}$, admitting the reflection symmetry $v \mapsto -v$, $u \mapsto -u$.

Case **ba.** $I = 0$, $L \neq 0$. Here if $P \neq 0$ we use it as above to remove sign anomalies, while if $P = 0$ compensation is impossible. This is the case of Fujita's diffusion equation $K(u) = 0$, $D(u) = 1/|\alpha u^2 + \beta u + \gamma|$, $\beta^2 - 4\alpha\gamma \neq 0$. Up to equivalence there are two cases

- $D(u) = |u|^{-1}$, $K(u) = 0$, admitting the hodograph symmetry (3.84).
- $D(u) = (1+u^2)^{-1}$, $K(u) = 0$, admitting the above reflection symmetry.

Interestingly, these cases were distinguished in the partial classification Tables 4.3, 4.4 for the same reasons.

Summary of invariant frames

So far we have the incomplete classification tree shown in Figure 4.2.

Completion to involutive system

For each of the five 'leaves' of the tree above we now complete the determining system to involutive form. This gives rise to a further hierarchy of branchings. Note the three common translation symmetries are always present. Hence the the solution space of the determining systems is always of dimension at least three. We therefore do not present results for any branch with a three dimensional solution space.

Case aa. $I \neq 0$, $L \neq 0$.

Here we are working on system (4.91).] of the frame Reid method gives a case splitting on $\Sigma_4 P$. If $\Sigma_4 P \neq 0$ only minimal translation symmetry is present. We pursue the case $\Sigma_4 P = 0$, that is, $P = \text{const}$. A further case splitting, on $\Sigma_4 Q$, arises. If $\Sigma_4 Q \neq 0$ we have only minimal symmetry. Hence we follow the branch $\Sigma_4 Q = 0$, $Q = \text{const}$. No further splittings arise, and the system is reduced to

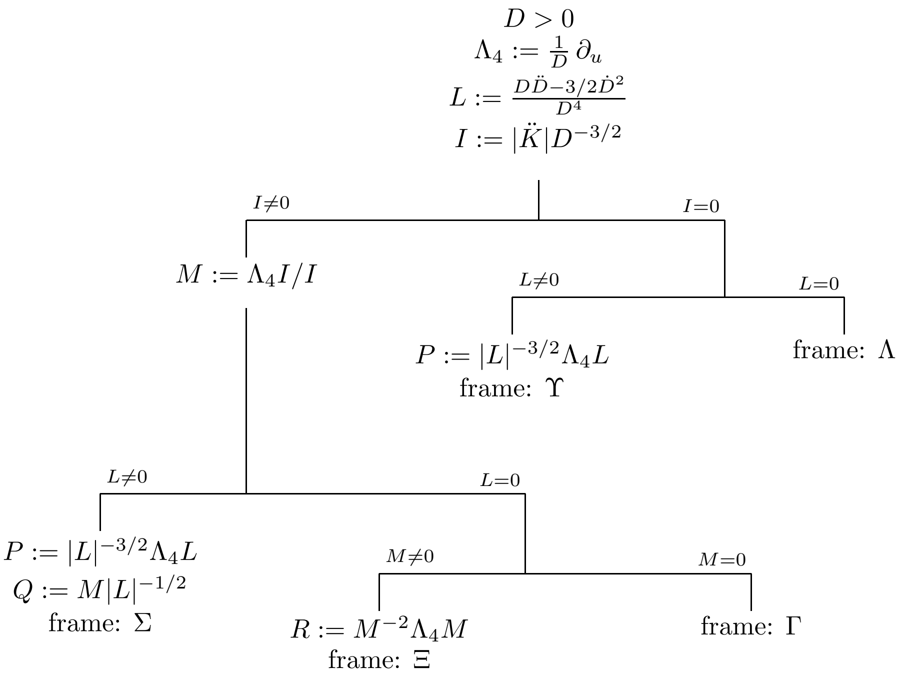

Figure 4.2: Preliminary classification tree for potential diffusion convection system (4.62). Branchings are on the basis of whether or not particular frames exist.

involutive form

$$
\begin{aligned}
\Sigma_1\beta^3 &= 0 & \Sigma_1^2\beta^2 &= 0 \\
\Sigma_2\beta^3 &= 0 & \Sigma_2\beta^2 &= -2Q\Sigma_1\beta^2 \\
\Sigma_3\beta^3 &= -2(2Q - \sigma P)\Sigma_1\beta^2 & \Sigma_3\beta^2 &= 0 \\
\Sigma_4\beta^3 &= (2Q - \sigma P)\beta^3 & \Sigma_4\beta^2 &= -\tfrac{1}{2}\beta^1 + Q\beta^2
\end{aligned}
$$

$$
\begin{aligned}
\Sigma_1\beta^1 &= -(2Q - \sigma P)\Sigma_1\beta^2 & \beta^4 &= 2\Sigma_1\beta^2 \\
\Sigma_2\beta^1 &= 2\sigma\Sigma_1\beta^2 \\
\Sigma_3\beta^1 &= 2\sigma\Sigma_1\beta^2 \\
\Sigma_4\beta^1 &= \tfrac{1}{2}(2Q - \sigma P)\beta^1 - \sigma\beta^2 - \sigma\beta^3
\end{aligned}
\tag{4.102}
$$

The four parametric quantities β^1, β^2, β^3, $\Sigma_1\beta^2$, give a four-parameter symmetry group.

Application of the method (Appendix B) for finding structure constants gives a Lie algebra of symmetry operators $\mathbf{Y}_1$, $\mathbf{Y}_2$, $\mathbf{Y}_3$, $\mathbf{Y}_4$ with commutation relations

	$\mathbf{Y}_1$	$\mathbf{Y}_2$	$\mathbf{Y}_3$	$\mathbf{Y}_4$
$\mathbf{Y}_1$	0	0	0	$-(2Q-\sigma P)\mathbf{Y}_1+\mathbf{Y}_2$
$\mathbf{Y}_2$		0	0	$2\sigma\mathbf{Y}_1-2Q\mathbf{Y}_2$
$\mathbf{Y}_3$			0	$2\sigma\mathbf{Y}_1-2(2Q-\sigma P)\mathbf{Y}_3$
$\mathbf{Y}_4$				0

Case aba. $I \neq 0$, $L = 0$, $M \neq 0$.

We apply the frame Reid method to system (4.96), giving a case splitting on $\Xi_4 R$. If $\Xi_4 R \neq 0$ there is minimal symmetry, and we present the case $\Xi_4 R = 0$, which terminates without further branching in involutive form:

$$
\begin{array}{lll}
\Xi_1\psi^3 = 0 & \Xi_1\psi^2 = -\frac{1}{2}\Xi_2\psi^2 & \Xi_1\psi^1 = (1-R)\Xi_2\psi^2 \\
\Xi_2\psi^3 = 0 & \Xi_2^2\psi^2 = 0 & \Xi_2\psi^1 = 0 \\
\Xi_3\psi^3 = 2(1-R)\Xi_2\psi^2 & \Xi_3\psi^2 = 0 & \Xi_3\psi^1 = -\Xi_2\psi^2 \\
\Xi_4\psi^3 = 2(1-R)\psi^3 & \Xi_4\psi^2 = -\frac{1}{2}\psi^1+\psi^2 & \Xi_4\psi^1 = (1-R)\psi^1-\psi^3 \\
 & \psi^4 = -\Xi_2\psi^2 &
\end{array}
\tag{4.103}
$$

The parametric quantities ψ^1, ψ^2, ψ^3, $\Xi_2\psi^2$ give a four-parameter symmetry group. The commutation relations of the symmetry algebra are

	$\mathbf{Y}_1$	$\mathbf{Y}_2$	$\mathbf{Y}_3$	$\mathbf{Y}_4$
$\mathbf{Y}_1$	0	0	0	$(1-R)\mathbf{Y}_1-\frac{1}{2}\mathbf{Y}_2$
$\mathbf{Y}_2$		0	0	$\mathbf{Y}_2$
$\mathbf{Y}_3$			0	$-\mathbf{Y}_1+2(1-R)\mathbf{Y}_3$
$\mathbf{Y}_4$				0

Case abb. $I \neq 0$, $L = 0$, $M = 0$.

This is the Burgers' equation case. There are no further splittings. Since this case is connected to the linear heat system by the Cole-Hopf transformation, an involutive system with infinitely many parametric derivatives results. We do not reproduce it here. If this linearization were not known, it is interesting to speculate whether it could be detected from the determining system. For this a frame version of the theory of Kumei and Bluman [41, 14] would be required.

Case ba. $I = 0$, $L \neq 0$.

Applying the frame Reid method to determining system (4.101), we find a case splitting on $\Upsilon_4 P$. If $\Upsilon_4 P \neq 0$ the system reduces to the involutive form

$$
\begin{aligned}
\Upsilon_1\omega^3 &= 0 & \Upsilon_1\omega^2 &= 0 & \Upsilon_1\omega^1 &= \tfrac{1}{2}\Upsilon_3\omega^3\\
\Upsilon_2\omega^3 &= 0 & (\Upsilon_2)^2\omega^2 &= \tfrac{1}{2}\Upsilon_3\omega^3 & \Upsilon_2\omega^1 &= 0\\
(\Upsilon_3)^2\omega^3 &= 0 & \Upsilon_3\omega^2 &= 0 & \Upsilon_3\omega^1 &= 0\\
\Upsilon_4\omega^3 &= \tfrac{1}{2}\sigma P\omega^3 & \Upsilon_4\omega^2 &= -\tfrac{1}{2}\omega^1 + \tfrac{3}{4}\sigma P\omega^2 & \Upsilon_4\omega^1 &= \tfrac{1}{4}\sigma P\omega^1 - \sigma\omega^2\\
& & \omega^4 &= 0 & &
\end{aligned}
\tag{4.104}
$$

The parametric derivatives ω^1, ω^2, ω^3, $\Upsilon_3\omega^3$ give a four-parameter symmetry group, representing the four symmetries common to all diffusion potential systems. The commutation relations of the symmetry algebra are

	$\mathbf{Y}_1$	$\mathbf{Y}_2$	$\mathbf{Y}_3$	$\mathbf{Y}_4$
$\mathbf{Y}_1$	0	0	0	$\frac{1}{2}\mathbf{Y}_1$
$\mathbf{Y}_2$		0	0	$\frac{1}{2}\mathbf{Y}_2$
$\mathbf{Y}_3$			0	$\mathbf{Y}_3$
$\mathbf{Y}_4$				0

