Symmetry Methods for Integrable Systems

Joseph Anthony Mulvey

A thesis submitted for the degree of Doctor of Philosophy
based on research carried out in the Department of
Mathematical Sciences, University of Durham.

May 1996

Abstract

This thesis discusses various properties of a number of differential equations which we will term "integrable". There are many definitions of this word, but we will confine ourselves to two possible characterisations — either an equation can be transformed by a suitable change of variables to a linear equation, or there exists an infinite number of conserved quantities associated with the equation that commute with each other via some Hamiltonian structure.

Both of these definitions rely heavily on the concept of the *symmetry* of a differential equation, and so Chapters 1 and 2 introduce and explain this idea, based on a geometrical theory of p.d.e.s, and describe the interaction of such methods with variational calculus and Hamiltonian systems.

Chapter 3 discusses a somewhat *ad hoc* method for solving evolution equations involving a series ansatz that reproduces well–known solutions. The method seems to be related to symmetry methods, although the precise connection is unclear.

The rest of the thesis is dedicated to the so–called *Universal Field Equations* and related models. In Chapter 4 we look at the simplest two–dimensional cases, the Bateman and Born–Infeld equations. By looking at their generalised symmetries and Hamiltonian structures, we can prove that these equations satisfy both the definitions of integrability mentioned above. Chapter Five contains the general argument which demonstrates the linearisability of the Bateman Universal equation by calculation of its generalised symmetries. These symmetries are helpful in analysing and generalising the Lagrangian structure of Universal equations. An example of a linearisable analogue of the Born–Infeld equation is also included. The chapter concludes with some speculation on Hamiltoian properties.

Statement of Originality

This thesis is comprised of research carried out in the Department of Mathematical Sciences, University of Durham, between October 1992 and October 1995. No part of it has been submitted to fulfil the requirements of any degree at this or any other university.

No claim of originality is made for the greater part of the material in Chapters 1 or 2 which contain background material. However the calculation of the symmetry generators of the Bateman equation (Table 1.1) is believed to be original. The work in Chapter 3 is based on research conducted jointly with Dr D.B. Fairlie and collected in the unpublished work [1], but many of the results appear to echo the work in [2]. Chapter 4 is comprised of original research carried out by the author with the exception of some elements of Section 4.1 which appear in [3], a joint work with Dr D.B. Fairlie. Chapter 5 is also the author's own research, apart from some material in Section 5.1 (which is introductory), and most of the material in Section 5.4 which appears in [3]. Parts of Chapter 4 are to be published [4], while some results from Chapter Five are available in the preprint [5]. Results due to other authors are cited appropriately.

Acknowledgements

A postgraduate's lot is not always an 'appy one, but the staff and my fellow students in the Department of Mathematical Sciences have conspired to make my stay in Durham a pleasant and memorable experience.

It is a pleasure to thank all those members of the academic staff who have given me help with technical matters over the past three years. I rarely found a door closed to my queries. Special thanks must go to James F. Blowey for help with LaTeX and for having the good taste to be an alumnus of Sussex University. Red Brickies Together!

It is customary to gush about the love and support one has received from one's peers during the course and this is a custom which I am very happy to observe. Special mention goes to:

- Alistair MacIntyre: I could hardly know him any better;
- Ben Jordan: for trash talk;
- Rachel–Louise Koktava: hope you go down a Bomb!
- Dr J. Martin Speight: for his infectious enthusiasm;
- Ramón Cóva: ay, muchacho!
- Andy Perella: for football and much else;
- Darren Wilkinson: for user–friendliness;
- Karen McGaul: for not talking about rowing *too* much.

Finally, and most importantly, I owe a huge debt of gratitude to my supervisor, David Fairlie. As well as benefitting enormously from his originality and expertise, I have found him to be a thoughtful and committed advisor who has kept me on the mathematical straight–and–narrow. It has been a pleasure to know him, and work

with him. Special thanks must also go to Lee Fairlie for her wonderful hospitality on those many happy occasions when I have been welcomed into their home.

My work was funded by a studentship from the Department of Education for Northern Ireland.

Contents

List of Figures

Chapter 1

Jet Bundles, Prolongation, and Symmetry

1.1 Introduction

Lie symmetry techniques have wide applicability in applied mathematics and physics. In this thesis, they will be used to find solutions of differential equations, study variational properties and conservation laws of non-standard classical field theories, and analyse the properties of infinite–dimensional Hamiltonian systems. This chapter introduces the methods and ideas to be applied in subsequent calculations.

However, in order to study the symmetries of a system of differential equations S using Lie groups and algebras, we need a geometric picture of S, and a corresponding language in which we can write our symmetry conditions. The most convenient formalism for the purpose is that of the *jet bundle*. While not wishing to wallow in pedantry, it will be helpful to review as much of this theory as is necessary to understand the context in which an applicable symmetry theory lies.

Roughly, the jet bundle theory tells us how and when to regard the derivatives of the fields in our classical equations as independent coordinates on a manifold. The differential equations of interest are viewed as algebraic equations defining a submanifold of this space and we can study the action of Lie group transformations on this

submanifold. In practice, it is the infinitesimal invariances due to the corresponding Lie algebra transformations that are calculated.

In this chapter, the aim is to introduce as much of jet space geometry as is necessary for future applications, and outline the methods used in real calculations. In particular, we will be interested in the idea of a *prolongation* of a geometrical object to the corresponding jet space. The strategy will be to introduce the first order jet spaces in some detail and then baldly state the results for the higher order spaces, as the reasoning behind them is essentially the same. Any subtleties arising from the transition to more derivatives will be left to the textbooks. We will then discuss the largely algorithmic procedures for finding the point symmetries of S and generating the corresponding similarity solutions. The chapter will conclude with a more cursory presentation of generalised symmetries, which will nonetheless be of great importance later.

The symmetry properties of variational (Lagrangian or Hamiltonian) equations will be reviewed in Chapter 2.

This presentation of these ideas owes much to the books by Saunders [6], Olver [7], and Hermann [8]. Additional matters are dealt with in [9] and [10]. Alternative texts on the methodology of finding and using symmetries are [11] and the review article [12].

1.2 The First Jet Bundle

Consider a smooth, real manifold M, of dimension m, and a bundle over M whose fibre is another smooth, real, manifold, U, of dimension n. The resulting construction we call (E, π, M), where E is the total space (locally isomorphic to $U \times M$) and π is a submersion which projects E to the base space M.

In this thesis, the base space will have coordinates x_i (the independent variables) where i runs from one to m, and the fibre will have coordinates u^α or ϕ^α (the dependent variables), with α ranging between one and n. This construction is too general

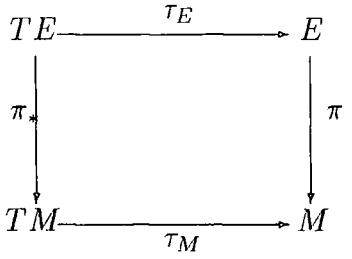

Figure 1.1: Tangent Bundles

for most of our applications, which will often deal with simple product manifolds $U \times M$ where U and M may be isomorphic to $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively.

Associated with this bundle, we can define the bundle of tangent spaces to both the base space and the total space. The tangent bundle to the total space is denoted (TE, τ_E, E), and the corresponding structure on the base space is denoted (TM, τ_M, M). Not surprisingly, the bundle τ_E projects to the bundle τ_M, as illustrated in the commutative diagram in Figure 1.1. The symbol π_* indicates the tangent mapping of π.

The tangent bundles can be endowed with local coordinates induced from those on M and E. On TM these can be written $(x_i, \dot{x}_i)$, and on TE we have $(x_i, \dot{x}_i, u^\alpha, \dot{u}^\alpha)$. Vector fields on M and E are sections of the respective tangent bundles and are related by π_*.

On this skeleton, we can build the first jet manifold. The definitions will all be expressed in terms of a local section $\gamma : W \to E$, where W is some open submanifold of M.

Definition 1.2.1 *Let (E, π, M) be a bundle as outlined above, and let $p \in M$. Two local sections γ_1 and γ_2 of π are said to be* one–equivalent *at p if*

$$\gamma_1(p) = \gamma_2(p),$$

and,

$$\left.\frac{\partial \gamma_1^\alpha}{\partial x_i}\right|_p = \left.\frac{\partial \gamma_2^\alpha}{\partial x_i}\right|_p,$$

where γ_1^α *and* γ_2^α *are the coordinate expressions of the local sections in terms of the coordinate system* (x_i, u^α) *on* E.

The equivalence class of such sections is called the one–jet *of (a generic section)* γ *at* p *and is denoted* $j^1_p\gamma$.

Notice that this is just a generalisation of the definition of a vector field. There, one considers the equivalence class of curves at a point with the same first derivative. The one-jet is an equivalence class of higher–dimensional objects sharing the same first derivatives at a point.

The set of one-jets gives us the first jet manifold.

Definition 1.2.2 *The* first jet manifold *of* π *is the set:*

$$\mathrm{J}^1\pi = \left\{ j^1_p\gamma : p \in M \right\},$$

for all possible local sections γ.

To show that this is a manifold requires a little extra discussion, which can be found in [6], to ensure that definitions and calculations are not dependent on coordinates. One can also demonstrate that $\mathrm{J}^1\pi$ is a bundle over both M and E [6] using two new projections, denoted π_1 and $\pi_{1,0}$ respectively.

$$\begin{aligned} \pi_1 : \mathrm{J}^1\pi &\to M, \\ j^1_p\gamma &\to p; \\ \pi_{1,0} : \mathrm{J}^1\pi &\to E, \\ j^1_p\gamma &\to \gamma(p). \end{aligned}$$

Both π_1 and $\pi_{1,0}$ are smooth, surjective submersions, and hence both $(\mathrm{J}^1\pi, \pi_1, M)$ and $(\mathrm{J}^1\pi, \pi_{1,0}, E)$ are bundles.

Coordinates for the first jet manifold can be induced from coordinates used for E. If there are local coordinates on E of the form $u = (x_i, u^\alpha)$, then there are corresponding local coordinates on $\mathrm{J}^1\pi$ of the form $u^1 = (x_i, u^\alpha, u^\alpha_i)$ where,

$$x_i\left(j^1_p\gamma\right) = x_i(p),$$

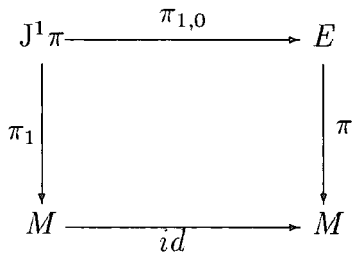

Figure 1.2: The First Jet Bundle

$$u^\alpha\left(j^1_p\gamma\right) = u^\alpha(\gamma(p)),$$
$$u^\alpha_i\left(j^1_p\gamma\right) = \left.\frac{\partial\gamma^\alpha}{\partial x_i}\right|_p.$$

First order differential equations have their natural home in this setting. We can make a rather pedantic definition of a familiar object.

Definition 1.2.3 *A* first order differential equation *on* (E,π,M) *is a closed embedded submanifold* S *of the first jet manifold* $\mathrm{J}^1\pi$. *A* solution *of the differential equation is a local section* s *defined on an open submanifold* $W \subset M$, *which satisfies* $j^1_p s \in S$ *for all* $p \in W$.

It remains to work out how to act on the solutions using Lie group transformations.

1.3 First Prolongations

An important tool for the symmetry theory is the *prolongation* of some geometrical object defined on the bundle π to the corresponding object on π_1. We will usually be prolonging vector fields expressing infinitesimal symmetries of our equations to the relevant jet space. Nevertheless, the basic definition is the following.

Definition 1.3.1 *If* γ *is a local section defined on an open submanifold* $W \subset M$, *its* first prolongation *is the local section of* π_1 *denoted* $\mathrm{pr}^{(1)}\gamma$ *and defined by:*

$$\mathrm{pr}^{(1)}\gamma(p) = j^1_p\gamma,\ p \in W.$$

It is straightforward to work out the coordinate expressions for our first prolongation by using the coordinate functions on $\mathrm{J}^1\pi$, u^α and u^α_i.

$$\begin{aligned} u^\alpha\left(\mathrm{pr}^{(1)}\gamma\right)(p) &= u^\alpha\left(j^1_p\gamma\right) \\ &= u^\alpha(\gamma(p)) \\ &= \gamma^\alpha(p), \end{aligned} \tag{1.3.1}$$

and,

$$\begin{aligned} u^\alpha_i\left(\mathrm{pr}^{(1)}\gamma\right)(p) &= u^\alpha_i\left(j^1_p\gamma\right) \\ &= \left.\frac{\partial\gamma^\alpha}{\partial x_i}\right|_p . \end{aligned} \tag{1.3.2}$$

So the coordinate representation of $\mathrm{pr}^{(1)}\gamma$ looks like $(\gamma^\alpha, \gamma^\alpha_i)$, where the i subscript denotes a partial derivative. The important point to bear in mind is that the prolongation process picks out a restricted class of local sections of π_1 — the most general local section of π_1 will have coordinates looking like $(\psi^\alpha, \psi^\alpha_i)$ but where the ψ^α_i bear no relation to the ψ^α.

The prolongation idea can be generalised from this definition to the notion of the prolongation of a bundle map between two bundles π and ρ, provided the map between the base spaces is a diffeomorphism.

Definition 1.3.2 *Suppose (E, π, M) and (H, ρ, N) are bundles and that $(f, \overline{f})$ is a bundle map between them, where $\overline{f}$ is a diffeomorphism: $\overline{f} : M \to N$. The* first prolongation of $(f, \overline{f})$ *is the map:*

$$\mathrm{pr}^{(1)}(f, \overline{f}) : \mathrm{J}^1\pi \to \mathrm{J}^1\rho,$$

defined by its action on sections $\gamma : W \subset M \to E$ by,

$$\mathrm{pr}^{(1)}(f, \overline{f})\left(j^1_p\gamma\right) = j^1_{\overline{f}(p)}\left(\tilde{f}(\gamma)\right),$$

where

$$\tilde{f} = f \circ \gamma \circ \overline{f}^{-1}\Big|_{\overline{f}^{-1}(W)} .$$

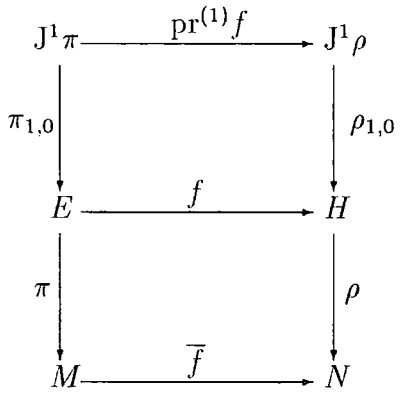

Figure 1.3: The Prolonged Bundle Map

For brevity, we will usually write $\mathrm{pr}^{(1)}f$ in place of the more cumbersome $\mathrm{pr}^{(1)}(f,\overline{f})$.

It should be noted [6] that the prolongation of $(f,\overline{f})$ induces new bundle maps $(\mathrm{pr}^{(1)}f, f) : \pi_{1,0} \to \rho_{1,0}$ and $(\mathrm{pr}^{(1)}f, \overline{f}) : \pi_1 \to \rho_1$, a fact summarised in the commutative diagram, Figure 1.3.

A straightforward calculation tells us the coordinate representation of the prolonged bundle map. Suppose that $\mathrm{J}^1\rho$ has local coordinates $(y_j, v^\beta, v^\beta_j)$, where the ranges of the indices are constrained in the obvious way by the dimensions of N and of the fibre of $\mathrm{J}^1\rho$. The bundle map f will have a coordinate expression in terms of these variables, such that,

$$f(x_i, u^\alpha) = (f_j, f^\beta). \tag{1.3.3}$$

Then it is found that:

$$\begin{aligned} y_j \circ \mathrm{pr}^{(1)}f &= f^j \\ v^\beta \circ \mathrm{pr}^{(1)}f &= f^\beta \\ v^\beta_j \circ \mathrm{pr}^{(1)}f &= D_{x_i}(f^\beta)\left(\frac{\partial\left(\overline{f}^{-1}\right)^i}{\partial y_j} \circ \overline{f}\right), \end{aligned} \tag{1.3.4}$$

where D_{x_i} is the total derivative with respect to x_i.

By defining the the prolonged bundle map, we have moved somewhat closer to finding a tool that can be used for studying symmetries, since one example of an admissible bundle map is a Lie group transformation of a differential equation S. The aim of the symmetry theory is to find particular group actions f that map

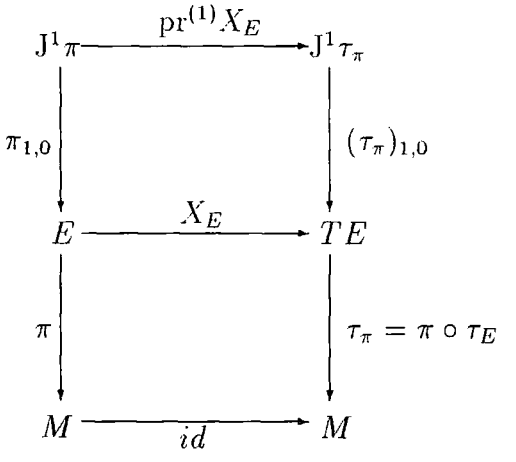

Figure 1.4: Prolonging a vector field

solutions s of S to other solutions $\tilde{f}(s)$. We will eventually define a symmetry of S to be any f such that

$$\mathrm{pr}^{(1)}f(S) = S.$$

However, when we come to actually calculate symmetries, it will be much more convenient to think in terms of infinitesimal transformations, and so we need an infinitesimal version of the preceding construction. In other words, we need to be able to prolong vector fields.

We are interested, therefore, in sections of $TJ^1\pi$. Geometrically, it turns out to be quite tricky to adequately define such objects. Motivated by our results on prolonging bundle maps, we might try the following strategy: remembering the definition of a section, we could view a vector field on E (denoted X_E) as defining a bundle map (X_E, id) from π to the composite bundle $(TE, \pi \circ \tau_E, M)$, denoted by τ_π. Then we can apply our results on bundle maps to this construction, and we end up with the situation summarised in Figure 1.4.

Notice, however, that the total space of the resulting jet manifold over TE is $J^1\tau_\pi$, not $TJ^1\pi$. Thinking in terms of local coordinates, we may readily see that these spaces are not identical. They do not even have the same dimension. If we give TE the usual coordinates $(x_i, \dot{x}_i, u^\alpha, \dot{u}^\alpha)$, then $J^1\tau_\pi$ has coordinates $(x_i, \dot{x}_i, u^\alpha, \dot{u}^\alpha, \dot{x}_{i,j}, u^\alpha_j, \dot{u}^\alpha_j)$ and $TJ^1\pi$ has coordinates $(x_i, \dot{x}_i, u^\alpha, \dot{u}^\alpha, u^\alpha_j, \dot{u}^\alpha_j)$. Counting dimensions, we find that

$J^1\tau_\pi$ has dimension m^2 more than $TJ^1\pi$ due to the presence of the $\dot{x}_{i,j}$ term.

The resolution of this problem can be found in [6]. Essentially, the idea is to find a surjective bundle map which takes elements of $J^1\tau_\pi$ to elements of $TJ^1\pi$.

We will employ a more utilitarian argument [7], suitable for future applications. It has already been mentioned that the results on prolongation of bundle maps is applicable to Lie group transformations. Since it is the infinitesimal generators of these transformations that are of interest, we can propose a working definition for their first prolongations.

Working Definition. Let G be a one-parameter Lie group whose elements g are written in terms of the infinitesimal generators $\mathbf{v}$ as $g = \exp(\varepsilon\mathbf{v})$. Then the first prolongation of $\mathbf{v}$ is determined by:

$$\mathrm{pr}^{(1)}\mathbf{v} = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathrm{pr}^{(1)}(\exp(\varepsilon\mathbf{v})). \tag{1.3.5}$$

Consider a generator $\mathbf{v}$ of the form,

$$\mathbf{v} = X^i(x,u)\frac{\partial}{\partial x_i} + U_\alpha(x,u)\frac{\partial}{\partial u^\alpha}, \tag{1.3.6}$$

and the equivalent vector field on M,

$$\mathbf{v}_M = X^i\frac{\partial}{\partial x_i}. \tag{1.3.7}$$

Then the prolonged vector field should look like,

$$X^i\frac{\partial}{\partial x_i} + U_\alpha\frac{\partial}{\partial u^\alpha} + U^{(i)}_\alpha\frac{\partial}{\partial u^\alpha_i},$$

since the zeroth prolongation must be the vector itself. The coefficients $U^{(i)}_\alpha$ must be calculated from our working definition.

To do this, we start off with a generator of the form (1.3.6), and a corresponding group transformation $g_\varepsilon = \exp(\varepsilon\mathbf{v}) : E \to \tilde{E}$. (The subscript is to remind us where the ε dependence lies.) In coordinates:

$$(\tilde{x},\tilde{u}) = g_\varepsilon \cdot (x,u). \tag{1.3.8}$$

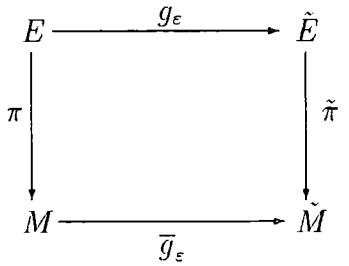

Figure 1.5: Group transformation

The projected group transformation acting on M is denoted $\overline{g}_\varepsilon$. The relevant sets and maps are sketched in Figure 1.5.

Now, consider a section $\gamma : M \to E$. This is transformed by the action of G to a section $\tilde{\gamma} : \tilde{M} \to \tilde{E}$ given by:

$$\tilde{\gamma} = \left[g_\varepsilon \circ \gamma \circ \overline{g}_\varepsilon^{-1} \right]. \tag{1.3.9}$$

To compute the coefficients $U_\alpha^{(i)}$ we need to work out the matrix of derivatives,

$$\left[\mathsf{J}\left[\tilde{\gamma}\right](\tilde{x}) \right]_k^\alpha = \frac{\partial \tilde{\gamma}^\alpha}{\partial \tilde{x}_k}, \tag{1.3.10}$$

where the J stands for "Jacobian". Acting with J on (1.3.9) tells us, via the chain rule, that,

$$\mathsf{J}\left[\tilde{\gamma}\right](\tilde{x}) = \mathsf{J}\left[g_\varepsilon \circ \gamma\right](x) \left\{ \mathsf{J}\left[\overline{g}_\varepsilon\right](x) \right\}^{-1}, \tag{1.3.11}$$

provided the inverse exists.

Now we differentiate (1.3.11) with respect to ε and set ε to zero to get expressions for U_i^α. It is useful to bear in mind that:

$$\begin{aligned} g_0 &= \mathrm{id}; \\ \overline{g}_0 &= \mathrm{id}; \\ \left.\frac{dg_\varepsilon}{d\varepsilon}\right|_{\varepsilon=0} &= \mathbf{v}; \\ \left.\frac{d\overline{g}_\varepsilon}{d\varepsilon}\right|_{\varepsilon=0} &= \mathbf{v}_M. \end{aligned}$$

Armed with these identities, we differentiate (1.3.11).

$$\begin{aligned} \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathsf{J}\left[\tilde{\gamma}(\tilde{x})\right] &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathsf{J}\left[g_\varepsilon \circ \gamma\right](x) - \mathsf{J}\left[\gamma\right](x) \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \mathsf{J}\left[\overline{g}_\varepsilon\right](x) \\ &= \mathsf{J}\left[\mathbf{v} \circ \gamma\right](x) - \mathsf{J}\left[\gamma\right](x) \mathsf{J}[\mathbf{v}_M](x). \end{aligned} \tag{1.3.12}$$

Reading off the (α, k) entry of this matrix tells us that:

$$U_\alpha^{(k)} = \frac{\partial}{\partial x_k}\left(U^\alpha(x, \gamma(x))\right) - \frac{\partial \gamma^\alpha}{\partial x_i}(x)\frac{\partial}{\partial x_k}\left(X^i(x, \gamma(x))\right). \tag{1.3.13}$$

We are now in a position to make a proper definition of the prolongation of a vector field.

Definition 1.3.3 *Let* $\mathbf{v} : E \to TE$ *be a vector field on* E *of the form (1.3.6). Then the* first prolongation of $\mathbf{v}$ *is defined by:*

$$\mathrm{pr}^{(1)}\mathbf{v} = X^i\frac{\partial}{\partial x_i} + U_\alpha\frac{\partial}{\partial u^\alpha} + \left(D_i(U_\alpha) - u_j^\alpha D_i(X^j)\right)\frac{\partial}{\partial u_i^\alpha}. \tag{1.3.14}$$

This definition works for all vector fields of the form (1.3.6), and does not rely on the group properties exploited here.

1.4 Higher Order Jets and Prolongations

As promised, the definitions presented in the preceding sections will be repeated here for higher order jet spaces without much discussion. Hopefully, the reader will have an intuitive understanding of these ideas from the discussion of the first jet manifold. Details may be found in [6].

One aspect of notation must be made clear, however. To discuss the higher jets, we must be able to write down derivatives of arbitrary order in a way that respects the equivalences of mixed partial differentiation. The standard language uses *multi-index notation.* A multi-index I of order k is an unordered k-tuple of numbers corresponding to k partial differentiations with respect to a given variable. The order of I is sometimes written $|I| = k$.

As an example, consider a real-valued function $f(x, y)$. The symbol

$$\frac{\partial^{|I|} f}{\partial x_I}(x, y)$$

can mean any of the following:

$$f(x,y),\ f_x(x,y),\ f_y(x,y),\ f_{xx}(x,y),\ f_{xy}(x,y),\ f_{yy}(x,y), \ldots,$$

and so on to arbitrarily high derivatives. By convention, capital letters will be used for multi-indices. Any refinements will be explained when they arise.

Equipped with this new notation, we proceed to the generalisation of the notion of one-equivalence. Once again, the discussion will be based on the bundle (E, π, M) and definitions will refer to a section $\gamma : M \to E$.

Definition 1.4.1 *Two local sections γ_1 and γ_2 of π are said to be* k-equivalent *at $p \in M$ if,*

$$\gamma_1(p) = \gamma_2(p),$$

and,

$$\left.\frac{\partial^{|I|}\gamma_1^\alpha}{\partial x_I}\right|_p = \left.\frac{\partial^{|I|}\gamma_2^\alpha}{\partial x_I}\right|_p,$$

for $1 \leq |I| \leq k$.

The equivalence class of sections k-equivalent to γ at p is called the k-jet *of γ at p, and is denoted $j_p^k\gamma$.*

As we expect, the set of such objects has the structure of a manifold.

Definition 1.4.2 *The* kth jet manifold *of π is the set,*

$$\mathrm{J}^k\pi = \left\{ j_p^k\gamma : p \in M \right\},$$

for all possible local sections γ.

We end up with an arrangement as shown in Figure 1.6, similar to that in Figure 1.2. The projections π_k and $\pi_{k,0}$ are the counterparts of the objects π_1 and $\pi_{1,0}$ on $\mathrm{J}^1\pi$.

Our notion of a first order differential equation can be generalised.

Definition 1.4.3 *A* kth order differential equation *is a closed, embedded submanifold S of $\mathrm{J}^k\pi$. A* solution *of S is a local section s defined on an open submanifold $W \subset M$, which satisfies $j_p^1(s) \in S$ for all $p \in W$.*

Prolongations of sections, bundle maps and vector fields all proceed without serious incident.

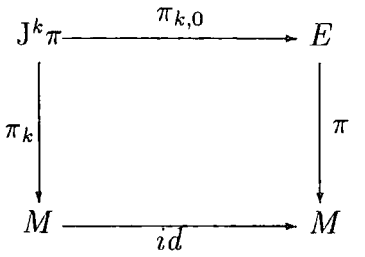

Figure 1.6: The k-Jet Bundle

Definition 1.4.4 *If γ is a local section of π, its* kth prolongation *is the local section of π_k denoted* $\mathrm{pr}^k\gamma$ *and defined by,*

$$\mathrm{pr}^k\gamma(p) = j^k_p\gamma,$$

for all $p \in W$.

The coordinate representation of $\mathrm{pr}^{(k)})\gamma$ is a straightforward generalisation of the first order case. It looks like:

$$\left(\gamma^\alpha, \frac{\partial\gamma^\alpha}{\partial x_i}, \ldots, \frac{\partial^{|I|}\gamma^\alpha}{\partial x_I}, \ldots\right),$$

for$1 \leq |I| \leq k$.

The definition of the prolonged bundle map follows from the definition of the prolongation of a section.

Definition 1.4.5 *Suppose (E, π, M) and (H, ρ, N) are bundles and that $(f, \overline{f})$ is a bundle map between them, where $\overline{f}$ is a diffeomorphism: $\overline{f} : M \to N$. The* kth prolongation of $(f, \overline{f})$ *is the map:*

$$\mathrm{pr}^{(k)}(f, \overline{f}) : J^k\pi \to J^k\rho,$$

defined by,

$$\mathrm{pr}^{(k)}(f, \overline{f})\left(j^k_p\gamma\right) = j^k_{\overline{f}(p)}\left(\tilde{f}(\gamma)\right),$$

where

$$\tilde{f} = f \circ \gamma \circ \overline{f}^{-1}\Big|_{\overline{f}^{-1}(W)}.$$

Again, the notation $\mathrm{pr}^{(k)}f$ will often be used as a shorthand for $\mathrm{pr}^{(k)}(f,\overline{f})$.