If $\Upsilon_4 P = 0$, we obtain an involutive form

$$
\begin{aligned}
\Upsilon_1\omega^3 &= 0 & \Upsilon_1^2\omega^2 &= 0 & \Upsilon_1\omega^1 &= \tfrac{1}{2}\Upsilon_3\omega^3\\
\Upsilon_2\omega^3 &= 0 & (\Upsilon_2)^2\omega^2 &= -\sigma P\Upsilon_1\omega^2 + \tfrac{1}{2}\Upsilon_3\omega^3 & \Upsilon_2\omega^1 &= 2\sigma\Upsilon_1\omega^2\\
(\Upsilon_3)^2\omega^3 &= 0 & \Upsilon_3\omega^2 &= 0 & \Upsilon_3\omega^1 &= 0\\
\Upsilon_4\omega^3 &= \tfrac{1}{2}\sigma P\omega^3 & \Upsilon_4\omega^2 &= -\tfrac{1}{2}\omega^1 + \tfrac{3}{4}\sigma P\omega^2 & \Upsilon_4\omega^1 &= \tfrac{1}{4}\sigma P\omega^1 - \sigma\omega^2\\
& & \omega^4 &= 2\Upsilon_1\omega^2 & &
\end{aligned}
\tag{4.105}
$$

The parametric quantities ω^1, ω^2, ω^3, $\Upsilon_3\omega^3$, $\Upsilon_1\omega^2$ give five symmetries. The structure of the symmetry algebra is

	$\mathbf{Y}_1$	$\mathbf{Y}_2$	$\mathbf{Y}_3$	$\mathbf{Y}_4$	$\mathbf{Y}_5$
$\mathbf{Y}_1$	0	0	0	$\frac{1}{2}\mathbf{Y}_1$	$\mathbf{Y}_2$
$\mathbf{Y}_2$		0	0	$\frac{1}{2}\mathbf{Y}_2$	$2\sigma\mathbf{Y}_1 - \sigma P\mathbf{Y}_2$
$\mathbf{Y}_3$			0	$\mathbf{Y}_3$	0
$\mathbf{Y}_4$				0	0
$\mathbf{Y}_5$					0

Case bb. $I = 0$, $L = 0$

There is no further case splitting. The involutive form is

$$\begin{array}{llll} \Lambda_1\lambda^3 = 0 & \Lambda_4\lambda^2 = -\frac{1}{2}\lambda^1 & \Lambda_1\lambda^1 = \frac{1}{2}\Lambda_3\lambda^3 & \\ \Lambda_2\lambda^3 = 0 & \Lambda_1^2\lambda^2 = \Lambda_3\lambda^2 & \Lambda_2\lambda^1 = 0 & \\ \Lambda_4\lambda^3 = 0 & \Lambda_2\Lambda_1\lambda^2 = -\frac{1}{2}\Lambda_3\lambda^1 & \Lambda_4\lambda^1 = 0 & (4.106) \\ (\Lambda_3)^2\lambda^3 = 0 & (\Lambda_2)^2\lambda^2 = \frac{1}{2}\Lambda_3\lambda^1 & (\Lambda_3)^2\lambda^1 = 0 & \\ & \Lambda_3\Lambda_2\lambda^2 = 0 & \lambda^4 = 2\Lambda_1\lambda^2 & \end{array}$$

The parametric quantities are λ^1, $\Lambda_3\lambda^1$, λ^3, $\Lambda_3\lambda^3$, $\Lambda_2\lambda^2$, and the two infinite sequences λ^2, $\Lambda_3\lambda^2$, $(\Lambda_3)^2\lambda^2$, ... and $\Lambda_1\lambda^2$, $\Lambda_3\Lambda_1\lambda^2$, $(\Lambda_3)^2\Lambda_1\lambda^2$, ... Hence this case admits an infinite-dimensional symmetry algebra. We do not attempt to find commutation relations for this case. The class consists of equations which can be mapped to the heat equation by an equivalence transformation, and we regard the symmetry properties as known.

Summary of classification

All in all the calculations of this section yield the classification tree shown in Figure 4.3. In this remarkably compact diagram is present all the information required to decide the symmetry properties of a diffusion convection potential system. The elegance and compactness of the result is apparent when compared with the output of Reid's [57] method. For instance, the case $\Lambda_4 P = \Lambda_4 Q = 0$, when written out in full, is

$$\begin{aligned} 12D\ddot{D}^3 - 6\dot{D}^2\ddot{D}^2 - 16D\dot{D}\ddot{D}\dddot{D} + 6\dot{D}^3\dddot{D} & \\ - 2D^2\ddddot{D}\ddot{D} + 3D\ddddot{D}\dot{D}^2 + 3D^2\dddot{D}^2 &= 0 \\ 4\ddot{K}\ddddot{K}D^2\ddot{D} - 6\ddot{K}\ddddot{K}D\dot{D}^2 + 8\ddot{K}\dddot{K}D\dot{D}\ddot{D} - 6\ddot{K}\dddot{K}\dot{D}^3 - 6\ddot{K}^2\ddot{D}^2D & \\ + 3\ddot{K}^2\dot{D}^2\ddot{D} - 4\dddot{K}^2D^2\ddot{D} + 6\dddot{K}^2D\dot{D}^2 - 2\ddot{K}\dddot{K}D^2\dddot{D} + 3\ddot{K}^2\dot{D}D\dddot{D} &= 0, \end{aligned} \quad (4.107)$$

which is the form in which Reid's method [57] returns the result.

In Figure 4.3, all the branchings are (by construction) invariant under the action of the equivalence group. Hence two equations connected by an equivalence transformation always occur on the same branch. This greatly cuts down on spurious branchings, and in fact *all* of the branches in Figure 4.3 discriminate symmetry properties. In contrast, Reid's method gives rise to a large number of apparently irrelevant branches. Note that equations occurring on different branches of the tree could be equivalent with respect to a transformation *not* in equivalence group (3.81): this in fact occurs, since the Burgers' and linear heat equation branches are connected by the Cole-Hopf transformation (3.89).

Ultimately one wishes to *solve* the classifying equations to find $D(u)$, $K(u)$, and to solve the determining system to find the symmetry operators. However, the count of symmetries in Figure 4.3 shows that the only cases with symmetry beyond that encountered in the partial classification of §4.2.3 are the linear heat and Burgers' branches. The symmetries of the heat system are well-known, and those of Burgers' system follow from these by the Cole-Hopf transformation of §3.4.2. Hence no further construction of symmetries is required.

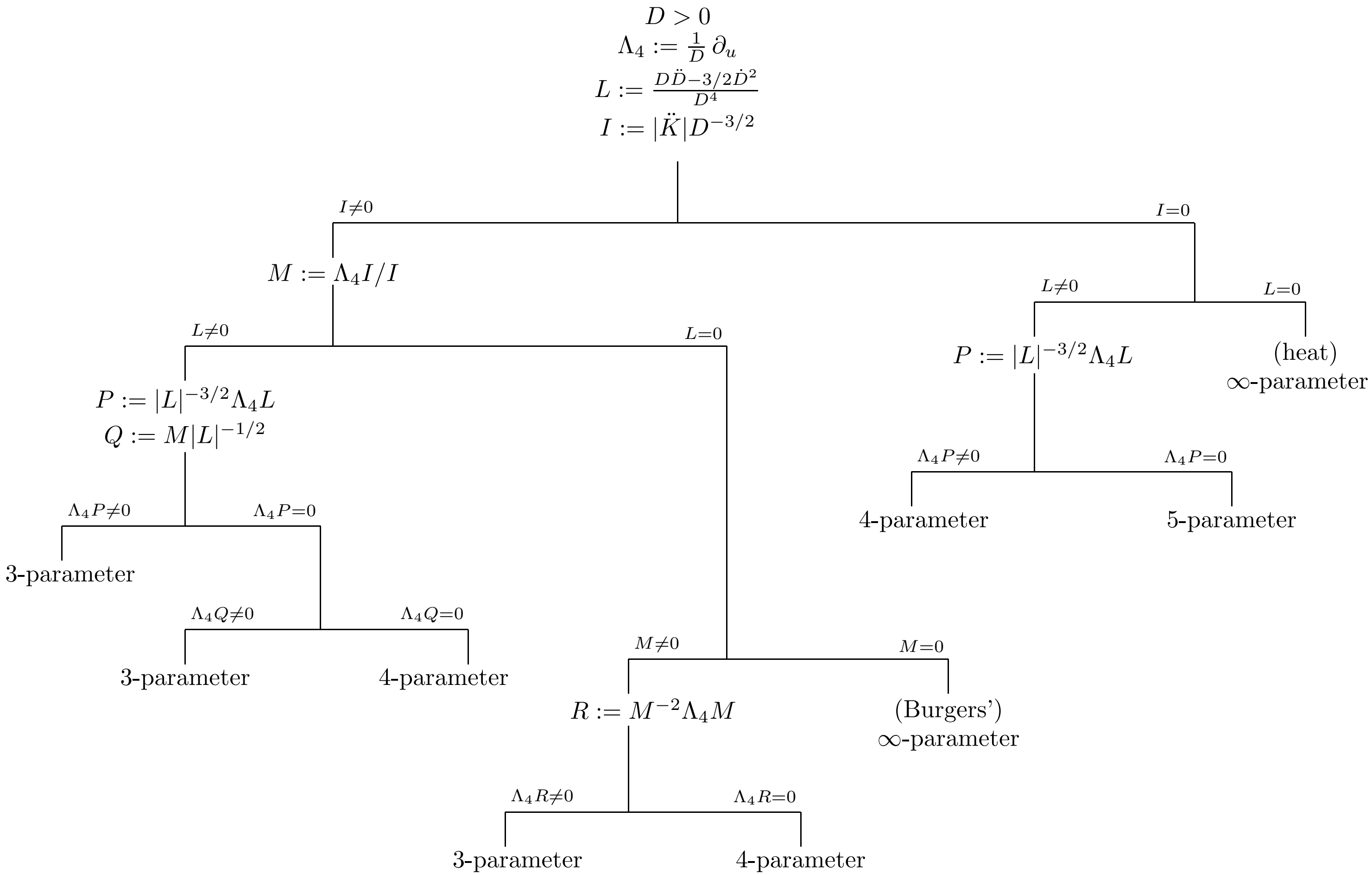

Figure 4.3: Complete symmetry classification tree for potential diffusion convection system (4.62).

Chapter 5

Conclusion

5.1 Further work

We now describe some further directions in which the equivalence methods introduced here could be pursued.