Finally, we prolong vector fields on E to the k-jet space. All the difficulties mentioned before remain in the higher order cases, but this time we will bypass all discussion of these problems and simply state the relevant definition.

Definition 1.4.6 *Let* $\mathbf{v} : E \to TE$ *be a vector field on* E *of the form (1.3.6). Then the* kth prolongation of $\mathbf{v}$ *is defined by,*

$$\mathrm{pr}^{(k)}\mathbf{v} = X^i\frac{\partial}{\partial x_i} + U_\alpha\frac{\partial}{\partial u^\alpha} + \sum_{|I|=1}^{k} U^I_\alpha\frac{\partial}{\partial u^\alpha_I}, \tag{1.4.1}$$

where,

$$U^I_\alpha = D_I\left(U_\alpha - u^\alpha_j X^j\right) + X^j u^\alpha_{I,j}.$$

This definition can be argued from induction on the discussion presented in the preceding section. See Olver [7].

It should be mentioned that in all of these definitions k can be taken to infinity. The resulting space, the infinite jet bundle $\mathrm{J}^\infty\pi$, is a properly defined, infinite–dimensional manifold [6], but we will not really need to know this. For our purposes, $\mathrm{J}^\infty\pi$ is shorthand for a k-jet bundle in which k is large enough to provide any order of derivatives we may require for a given application. By convention, the infinite prolongation of a section, for example, will be written $\mathrm{pr}\,(\gamma)$, rather than $\mathrm{pr}^\infty(\gamma)$. We will meet it in our discussions of generalised symmetries and of variational calculus.

1.5 Point Symmetries

Now that we are clear on how to prolong the relevant tools, we can proceed to the techniques used in calculations. By far the best way to understand these methods is through worked examples. We will look at two simple calculations that will be of use later. The reader will find plenty of other examples in the books by Olver [7] and Stephani [11]. Consequently, this we will leave the cited texts to supply further practical details. In particular, the review article [12] is recommended as a very applications–oriented "users' guide" to symmetry methods.

We begin by formally reiterating the definition of a symmetry mentioned in Section 1.3.

Definition 1.5.1 *A* point symmetry *of a differential equation $S \subset \mathrm{J}^k\pi$ is a bundle map which takes π to itself, such that a solution s is mapped to a new solution $\tilde{f}(s)$. In other words,*

$$\mathrm{pr}^{(k)} f(S) = S.$$

Recall that in order for the prolongation of the map to exist, the bundle map must project to a diffeomorphism.

More prosaically, we are looking for transformations,

$$x \to \tilde{x}(x,u),$$
$$u \to \tilde{u}(x,u),$$

mapping solutions of S to other solutions of S. We will be usually be discussing the case when these mappings are induced by the action of a Lie group. So we define a *symmetry group* of S to be a local group of transformations which are point symmetries of S. Our criterion of invariance is that the prolonged group action leaves the submanifold defined by S invariant. There are coordinate–free formulations of the method, but even for non-trivial bundles, the local coordinate method presented here is the least troublesome. See [11] for the interesting example of the symmetries of the covariant Hamilton–Jacobi equation.

The crucial step in being able to calculate symmetry groups is that in many cases we can equivalently discuss the infinitesimal transformations due to the generators of the Lie groups. This has the computational advantage of replacing a possibly nonlinear operation by a linear one. Nevertheless, to be confident of this result, we must quantify what constitutes "most cases". To do this, we write S as a system of equations $\Delta = 0$ and define its *rank* to be the rank of the Jacobian matrix:

$$[\mathrm{J}] = \left[\frac{\partial \Delta_\nu}{\partial x_i}, \frac{\partial \Delta_\nu}{\partial u^\alpha_J}\right]. \tag{1.5.1}$$

Then the condition for the equivalence of the finite and infinitesimal approaches can be expressed in a simple theorem.

Proposition 1.5.1 *Let G be a local Lie group of transformations acting on some open subset of E. If the kth order system of differential equations S has maximal rank on that subset, and*

$$\mathrm{pr}^{(k)}\mathbf{v}\,[\Delta] = 0 \tag{1.5.2}$$

on solutions of S for all generators $\mathbf{v}$ of G, then G is a point symmetry group of S.

The proof is not difficult and may be found in [7]. Essentially, the maximum rank condition ensures that there are no vector fields that vanish trivially on solutions of S.

Notice that this only proves the sufficiency of discussing the infinitesimal generator. To prove that the equivalence of the two requires the additional assumption of *local solvability*. A full discussion may be found in [7]. We only need to assume that that the maximal rank condition holds, in which case the set of vector fields obeying (1.5.2) forms a Lie algebra from which we can reconstruct the corresponding group.

We have already worked out how to prolong a vector field, and so we are in a position to work out the algorithm for the calculation of a symmetry group. Proceed as follows:

- The generators of the infinitesimal symmetries are postulated to be of the form (1.3.6).
- Using the formula (1.4.1), the prolonged vector field is applied to S.
- After taking account of the dependencies due to the equation of motion, we impose (1.5.2).
- Requiring the coefficients of all the derivatives of u^α to vanish generates a large number of (hopefully easy) linear p.d.e.s in the variables X_i and U_α.

The set of solutions of these equations defines the possible infinitesimal symmetry generators.

One advantage of the algorithmic nature of this process is that it can be largely automated. In this work, point symmetries are calculated with the aid of the SPDE Lie-symmetry package which is part of the REDUCE symbolic manipulation program, although the reliability of this package is questionable. The software seems to work best for finding the symmetries of evolution equations. The author has noted some possible sign errors in the calculation of symmetry generators for non-evolution equations. Further, the program's capabilities for solving the determining equations are limited, and may leave the researcher with some paperwork left to do. Finally, as a minor point, the user interface is not particularly flexible. The author has not had sufficient exposure to other software to recommend anything else, but the reader should be aware that there are a number of packages available, including a MAPLE library routine which uses an alternative approach involving differential forms. The reader will find[1] a comprehensive survey of the available software in [13].

If the maximal rank condition holds, these infinitesimal symmetry generators can be exponentiated to finite transformations which form a group. This amounts to solving a set of differential equations,

$$\begin{aligned} \frac{d\tilde{x}^i}{d\varepsilon} &= X^i(\tilde{x}(\varepsilon), \tilde{u}(\varepsilon)), \text{ with } \tilde{x}|_{\varepsilon=0} = x \\ \frac{d\tilde{u}^\alpha}{d\varepsilon} &= U^\alpha(\tilde{x}(\varepsilon), \tilde{u}(\varepsilon)), \text{ with } \tilde{u}|_{\varepsilon=0} = u, \end{aligned} \tag{1.5.3}$$

sufficiently close to the starting point $\varepsilon = 0$.

At this point, we will look at a pair of practical examples which will be of use later.

[1]The author is grateful to Professor P. Clarkson for pointing this out

1.5.1 Example: the KdV equation

The Korteweg–de Vries (KdV) equation will be discussed in Chapter 3. It has the form:

$$\Delta_1 = u_t - 6uu_x + u_{xxx} = 0. \tag{1.5.4}$$

Following the prescription above, we postulate that the symmetry generator looks like:

$$\mathbf{v} = X(x,t,u)\frac{\partial}{\partial x} + T(x,t,u)\frac{\partial}{\partial t} + U(x,t,u)\frac{\partial}{\partial u}. \tag{1.5.5}$$

Since the equation is third order we need to work out the coefficients of the third prolongation. If we apply a prolonged vector field of the form (1.4.1) to (1.5.4), the resultant expression is,

$$\mathrm{pr}^{(3)}\mathbf{v}(\Delta_1) = U^t - 6u_xU - 6uU^x + U^{xxx}, \tag{1.5.6}$$

where,

$$\begin{aligned}
U^x &= D_x(U) - u_xD_x(X) - u_tD_x(T),\\
U^t &= D_t(U) - u_xD_t(X) - u_tD_t(T),\\
U^{xxx} &= D_x^3(U) - u_xD_x^3(X) - u_tD_x^3(T) - 3u_{xx}D_x^2(X) - 3u_{xt}D_x^2(T)\\
&\quad -3u_{xxx}D_x(X) - 3u_{xxt}D_x(T),
\end{aligned} \tag{1.5.7}$$

as defined in the second part of (1.4.1).

To find the symmetry condition, we have to expand (1.5.6), make the substitution,

$$u_t = 6uu_x - u_{xxx}, \tag{1.5.8}$$

and, in accordance with (1.5.2), set the resulting expression equal to zero.

To derive the associated conditions on X, T and U, we successively equate coefficients of derivatives of u to zero. The first one to look at is that of u_{xxt} which happens to be equal to $D_x(T)$. From it we deduce that T is a function of t alone. Likewise, since the coefficient of u_{xx}^2 is X_u (subscript indicates partial derivative), we find that X is independent of u. The coefficient of u_{xxx} tells us that,

$$\frac{dT}{dt} = 3\frac{\partial X}{\partial x}, \tag{1.5.9}$$

and hence that,

$$X = \frac{x}{3}\frac{dT}{dt} + f_1(t), \tag{1.5.10}$$

where f_1 is an arbitrary function. Next, we find from the coefficient of u_{xx} that,

$$\frac{\partial^2 U}{\partial x \partial u} = \frac{\partial^2 U}{\partial u^2} = 0, \tag{1.5.11}$$

so we can write,

$$U = f_2(t)u + f_3(x,t), \tag{1.5.12}$$

where f_2 and f_3 are again arbitrary. The remaining terms yield:

$$6U - \frac{\partial X}{\partial t} - 6u\frac{\partial T}{\partial t} + 6u\frac{\partial X}{\partial x} = 0, \tag{1.5.13}$$

and,

$$\frac{\partial^3 U}{\partial x^3} - 6u\frac{\partial U}{\partial x} + \frac{\partial U}{\partial t} = 0. \tag{1.5.14}$$

These conditions are all solved by,

$$\begin{aligned} X &= c_1 + 6c_2 t + c_3 x, \\ T &= c_4 + 3c_3 t, \\ U &= -(c_2 + 2c_3 u). \end{aligned} \tag{1.5.15}$$

The c's are arbitrary constants. Hence the infinitesimal symmetry algebra is spanned by:

$$\begin{aligned} \mathbf{v}_1 &= \partial_x, \\ \mathbf{v}_2 &= \partial_t, \\ \mathbf{v}_3 &= 6t\partial_x - \partial_u, \\ \mathbf{v}_4 &= x\partial_x + 3t\partial_t - 2u\partial_u. \end{aligned} \tag{1.5.16}$$

These are respectively: space and time translation, a Galilean boost, and a scaling symmetry.

1.5.2 Example: the Bateman equation

The nonlinear p.d.e.,

$$\Delta_2 = u_x^2 u_{tt} - 2u_x u_t u_{xt} + u_t^2 u_{xx} = 0, \tag{1.5.17}$$

will be called the Bateman equation in this thesis. It will occupy much of our attention in Chapter 4. For the moment, we will confine ourselves to study of its point symmetries.

The first stage of our algorithm is straightforward. We consider an ansatz like (1.5.5) for the generator. Then, the prolongation formula (1.4.1) is used with $k = 2$ to calculate the characteristic of the prolonged vector, which will look like,

$$\mathrm{pr}^{(2)}\mathbf{v} = X\frac{\partial}{\partial x}+T\frac{\partial}{\partial t}+U\frac{\partial}{\partial u}+U^x\frac{\partial}{\partial u_x}+U^t\frac{\partial}{\partial u_t}+U^{xx}\frac{\partial}{\partial u_{xx}}+U^{xt}\frac{\partial}{\partial u_{xt}}+U^{tt}\frac{\partial}{\partial u_{tt}}, \tag{1.5.18}$$

with the functions $U^x, U^t, \ldots$ given by the second part of (1.4.1). For example,

$$U^x = \frac{\partial U}{\partial x} + \frac{\partial U}{\partial u}u_x - u_x\frac{\partial X}{\partial x} - u_x^2\frac{\partial X}{\partial u} - u_t\frac{\partial T}{\partial x} - u_t u_x\frac{\partial T}{\partial u}. \tag{1.5.19}$$

The second order coefficients are more lengthy, but are not particularly difficult to calculate.

Applied to Δ_2, we find that the prolonged transformation is given by:

$$\mathrm{pr}^{(2)}\mathbf{v}(\Delta_2) = 2U^x(u_x u_{tt} - u_t u_{xt}) + 2U^t(u_t u_{xx} - u_x u_{xt}) + U^{xx}u_t^2 - 2u_x u_t U^{xt} + u_x^2 U^{tt}. \tag{1.5.20}$$

Before imposing the symmetry condition, we take account of the equation of motion itself by making the substitution,

$$u_{tt} = \frac{2u_x u_t u_{xt} - u_t^2 u_{xx}}{u_x^2}, \tag{1.5.21}$$

into Equation (1.5.20). Then, expanding everything out and doing large amounts of tedious algebra, we find the following symmetry condition:

$$\begin{aligned} &2u_t u_x^4\frac{\partial^2 X}{\partial x\partial t} - u_x^5\frac{\partial^2 X}{\partial t^2} - u_t^2 u_x^3\frac{\partial^2 X}{\partial x^2} + 2u_t^2 u_x^3\frac{\partial^2 T}{\partial x\partial t} - u_t u_x^4\frac{\partial^2 T}{\partial t^2} - u_t^3 u_x^2\frac{\partial^2 T}{\partial x^2} \\ &+ 2u_t u_x^2 u_{xt}\frac{\partial U}{\partial x} + 2u_t u_x^2 u_{xx}\frac{\partial U}{\partial t} - 2u_x u_t^2 u_{xx}\frac{\partial U}{\partial x} - 2u_x^3 u_{xt}\frac{\partial U}{\partial t} + u_t^2 u_x^2\frac{\partial^2 U}{\partial x^2} \\ &+ u_x^4\frac{\partial^2 U}{\partial t^2} - 2u_t u_x^3\frac{\partial^2 U}{\partial x\partial t} = 0. \end{aligned} \tag{1.5.22}$$

It remains to set each independent term to zero to derive the conditions on X, T and U.

To begin with, the two terms involving u_{xx} tell us that U is independent of x and t. So we write $U = \eta(u)$. Then, we see that the term in u_x^5 is the second t-derivative of X, meaning that,

$$X = f_1(x,u)t + f_2(x,u), \tag{1.5.23}$$

where f_1 and f_2 are arbitrary functions. Likewise, from the $u_t^3 u_x^2$ term we deduce that,

$$T = f_3(t,u) + f_4(t,u). \tag{1.5.24}$$

Finally, the last two monomials with nonzero coefficients are $u_t^2 u_x^3$ and $u_t u_x^4$. They yield:

$$2\frac{\partial^2 T}{\partial x \partial t} - \frac{\partial^2 X}{\partial x^2} = 0; \tag{1.5.25}$$

$$2\frac{\partial^2 X}{\partial x \partial t} - \frac{\partial^2 T}{\partial t^2} = 0. \tag{1.5.26}$$

By differentiating (1.5.25) with respect to x and (1.5.26) with respect to t, and using the conditions we have already derived, these conditions are seen to imply,

$$\begin{aligned} \frac{\partial^3 X}{\partial x^3} &= 0, \\ \frac{\partial^3 T}{\partial t^3} &= 0. \end{aligned} \tag{1.5.27}$$

These are easily turned into a set of conditions on the functions $f_1, \ldots, f_4$. We find that they must be of the form:

$$\begin{aligned} f_1(x,u) &= \frac{1}{2}\xi_1 x^2 + \xi_2 x + \xi_3 \\ f_2(x,u) &= \frac{1}{2}\xi_4 x^2 + \xi_5 x + \xi_6 \\ f_3(t,u) &= \frac{1}{2}\xi_7 t^2 + \xi_8 t + \xi_9 \\ f_4(t,u) &= \frac{1}{2}\xi_{10} t^2 + \xi_{11} t + \xi_{12} \end{aligned} \tag{1.5.28}$$

$\mathbf{v}_1$	=	$\eta\ \partial_u$	(Scalings, diffeomorphisms)
$\mathbf{v}_2$	=	$\xi_6\ \partial_x$	(Translations)
$\mathbf{v}_3$	=	$\xi_{12}\ \partial_t$	(Translations)
$\mathbf{v}_4$	=	$\xi_5\ x\ \partial_x$	(Scalings)
$\mathbf{v}_5$	=	$\xi_{11}\ t\ \partial_t$	(Scalings)
$\mathbf{v}_6$	=	$\xi_9\ x\ \partial_t$	(Rotations/boosts)
$\mathbf{v}_7$	=	$\xi_3\ t\ \partial_x$	(Rotations/boosts)
$\mathbf{v}_8$	=	$\xi_8\ x\ (x\ \partial_x + t\ \partial_t)$	(Conformal transformations)
$\mathbf{v}_9$	=	$\xi_2\ t\,(x\partial_x + t\partial_t)$	(Conformal transformations)

Table 1.1: Point Symmetries of the Bateman Equation

The ξ's are all functions of u alone. If we substitute the resulting solutions for X and T back into (1.5.25) and (1.5.26), we find some constraints on these, namely that,

$$\begin{aligned} \xi_1 &= \xi_7 = 0, \\ \xi_4 &= 2\xi_8, \\ \xi_{10} &= 2\xi_2. \end{aligned} \tag{1.5.29}$$

With this information, we can write down a set of vectors that span the infinitesimal symmetries. They are listed in Table 1.1.

In the table, η and the ξ's are all arbitrary functions of u alone. This means that the number of independent basis vectors is effectively infinite. The vectors clearly include all possible $GL(2,\mathbb{R})$ invariances and the diffeomorphism symmetry of the equation.

1.6 Deducing Solutions from Group Properties

Knowledge of the point symmetries of an arbitrary S is usually not sufficient to determine a general solution. However, large classes of solutions which are invariant under restricted subgroups of transformations can be deduced from the symmetry structure. This technique will not be widely used in this thesis, so the treatment will be restricted to a brief description, illustrated with some examples.

The aim is to study particular solutions given knowledge of the symmetries. One possible method is clear from the definition of a point symmetry. If we know a par-

ticular solution of the equation, we can apply group transformations to it to produce other solutions, hopefully independent of the original solution. Examples of this approach can be found in [11]. The idea is the same as that underlying Bäcklund transformations, but these are more usually associated with generalised symmetries.

The second idea is to use the symmetries of the equations to find variables in which to express group-invariant solutions, and thereby reduce the number of variables, and the order of the system. A common example is the use of polar variables to find spherically symmetric solutions of the Laplace equation. This amounts to replacing the partial differential equation with an ordinary differential equation in the radial variable.

The difficulty in the method is mostly confined to finding the correct "similarity" variables. For a start, there is no unique choice of these group-specific coordinates, which we shall denote $\Omega = (\xi, \eta)$. They can sometimes be found directly from the requirement that,

$$\mathrm{pr}^{(k)}\mathbf{v}(\Omega) = 0. \tag{1.6.1}$$

The method of characteristics usually allows solutions of this equation to be found, otherwise some more *ad hoc* method might be required.

A more systematic approach is to use the finite group transformations calculated using (1.5.3). Usually, one of the equations (1.5.3) can be inverted to provide an expression for ε in terms of $x, u, \tilde{x}$,and $\tilde{u}$. This is used to define the orbits of the group action nonparametrically. This representation of the orbits should provide expressions for the new variables, ξ and $\eta(\xi)$. Substitution into S will provide a reduced set of equations for η. Once again, the reader will find many examples in [11] and also in [7].

1.6.1 Example: reductions of the KdV equation

As a simple, accessible case, consider the problem of finding the invariant solutions of the KdV equation, which as we saw has four separate infinitesimal symmetries: $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, $\mathbf{v}_4$ in (1.5.16).

We note that the KdV has the translation vector fields $\mathbf{v}_1 = \partial_x$ and $\mathbf{v}_2 = \partial_t$ as symmetry generators. These can be combined to provide the travelling wave solutions. What are the correct variables to use here? Consider a small transformation,

$$(x, t, u) \to (x + c\varepsilon, t + \epsilon, u),$$

generated by the linear combination $c\partial_x + t\partial_t$. We can work out by the method of characteristics that any function f depending only on the characteristic variable $y = x - ct$ is invariant under this transformation. Substituting this into (1.5.4) we find,

$$-cf_y = 6ff_y - f_{yyy}, \tag{1.6.2}$$

which integrates to:

$$-cf = 3f^2 - f_{yy} + k, \tag{1.6.3}$$

where k is an arbitrary constant. After multiplying across by f_y, we can integrate again:

$$-\frac{1}{2}cf^2 = f^3 - \frac{1}{2}f_y^2 + kf + l, \tag{1.6.4}$$

with l another constant. This equation has a general solution in terms of elliptic functions, but an interesting class of solutions is obtained when the wave is constrained to go to zero at positive and negative infinity. Then k and l are zero, and there are solutions,

$$u = -c\,\mathrm{sech}^2\left(\frac{\sqrt{c}}{2}(x - ct) + \delta\right), \tag{1.6.5}$$

where δ is some phase change. These are the one–soliton solutions.

Proceeding to the Galilean symmetry, it is clear that $y = t$ is an invariant of this transformation. Also, using the method of characteristics on (1.6.1) yields,

$$\frac{dx}{6t} = -du, \tag{1.6.6}$$

implying that $v = x + 6tu$ is also invariant. Rearranging, and working out the expressions for the derivatives of u, we find that (1.5.4) becomes simply,

$$\frac{dv}{dy} = 0. \tag{1.6.7}$$

Solving this, and using the expression for v, gives the class of solutions,

$$u = \frac{k - x}{6t}, \tag{1.6.8}$$

where k is an arbitrary constant.

An interesting further class of solutions can be found if this Galilean symmetry is combined with a time translation. This yields a so-called *Painlevé transcendent.* Details can be found in Olver [7].

Finally, we look at the scale invariant solutions generated by $\mathbf{v}_4$. Expanding out (1.6.1) and using the method of characteristics gives two invariants:

$$y = \frac{x}{t^{1/3}} \text{ and } w = ut^{2/3}. \tag{1.6.9}$$

After finding the expressions for the derivatives of $u(x,t)$ in terms of those of $w(y)$, the reduced equation becomes:

$$18ww_y + 2w + yw_y - w_{yyy} = 0. \tag{1.6.10}$$

This can be solved using a Miura transformation to eventually yield another Painlevé equation. Again, the details are presented in Olver.

1.7 Generalised Symmetries

The final topic in our review of symmetry techniques is the concept cf generalised symmetries and how to find them. Briefly, the idea is to extend the nction of what constitutes an admissible generator for a symmetry by allowing vector fields on E to depend on derivatives of the dependent variables. The resulting symmetries will contain information about the integrability of the equations (possibly by a linearisation mechanism), and their Hamiltonian or Lagrangian properties, if there are any.

We begin by giving a practical definition of a generalised vector field in local coordinates.

Definition 1.7.1 *A* generalised vector field *is a vector field over E with the coordinate expression,*

$$\mathbf{v} = X^i[u]\frac{\partial}{\partial x_i} + U_\alpha[u]\frac{\partial}{\partial u^\alpha}, \tag{1.7.1}$$

where the coefficients X^i and U_α can depend on derivatives of u up to some specified order.

A more rigorous definition (to be found in Saunders [6]) would characterise a generalised vector field as a vector field along the bundle $\pi_{k,0}$.

The prolongation formula (1.4.1) holds unchanged for these new generators, with the same formula for the coefficients of the higher order terms. Very often the most appropriate prolongation to use is the infinite one, denoted simply pr $\mathbf{v}$. Often, we will not know *a priori* which order of derivatives is important in a generalised symmetry or conservation law, and the infinite prolongation covers all eventualities.

With this in mind, we define a generalised symmetry.

Definition 1.7.2 *A generalised vector field of the form (1.7.1) is said to be an* infinitesimal generalised symmetry *of a system of differential equations S if*

$$\text{pr } \mathbf{v}\,[\Delta] = 0, \tag{1.7.2}$$

on solutions of S.

We can impose the same maximal rank conditions as before to enable construction of the finite group transformations.

One pleasing feature of generalised vector fields is that simplified forms may be admissible for symmetry applications. In particular, given $\mathbf{v}$ as in (1.7.1), its *evolutionary* or *vertical* representative is

$$\mathbf{v}_{U'} = U'_\alpha\frac{\partial}{\partial u^\alpha}, \tag{1.7.3}$$

where,

$$U' = U_\alpha - u^\alpha_j X^j.$$

(U' is called the *characteristic* of the evolutionary representative.) It is not too difficult to show that the corresponding prolongation of S,

$$\text{pr } \mathbf{v}_{U'}[\Delta] = D_I U'_\alpha \frac{\partial \Delta}{\partial u^\alpha_I}, \tag{1.7.4}$$

is equivalent to pr $\mathbf{v}[\Delta]$ on solutions of S, and so the symmetry condition (1.7.2) may be restated as,

$$\text{pr } \mathbf{v}_{U'}[\Delta] = 0. \tag{1.7.5}$$

It is this form of Definition 1.7.2 that provides the algorithm for finding generalised symmetries. The method is as follows:

- decide on the maximum order of the derivatives on which the symmetries can depend;
- postulate an ansatz for the evolutionary representative of the symmetry (this will be a vector field looking like (1.7.3));
- impose dependencies due to S;
- impose the symmetry condition (1.7.5);
- solve the resulting equation, possibly by equating to zero the coefficients of any derivatives of order higher than those in the symmetry characteristic.

As a very simple example (which appears in [7]), consider the second order symmetries of the nonlinear wave equation,

$$u_t = uu_x. \tag{1.7.6}$$

Assuming a symmetry generator of the form (1.7.3) exists, the symmetry condition is,

$$D_t(Q) = uD_x(Q) + u_x Q, \tag{1.7.7}$$

where Q is the characteristic of the symmetry. This is expanded and all the t-derivatives of u are replaced using the equation of motion. Then (1.7.7) reduces

to:

$$\frac{\partial Q}{\partial t} - u\frac{\partial Q}{\partial x} + u_x^2\frac{\partial Q}{\partial u_x} + 3u_x u_{xx}\frac{\partial Q}{\partial u_{xx}} = u_x Q. \tag{1.7.8}$$

This is solved by a standard technique (see [7]) to give the result that any Q of the form,

$$Q = u_x R\left(x + tu, u, t + \frac{1}{u_x}, \frac{u_{xx}}{u_x^3}\right), \tag{1.7.9}$$

is the characteristic of a generalised symmetry of (1.7.6).

Another worked example is presented in Section 4.2. A slightly different kind of analysis appears in Section 5.2.

The task of finding large numbers of generalised symmetries can be greatly eased if one can find a *recursion operator* for the system. This is a linear pseudodifferential operator $\mathcal{R}$ with the property that, given an evolutionary generalised symmetry $\mathbf{v}_Q$ of Δ, the function $\tilde{Q} = \mathcal{R}Q$ is the characteristic of a new symmetry $\mathbf{v}_{\tilde{Q}}$. With luck, repeated application of the recursion operator on Q will produce large numbers of generalised symmetries, although there is no guarantee that the sequence will not terminate after a finite number of iterations. The recursion operator of the KdV equation is,

$$\mathcal{R} = D_x^3 + \frac{2}{3}u + \frac{1}{3}u_x D_x^{-1}. \tag{1.7.10}$$

(The inverse differential operator is just defined by its action on D_x.) The reader may verify that operating on the KdV symmetry with characteristic u_x with (1.7.10) gives the symmetry with characteristic $u_{xxx} + uu_x$.

Without extra knowledge (e.g. details of the Hamiltonian structure) there is no sure–fire way of constructing recursion operators. However, there are a number of criteria which $\mathcal{R}$ must meet in order to work. In particular, it is known [7] that a system of evolution equations of the form,

$$\mathbf{u}_t = \mathbf{K}[\mathbf{u}], \tag{1.7.11}$$

which admits a recursion operator, admits a Lax–type representation,

$$\mathcal{R}_t = [\mathcal{A}, \mathcal{R}], \tag{1.7.12}$$

if the $\mathcal{R}$ is the recursion operator and $\mathcal{A}$ is the Fréchet derivative:

$$\mathcal{A} = \mathsf{D}_{\mathbf{K}} = \left(\frac{d}{d\varepsilon} \mathbf{K}[\mathbf{u}_\varepsilon] \right) \Big|_{\varepsilon=0} . \tag{1.7.13}$$

Here, $\mathbf{u}_\varepsilon$ denotes a one–parameter (ε) family of perturbations of $\mathbf{u}$. See [7] for details. See also [14] for a discussion of how the existence of a recursion operator affects integrability properties.

Chapter 2

Lagrangian and Hamiltonian Methods

2.1 Introduction

It is a religiously held belief that the most fundamental descriptions of physics (both classical and quantum) arise from Lagrangian and Hamiltonian field theories. Fittingly, the symmetries of such problems also have a distinguished role in the toolbox of analytic techniques open to physicists, as Noether's celebrated theorem links conservation laws in physical theories to symmetries of their respective actions.