First we describe some possible modifications to the definition of equivalence transformations in Chapter 3. The conditions we placed on transformations there, realizing them as projectable coordinate transformations on an augmented space, and requiring that they act on every solution of every equation in the class, are extremely restrictive. The most obvious generalization is to permit the variables (x', u') to depend on arbitrary elements $a = \phi(w)$ in some way. Permitting dependence on the value a alone is insufficient, and it seems that the appropriate generalization is to allow Lie-Bäcklund transformations for a as a function of w. Another modification is to relax the requirement that equivalence transformations act on every equation in the class. This would lead to an 'equivalence classification', with a hierarchy of subclasses admitting richer equivalences than usual. The examples in §1.2 already show the need for both these kinds of generalization.

Apart from the intrinsic interest of transformation properties, a principal motivation for using equivalence methods is for assisting in the symmetry classification for a class of p.d.e.'s. We have shown in §4 how useful the equivalence group can be for such problems. However, further computational experience with the frame classification method of §4.5 is required, especially on difficult examples with an infinite-parameter equivalence group. Our treatment also has some theoretical gaps, the vital one being the absence of a proof of the frame Riquier Conjecture 4.3.17. Our use of frame involutive form to obtain the dimension and structure of symmetry groups relies directly on the frame Riquier conjecture. Hence it is essential to establish this result, so that the method for counting symmetries and finding commutation relations can be placed on a sound footing.

An interesting issue raised when Reid's method is referred to an arbitrary moving frame is the ordering of frame derivatives. Because frame operators do not commute, the ordering process is more subtle than in the classical p.d.e. case. We simply demanded that the ordering imposed be a Janet ordering on equivalence classes of derivatives, leaving the relative order of $\Delta_{12}\theta$ and $\Delta_{21}\theta$ unresolved. Certainly this can be extended to an ordering on all frame derivatives by assigning some ordering within each equivalence class. However, it is not clear that this is the most general ordering possible. Characterization of permissible orderings may be more difficult than in the p.d.e. case.

Lastly, an important task is to develop a computer implementation of the frame Reid method. The general program structure used by Reid for the p.d.e. case certainly carries over. However every procedure, and even the data structure for derivatives, requires modification to account for noncommuting operators, so this is a nontrivial undertaking.

5.1.1 Isovector method for frame determining system

Our method for writing and manipulating frame determining systems is only partially geometric. The frame derivatives $\Delta_i\theta^j$ of components θ^j of a vector field are of no particular geometric importance: in particular they do not constitute a tensor. The important geometric quantities are the covariant derivatives $\theta^j_{;i} = \Delta_i\theta^j + \gamma^j_{ik}\theta^k$ with respect to a connection defined by the frame Δ. It is not clear whether any computational advantage accrues from using covariant derivatives in preference to simple frame derivatives. The approach actually presented has the advantage of more closely parallelling Reid's method.

A second way in which our method fails to be geometric is that we derive determining equations using coordinate-based calculations. Only later is the determining system given a geometric formulation by referring it to a moving frame. In this we are the opposite of Harrison and Estabrook's isovector method [29]. There the original p.d.e.'s are formulated geometrically as an ideal $\mathcal{I}$ of differential forms. The symmetry vector fields $\mathbf{Y}$ are found by requiring vanishing of the Lie derivative $\mathcal{L}_{\mathbf{Y}}\mathcal{I} = 0$, $(\mathrm{mod}\,\mathcal{I})$. Once they derive their invariance condition, geometric formalism is abandoned, and determining equations are treated simply as systems of p.d.e.'s. Both the Liesymm package in the Maple V symbolic language [20], and the program of Kersten [40], use the Harrison-Estabrook formalism to derive determining equations. Everything is referred to basis vector fields ∂_{x^i} and basis forms dx^i, and the method amounts to a recondite procedure for constructing determining equations.

However, a completely geometric formulation of determining equations is possible, referring every step of the calculation to a moving frame. First we formulate the original p.d.e.'s as a collection of differential forms Ω^i, collectively denoted by $\mathcal{I}$. Similarly reformulate the auxiliary system satisfied by the arbitrary elements as a collection of forms Θ^i. Next, by applying the Cartan equivalence method to the equivalence group, we construct an invariant coframe ω, along with its structure relations, which here take the form $d\omega^k = \gamma^k_{ij}\,\omega^i \wedge \omega^j$. Rewrite the forms Ω^i representing the p.d.e.'s in terms of the coframe ω. Similarly rewrite Θ^i representing the auxiliary system in terms of the coframe ω. The Harrison-Estabrook isovector procedure can now be applied. Write $\mathbf{Y} = \zeta^i\Delta_j$, where Δ is the invariant frame dual to the coframe ω. The isovector condition

$$\mathcal{L}_{\mathbf{Y}}\Omega^j = 0 \qquad (\mathrm{mod}\,\mathcal{I})$$

is then stated entirely in terms of frames, and yields a collection of differential forms representing the determining system. This can be broken up by picking off coefficients of the basis forms ω, to yield a frame determining system, which can be reduced using the frame Reid method. Alternatively the Cartan-Kähler theory might be used to count the dimension of the solution space, i.e., the number of symmetries. Indeed this version of our method would be theoretically more satisfactory, in that the integration theory used to count the symmetries already exists.

We sketch an example of this method, rederiving some of the results of Section 4.5.2 for the potential diffusion convection system. First we reformulate the system

$$\begin{aligned} v_x &= u \\ v_t &= au_x - b \end{aligned}$$

where $a = D(u)$ and $b = K(u)$ as the one forms

$$\begin{aligned} du - p\,dx - q\,dt \\ dv - u\,dx - (ap - b)\,dt \end{aligned}$$

where p, q represent u_x, u_t respectively. The auxiliary system $a_x = 0$ etc. is written as the forms

$$da = \dot{a}\,du \qquad\qquad db = \dot{b}\,du.$$

Thus we are working on a six dimensional space (x, t, u, v, p, q).

Constructing an invariant coframe with respect to the equivalence group, we come upon branchings similar to those in Figure 4.2. Suppose the conditions $I \neq 0$, $L \neq 0$ are satisfied, so that the invariants are P, Q. Because here we are working on a larger space, there are additional invariants F, G expressed in terms of p, q respectively and various derivatives a, $\dot{a}$, etc. We construct an invariant coframe ω whose structure relations are

$$\begin{aligned} d\omega^1 &= 0 \\ d\omega^2 &= Q\omega^1 \wedge \omega^2 - \omega^1 \wedge \omega^3 \\ d\omega^3 &= -\tfrac{1}{2}\sigma\omega^1 \wedge \omega^2 + \tfrac{1}{2}(2Q - \sigma P)\omega^1 \wedge \omega^3 - \sigma\omega^1 \wedge \omega^4 \\ d\omega^4 &= (2Q - \sigma P)\omega^1 \wedge \omega^4 \\ d\omega^5 &= (-Q + \sigma P)\omega^1 \wedge \omega^5 \\ d\omega^6 &= \sigma(1 + F)\omega^1 \wedge \omega^5 + (-2Q + 3/2\sigma P)\omega^1 \wedge \omega^6 \end{aligned}$$

The diffusion convection system referred to this coframe is generated by forms Ω^1, Ω^2:

$$\begin{aligned} \Omega^1 &= \omega^1 - \sigma F\omega^3 - \sigma G\omega^4 \\ \Omega^2 &= \omega^2 - F\omega^4. \end{aligned} \tag{5.1}$$

The frame derivatives of the invariants P, Q, F, G are not arbitrary, but are constrained to satisfy the auxiliary system

$$\begin{aligned} dP &= P_{,1}\omega^1 \\ dQ &= P_{,1}\omega^1 \\ dF &= F(-Q + \sigma P)\omega^1 + \omega^5 \\ dG &= \Big(\tfrac{1}{2}\sigma F(2 + F) + G(-2Q + 3/2\sigma P)\Big)\omega^1 + \omega^6 \end{aligned}$$

where $P_{,1} = \Delta_1 P$ where Δ is the frame dual to the coframe ω.

We seek symmetry vector fields $\mathbf{Y}$ which we suppose referred to the frame Δ: $\mathbf{Y} = \theta^i \Delta_i$. The symmetry condition is that

$$\mathcal{L}_{\mathbf{Y}}\Omega = 0 \qquad (\operatorname{mod} \Omega)$$

where $\mathcal{L}$ is the Lie derivative. That is, there are functions λ_1^1, λ_1^2, λ_2^1, λ_2^2 such that

$$\begin{aligned} \mathcal{L}_{\theta^i\Delta_i}\Omega^1 &= \lambda_1^1\Omega^1 + \lambda_2^1\Omega^2 \\ \mathcal{L}_{\theta^i\Delta_i}\Omega^2 &= \lambda_1^2\Omega^1 + \lambda_2^2\Omega^2 \end{aligned}$$

Replacing Ω^1, Ω^2 by their expressions (5.1) in terms of the coframe ω, and using the identity

$$\mathcal{L}_{\mathbf{Y}}\Omega = d\langle\Omega \mid \mathbf{Y}\rangle + \mathbf{Y} \rfloor d\Omega,$$

we find for example from the Ω^2 equation,

$$\begin{aligned} &d\theta^2 - F\,d\theta^4 + ((2Q - \sigma P)F\theta^4 + \theta^3 - Q\theta^2)\omega^1 + Q\theta^1\omega^2 - \theta^1\omega^3 \\ &\quad - (\theta^5 + QF\theta^1)\omega^4 = \lambda_1^2(\omega^1 - \sigma F\omega^3 - \sigma G\omega^4) + \lambda_2^2(\omega^2 - F\omega^4). \end{aligned}$$

Picking off coefficients gives a linear homogeneous frame system for θ^i and λ_j^i, which is in fact the determining system. Rendering it involutive using the frame Reid method gives the results derived in Section 4.5.2.

This procedure is elegant and geometric throughout. It combines the Cartan, Harrison-Estabrook, and Reid procedures into one vast algorithm for symmetry classification. The approach follows on naturally from the Cartan equivalence method, as opposed to our treatment in Section 4.3, which was designed to tie in naturally with Reid's method. There are disadvantages in the process: the space (x, t, u, v, p, q) on which we construct the invariant coframe is of dimension six; the derivative coordinates p, q would seem to be of lesser importance, and did not occur in our previous formulation. This leads to more intensive calculations.