The aim of this chapter is to present a review of these techniques using the language and tools of Chapter 1 — principally the fibre–bundle (E, π, M) and its infinite jet–bundle $\mathrm{J}^\infty\pi$. To begin with, we will discuss the calculus of variations and use our prolongation techniques to describe the interaction of the variational calculus and the symmetries of variational problems, in which the primary result is Noether's theorem. We will also look at the unified geometrical structure underlying much of this formalism, namely the *variational bicomplex*, which plays a similar role to the de Rham complex in differential geometry. We will also review the basic facts surrounding Hamiltonian evolution equations and their symmetries, including the construction of recursion operators from Hamiltonian structures.

The basic references for this chapter are [7], [15] and the review material in [16].

2.2 The Calculus of Variations

We start by defining the basic object of the variational calculus.

Definition 2.2.1 *A* kth order Lagrangian density *on* π *is a* C^∞ *function* $\mathcal{L} : \mathrm{J}^k\pi \to \mathbb{R}$. *A* kth order Lagrangian *is an* m*-form* $\lambda = \mathcal{L}\omega$ *where* ω *is the volume form on* M.

From now on, however, rather than specify k, we will just use the infinite prolongation, so that our arguments apply to Lagrangians of any order. The Lagrangian determines the action corresponding to γ by the formula,

$$\mathcal{S}(\gamma) = \int_N [\mathrm{pr}\,(\gamma)]^*\,\lambda, \tag{2.2.1}$$

where N is some compact m-dimensional submanifold of M. Given a vertical vector field $\mathbf{v}_Q$, with flow ψ_ε, we can define the *variation* of γ due to $\mathbf{v}_Q$ as the one–parameter family of local sections $\tilde{\psi}_\varepsilon = \psi_\varepsilon \circ \gamma$. If, in addition, $\mathbf{v}_Q$ is constrained to disappear on the boundary of N, we can pose the basic variational problem.

Definition 2.2.2 *The local section* γ *is called an* extremal *of* $\mathcal{S}$ *if,*

$$\left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \int_N \left[\mathrm{pr}\,(\tilde{\psi}_\varepsilon)\right]^*\,\lambda = 0. \tag{2.2.2}$$

In other words γ *determines a local maximum or minimum of the action.*

The business of variational calculus is to convert this definition into a condition on λ.

To do this, notice that if we swap the order of the integral and the ε derivation in (2.2.2), the resulting integrand is basically a Lie derivative. A short calculation tells us that:

$$\int_N \left.\frac{\partial}{\partial\varepsilon}\right|_{\varepsilon=0} \left[\mathrm{pr}\,(\tilde{\psi}_\varepsilon)\right]^*\,\lambda = \int_N [\mathrm{pr}\,(\gamma)]^*\,\mathsf{L}_{\mathrm{pr}\,(\mathbf{v}_Q)}(\lambda). \tag{2.2.3}$$

The symbol L_X stands for a Lie derivative along X. So a necessary condition for the extremality of γ is that:

$$\int_N [\mathrm{pr}\,(\gamma)]^* \mathsf{L}_{\mathrm{pr}\,(\mathbf{v}_Q)}(\lambda) = 0. \tag{2.2.4}$$

Of course it would be nice to have this result in some usable coordinate language. Since Definition 2.2.1 specifies that the Lagrangian is defined in terms of a fixed volume form we can do basically the same calculation with the differential functions rather than the m-forms. We will need an object similar to a Lie derivative designed for use with differential functions. We have already met the relevant tool in Section 1.7. If $R : \mathrm{J}^k\pi \to \mathbb{R}^n$ is a differential function, its *Fréchet derivative* is the differential operator denoted D_R and defined by the formula,

$$\mathsf{D}_R(Q) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} R\,[u + \varepsilon Q[u]]\,. \tag{2.2.5}$$

See Olver [7] for a full description of the properties of the Fréchet derivative.

Sorting out the coordinate expressions, we see that (2.2.2) comes down to,

$$\int_N \mathsf{D}_{\mathcal{L}}(Q)\omega = 0, \tag{2.2.6}$$

where Q is the characteristic of $\mathbf{v}_Q$. We can formally integrate the left hand side by parts to find:

$$\int_N \mathsf{D}_{\mathcal{L}}(Q)\omega = \int_N \{Q\mathsf{D}^*_{\mathcal{L}}(1) + \mathrm{Div}\,V\}\,\omega. \tag{2.2.7}$$

$\mathsf{D}^*_{\mathcal{L}}$ indicates the adjoint of the operator $\mathsf{D}_{\mathcal{L}}$ and V is a vector that depends on Q and the details of $\mathcal{L}$. The local statement,

$$\mathsf{D}_{\mathcal{L}}(Q) = Q\mathsf{D}^*_{\mathcal{L}}(1) + \mathrm{Div}\,V, \tag{2.2.8}$$

is called the *first variational formula.* We will meet it when we discuss Noether's theorem and the bicomplex. The expression $\mathsf{D}^*_{\mathcal{L}}(1)$ is called the *Euler variation* of $\mathcal{L}$, and is usually written $\mathcal{E}(\mathcal{L})$. It has the coordinate expression:

$$\mathcal{E}\mathcal{L} = (-D)_J \frac{\partial \mathcal{L}}{\partial u_J}. \tag{2.2.9}$$

Notice that if $\mathcal{L}$ is a divergence, (2.2.8) implies that $\mathcal{E}(\mathcal{L})$ is identically zero. When we impose (2.2.2), we find that the boundary conditions kill the divergence and so we recover the familiar necessary condition for a section to be extremal:

$$\mathcal{E}(\mathcal{L}) = 0. \tag{2.2.10}$$

The statement in the language of forms is that the *Euler–Lagrange form* $\mathcal{E}(\lambda) = \mathcal{E}(\mathcal{L})du^\alpha \wedge \omega$ is zero.

Finally, using the Fréchet derivative definition of the Euler variation, we can work out a handy formula for the Euler derivative of a product of two functions. If $\mathcal{F}_1$ and $\mathcal{F}_2$ are smooth, real–valued functions, then the Euler derivative of their product is,

$$\mathcal{E}(\mathcal{F}_1\mathcal{F}_2) = \mathsf{D}^*_{\mathcal{F}_1}(\mathcal{F}_2) + \mathsf{D}^*_{\mathcal{F}_2}(\mathcal{F}_1). \tag{2.2.11}$$

The proof is pretty straightforward from the definition of $\mathcal{E}$ and the derivation properties of the Fréchet operator, or directly from the components of (2.2.9)

2.3 Symmetries of Variational Problems

The symmetries of variational problems are particularly interesting because of the theorem of Noether linking conservation laws to symmetries of the action. To discuss these ideas, we need to define what is meant by a "symmetry of the action". We will only skim lightly over these topics, motivated by future needs, and leave the reader to consult the cited texts for further enlightenment. The presentation follows [16], although the material can be found in slightly different language in [7].

We are most interested in investigating the Euler–Lagrange equations of an action which is invariant under some group transformation. In other words, given some action $\mathcal{S}$, we want to look at point transformations g with the property that,

$$\mathcal{S}(\gamma) = \mathcal{S}(g(\gamma)), \tag{2.3.1}$$

where $\gamma : M \to E$ is any local section. The corresponding infinitesimal statement is that the corresponding Lagrangian m-form is invariant under the flow induced by the

infinitesimal generator of g, $\mathbf{v}$; in other words, if

$$\mathsf{L}_{\mathrm{pr}\,\mathbf{v}}(\lambda) = 0, \tag{2.3.2}$$

where L_X again denotes the Lie derivative along X. If $\mathbf{v}$ is of the form (1.3.6), then (2.3.2) reduces to,

$$\mathrm{pr}\,\mathbf{v}(\mathcal{L}) + \mathcal{L}\,\mathrm{Div}\,X = 0. \tag{2.3.3}$$

If we notice that the Euler variation commutes with the Lie derivative (not difficult to prove), we find that:

$$\mathsf{L}_{\mathrm{pr}\,\mathbf{v}}(\mathcal{E}(\lambda)) = 0. \tag{2.3.4}$$

Compare this with the standard definition of the symmetry algebra of a system of differential equations Δ (1.7.2):

$$\mathrm{pr}\,\mathbf{v}(\Delta) = 0.$$

This is a much less restrictive condition than (2.3.4) and so it is profitable to make a distinction between symmetries which obey (2.3.4) and those that do not.

Definition 2.3.1 *A* distinguished symmetry, $\mathbf{v}$, *of a (system of) differential equations Δ is a vector field satisfying,*

$$\mathsf{L}_{\mathrm{pr}\,\mathbf{v}}(\Delta) = 0. \tag{2.3.5}$$

The standard example of a symmetry that is not a distinguished symmetry is provided by the scaling vector field $u\partial_u$.

Corresponding to generalised vector fields, we can define *variational* symmetries. Essentially, the condition (2.3.3) is replaced by,

$$\mathrm{pr}\,\mathbf{v}(\mathcal{L}) + \mathcal{L}\,\mathrm{Div}\,X = \mathrm{Div}\,B, \tag{2.3.6}$$

for some suitable B. If $\mathbf{v}$ is of the form (1.7.1), there is a handy equivalent statement in terms of its evolutionary representative; it is not hard to prove that (2.3.3) can be written:

$$\mathrm{pr}\,\mathbf{v}_Q(\mathcal{L}) = \mathrm{Div}\,\tilde{B}. \tag{2.3.7}$$

where $\mathbf{v}_Q$ is the evolutionary representative of $\mathbf{v}$ and $\tilde{B} = B - \mathcal{L}X$. It is a theorem that a generalised variational symmetry of the action defines a generalised symmetry of the Euler–Lagrange equations. Variational symmetries can be determined by their action on the equations themselves. The generalised vector field $\mathbf{v}_Q$ is a variational symmetry of the a set of Euler–Lagrange equations Δ if and only if:

$$\mathsf{D}_\Delta(Q) + \mathsf{D}^*_Q(\Delta) = 0. \tag{2.3.8}$$

The study of the symmetries of a variational problem inevitably leads to discussion of conservation laws.

Definition 2.3.2 *A* conservation law *for a system of p.d.e.s, S, is a statement of the form:*

$$\mathrm{Div}\, P = 0, \tag{2.3.9}$$

which is true on all solutions of S.

Given some assumptions about S having maximal rank (in the sense of (1.5.1)), this is equivalent to the statement,

$$\mathrm{Div}\, P = P^J D_J(\Delta), \tag{2.3.10}$$

the P^J being some continuous, real–valued functions on $\mathrm{J}^\infty\pi$. Integrating by parts, (2.3.10) is simply,

$$\mathrm{Div}\, P' = P\ \Delta, \tag{2.3.11}$$

where,

$$P = (-D)_J P^J, \tag{2.3.12}$$

and P' is just some vector which vanishes on solutions of S as P does. The function P is called the *characteristic* of the conservation law. Using the form (2.3.11) and the variational product formula (2.2.11), we can derive a simple condition for a differential function Q to be the characteristic of a conservation law:

$$\mathsf{D}^*_\Delta(Q) + \mathsf{D}^*_Q(\Delta) = 0, \tag{2.3.13}$$

which in turn implies,

$$\mathsf{D}^*_\Delta(Q) = 0 \tag{2.3.14}$$

in the space of solutions of S. There are various ways in which conservation laws can be trivial, or trivially related to some other law. A list of such properties, and how they are characterised can be found in Olver [7]. Understanding the nature of such laws is helpful in motivating the variational complex which we are about to meet.

The Noether theorem arises when we notice that the condition (2.3.3), or the generalised version (2.3.6) make the left hand side of (2.2.8) zero, and so we have:

$$\mathrm{Div}\, V = Q\mathcal{E}(\lambda), \tag{2.3.15}$$

showing that Q is the characteristic of a conservation law for the Euler–Lagrange equation.

2.4 The Variational Bicomplex

The unified theory of all this variational formalism is the variational complex. Its most immediate application is to the so–called *inverse problem* of variational calculus — namely, given a differential equation S, can we determine if it is the Euler–Lagrange equation for extremising some unknown action? The problem can be tackled at various levels of generality. The simplest version, determining if S is exactly an Euler–Lagrange expression, is largely solved. Questions remain about more general problems, such as deciding whether S is *equivalent* to a variational equation via some multiplicative factor.

In the present exposition, we will motivate the existence of the bicomplex by looking at the inverse problem and its cohomological solution, and then state some lore surrounding it that will be useful later on. The reader will find excellent introductions to the topic in the review articles by Anderson [15, 17] and, of course, the ubiquitous Olver [7]. More detail may be found in [18].

We start with the simplest inverse problem in the case of scalar fields. Given a Lagrangian density $\mathcal{L}[u]$, we have the basic variational formula (2.2.8). As remarked

in [15], this uniquely characterises the Euler–Lagrange equation. In other words, if,

$$\mathsf{D}_{\mathcal{L}}(Q) = \Delta[u]Q + \operatorname{Div} V[u], \tag{2.4.1}$$

then, for some suitable $V[u]$, $\Delta[u]$ is the Euler–Lagrange expression corresponding to $\mathcal{L}$.

This is the property we use to answer the simplest version of the inverse problem: "When is $\Delta[u]$ the Euler–Lagrange expression for some unknown Lagrangian $\mathcal{L}$?" If we operate on (2.4.1) with $\mathcal{E}$, and use the fact that the variation annihilates divergences, we deduce:

$$\begin{aligned} \mathcal{E}\left(\Delta[u]\right)Q &= \mathcal{E}\left(\mathsf{D}_{\mathcal{L}}(Q)\right), \\ \Rightarrow \mathsf{D}^{*}_{\Delta[u]}(Q) &= \mathsf{D}^{*}_{\mathsf{D}_{\mathcal{L}}(Q)}(1), \\ &= \mathsf{D}_{\mathcal{E}\mathcal{L}}(Q), \\ &= \mathsf{D}_{\Delta[u]}(Q), \end{aligned} \tag{2.4.2}$$

if Δ is indeed the Euler–Lagrange expression for $\mathcal{L}$. This restriction, that the Fréchet derivative of Δ be self–adjoint, is called the *Helmholtz condition.* If we define the operator,

$$\mathrm{H}(\Delta) = \mathsf{D}_{\Delta[u]} - \mathsf{D}^{*}_{\Delta[u]}, \tag{2.4.3}$$

then (2.4.2) is simply $\mathrm{H}(\Delta) = 0$.

At this point, it is possible to summarise of what we have learned so far. We can write down a chunk of what looks like a formal cochain complex:

$$\left[\begin{array}{c}\text{Vectors on}\\ \mathrm{J}^{\infty}\pi\end{array}\right] \xrightarrow{\text{Div}} \left[\begin{array}{c}\text{Functions}\\ \mathrm{J}^{\infty}\pi \to \mathbb{R}\end{array}\right] \xrightarrow{\mathcal{E}} \left[\begin{array}{c}\text{Functions}\\ \mathrm{J}^{\infty}\pi \to \mathbb{R}\end{array}\right] \xrightarrow{H} \left[\begin{array}{c}\text{Diff. Operators}\\ \text{on } \mathrm{J}^{\infty}\pi\end{array}\right] \tag{2.4.4}$$

The composition of successive maps is zero. The exactness of this pseudo-sequence is the statement that the Euler variation annihilates divergences, and that Euler–Lagrange forms satisfy the Helmholtz condition. The variational bicomplex arises from formalising and generalising this construction to provide a *bona fide* bigraded complex on $\mathrm{J}^{\infty}\pi$.

To define this object, we need to discuss the algebra of forms on $J^\infty\pi$ which is denoted $\Omega^*(J^\infty\pi)$. As usual, this algebra is generated by the one-forms through a wedge operation. In coordinates these one–forms are defined by their actions on vector fields in $TJ^\infty\pi$:

$$\begin{aligned}\left\langle dx^i; \frac{\partial}{\partial x^j}\right\rangle &= \delta^i_j,\\ \left\langle du^\alpha; \frac{\partial}{\partial u^\beta}\right\rangle &= \delta^\alpha_\beta,\\ \left\langle du^\alpha_I; \frac{\partial}{\partial u^\beta_J}\right\rangle &= \delta^I_J\,\delta^\alpha_\beta.\end{aligned} \tag{2.4.5}$$

Among all differential forms on $J^\infty\pi$, there is a distinguished set, the *contact forms.*

Definition 2.4.1 *A* contact form *is a form σ on $J^\infty\pi$ which satisfies*

$$[\mathrm{pr}\,(\gamma)]^*\,(\sigma) = 0,$$

for all local sections γ of E.

The contact forms define a differential ideal $\mathcal{C}(J^\infty\pi)$ of the algebra $\Omega^*(J^\infty\pi)$. This ideal is generated by contact one-forms having the coordinate representation,

$$\theta^\alpha_{i_1\ldots i_k} = du^\alpha_{i_1\ldots i_k} - u^\alpha_{i_1\ldots i_k j}dx^j,\ k = 0,1,2,\ldots,$$

which satisfy,

$$d\theta^\alpha_{i_1\ldots i_k} = dx^j \wedge \theta^\alpha_{i_1\ldots i_k j}.$$

As far as we are concerned, the importance of these one forms is that they form a local basis of $\Omega^*(J^\infty\pi)$.

The variational bicomplex arises when we bigrade $\Omega^*(J^\infty\pi)$. The definition of the bigradation is rather convoluted.

Definition 2.4.2 *A p-form ρ on $J^\infty\pi$ is said to be* of type (r,s), *where $r+s=p$ if, at each point of $J^\infty\pi$,*

$$\rho\,(\mathbf{v}_1,\ldots,\mathbf{v}_p) = 0,$$

whenever (i) *more than* s *of the vectors* $\mathbf{v}_1, \ldots, \mathbf{v}_p$ *are* π_∞*-vertical, or* (ii) *more than* r *of them annihilate all contact one–forms.*

In coordinates, this just means that an (r, s)-form ρ looks like:

$$\rho = f[u]\, dx^{i_1} \wedge \ldots \wedge dx^{i_r} \wedge \theta^{\alpha_1}_{J_1} \wedge \ldots \wedge \theta^{\alpha_s}_{J_s}.$$

The space of (r, s)-forms is called $\Omega^{(r,s)}(\mathrm{J}^\infty\pi)$ and we can write:

$$\Omega^p(\mathrm{J}^\infty\pi) = \bigoplus_{r+s=p} \Omega^{(r,s)}(\mathrm{J}^\infty\pi). \tag{2.4.6}$$

The next step is to define the mappings between these spaces. The graded algebra of forms on $\mathrm{J}^\infty\pi$ comes with the differential,

$$d : \Omega^p(\mathrm{J}^\infty\pi) \to \Omega^{p+1}(\mathrm{J}^\infty\pi).$$

This can be decomposed into "vertical" and "horizontal" pieces, $d = d_H + d_V$, where:

$$d_H : \Omega^{r,s}(\mathrm{J}^\infty\pi) \to \Omega^{r+1,s}(\mathrm{J}^\infty\pi)$$
$$d_V : \Omega^{r,s}(\mathrm{J}^\infty\pi) \to \Omega^{r,s+1}(\mathrm{J}^\infty\pi).$$

The property $d^2 = 0$ implies that,

$$\begin{aligned} d_H^2 &= 0 \\ d_V^2 &= 0 \\ d_H\, d_V &= -d_V\, d_H. \end{aligned} \tag{2.4.7}$$

In local coordinates, the vertical and horizontal differentials of a function ($(0,0)$-form) are given by:

$$\begin{aligned} d_H f &= D_i(f[u])\, dx^i, \\ d_V f &= \sum_{|J|=0}^{\infty} \frac{\partial f}{\partial u^\alpha_J} \theta^\alpha_J. \end{aligned} \tag{2.4.8}$$

The action on one–forms is straightforward:

$$\begin{aligned} d_H(dx^i) &= 0; \qquad & d_V(dx^i) &= 0; \\ d_H(\theta^\alpha_I) &= dx^i \wedge \theta^\alpha_{Ij}; \qquad & d_V(\theta^\alpha_I) &= 0. \end{aligned} \tag{2.4.9}$$

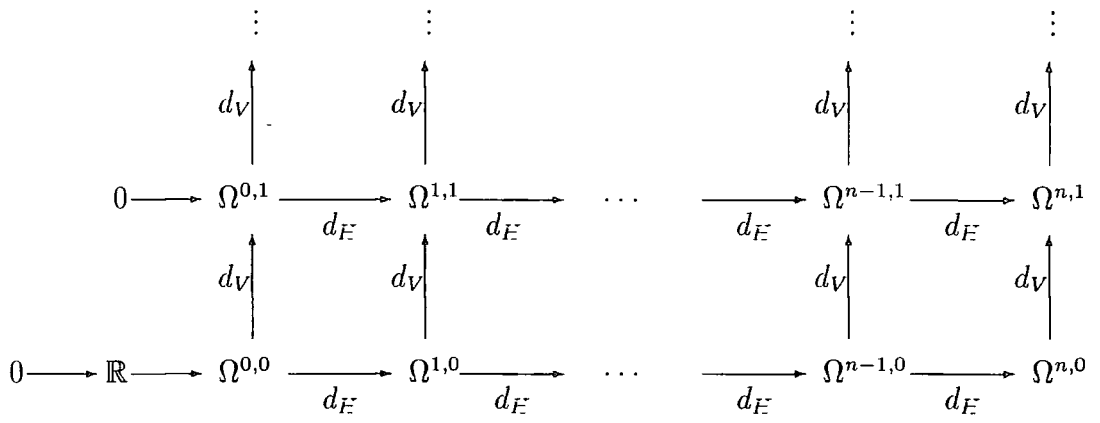

Figure 2.1: The Variational Bicomplex

The vertical differential will be useful in our variational calculus. Indeed, given a (generalised) vector field $\mathbf{v}_Q = Q\partial_u$, the basic variational formula (2.2.8) can be rewritten,

$$\mathsf{D}_{\mathcal{L}}(Q) = \mathrm{pr}\,\mathbf{v}_Q \lrcorner\, d_V\mathcal{L}, \tag{2.4.10}$$

where, as usual, γ is a section of E.

We end up with the commutative diagram Figure 2.1. Forms of type (n, s) are automatically d_E closed, but need not be d_E exact.

As we saw in Section 2.2, Lagrangians λ are forms in $\Omega^{n,0}$. The corresponding action functional for a section $\gamma : M \to E$ is

$$S[\gamma] = \int_M (\mathrm{pr}\,(\gamma))^*(\lambda).$$

The Euler–Lagrange form lies in the space $\Omega^{(n,1)}$, but to reconstruct our mini-complex (2.4.4) we need to refine our structure with a new mapping.

Define the *interior Euler operator*, an endomorphism $I : \Omega^{(n,s)} \to \Omega^{(n,s)}$, with the local coordinate expression:

$$I(\omega) = \frac{1}{s}\theta^\alpha \wedge \left[\left(\frac{\partial}{\partial u^\alpha} \lrcorner\, \omega\right) - D_i\left(\frac{\partial}{\partial u^\alpha_i} \lrcorner\, \omega\right) + \ldots + (-1)^{|J|} D_J\left(\frac{\partial}{\partial u^\alpha_J} \lrcorner\, \omega\right) + \ldots\right]. \tag{2.4.11}$$

I is also a projection operator: $I^2 = I$. A form in $\Omega^{(n,s)}$ which lies in the image of I is called a *source form*, and the set of all such forms is denoted $\mathcal{F}^s$.

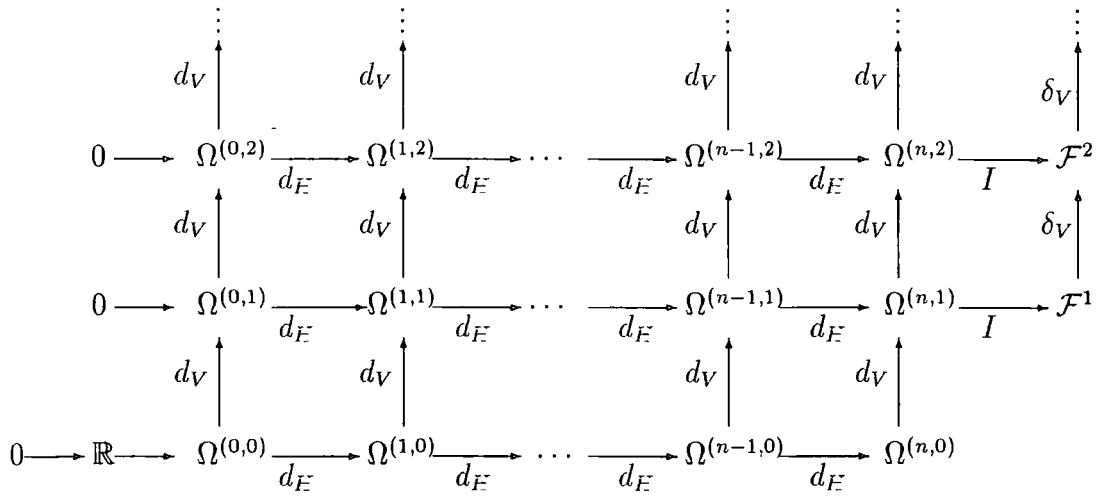

Figure 2.2: The Augmented Bicomplex

The importance of the interior Euler operator becomes clear when we notice that the Euler–Lagrange operator can now be written as,

$$\mathcal{E}(\lambda) = I(d_V \lambda). \tag{2.4.12}$$

We can now restate the Helmholtz condition by saying that an $(n,1)$-form $\Delta = \Delta_\alpha[u]\theta^\alpha \wedge dx^1 \wedge \ldots \wedge dx^n$ is an exact Euler–Lagrange form if and only if,

$$I(d_V \Delta) = 0. \tag{2.4.13}$$

Given that much of the work in this thesis is concerned with variational equations, we will often abuse the terminology "source form"to mean the form of any differential equation before it is set to zero.

The statement (2.4.13), and its generalisations to higher source forms, can be formalised by augmenting the bicomplex using a new differential, $\delta_V : \mathcal{F}^s \to \mathcal{F}^{s+1}$, defined by,

$$\delta_V(\omega) = I(d_V(\omega)). \tag{2.4.14}$$

This squares to zero as required, and we end up with the augmented bicomplex shown in Figure 2.2.

Our prototype complex (2.4.4) lies in the bottom–right edge of this larger structure, and this is the part of the bicomplex which interests us most.

$$0 \longrightarrow \mathbb{R} \xrightarrow{d_E} \Omega^{(0,0)} \xrightarrow{d_E} \ldots \xrightarrow{d_E} \Omega^{(n,0)} \xrightarrow{\mathcal{E}} \mathcal{F}^1 \xrightarrow{\delta_V} \ldots$$

Figure 2.3: The Euler–Lagrange Complex

Definition 2.4.3 *The* Euler–Lagrange complex $\mathcal{E}^*(\mathrm{J}^\infty\pi)$ *for the bundle* (E, π, M) *is the edge complex of the augmented diagram Figure 2.2. See Figure 2.3.*

All the variational calculus we have discussed is subsumed into the geometry of this complex. The variation formula (2.4.10) can be applied to each stage of the sequence. Given a vertical vector field $\mathbf{v}_Q$ on E, and $\omega \in \Omega^{(r,s)}$,

$$\mathrm{pr}\,\mathbf{v}_Q \lrcorner\, d_E(\omega) = -d_E(\mathrm{pr}\ \mathbf{v}_Q \lrcorner\, \omega), \tag{2.4.15}$$

or,

$$\mathsf{L}_{\mathrm{pr}\,\mathbf{v}_Q}(\omega) = d_V(\mathrm{pr}\,\mathbf{v}_Q \lrcorner\, \omega) + \mathrm{pr}\ \mathbf{v}_Q \lrcorner\, d_V(\omega). \tag{2.4.16}$$

Conservation laws are written in terms of forms,

$$\nu = \sum_{i=1}^{m} (-1)^i P^i dx^1 \wedge \ldots dx^{i-1} \wedge dx^{i+1} \ldots dx^m, \tag{2.4.17}$$

of type $(m-1, 0)$, in which case (2.3.11) is written:

$$d_E \nu = X_P \lrcorner\, \Delta\theta \wedge \omega, \tag{2.4.18}$$

where X_P is the vector field with characteristic P.

The Euler complex is exact when π is the trivial bundle $\pi : \mathbb{R}^{n+m} \to \mathbb{R}^m$, and this fact provides us with all the information we need about the inverse problem. (The homotopy operators that demonstrate exactness are constructed in [7]). Also, note that we can construct "functional forms" out of the elements of $\Omega^*(\mathrm{J}^\infty\pi)$ by a formal integral over the relevant forms, provided we incorporate the boundary effects and equivalences due to the kernels of the relevant differentials [7]. The important point to note is that the wedge product is *only* well defined on the differential forms, not on their functional versions.