5.2 Conclusions

In this dissertation we have endeavoured to give a systematic and detailed method for finding equivalence transformations for a class of differential equations. Although the construction is not difficult, and has been available in albeit sketchy form for almost a decade, it appears to have been little used. The systematic use of equivalence transformations appears to have been confined to the Cartan equivalence method, which is hampered by its insistence on extracting only transformation information contained in a given group acting on the class of d.e.'s. The examples treated by the Cartan method have tended to be classes of geometric objects under the action of some natural transformation group. Since physical classes of equations rarely represent a geometrically natural class, and generally do not come provided with a transformation group attached, application of the Cartan method to physically significant problems has been seriously hampered.

The method we have described does not usually yield exhaustive transformation information on the class of differential equations under consideration. Nevertheless, we have shown by example that the information contained in the equivalence group is nontrivial, and can give significant insight into relationships between various equations in the class. Some of these relationships were explored for examples in §3.4.

It appears to us that one of the most important uses of the equivalence group is in systematically ordering the calculations and results of symmetry classification. This is extremely important, since application of similarity or other methods based on a symmetry approach requires first the construction of a symmetry group. The multitude of cases which arise when performing a symmetry classification, and the multitude of parameters occurring in each of the cases, can lead to difficulty in stating symmetry classification results. Order is usually imposed by an ad hoc parameter removal, which the equivalence group makes more complete and systematic. Certain symmetry classification information is easily available using by the methods described in §4.2. In fact the symmetries contained in the equivalence group are in some sense the 'predictable' symmetries. Their construction and classification pose few problems for a finite-parameter equivalence group, and for the infinite case the Cartan equivalence method is available. The symmetries not contained in the equivalence group are not predictable by our methods. Despite this (or perhaps because of this), such symmetries are of great interest, and it is unacceptable to confine our attention solely to the symmetries from the equivalence group. Hence we have been led to the method of Section 4.3, which enables a complete symmetry classification, while taking full account of the (necessarily partial) information contained in the equivalence group. In this it combines the best features of the Cartan equivalence method (utilizing transformation information) and Reid algorithm (giving a complete classfication). The geometric formulation of determining equations in terms of moving frames has an intrinsic elegance which is reflected in the nature of the classification tree produced (Figure 4.3). Our method uses the results produced by the Tresse/Cartan equivalence method as an *input* to the frame Reid method of §4.3.2, thus providing a bridge between the *geometric* method of Tresse/Cartan and the *analytic* method of Reid. In this way it synthesizes a significant portion of symmetry theory for d.e.'s.

Appendix A

Algorithms for Frame Systems

In this appendix we present algorithms for the frame systems of §4.3—§4.5. First we cover reduction to involutive form, then we consider the classification case with arbitrary elements present. It is not necessary for the frame system to be a determining system for symmetries, although this is the only way in which we use frame systems.

A.1 Reduction to frame involutive form

We give a sequence of procedures, starting with the most elementary, and culminating in a procedure for reduction to involutive form.

A.1.1 Orthonomic form

We assume that the following elementary procedures are available:

maxorder(S)
    input: A finite set S of frame derivatives $\Delta_J\theta^j$
    output: The element $s \in S$ highest in the ordering

removepermutations(R)
    input: A set R of frame equations.
    action: For each equation $r \in R$, check for presence in r of derivatives $\Delta_I\theta^j$, $\Delta_J\theta^j$ which are permutations of one another.
    If present, use structure relations to write them all in terms of one among them:
        $\Delta_I\theta^j = \Delta_J\theta^j +$ permutation terms.
    output: Equations R with permutations removed.

leadingderiv(eqn)
    input: A frame equation *eqn*.
    output: The derivative of highest order occurring in *eqn*:
    $leadingderiv := maxorder\{\Delta_J\theta^j \mid \Delta_J\theta^j \text{ present in } eqn\}$.

solve(*eqn*,*deriv*)
input: A frame equation *eqn*.
A frame derivative $deriv = \Delta_J \theta^j$ occurring in *eqn*.
output: *eqn* rewritten in the form $\Delta_J \theta^j = \text{rhs}$.

subst(*eqn*,U)
input: An equation *eqn* in solved form $\Delta_J \theta^j = \text{rhs}$.
A set U of frame equations.
action: Replace each occurrence of $\Delta_I \theta^j$ where I is a permutation of J by $\Delta_I \theta^j = \text{rhs} + \text{permutation terms}$
output: U with $\Delta_J \theta^j$ substituted out.

With these defined, orthonomic form (Definition 4.3.7) may be achieved by the following 'Gaussian elimination' algorithm:

Algorithm A.1.1. (orthonomic)

function *orthonomic*(DQ)
  unsolved := DQ
  solved := $\emptyset$
  repeat
    unsolved := *removepermutations*(*unsolved*)
    maxderiv := *maxorder*{*leadingderiv*(*eqn*) | *eqn* $\in U$}
    maxset := {*eqn* $\in$ *unsolved* | *leadingderiv*(*eqn*) = *maxderiv*}
    nexteqn := (any element in *maxset*)
    nexteqn := *solve*(*nexteqn*,*maxderiv*)
    solved := *subst*(*nexteqn*,*solved*) $\cup$ {*nexteqn*}
    unsolved := *subst*(*nexteqn*,*unsolved* \ {*nexteqn*})
  until *unsolved* = $\emptyset$
  orthonomic := *solved*
end

We note

1. Substitution of *nexteqn* into the set *solved* cannot cause these to lose solved form, since the derivative being substituted is lower in the ordering than all the leading derivatives in *solved*.

2. The number of equations in *unsolved* decreases by one for each pass through the loop, so the procedure terminates after a finite number of steps.

A.1.2 Reduced orthonomic form

The process of implicit substitution was described in §4.3.2. We denote implicit substitution of a leading derivative $\Delta_I \theta^j$ into $\Delta_J \theta^j$ throughout a system by

$$\textit{implicit\_subst}(\Delta_I \theta^j,\ \Delta_J \theta^j,\ \textit{system})$$

With this, a system may be brought to reduced orthonomic form (Definition 4.3.9) as follows:

Algorithm A.1.2. (reduce)

function *reduce*(*system*)
  repeat
    system := *orthonomic*(*system*)
    while exist $\Delta_J\theta^j$, derivative of leading $\Delta_I\theta^j$
      do
        system := *implicit_subst* ($\Delta_I\theta^j$, $\Delta_J\theta^j$, *system*)
      od
  until *system* is orthonomic
  reduce := *system*
end

We note

1. Carrying out all implicit substitutions is a finite process. Implicit substitution strictly *lowers* the order of derivatives occurring. Because the derivatives originally occurring are of finite order, it is impossible for there to be infinitely many implicit substitutions.

2. In this procedure, reduction to orthonomic form is executed a finite number of times. Each iteration generates an orthonomic system. If implicit substitutions in this system affect only the right hand sides, the system remains orthonomic, and we exit. If implicit substitutions also affect leading derivatives, orthonomic form is lost, and the loop is entered again. However, the order of leading derivatives has now strictly decreased. Sufficiently many iterations would therefore cause the system to disappear altogether. Hence the loop can only be traversed a finite number of times.

A.1.3 Involutive form

The process of computing compatibility conditions was described in §4.3.2. All compatibility conditions of a reduced orthonomic frame system DQ are generated, then simplified by implicit substitution from DQ. We denote the result of this process *compatibility(system)*. Involutive form (Definition 4.3.12) is achieved as follows:

Algorithm A.1.3. (involutive)

function *involutive*(*system*)
  repeat
    system := *reduce*(*system*)
    compat := *compatibility*(*system*)
    system := *system* $\cup$ *compat*
  until *compat* $= \emptyset$
  involutive := *system*
end

The argument which proves this process must be finite is originally due to Tresse [68] (see also [56]).

A.2 Group classification

To effect a classification for a frame system containing arbitrary elements, the procedures detailed in §A.1 must be modified. The changes are as follows:

1. A classifying (frame) system CQ is now present. This starts out as the frame auxiliary system, and has classifying equations appended to it as the calculation proceeds. Every time CQ is modified, we reduce it to frame involutive form.

2. The classifying system CQ and classifying inequalities CI are made available to each procedure, notably *solve*, *orthonomic*, *reduce*, and *involutive*.

3. All possible implicit substitutions from from the classifying system CQ must be carried out at each step. This keeps coefficients in the determining system as simple as possible. In particular, when *orthonomic* determines the leading derivative in an equation, the coefficient of this leading derivative will already have been simplified subject to CQ, and therefore will not vanish as a consequence of CQ.

4. The procedures *solve*, *orthonomic*, *reduce*, and *involutive* run to completion only if all divisions were unequivocally possible (i.e., the coefficients divided by are constants or are nonzero by virtue of inequalities CI). If *solve* cannot resolve division by the coefficient required, this coefficient must be noted as a *pivot*, and *solve* returns without effecting the division. Then *orthonomic*, *reduce*, and ultimately *involutive* are also unsuccessful, and return with only a partially processed determining system along with the unresolved *pivot*. A branching is now carried out, with involutive form being sought separately for the two cases $pivot \neq 0$ and $pivot = 0$.

We redefine the function *involutive* as follows.

Algorithm A.2.1. (involutive)

```
function involutive(DQ,CQ,CI,var pivot)
  INPUT:   DQ ...  frame determining system
           CQ ...  frame classifying system
           CI ...  classifying frame inequalities
  OUTPUT: (pivot = null)   divisions were successful
                           involutive gives involutive form of DQ
          (pivot non-null) a pivot was encountered
                           involutive is partial progress towards involutive form

    repeat
        DQ := reduce(DQ, CQ, CI, pivot)
        if pivot = (null) then
            compat := compatibility(DQ,CQ)
            DQ := DQ ∪ compat
        else
            involutive := DQ
            RETURN
        fi
    until compat = ∅
    involutive := DQ
end
```

The important feature is the 'unsuccessful' return, which returns *pivot* through the parameter list. Procedures *reduce*, *orthonomic* and *solve* are modified similarly: a successful return is indicated by a null *pivot* (all divisions were resolved); an unsuccessful return being accompanied by a nontrivial *pivot* expression.

As noted in §4.5, these modifications permit generation of a binary *classification tree* by a recursive procedure, which we reproduce here for convenience.

Algorithm A.2.2. (classify)

function *classify*(DQ,CQ,CI)
  INPUT: DQ ... frame determining system
  CQ ... frame classifying system
  CI ... classifying frame inequalities
  OUTPUT: Nothing
  SIDE EFFECT: Involutive form and corresponding classifying systems and inequalities for each leaf of the tree are printed out.