2.5 Hamiltonian Evolution Equations

The Lagrangian formalism of the previous sections is not applicable to all equations of physical interest. In particular, no evolution equation of the form,

$$u_t = K[u], \tag{2.5.1}$$

admits an Euler–Lagrange representation, a fact that follows from the Helmholtz condition. To deal with such objects we will use the Hamiltonian picture of evolution equations. The infinite–dimensional Hamiltonian formalism is based on the finite–dimensional Hamiltonian theory of classical dynamical systems, details of which can be found in [19] or [7].

To start the infinite–dimensional analysis, we introduce the two ingredients necessary for a Hamiltonian structure. First, we require a *Hamiltonian functional* $\mathcal{H}[u]$ which we define in terms of an integral over space of a Hamiltonian density H:

$$\mathcal{H}[u] = \int \mathsf{H} dx. \tag{2.5.2}$$

Second, we need a matrix differential operator, the *Poisson operator* $\mathcal{J}$ such that (2.5.1) can be written,

$$u_t = \mathcal{J}\delta\mathcal{H}[u]. \tag{2.5.3}$$

$\mathcal{J}$ must be a skew-adjoint differential operator. In this context, "skew–adjoint" means that,

$$\int_N A.\mathcal{J}Bdx = \int_N B.\mathcal{J}^*Adx = -\int_N B.\mathcal{J}Adx, \tag{2.5.4}$$

assuming that the support of A and B on the region N allows formal integration by parts. (This will be a constant assumption throughout this thesis.) The associated *Poisson bracket* of two functionals, $\mathcal{P}$ and $\mathcal{Q}$, is the bilinear operation:

$$\{\mathcal{P}, \mathcal{Q}\} = \int \delta\mathcal{P}.\mathcal{J}\delta\mathcal{Q}dx. \tag{2.5.5}$$

This bracket must satisfy the Jacobi identity:

$$\{\{\mathcal{P}, \mathcal{Q}\}, \mathcal{R}\} + \{\{\mathcal{R}, \mathcal{P}\}, \mathcal{Q}\} + \{\{\mathcal{Q}, \mathcal{R}\}, \mathcal{P}\} = 0, \tag{2.5.6}$$

for all functionals $\mathcal{P}, \mathcal{Q}$ and $\mathcal{R}$.

Comparing with the finite–dimensional definition of a Hamiltonian system, we see that we have dropped the stipulation that the Poisson bracket obeys a Leibniz rule. This is largely because there is no acceptable method to define the multiplication of two functionals. Nevertheless, there is a well–defined Hamiltonian vector field, $\mathbf{v}_{\mathcal{H}}$, satisfying,

$$\operatorname{pr} \mathbf{v}_{\mathcal{H}}(\mathcal{P}) = \{\mathcal{P}, \mathcal{H}\}, \tag{2.5.7}$$

with characteristic $\mathcal{J}\delta\mathcal{H}$. A Lie bracket [. , .] between Hamiltonian vector fields can be defined by:

$$\operatorname{pr} [\mathbf{v}_{\mathcal{G}}, \mathbf{w}_{\mathcal{H}}](P) = \operatorname{pr} \mathbf{v}_{\mathcal{G}} (\operatorname{pr} \mathbf{w}_{\mathcal{H}}(P)) - \operatorname{pr} \mathbf{w}_{\mathcal{H}} (\operatorname{pr} \mathbf{v}_{\mathcal{G}}(P)). \tag{2.5.8}$$

Essentially the Lie bracket defines a new Hamiltonian vector field with characteristic,

$$\operatorname{pr} \mathbf{v}_{\mathcal{G}}(\mathcal{J}\delta\mathcal{H}) - \operatorname{pr} \mathbf{w}_{\mathcal{H}}(\mathcal{J}\delta\mathcal{H}).$$

There is a basic theorem, which is useful when discussing symmetries, stating that,

$$\operatorname{pr} \mathbf{v}_{\{\mathcal{P},\mathcal{Q}\}} = -[\mathbf{v}_{\mathcal{P}}, \mathbf{w}_{\mathcal{Q}}], \tag{2.5.9}$$

for functionals $\mathcal{P}, \mathcal{Q}$.

The main difficulty in analysing Hamiltonian evolution equations is verifying the Jacobi identity (2.5.6). The full calculation using the bracket definition (2.5.5) usually involves a horrendous amount of labour. Therefore, in this thesis, we will use a more efficient method due to Olver [20] and described in his book [7]. The technique is based on an observation from the finite–dimensional Hamiltonian formalism. If we define the bivector,

$$\Theta = \frac{1}{2} J^{ij} \partial_i \wedge \partial_j, \tag{2.5.10}$$

where J^{ij} is the structure matrix of some Hamiltonian system, then it can be shown that there is a binary operation called the *Schouten bracket* [. , .]$_S$ such that the statement,

$$[\Theta, \Theta]_S = 0, \tag{2.5.11}$$

is equivalent to the Jacobi identity. The presentation here is geared towards describing the method for later use — the reader is referred to the cited texts for further motivation and justification.

The basic objects are univectors denoted Φ^α_J which can be expressed in coordinates as $\frac{\partial}{\partial u^\alpha_J}$. We will construct functional versions of them and multivectors constructed from them in a manner dual to that of the forms discussed in Section 2.4. These univectors have the property,

$$\langle \Phi^\alpha_J; P \rangle = D_J P^\alpha, \tag{2.5.12}$$

where P is some vertical form whose coefficients P^α are differential functions. Then a *functional k-vector* looks like,

$$\Theta = \int R^{\alpha'}_J[u] \Phi^{\alpha_1}_{J_1} \wedge \ldots \wedge \Phi^{\alpha_k}_{J_k}, \tag{2.5.13}$$

where the multi-indices α' and J have the k components shown.

Using the blanket "integration by parts" assumption, it can be shown that any functional bivector can be written in the canonical form,

$$\Theta = \frac{1}{2} \int \{\Phi \wedge \mathcal{D}\Phi\}\, dx, \tag{2.5.14}$$

where $\mathcal{D}$ is a skew-adjoint differential operator. Θ then determines an inner product corresponding to the Poisson bracket.

Correspondingly, the canonical trivector is,

$$\Psi = \frac{1}{2} \int \{\Phi \wedge \mathrm{pr}\, \mathbf{v}_{\mathcal{D}\Phi}(\mathcal{D}) \wedge \Phi\}\, dx, \tag{2.5.15}$$

where $\mathbf{v}_{\mathcal{D}\Phi}$ is an evolutionary vector field with characteristic $\mathcal{D}\Phi$. This can also be written,

$$\Psi = -\mathrm{pr}\, \mathbf{v}_{\mathcal{D}\Phi}(\Theta) = -\frac{1}{2} \int \{\mathrm{pr}\, \mathbf{v}_{\mathcal{D}\Phi}(\Phi \wedge \mathcal{D}\Phi)\}\, dx, \tag{2.5.16}$$

where Θ is the canonical bivector in (2.5.14). The trivector determines the inner product that generates the Jacobi expression. So the basic result is the following.

Proposition 2.5.1 *The Jacobi identity determined by the skew–adjoint operator $\mathcal{D}$ is satisfied if and only if:*

$$\frac{1}{2}\int \{\mathrm{pr}\, \mathbf{v}_{\mathcal{D}\Phi}(\Phi \wedge \mathcal{D}\Phi)\}\, dx = 0. \tag{2.5.17}$$

Then $\mathcal{D}$ defines a Poisson structure.

This procedure greatly speeds up the verification of the Jacobi identity. Nevertheless, there is still a certain amount of work to be done. Appendix C contains REDUCE code to calculate the integrand of (2.5.17) for a particular example to be met in Chapter 4. It is easily adaptable to other Poisson operators. When the integrand of (2.5.17) has been found, it only remains to check that it is a total derivative in order to satisfy the Jacobi identity.

2.6 Symmetries of Hamiltonian Systems

There is a great deal to be said about the interaction of symmetry algebras and Hamiltonian equations. A comprehensive study can be found in the book by Marsden and Ratiu [19]. We will merely skim over a few essential points concerning conservation laws. Of much more use will be the theory of biHamiltonian systems and Magri's theorem that provides a means of demonstrating the complete integrability of a system of evolution equations.

The theory of the generalised symmetries and conservation laws of a Hamiltonian equation is essentially the same as explained in Chapter 1. However, there are a number of additional features introduced by the Hamiltonian structure. To begin with, we must define the analogue of the Casimir functions of finite–dimensional Hamiltonian systems. Our presentation follows Olver [7].

Definition 2.6.1 *Suppose $\mathcal{J}$ is a Poisson operator. A* distinguished functional *of $\mathcal{J}$ is a functional C such that $\mathcal{J}\delta\mathcal{C} = 0$ for all x and u.*

These distinguished functionals are always conserved.

The question of less trivial conserved quantities requires a Noether–like link between symmetries and charges. The relevant theorem is given in [7].

Theorem 2.6.1 *Consider a system of Hamiltonian equations of the form (2.5.3). A Hamiltonian vector field* $\mathbf{v}_{\mathcal{P}}$ *(with characteristic* $\mathcal{J}\delta\mathcal{P}$*) determines a generalised symmetry of the system if and only if there is a conserved functional* $\tilde{\mathcal{P}}$ *equivalent to* $\mathcal{P}$ *by,*

$$\tilde{\mathcal{P}} = \mathcal{P} + \mathcal{C},$$

where $\mathcal{C}$ *is a distinguished functional.*

The importance of this theorem lies in its application to the theory of biHamiltonian systems; in other words the situation when the evolution equation has two distinct Hamiltonian formulations. This fact and its significance were first noted by Magri [21]. The argument followed here is the form due to Olver [7].

As an illustrative example, consider the Korteweg–de Vries (KdV) equation in the form:

$$u_t = u_{xxx} + uu_x. \tag{2.6.1}$$

This evolution equation can be expressed as a Hamilton equation like (2.5.3) in two independent ways. The first structure follows from the Hamiltonian,

$$\mathcal{H}_1[u] = \int \left[-\frac{1}{2}u_x^2 + \frac{1}{6}u^3\right] dx, \tag{2.6.2}$$

and the Poisson structure,

$$\mathcal{J}_0 = D_x. \tag{2.6.3}$$

The second Hamiltonian formulation is defined by,

$$\mathcal{H}_0[u] = \int \frac{1}{2}u^2 dx, \tag{2.6.4}$$

and,

$$\mathcal{J}_1 = D_x^3 + \frac{2}{3}uD_x + \frac{1}{3}u_x. \tag{2.6.5}$$

Both of these structures are correctly skew–symmetric and obey the Jacobi identity.

The four point symmetries of the KdV equation which were found in Section 1.5 are equivalent to evolutionary generalised symmetries with characteristics,

$$\begin{aligned} Q_1 &= u_x, \\ Q_2 &= u_{xxx} + uu_x, \\ Q_3 &= 1 + tu_x, \\ Q_4 &= 2u + xu_x + 3t(u_{xxx} + uu_x), \end{aligned} \tag{2.6.6}$$

taking into account the dependency due to (2.6.1).

Using the characterisation (2.5.7) of Hamiltonian vector fields, we see that Q_1, Q_2 and Q_3 are Hamiltonian symmetries with respect to the structure $\mathcal{J}_0$, and the corresponding conserved densities are easily calculated, denote them by $\mathcal{C}_1$, $\mathcal{C}_2$, and $\mathcal{C}_3$. In addition, there is one distinguished functional, $\mathcal{C}_0$. The question is, does the second Hamiltonian structure $\mathcal{J}_1$ give the same results? Repeating the analysis, it is found that there are no distinguished functionals with respect to $\mathcal{J}_1$ and that Q_1, Q_2 and Q_4 are the Hamiltonian characteristics. However, two of the corresponding conserved functionals are identical to $\mathcal{C}_1$ and $\mathcal{C}_3$, and the other is $\mathcal{C}_0$. So we conclude that for the structure defined by $\mathcal{J}_1$, the conserved functional $\mathcal{C}_2$ must follow from a true generalised symmetry, denoted Q_5. Theorem 2.6.1 guarantees that this is Hamiltonian with respect to $\mathcal{J}_1$.

The gateway to the theory of biHamiltonian systems is opened when we make the crucial observation that Q_5 is also Hamiltonian with respect to $\mathcal{J}_0$, and that this leads to a further conserved functional $\mathcal{C}_5$. Continuing this recursive analysis between the two Hamiltonian structures, we end up with a procedure for generating an infinite tower of conserved quantities.

- Given a conserved functional $\mathcal{C}_k$ with respect to, say, $\mathcal{J}_0$, find the (necessarily existent) Hamiltonian characteristic Q_{k+1} such that $\mathcal{C}_k$ is its conserved functional with respect to $\mathcal{J}_1$.

- Find the new (necessarily existent) functional $\mathcal{C}_{k+1}$ which generates Q_{k+1} as a Hamiltonian characteristic with respect to $\mathcal{J}_0$.

- Repeat this procedure *ad nauseam.*

The result is, hopefully, an infinite list of mutually commuting conserved functionals and the corresponding symmetries.

We can summarise the procedure outlined above by saying we are looking for a sequence of characteristics Q_k and and conserved functionals $\mathcal{C}_k$ satisfying the recurrence,

$$Q_k = \mathcal{J}_0 \delta \mathcal{H}_k = \mathcal{J}_1 \delta \mathcal{H}_{k-1}, \tag{2.6.7}$$

and so find a new characteristic $Q_{k+1} = \mathcal{J}_1 \delta \mathcal{H}_k$. We can formally solve this by defining $\mathcal{R} = \mathcal{J}_1 \circ \mathcal{J}_0^{-1}$, whence:

$$Q_{k+1} = \mathcal{R} Q_k. \tag{2.6.8}$$

Clearly, $\mathcal{R}$ is a recursion operator for the system. Using (2.6.3) and (2.6.5), the reader may verify that $\mathcal{J}_1 \circ \mathcal{J}_0^{-1}$ for the KdV equation does in fact reproduce the recursion operator (1.7.10).

The rigorous implementation of this procedure is based on a series of theorems stemming from the work of Magri [21]. The basic definition imposes a subsidiary condition on the Hamiltonian structures.

Definition 2.6.2 *Two matrix differential operators $\mathcal{J}_0$ and $\mathcal{J}_1$ are said to be a* Hamiltonian pair *if every linear combination $a\mathcal{J}_0 + b\mathcal{J}_1$ $(a, b \in \mathbb{R})$ is a Hamiltonian operator as defined in Equations (2.5.3...2.5.6). If a system of evolution equations can be written as in the form (2.5.3) using both $\mathcal{J}_0$ and $\mathcal{J}_1$, it is said to be* biHamiltonian.

In Olver [7], it is proved that it is sufficient to show that $\mathcal{J}_0$, $\mathcal{J}_1$ and $\mathcal{J}_0 + \mathcal{J}_1$ are all Hamiltonian in order for the pair to be compatible.

Magri's main theorem can now be stated.