  $DQ := involutive(DQ, CQ, CI, pivot)$
  if $pivot =$ (null) **then**
    print(DQ,CQ,CI)
  else
    $classify(DQ, CQ, CI \cup \{pivot \neq 0\})$
    $classify(DQ, CQ \cup \{pivot = 0\}, CI)$
  fi
end

Appendix B

Structure Constants

Consider a frame determining system DQ for an algebra of Lie symmetry operators. A finite-dimensional Lie algebra is characterized by its structure constants C^k_{ij}. Our goal here is to show how C^k_{ij} can be found directly from the determining system DQ, *without* knowing its solutions. The method is an improvement of that originally given by Reid [57], which was based on Taylor expansion of the solution about an initial data point. This improvement has been implemented in [58] for the coordinate case.

Determining equations are linear and homogeneous, which implies that their solutions constitute a vector space. In addition, they have the property of being 'closed under commutation'. Suppose DQ are referred to a coordinate system. Let ξ, η be two solutions of DQ, so that $\mathbf{X} = \xi^i \partial_{w^i}$ and $\mathbf{Y} = \eta^i \partial_{w^i}$ are symmetry vector fields. Since $[\mathbf{X}, \mathbf{Y}]$ is also a symmetry vector field, it follows that if $\zeta^i = \xi^j \partial_{w^j} \eta^i - \eta^j \partial_{w^j} \xi^i$, then ζ is also a solution of DQ. Thus the commutator bracket on vector fields induces a commutator bracket on solutions of DQ, and we write $\zeta = [\xi, \eta]$.

Now suppose the determining equations are referred to a moving frame. The system DQ is certainly linear and homogeneous in its dependent variables θ^i. Once again the commutator bracket on vector fields induces a bracket on solutions:

Proposition B.0.3. *Let DQ be a differential system for the components of a symmetry vector field referred to a moving frame Δ with structure relations $[\Delta_i, \Delta_j] = \gamma^k_{ij} \Delta_k$. Let χ, ψ be two solutions of DQ. Then if*

$$\omega^l = \chi^j \Delta_j \psi^l - \psi^j \Delta_j \chi^l + \gamma^l_{ij} \chi^i \psi^j \tag{B.1}$$

then ω is also a solution of DQ. We write $\omega = [\chi, \psi]$.

Proof. Corresponding to the solutions χ, ψ are the symmetry vector fields $\mathbf{X} = \chi^i \Delta_i$ and $\mathbf{Y} = \psi^i \Delta_i$. Their commutator $\mathbf{Z} = [\mathbf{X}, \mathbf{Y}]$ is also a symmetry vector field. Writing $\mathbf{Z} = \omega^i \Delta_i$, it follows that ω is a solution of DQ. Computing components of ω yields (B.1). □

Suppose the solution space of DQ is of finite dimension r, i.e., the Lie symmetry algebra corresponds to an r-parameter group. If θ_i is a basis for the solutions of DQ[‡], it follows that $[\theta_i, \theta_j] = C^k_{ij} \theta_k$ for some constants C^k_{ij}, which are the structure

‡Indices θ_i which are 'down' count which solution, indices ψ^j which are 'up' count components of solutions.

constants of the Lie algebra. Of course, if solutions θ_i are known explicitly, the structure constants C^k_{ij} may be found by explicitly computing the commutator $[\theta_i, \theta_j]$. Construction of C^k_{ij} *without* knowing solutions is based on the following results.

Lemma B.0.4. *Let χ, ψ be solutions of the determining system DQ, and let $P(\chi)$, $P(\psi)$ represent the parametric derivatives of χ, ψ respectively. Thus $P^i(\chi)(w) = \Delta_J \chi^j(w)$ for some indices j, J. Let $\omega = [\chi, \psi]$ be the commutator of χ, ψ. Then each parametric derivative $P^l(\omega)$ is a skew symmetric bilinear function B^l of $P(\chi)$, $P(\psi)$:*

$$P^l(\omega)(w) = B^l\Big(P(\chi)(w), P(\psi)(w)\Big)$$

at each point w.

Proof. Equation (B.1) expresses the component ω^i as a skew symmetric bilinear function of $(\chi, \Delta_j \chi)$ and $(\psi, \Delta_j \psi)$. Applying frame derivative operators to (B.1), it is found that every derivative $\Delta_I \omega^i$ is a skew symmetric bilinear function of derivatives $(\Delta_J \chi^j)$ and $(\Delta_J \psi^j)$ of order $|J| = 0, 1, \ldots, |I| + 1$. This is true in particular for parametric derivatives of the commutator ω: each $P^l(\omega)$ is a skew symmetric bilinear function of derivatives $(\Delta_J \chi^j)$ and $(\Delta_J \psi^j)$.

Now both χ, ψ are solutions of the determining system DQ. Suppose the highest order derivatives in the bilinear functions for $P^l(\theta)$ are of order K. We prolong the determining system DQ to order K, and write it as $L(\theta) = AP(\theta)$, where $L(\theta)$ are the leading derivatives of θ up to order K, and A is a coefficient matrix. (All the terms here are functions of w.) The substitutions $L(\chi) = A\,P(\chi)$ and $L(\psi) = A\,P(\psi)$ (same matrix A in each case) in the expressions for $P^l(\theta)$ preserve the properties of bilinearity and skew symmetry, so we have

$$P^l(\omega) = B^l(P(\chi), P(\psi)).$$

where B^l is skew symmetric and bilinear. □

All of the above holds pointwise; we have suppressed the arguments $B^l(w)$, $P(\psi)(w)$, $A(w)$ etc. which occur throughout.

We now establish the main result.

Theorem B.0.5. *The parametric derivatives $P^l(\omega)$ of the commutator $\omega = [\chi, \psi]$ of two solutions χ, ψ of the determining system DQ are given by*

$$P^l(\omega)(w) = B^l_{ij}(w) P^i(\chi)(w) P^j(\psi)(w) \tag{B.2}$$

for some skew symmetric coefficients $B^l_{ij}(w) = -B^l_{ji}(w)$. At each point w_0 where initial data can be posed, the coefficients $B^l_{ij}(w_0)$ are structure constants C^l_{ij} of the Lie symmetry algebra with respect to some basis.

Proof. Equation (B.2) is just a restatement of the lemma above. Choose a basis of solutions $\theta_1, \theta_2, \ldots, \theta_r$ of DQ as follows. Pose as initial data for θ_i

$$P^j(\theta_i)(w_0) = \delta^j_i \qquad j = 1, 2, \ldots, r$$

where δ^j_i is the Kronecker delta. The frame Riquier conjecture (4.3.17) assures us that for each $i = 1, 2, \ldots, r$, this choice of initial data gives a corresponding unique solution $\theta_i(w)$ in some neighbourhood of w_0. With respect to this basis the structure constants are defined by

$$[\theta_i, \theta_j] = C^l_{ij}\theta_l.$$

Hence

$$\begin{aligned} P^l([\theta_i, \theta_j])(w_0) &= P^l(C^m_{ij}\theta_m)(w_0) \\ &= C^m_{ij} P^l(\theta_m)(w_0) \\ &= C^l_{ij} \end{aligned}$$

due to our choice of initial data. Thus C^l_{ij} is interpreted as the l-th piece of initial data for the commutator of solution i with solution j.

However, writing (B.2) with $\chi \equiv \theta_i$, $\psi \equiv \theta_j$, we have

$$\begin{aligned} P^l([\theta_i, \theta_j])(w_0) &= B^l_{mn}(w_0) P^m(\theta_i)(w_0) P^n(\theta_j)(w_0) \\ &= B^l_{ij}(w_0) \end{aligned}$$

again due to our choice of initial data. Hence for the basis chosen above we have $C^l_{ij} = B^l_{ij}(w_0)$. This holds for any suitable initial data point w_0 with its associated choice of basis. □

Hence to find structure constants, we follow the calculations described in the derivation of Lemma B.0.4, expressing the parametric derivatives of the commutator $[\chi, \psi]$ in terms of parametric derivatives of χ and of ψ. Picking off coefficients and evaluating at any convenient point w_0 yields suitable C^l_{ij}.

This process generalizes the process described by Reid, et al. [58] to frame determining systems.

Example B.0.6. Consider the involutive determining system

$$\begin{array}{lll} \Sigma_1\beta^3 = 0 & (\Sigma_1)^2\beta^2 = 0 & \Sigma_1\beta^1 = P\Sigma_1\beta^2 \\ \Sigma_2\beta^3 = 0 & \Sigma_2\beta^2 = 0 & \Sigma_2\beta^1 = 2\Sigma_1\beta^2 \\ \Sigma_3\beta^3 = 2P\Sigma_1\beta^2 & \Sigma_3\beta^2 = 0 & \Sigma_3\beta^1 = 2\Sigma_1\beta^2 \\ \Sigma_4\beta^3 = -P\beta^3 & \Sigma_4\beta^2 = -\frac{1}{2}\beta^1 & \Sigma_4\beta^1 = -\frac{1}{2}P\beta^1 - \beta^2 - \beta^3 \\ & & \beta^4 = 2\Sigma_1\beta^2, \end{array} \tag{B.3}$$

which is a special case of (4.102), obtained by setting $Q = 0$, $\sigma = 1$. The system is for components β^i of a symmetry vector field $\mathbf{Y} = \beta^i\Sigma_i$, referred to a frame Σ with structure relations (4.89)

$$\begin{array}{lll} [\Sigma_1, \Sigma_2] = 0 & [\Sigma_1, \Sigma_3] = 0 & [\Sigma_1, \Sigma_4] = -\frac{1}{2}P\Sigma_1 - \frac{1}{2}\Sigma_2 \\ & [\Sigma_2, \Sigma_3] = 0 & [\Sigma_2, \Sigma_4] = -\Sigma_1 \\ & & [\Sigma_3, \Sigma_4] = -\Sigma_1 - P\Sigma_3, \end{array}$$

where P is a constant. The parametric derivatives are β^1, β^2, β^3, $\Sigma_1\beta^2$. Let χ, ψ be two solutions, and let $\omega = [\chi, \psi]$. After using the structure relations, we find, for instance

$$\begin{aligned} \omega^1 &= (\chi^1\Sigma_1\psi^1 - \psi^1\Sigma_1\chi^1) + (\chi^2\Sigma_2\psi^1 - \psi^2\Sigma_2\chi^1) \\ &\quad + (\chi^3\Sigma_3\psi^1 - \psi^3\Sigma_3\chi^1) + (\chi^4\Sigma_4\psi^1 - \psi^4\Sigma_4\chi^1) \\ &\quad + -\tfrac{1}{2}P(\chi^1\psi^4 - \psi^1\chi^4) - (\chi^2\psi^4 - \psi^2\chi^4) - (\chi^3\psi^4 - \psi^3\chi^4). \end{aligned}$$

We may now substitute for the principal derivatives $\Sigma_1\psi^1 = P\Sigma_1\psi^2$ etc. from the determining system (B.3), to obtain

$$\omega^1 = P(\chi^1\Sigma_1\psi^2 - \psi^1\Sigma_1\chi^2) + 2(\chi^2\Sigma_1\psi^2 - \psi^2\Sigma_1\chi^2) + 2(\chi^3\Sigma_1\psi^2 - \psi^3\Sigma_1\chi^2).$$

Similarly, we find

$$\begin{aligned} \omega^2 &= (\chi^1\Sigma_1\psi^2 - \psi^1\Sigma_1\chi^2) \\ \omega^3 &= 2P(\chi^3\Sigma_1\psi^2 - \psi^3\Sigma_1\chi^2) \\ \Sigma_1\omega^2 &= 0. \end{aligned}$$

The coefficients in these bilinear expressions are the structure constants C^l_{ij}. Hence the commutation relations for the Lie symmetry algebra are

$$\begin{array}{lll} [\mathbf{X}_1, \mathbf{X}_2] = 0 & [\mathbf{X}_1, \mathbf{X}_3] = 0 & [\mathbf{X}_1, \mathbf{X}_4] = P\mathbf{X}_1 + \mathbf{X}_2 \\ & [\mathbf{X}_2, \mathbf{X}_3] = 0 & [\mathbf{X}_2, \mathbf{X}_4] = 2\mathbf{X}_1 \\ & & [\mathbf{X}_3, \mathbf{X}_4] = 2\mathbf{X}_1 + 2P\mathbf{X}_3 \end{array}$$

Note that we use only existence of the basis θ_i: explicit construction of the solution basis θ_i is not necessary.

Appendix C

Similarity Solution for Nonlinear Diffusion

We examine the system of o.d.e.'s (3.59) for Boltzmann's similarity solution of the nonlinear diffusion equation (3.57):

$$\begin{aligned} \frac{dy}{du} &= \frac{z}{2} \\ \frac{dz}{du} &= -\frac{D(u)}{y}. \end{aligned} \tag{C.1}$$

This system is subject to the boundary conditions

$$\begin{aligned} &z = 0 \quad u = u_0 \\ &z \to \infty \quad u = u_i \quad y = 0\,. \end{aligned} \tag{C.2}$$

The problem (C.1, C.2) is important not only because it is mathematically simple, but because the boundary conditions are easy to realize experimentally [21]. Note that the affine equivalence transformations generated by (3.69)

$$\begin{aligned} z' &= \lambda z \\ u' &= \alpha u + \beta \\ y' &= \lambda \alpha y \\ a' &= \lambda^2 a \end{aligned} \tag{C.3}$$

are available, and can be used to rescale boundary conditions (C.2) to

$$z = 0 \quad u = 1 \tag{C.4a}$$

$$z \to \infty \quad u = 0 \quad y = 0\,. \tag{C.4b}$$

C.1 Power law diffusivity

For arbitrary $D(u)$, the problem (C.1, C.4) cannot be further simplified, and solutions are obtained numerically by shooting [62]. For several diffusivities, additional simplification is possible. The general analytical reduction for these cases

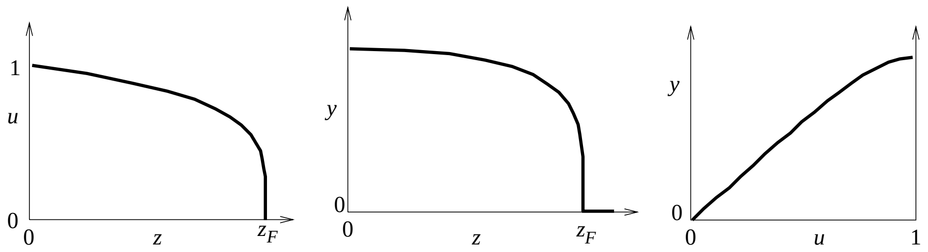

Figure C.1: Relation between concentration u, flux y and spatial coordinate z for the diffusion problem (C.1, C.6, C.4). At the singularity $u = 0$, $y = 0$, $z = z_F$, the flux and concentration are both continuous.

was given by Lisle and Parlange [45], from which all of the following results are drawn.

When $D(u)$ obeys a power law

$$D(u) = (au + b)^m \tag{C.5}$$

a scaling symmetry becomes available. The constant a can be scaled to any convenient value (e.g. unity), but in general b cannot be removed by the transformations (C.3) without disturbing the rescaled boundary conditions (C.4). The symmetry is the basis of several analytical methods for reducing the boundary value problem. The greatest simplification is achieved when $b = 0$, $a > 0$ and $m > 0$ in (C.5), which is a case of physical significance. In this case, we write the diffusivity as

$$D(u) = (m + 1)u^m \tag{C.6}$$

where a scaling from (C.3) has been used to enforce the condition $\int_0^1 D(u)\,du = 1$. General results [62, 63, 6] show the solutions have singular properties. The diffusant penetrates only as far as a finite 'front' i.e., there exists a value z_F of z such that $u(z) = 0$, $y(z) = 0$ for $z \geq z_F$. Both concentration u and flux y vanish at the front, so that these functions are continuous, but their derivatives may not be (Figure C.1). This remarkable situation is associated with the fact that when $D(0) = 0$, the coefficient of u_{xx} in the original partial d.e. (3.65) vanishes, and the equation is not parabolic in the neighbourhood of this point.

Several methods for dealing with the boundary value problem (C.1, C.4, C.6) are available. All of them rely for their success on a symmetry transformation

$$\begin{aligned} u' &= c^2 u \\ z' &= c^m z \\ y' &= c^{m+2} y, \end{aligned} \tag{C.7}$$

which is inherited from the equivalence subgroup (C.3). This scaling symmetry maps the front $u = 0$, $z = z_F$, $y = 0$ to a front $u' = 0$, $z' = z'_F$, $y' = 0$; and maps the surface $z = 0$ to the surface $z' = 0$.

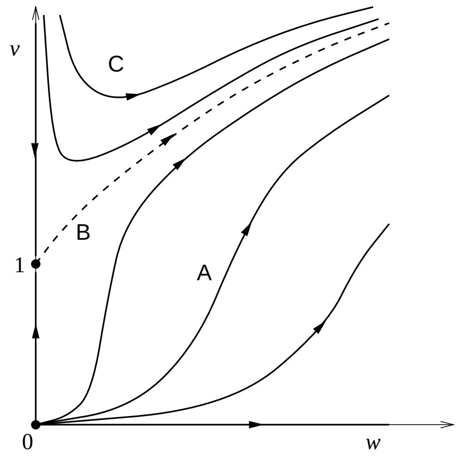

Figure C.2: Phase portrait for first quadrant of (w, v) plane for (C.9). The three classes A, B, C of trajectories represent respectively smooth solutions, the singular solutions shown in Figure C.1, and singular solutions with discontinuous flux and concentration.

C.1.1 Phase reduction

The existence of symmetry (C.7) allows (C.1, C.6) to be reduced to the phase plane. Introduce the invariants

$$\begin{aligned} w &= u^m z^{-2} \\ v &= 2y/(uz) \,. \end{aligned} \tag{C.8}$$

of the transformation group (C.7). Standard theory [13, §3.3], [47, §2.5] shows that system (C.1, C.6) is reduced to the single first order equation

$$\frac{dv}{dw} = \frac{v}{w}\,\frac{1 - v + 2w}{mv + 4w} \tag{C.9}$$

in the half-plane $w \geq 0$. There are two finite critical points: the node-like quadratic singularity at $w = 0$, $v = 0$, representing solutions $u = \text{const}$; and a saddle at $w = 0$, $v = 1$. The fourth quadrant is irrelevant for our purposes. The phase portrait of (C.9) for the first quadrant of the w, v plane is sketched in Figure C.2. The trajectory of interest is the separatrix emanating from the saddle and exiting to infinity with $v \sim qw^{1/2}$, $q > 0$.

The phase equation (C.9) can be (numerically) solved for $v(w)$, where the initial condition $w = 0, v = 1$ is enforced. A starting value for the derivative dv/dw at this singular point is furnished by Taylor expansion. Once $v(w)$ is known, the concentration $u(w)$ is recovered by the quadrature

$$\int_w^\infty \frac{v(\bar{w})\, d\bar{w}}{\bar{w}(mv(\bar{w}) + 4\bar{w})} = \int_u^1 \frac{d\bar{u}}{\bar{u}} \tag{C.10}$$

after which y, z follow from (C.8):

$$\begin{aligned} z &= u^{m/2} w^{-1/2} \\ y &= \tfrac{1}{2} v(w) u(w) z(w) \end{aligned} \tag{C.11}$$

Equations (C.10, C.11) yield the solution u, y, z parametrically as a function of w. Note that this procedure effectively reduces the *boundary* value problem (C.1, C.6, C.4) to an *initial* value problem. The solution method of Parlange, et al. [53] explicitly used this reduction.

C.1.2 Exact shooting

Rather than reducing the equation to the phase plane, it is possible to directly use symmetry (C.7) to map a numerical solution of the original system (C.1, C.6) to the desired solution satisfying the correct boundary conditions (C.4). The procedure, due to Shampine [64] is as follows:

1. Guess a value of the front location z_F, and integrate (C.1, C.6) with the initial condition $z = z_F$, $u = 0$, $y = 0$. It may be advantageous to write the system with u as the independent variable to begin the integration. Taylor expansion is necessary to resolve the indeterminacy of the system at the initial point.

2. Terminate the integration at the surface $z = 0$. Let the value of u at this point be $\hat{u}$.

3. Compute $c = \hat{u}^{-1/2}$. Perform transformation (C.7) with this value of c on the solution thus constructed. The functions $u'(z')$, $y'(z')$ satisfy the o.d.e. (C.1, C.6) *and* boundary conditions (C.4), and hence specify the exact solution of the boundary value problem.

C.1.3 Series solution

Heaslet and Alksne [30] found a formal series solution of the boundary value problem (C.1, C.6, C.4), of the form

$$u = (z - z_F)^{1/m} \sum_{k=1}^{\infty} a_k (z - z_F)^k \tag{C.12}$$

For our purposes it is more convenient to recast the series with u as the independent variable:

$$y = \tfrac{1}{2} z_F u \left(1 - \sum_{k=1}^{\infty} p_k(m) \left[\frac{2u^m}{m z_F^2} \right]^k \right) \tag{C.13}$$

and

$$z = z_F \left(1 - \sum_{k=1}^{\infty} (1 + mk) p_k(m) \left[\frac{2u^m}{m z_F^2} \right]^k \right) . \tag{C.14}$$

Boundary condition (C.4b) is automatically satisfied by a series of this form. The coefficients $p_k(m)$ are found from an explicit recurrence obtained by substituting the series (C.13, C.14) into the system (C.1, C.6). The front location z_F is then found by enforcing boundary condition (C.4a).

C.2 Modified power law diffusivity

By applying equivalence transformations (3.61), the power law diffusivity (C.6) can be mapped to the diffusivity

$$D(u') = (m+1)\frac{1}{(\gamma u' + \delta)^2}\left(\frac{\alpha u' + \beta}{\gamma u' + \delta}\right)^m . \tag{C.15}$$

We seek the solution of the boundary value problem for the 'dashed' system (C.1, C.15, C.4), concentrating on those parameter values obtainable by transformation of the singular problem treated above for the power law. Using affine transformations (C.3) we map diffusivity (C.15) to the form

$$D(u) = (m+1)\frac{1-\mu}{(1-\mu u)^2}\left[\frac{(1-\mu)u}{1-\mu u}\right]^m \tag{C.16}$$

where $\mu \in (-\infty, 1)$, $m > 0$, and the awkward looking scaling is chosen so that $\int_0^1 D(u)\,du = 1$. Despite the singular behaviour at the point $u = \mu^{-1}$, this diffusivity is of some physical interest [2]. It might be supposed that the boundary value problem (C.1, C.6, C.4) for the power law can be mapped to the corresponding problem (C.1, C.16, C.4) for diffusivity (C.16), but this is not so. The *solution curves* of the two boundary value problems are in correspondence, but the surface boundary condition (C.4a) maps to a *nonzero* value of z if $\mu \neq 0$ i.e. to a *moving* boundary. Although this moving boundary problem may therefore be easily solved, we are most interested in a *fixed* boundary condition for the new diffusivity, corresponding to a moving boundary for the power law diffusivity. This makes the mapping process slightly awkward. Rather than explicitly carrying this out, Lisle and Parlange [45] use the mapping between solution curves to map the *methods* across from the power law to the new case.

The symmetry (C.7) for the power law case (C.6) maps to the symmetry transformation

$$\begin{aligned} u' &= \frac{\tau u}{1-(1-\tau)\mu u} \\ y' &= \tau^{m/2}\frac{\tau y}{1-(1-\tau)\mu u} \\ z' &= \tau^{m/2}[\, z + (1-\tau)\mu(2y - zu)\,] \end{aligned} \tag{C.17}$$

for diffusivity (C.16), where τ is the group parameter.

C.2.1 Phase reduction

The new diffusivity (C.16) can be mapped to the power law (C.6) and thence reduced to the *same* equation (C.9) in the phase plane. The same separatrix trajectory is required, but instead of taking w to infinity, the integration is stopped at the value $w = w_0$ such that $v(w_0) = \mu^{-1}$. The quadrature (C.10) is replaced by

$$\int_w^{w_0} \frac{v(\bar{w})\,d\bar{w}}{\bar{w}(mv(\bar{w}) + 4\bar{w})} = \int_{u'}^1 \frac{d\bar{u}}{\bar{u}(1-\mu\bar{u})} . \tag{C.18}$$

Note that instead of having a termination point fixed a priori as in the power law case, the point w_0 at which integration is terminated is determined in the course of the calculation. Nevertheless no iteration is required, and the problem is effectively reduced to an initial value problem.

C.2.2 Exact shooting

Direct use of symmetry transformation (C.17) allows the numerical solution of the boundary value problem (C.1, C.16, C.4) to be simplified. The procedure is as follows:

1. Guess a value of the front location z_F, and integrate (C.1, C.16) with the initial condition $z = z_F$, $u = 0$, $y = 0$.

2. Terminate the integration when values $(\bar{z}, \bar{u}, \bar{f})$ of (z, u, f) are encountered satisfying

$$\frac{\bar{z}\bar{u}}{2\bar{f}} = \frac{\mu(1-\bar{u})}{1-\mu\bar{u}} \tag{C.19}$$

3. Compute $\bar{\tau} = (1-\mu)\bar{u}/(1-\mu\bar{u})$. Perform the transformation (C.17) with $\tau = \bar{\tau}$ on the solution thus constructed. The functions $u'(z')$, $y'(z')$ satisfy the o.d.e. (C.1, C.16) and boundary conditions (C.4), and hence are the exact solution. The exact front location is $z'_F = \bar{\tau}^{-m/2} z_F$.

This generalizes Shampine's [64] method for the power law (C.6), showing that boundary value problem (C.1, C.16, C.4) may be reduced to solving an *initial* value problem.

C.2.3 Series solution

Applying transformation (3.61) to the series solution (C.13, C.14) found above for the power law shows that diffusivity (C.16) admits a solution with y/u expanded in powers of $x(u)$, where

$$x(u) = \left[\frac{(1-\mu)u}{1-\mu u}\right]^m . \tag{C.20}$$

This is conveniently written

$$y(u) = \tfrac{1}{2} z_F u \left(1 - \sum_{k=1}^{\infty} p_k(m)\ [\,\bar{t}\, x(u)]^k\right) \tag{C.21}$$

where z_F is the (as yet unknown) location of the front;

$$\bar{t} = 2(1-\mu)/(m z_F^2)\ ; \tag{C.22}$$

and the $p_k(m)$ are obtained from the same recurrence as for the power law:

$$\begin{aligned} p_1 &= 1 \\ p_k &= \frac{1}{k(1+mk)} \sum_{j=1}^{k-1} j(1+mj)\, p_j\, p_{k-j}\ , \qquad k \geq 2\ . \end{aligned} \tag{C.23}$$

Differentiating once shows

$$z(u) = z_F \left(1 - \sum_{k=1}^{\infty} a_k(m, \mu) \; [\,\bar{t}\, x(u)]^k \right) . \tag{C.24}$$

where

$$a_k(m, \mu) = p_k(m) \left(1 + \frac{mk}{1-\mu} \right) , \tag{C.25}$$

Boundary condition (C.4b) is automatically satisfied when y, z have the form (C.21, C.24). The other boundary condition (C.4a) is to be satisfied by choosing the value z_F.

Assuming series (C.24) is valid to $u = 1$, $\bar{t}$ must satisfy the equation

$$\sum_{k=1}^{\infty} a_k(m, \mu) \bar{t}^{\,k} = 1 \tag{C.26}$$

so that $\bar{t}$ is a function of m, μ. The series solution is given by (C.13–C.26).

The obvious way to solve (C.26) approximately is to truncate the series after a finite number of terms and solve the resulting polynomial equation. An alternative, more explicit, and more accurate method is to use reversion of series on (C.26). Let

$$r(t) = \sum_{k=1}^{\infty} a_k t^k$$

and revert this series to obtain

$$t = \sum_{l=1}^{\infty} b_l \, r^l \; . \tag{C.27}$$

The first few b_l are given by [1, 3.6.25]. The value $\bar{t}$ is found by evaluating (C.27) at $r = 1$, so that

$$\bar{t} = \sum_{l=1}^{\infty} b_l(m, \mu) \; . \tag{C.28}$$

The procedure is valid provided $\varepsilon = (1 - \mu) m^{-1}$ is sufficiently small. Some numerical results using both the series and exact shooting are given by Lisle and Parlange [45]. The series is exceptionally accurate for ε less than about 0.25, and loses its usefulness only when ε is larger than about 2. The accuracy properties are essentially independent of the value of m.

C.3 Discussion

The above methods for dealing with the new diffusivity (C.16) all result from the enlargement of the equivalence group (3.61) which results by considering the system form (C.1) of the d.e.'s. All the well-known solution methods for dealing with the power law (C.6) carry over to the new case. Of course in principle the methods detailed above for dealing with the new case could be directly derived from symmetry properties of the equation. However, the form (C.1), as a system of first order ordinary d.e.'s, cannot be completely group analyzed [52]. Eliminating z, one obtains the scalar equation

$$y\frac{d^2y}{du^2} = -\frac{1}{2}D(u) \ . \tag{C.29}$$

This *can* be group analyzed, and the symmetry (C.17) found: this leads to the phase reduction and exact shooting methods described above. However the correct form (C.21) of the series solution is far from obvious if one does not map from the power law case (C.13). Moreover, mapping from the power law unifies the two cases; without the availability of the equivalence transformation (3.61) the reductions would appear similar but unrelated.

The symmetry properties used here are inherited from the equivalence group of the o.d.e. system (C.1). As described in §3.4.1, this equivalence group is in turn inherited from the p.d.e. potential system (3.57). Hence all of the results given here follow from equivalence analysis.

Bibliography

[1] Abramowitz, M. and I.A. Stegun. 1964. *Handbook of Mathematical Functions.* Dover, New York.

[2] Ahuja, L.R. and D. Swartzendruber. 1972. An improved form of soil-water diffusivity function. *Soil Sci. Soc. Am. Proc.* **36**: 9–14.

[3] Akhatov, I.Sh., R.K. Gazizov and N.Kh. Ibragimov. 1987. Group classification of the equations of nonlinear filtration. *Soviet Math. Dokl.* **35** : 384–386.

[4] Akhatov, I.Sh., R.K. Gazizov and N.Kh. Ibragimov. 1991. Nonlocal symmetries: heuristic approach. *J. Soviet Math.* **55**: 1401–1450.

[5] Arnold, V.I. 1990 (2nd ed.) *Mathematical Methods of Classical Mechanics.* Springer, New York.

[6] Atkinson, F.V. and L.A. Peletier. 1971. Similarity profiles of flows through porous media. *Arch. Rat. Mech. Anal.* **42**: 369–379.

[7] Bluman, G.W. 1967. *Construction of solutions to partial differential equations by the use of transformation groups.* Ph.D. thesis, California Institute of Technology.

[8] Bluman, G.W. and J.D. Cole. 1969. The general similarity solution of the heat equation. *J. Math. Mech.* **18**: 1025–1042.

[9] Bluman, G.W. and J.D. Cole. 1974. *Similarity Methods for Differential Equations.* Springer, New York.

[10] Bluman, G.W. and S. Kumei. 1980. On the remarkable nonlinear diffusion equation $\frac{\partial}{\partial x}[a(u+b)^{-2}\frac{\partial u}{\partial x}] - \frac{\partial u}{\partial t} = 0$. *J. Math. Phys.* **21**: 1019–1023.

[11] Bluman, G.W. and S. Kumei. 1987. Invariance properties of the wave equation. *J. Math. Phys.* **28**: 307–312.

[12] Bluman, G.W. and S. Kumei. 1988. Exact solutions for wave equations of two-layered media with smooth transition. *J. Math. Phys.* **29**: 86–96.

[13] Bluman, G.W. and S. Kumei. 1989. *Symmetries and Differential Equations.* Springer, New York.

[14] Bluman, G.W. and S. Kumei. 1990. Symmetry-based algorithms to relate partial differential equations: I. Local symmetries. *Europ. J. Appl. Math.* **1**: 189–216.

[15] Bluman, G.W., S. Kumei and G.J. Reid. 1988. New classes of symmetries for partial differential equations. *J. Math. Phys.* **29**: 806–811; Erratum, *J. Math. Phys.* **29**: 2320.

[16] Bluman, G.W. and G.J. Reid. 1988. New symmetries for ordinary differential equations. *IMA J. Appl. Math.* **40**: 87–94.

[17] Bryant, R.L. 1987. On notions of equivalence of variational problems with one independent variable, in *Differential Geometry* (M. Luksic, C. Martin and W.F. Shadwick eds.) AMS, Providence, RI.

[18] Cartan, E.J. 1908. Les sous-groupes des groupes continus de transformations. *Oeuvres Complètes* Part II, Vol. 2, p.719–856. Gauthier-Villars, Paris.

[19] Cartan, E.J. 1910. Les systèmes de Pfaff à cinq variables et les équations aux dérivées partielles du second ordre. *Oeuvres Complètes* Part II, Vol. 2, p.927–1010.

[20] Char, B.W., K.O. Geddes, G.H. Gonnet, M.B. Monagan, and S.M. Watt. 1990. (3rd ed.) MAPLE *Reference Manual.* Watcom, Waterloo.

[21] Crank, J. 1956. *The Mathematics of Diffusion.* Cambridge Univ. Press, Cambridge.

[22] Eisenhart, L.P. 1933. *Continuous Groups of Transformations.* Princeton Univ. Press, Princeton, NJ.

[23] Fujita, H. 1954. The exact pattern of a concentration-dependent diffusion in a semi-infinite medium, Part III. *Text. Res. J.* **24**: 234–240.

[24] Fokas, A.S. and Y.C. Yortsos. 1982. On the exactly solvable equation $S_t = [(\beta S+\gamma)^{-2}S_x]_x+\alpha(\beta S+\gamma)^{-2}S_x$ occurring in two phase flow in porous media. *SIAM J. Appl. Math.* **42**: 318–332.

[25] Gardner, R.B. 1983. Differential Geometric Methods Interfacing Control Theory, in *Differential Geometric Control Theory.* (R. Brockett, R. Millman and H. Sussman eds.) Birkhauser, Boston.

[26] Gardner, R.B. 1989. *The Method of Equivalence and Its Applications.* SIAM, Philadelphia.

[27] Gardner, R.B. and W.F. Shadwick. 1987. Overdetermined equivalence problems with an application to feedback equivalence, in *Differential Geometry.* (Luksic, M., C. Martin and W.F. Shadwick eds.) AMS, Providence, RI.

[28] Goldstein, H. 1980. (2nd ed.) *Classical Mechanics.* Addison-Wesley, Reading, MA.

[29] Harrison, B.K. and F.B. Estabrook. 1971. Geometric approach to invariance groups and solution of partial differential equations. *J. Math. Phys.* **12**: 653–666.

[30] Heaslet, M.A. and A. Alksne. 1961. Diffusion from a fixed surface with a concentration-dependent coefficient. *J. SIAM* **9**: 584–596.

[31] Hsu, L., N. Kamran and P.J. Olver. 1989. Equivalence of higher order Lagrangians. II. The Cartan form for particle Lagrangians. *J. Math. Phys.* **30:** 902–906.

[32] Ibragimov, N.H., M. Torrisi, and A. Valenti. 1991. Preliminary group classification of equations $v_{tt} = f(x, v_x)v_{xx} + g(x, v_x)$. *J. Math. Phys.* **32**: 2988–2995.

[33] Janet, M. 1920. Sur les systèmes d'équations aux dérivées partielles. *J. de math* **3**: 65–151.

[34] Kamran, N. 1989. Contributions to the study of the equivalence problem of Élie Cartan and its applications to partial and ordinary differential equations. Preprint, Institute for Advanced Study, Princeton Univ., Princeton.

[35] Kamran, N., K.G. Lamb, and W.F. Shadwick. 1985. The local equivalence problem for $d^2y/dx^2 = F(x, y, dy/dx)$ and the Painlevé transcendents. *J. Diff. Geom.* **22**: 139–150.

[36] Kamran, N. and P.J. Olver. 1989. Equivalence problems for first order Lagrangians on the line. *J. Diff. Eq.* **80:** 32–78.

[37] Kamran, N. and P.J. Olver. 1989. Equivalence of higher order Lagrangians. I. Formulation and reduction. IMA Preprint #494, Univ. of Minnesota.

[38] Kamran, N. and P.J. Olver. 1989. Equivalence of differential operators. *SIAM J. Math. Anal.* **20:** 1172–1185.

[39] Kamran, N. and W.F. Shadwick. 1987. Cartan's method of equivalence and the classification of second order ordinary differential equations, in *Differential Geometry*. (Luksic, M., C. Martin and W.F. Shadwick eds.) AMS, Providence, RI.

[40] Kersten, P.H.M. 1987. *Infinitesimal symmetries: a computational approach.* Centrum voor Wiskunde en Informatica, Amsterdam.

[41] Kumei, S. and G.W. Bluman. 1982. When nonlinear differential equations are equivalent to linear differential equations. *SIAM J. Appl. Math.* **42**: 1157–1173.

[42] Landau, L.D. and E.M. Lifshitz. 1968. *Mechanics.* Pergamon Press, Oxford.

[43] Lie, S. 1884. Über Differentialinvarianten. *Math. Annalen* **24**: 537–578. (English translation in R. Hermann, 1976. *Sophus Lie's 1884 differential invariant paper.*, MathSci Press, Brookline, MA.)

[44] Lisle, I.G. 1983. Group properties of differential equations. Masters project, University of Queensland.

[45] Lisle, I.G. and J.-Y. Parlange. 1992. Analytical reduction for a concentration dependent diffusion problem. *ZAMP* (in press).

[46] Ma, A. 1990. *Extended Group Analysis of the Wave Equation.* M.Sc. Thesis, University of British Columbia.

[47] Olver, P.J. 1986. *Applications of Lie Groups to Differential Equations.* Springer, New York.

[48] Olver, P.J. and P. Rosenau. 1987. Group-invariant solutions of differential equations. *SIAM J. Appl. Math.* **47**: 263–278.

[49] Oron, A. and P. Rosenau. 1986. Some symmetries of the nonlinear heat and wave equations. *Phys. Lett.* **A 118:** 172–176.

[50] Ovsiannikov, L.V. 1962. *Group Properties of Differential Equations.* Novosibirsk. (in Russian: English translation by G.W. Bluman, unpublished.)

[51] Ovsiannikov, L.V. 1974. Some problems arising in group analysis of differential equations, in *Symmetry, Similarity and Group Theoretic Methods in Mechanics.* (P.G. Glockner and M.C. Singh eds.) Amer. Acad. of Mech.

[52] Ovsiannikov, L.V. 1982. *Group Analysis of Differential Equations.* Academic Press, New York.

[53] Parlange, J.-Y., R.D. Braddock and B.T. Chu. 1980. First integrals of the diffusion equation; an extension of the Fujita solutions. *Soil Sci. Soc. Am. J.* **44**: 908–911.

[54] Philip, J.R. 1970. Flow in porous media. *Ann. Rev. Fluid Mech.* **2** : 177–204.

[55] Reid, G.J. 1990. A triangularization algorithm which determines the Lie symmetry algebra of any system of PDEs. *J. Phys.* **A23**: L853–859.

[56] Reid, G.J. 1991. Algorithms for reducing a system of PDEs to standard form, determining the dimension of its solution space and calculating its Taylor series solution. *Euro. J. Appl. Maths.* **2**: 293–318.

[57] Reid, G.J. 1991. Finding abstract Lie symmetry algebras of differential equations without integrating determining equations. *Euro. J. Appl. Maths.* **2**: 319–340.

[58] Reid, G.J., I.G. Lisle, A. Boulton and A. D. Wittkopf. 1992. Algorithmic determination of commutation relations for Lie symmetry algebras of PDEs. *ISSAC '92 Conference Proceedings* (to appear).

[59] Riquier, C. 1910. *Les Systèmes d'Équations aux Dérivées Partielles.* Gauthier-Villars, Paris.

[60] Rosen, G. 1979. Nonlinear heat conduction in solid H_2. *Phys. Rev.* **B 19**: 2398–2399.

[61] Shadwick, W.F. 1987. Cartan's method of equivalence and the calculus of variations, in *Differential Geometry and its Applications.* (D. Krupka and A. Švec eds.) D. Reidel, Dordrecht.

[62] Shampine, L.F. 1973. Concentration-dependent diffusion. *Quart. Appl. Math.* **30**: 441–452.

[63] Shampine, L.F. 1973. Concentration-dependent diffusion. II. Singular problems. *Quart. Appl. Math.* **31**: 287–293.

[64] Shampine, L.F. 1973. Some singular concentration dependent diffusion problems. *ZAMM* **53**: 421–422.

[65] Spivak, M.D. 1979. (2nd ed.) *A Comprehensive Introduction to Differential Geometry* Vols. I, II. Publish or Perish, Houston, Texas.

[66] Storm, M.L. 1951. Heat conduction in simple metals. *J. Appl. Phys.* **22**: 940–951.

[67] Thomas, J.M. 1929. Riquier's existence theorems. *Annals of Maths.* **30**: 285–310.

[68] Tresse, A. 1894. Sur les invariants différentiels des groupes continus de transformations. *Acta Mathematica* **18**: 1–88. (English translation I. Lisle 1989, available from the author)

[69] Tresse, A. 1896. Détermination des invariants ponctuels de l'équation différentielle ordinaire du second ordre $y'' = \omega(x, y, y')$. S. Hirzel, Leipzig.

[70] Varley, E. and B.R. Seymour. 1985. Exact solutions for large amplitude waves in dispersive and dissipative systems. *Stud. Appl. Maths.* **72**: 241–262.