Theorem 2.6.2 *Given a system of evolution equations of the form (2.5.1) with bi-Hamiltonian structure defined by $(\mathcal{J}_0, \mathcal{H}_1)$ and $(\mathcal{J}_1, \mathcal{H}_0)$, define the recursion operator,*

$$\mathcal{R} = \mathcal{J}_1 \circ \mathcal{J}_0^{-1},$$

and the basic characteristics,

$$Q_0 = \mathcal{J}_0 \delta \mathcal{H}_0,$$

$$Q_1 = K[u].$$

Assume that for all $k \geq 1$, *we can recursively define,*

$$Q_k[u] = \mathcal{R} Q_{k-1}[u],$$

then there exists a sequence of functionals $\mathcal{H}_0, \mathcal{H}_1, \mathcal{H}_2, \ldots$ *such that:*
(1) for all $k \geq 1$, *the evolution equations,*

$$u_t = Q_k[u] \tag{2.6.9}$$

are biHamiltonian with structures $(\mathcal{J}_0, \mathcal{H}_k)$ *and* $(\mathcal{J}_1, \mathcal{H}_{k-1})$;
(2) the symmetries $\mathbf{v}_{Q_k}$ *all commute,*

$$[\mathbf{v}_{Q_k}, \mathbf{v}_{Q_l}] \quad \forall\ k, l \geq 0;$$

(3) the functionals $\mathcal{H}_k$ *are all in involution with respect to either structure,*

$$\{\mathcal{H}_k, \mathcal{H}_l\}_{\mathcal{J}_0} = \{\mathcal{H}_k, \mathcal{H}_l\}_{\mathcal{J}_1} = 0 \quad \forall\ k, l \geq 0,$$

and all are conserved functionals for all of the biHamiltonian equations of the hierarchy.

Two remarks must be made. First, we have failed to mention a mild nondegeneracy condition on the structure $\mathcal{J}_0$, discussion of which can be found in [7]. Suffice to say that all examples studied in this thesis satisfy the condition. Second, although the theorem is very powerful, it relies on the assumption that the recursive definition of the Q_k carries on *ad infinitum.* See [22] and Chapter 4 for examples where this fails.

Chapter 3

Series Solutions of Evolution Equations

3.1 Introduction

This chapter is concerned with a technique discussed in [23, 1] for finding solutions to a broad class of nonlinear partial differential equations. We use a series expansion, of a type originally proposed by Stokes [24], to solve problems like the Korteweg–de Vries Equation by a perturbation method, using the linearised equations as the initial approximations. After demonstrating the basic idea with an example, we will use the method to provide solutions of a couple of harder nonlinear problems. The process itself is based on a straightforward but computationally complicated algorithm which is best implemented with the aid of a computer algebra package such as REDUCE or MAPLE. However, the extraction of solutions from these computations still relies heavily on intuitive inspection of the results, and the fundamental reasons behind the technique remain mysterious. Nonetheless, this approach has yielded some well known results, exact solutions of the equations concerned. Speculation on the underlying theory is reserved until later.

Although this work was carried out independently, many of the results of this

chapter appear in a paper by R.R. Rosales[1] [2]. Rosales' work goes substantially beyond what appears here.

We will examine only equations in (1+1) dimensions with a single time derivative and a simple quadratic nonlinearity since these will prove to be most amenable to the method. The generalisation to problems that are of second order in time (e.g. the Boussinesq equation [25]) or have higher degrees of nonlinearity (e.g. the Modified KdV equation [26]) is superficially straightforward but messy. We found no solutions for these equations by this procedure, but Rosales demonstrates the soliton solutions of the Boussinesq equation. The generalisations to higher dimensions are not considered (although, again, Rosales considers some examples) although a step has been taken in this direction in [23]. Problems of interest in higher dimensions include the Navier–Stokes equations and Polyakov's theory of conformal turbulence [27], discussed in [23].

Good examples of the sort of equation in which we are interested are the simplest nonlinear wave (NLW) equation (not considered in [2]),

$$u_t = uu_x, \tag{3.1.1}$$

the well-known Burgers equation,

$$u_t = -uu_x + \nu u_{xx}, \tag{3.1.2}$$

and the Korteweg–de Vries (KdV) equation,

$$u_t = 6uu_x - u_{xxx}. \tag{3.1.3}$$

As always, u is a function of the two independent variables x and t and, in (3.1.2), ν is a constant. Equation (3.1.2) arises in the study of sound waves in viscous substances, and as a model of magnetohydrodynamic phenomena in media with finite electrical conductivity. The KdV equation famously describes long waves in shallow Scottish canals, and other weakly nonlinear wave motion.

[1]The author is grateful to Prof. P.A. Clarkson for bringing this paper to his attention.

All of these are soluble to varying extents. The NLW equation has an implicit solution given by the method of characteristics [25],

$$u = G(x + ut), \tag{3.1.4}$$

where G is an arbitrary function. The general solution to the Burgers equation is found by the ingenious Hopf–Cole transformation [25],

$$u = -2\nu \frac{\partial}{\partial x} \log \phi(x,t) \tag{3.1.5}$$

which transforms (3.1.2) to the heat equation,

$$\phi_t = \nu \phi_{xx} \tag{3.1.6}$$

which can be tackled by a number of methods. The KdV equation may be integrated by the inverse–scattering technique [28, 26] (see Appendix B) or the direct method [29].

In these special cases, the power series solutions which we will find can be summed formally to give closed solutions and the summation reproduces some of the answers mentioned above. In particular, the result for the KdV equation is found to be closely related to both the results of inverse–scattering technique and the "trace" representation of the soliton solutions discovered by Wadati and Sawada [30]. The connection is made explicit in Ref. [2].

3.2 Stokes' Method and its Application

Stokes expansion, as found in [25], is a method in its own right, but we will use it mainly as a tool for use with a more general procedure. First though, it is instructive to study its use to find a solution to a simple nonlinear problem.

Our chosen example is the NLW equation. As an ansatz for u we use:

$$u = u_0 + u_1 e^{ax} + u_2 e^{2ax} + \ldots + u_k e^{kax} + \ldots, \tag{3.2.1}$$

where the u_i are functions of t only. (The subscripts label distinct functions, not derivatives of u.) The constant a may be complex so this is a general Fourier or Dirichlet series. It is substituted back into equation (3.1.1) and the u_i are found successively by comparing the coefficients of like powers of e^x. This means solving a nested sequence of easy ordinary differential equations. Each time a differential equation is solved an arbitrary constant of integration is introduced.

On substituting the ansatz (3.2.1) into equation (3.1.1) and comparing coefficients we get:

$$u_{0_t} = 0,$$

telling us that u_0 is a constant which we shall call c. Continuing the process gives:

$$\begin{aligned} u_{1_t} &= cu_1, \\ u_{2_t} &= u_1^2 + 2cu_2, \\ u_{3_t} &= 3u_1u_2 + 3cu_3, \\ &\cdots \end{aligned} \tag{3.2.2}$$

and the differential recurrence relation for the general term:

$$u_{k_t} = kcu_k + \sum_{i=1}^{k-1}(k-i)u_iu_{k-i}. \tag{3.2.3}$$

Solving the equations (3.2.2) in turn yields:

$$\begin{aligned} u_1 &= c_1e^{ct}, \\ u_2 &= \left(c_1^2t + c_2\right)e^{2ct}, \\ u_3 &= \left(\frac{3}{2}c_1^3t^2 + 3c_1c_2t + c_3\right)e^{3ct}, \\ &\cdots \end{aligned} \tag{3.2.4}$$

where the c_i are the arbitrary constants of integration. In Appendix A the reader will find a MAPLE routine to calculate an arbitrary number of these terms. Clearly there is an emergent pattern in the sequence. Each u_k is a polynomial of order $k-1$ in t multiplied by e^{kct}. Further analysis (guessing!) reveals that the solution can best

be written as

$$u(x,t) = c + \sum_{N=1}^{\infty} \sum_{n=1}^{N} \sum_{\text{all possible } n_j} \prod_{\text{all possible } j} \frac{c_j^{n_j} (Nt)^{n-1} \exp\left(N(ax+ct)\right)}{n_j!}, \quad (3.2.5)$$

where "all possible n_j" and "all possible j" mean all n_j and j such that

$$\sum_j j n_j = N,$$

and

$$\sum_j n_j = n.$$

This form is not especially useful as it involves a sum over all partitions of the integer n. We can generate this partition by a formal multinomial to give a nicer form:

$$u(x,t) = c + \sum_{n=1}^{\infty} \left(\frac{t}{c}\right)^{n-1} \frac{1}{n!} \frac{\partial^{n-1}}{\partial x^{n-1}} \left(\sum_k c_k \exp\left(k(ax+ct)\right) \right)^n. \quad (3.2.6)$$

It is easy to verify that this can be expressed as a contour integral:

$$u(x,t) = c + \sum_{n=1}^{\infty} \frac{1}{2\pi i} \oint \frac{c}{nt} \left(c_1 \exp(ax+z) + c_2 \exp(2(ax+z)) + \cdots\right)^n \left(\frac{t}{c(z-ct)}\right)^n dz. \quad (3.2.7)$$

We can use any closed contour that loops around the multiple pole at $z = ct$. Performing the sum yields:

$$u(x,t) = c - \frac{1}{2\pi i} \oint \frac{1}{ct} \log\left(1 - \frac{t}{c(z-ct)} \sum_k c_k \exp\left(k(ax+z)\right)\right) dz, \quad (3.2.8)$$

an explicit formal solution of (3.1.1).

Furthermore, the initial value problem can now be solved directly. If we Fourier or Dirichlet analyse the initial profile of u in the manner of (3.2.1) we get,

$$u(x,0) = F(ax) = c + c_1 e^{ax} + c_2 e^{2ax} + c_3 e^{3ax} + \ldots, \quad (3.2.9)$$

where the c's are identical to those in (3.2.4 ... 3.2.8). This allows (3.2.8) to be rewritten by observing that the sum in the logarithm is actually $F(ax+ct)$ giving,

$$u(x,t) = c - \frac{1}{2\pi i} \oint \frac{1}{ct} \log\left(1 - \frac{t\left(F(ax+z) - c\right)}{c(z-ct)}\right) dz. \quad (3.2.10)$$

There is in fact a much more direct form of solution of the NLW equation[2]. The trick is to expand the solution in a Taylor expansion,

$$u(x,t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \frac{\partial^n}{\partial x^n} F(x), \tag{3.2.11}$$

and make use of the conservation laws,

$$(u^n)_t = \frac{n}{n+1} \left(u^{n+1}\right)_x. \tag{3.2.12}$$

It is then straightforward to prove that:

$$u(x,t) = \sum_{n=0}^{\infty} \frac{t^n}{(n+1)!} \frac{\partial^n}{\partial x^n} F(x)^{n+1}, \tag{3.2.13}$$

an explicit, calculable solution in terms of the initial data. It is a special feature of the NLW equation that the initial value problem can be solved so directly. This is not true in any other example which we have examined.

In passing we note that the ansatz used by Rosales [2] is somewhat more general than ours, and is in some ways closer to Stokes' original expansion. In particular, he assumes no separation of the dependence on x and t, requiring only that the solution is decomposed into a sum of constituent functions which depend on both x and t. We will return to this point in the conclusion.

3.3 A Perturbation Method Using Stokes' Expansion

Let us consider the Burgers equation as an example. Substitute (3.2.1) into (3.1.2) (with $a = 1$ for convenience) to get,

$$u_{1_t} = u_1, \tag{3.3.1}$$

and the recurrence,

$$u_{k_t} = \sum_{i=1}^{k-1} (k-i) u_i u_{k-i} - k^2 u_k, \tag{3.3.2}$$

[2]This observation is due to D.B. Fairlie.

which can be written

$$\left(u_k e^{k^2 t}\right)_t = \sum_{i=1}^{k-1} (k-i) u_i u_{k-i}. \tag{3.3.3}$$

We can neglect u_0 here as it is only comes in to the expansion in a trivial manner that allows it to be redefined away. Then the first few terms are:

$$\begin{aligned} u_1 &= c_1 \exp(\nu t), \\ u_2 &= c_2 \exp(4\nu t) - \frac{1}{2\nu} c_1^2 \exp(2\nu t), \\ u_3 &= c_3 \exp(9\nu t), -\frac{3}{4\nu} c_1 c_2 \exp(5\nu t) + \frac{1}{4\nu^2} c_1^3 \exp(3\nu t) \\ &\cdots . \end{aligned} \tag{3.3.4}$$

It is clear that the structure of the u_k will not permit a tidy connection to the initial value problem, because they are not polynomials in t. It is also more difficult to find a closed expression for the u_k.

The challenge is to understand the recurrence (3.3.3). If we study the structure of the sequence (3.3.4), we see that all the terms in the u_k are constructed out of elements like $c_i \exp(i^2 \nu t)$. We can generate these elements as solutions of the heat equation (3.1.6). If the ansatz (3.2.1) is applied to it, the heat equation in u generates an expansion,

$$\sum_k u_{k_t} \exp(kx) = \nu \sum k^2 u_k \exp(kx), \tag{3.3.5}$$

which clearly generates a series of ODEs which are homogeneous versions of (3.3.3), and whose solutions contain the elements we require. The higher order terms are built up from interactions of these elements. The revised idea, then, is to solve the linear problem (the heat equation) and then build up higher order corrections from the nonlinear driving terms.

So, if we introduce a new set of auxiliary variables w_k, the first stage is to solve,

$$w_{1_t} = \nu w_{1_{xx}} \tag{3.3.6}$$

by a Stokes expansion to get,

$$w_1(x,t) = \sum_{k_j} c_{k_j} \exp\left(k_j x + \nu k_j^2 t\right), \tag{3.3.7}$$

i.e. w_1 is a sum of all terms with just one arbitrary constant. Then use this to perturb the next order equation,

$$w_{2_t} - \nu w_{2_{xx}} = -w_1 w_{1_x}, \tag{3.3.8}$$

again using a Stokes expansion as an ansatz for w_2. This gives the sum of all terms containing two arbitrary constants. Continuing the process gives a sum of terms containing n constants with a general term:

$$(n_1k_1 + n_2k_2 + \ldots..n_rk_r)\prod_j \frac{1}{n}\left(\frac{-1}{2\nu}\right)^{n-1}\frac{n!}{n_j!}\left(\frac{c_{k_j}}{k_j}e^{k_jx+\nu k_j^2t}\right)^{n_j}, \tag{3.3.9}$$

which can be written,

$$\frac{\partial}{\partial x}\prod_j \frac{1}{n}\left(\frac{-1}{2\nu}\right)^{n-1}\frac{n!}{n_j!}\left(\frac{c_{k_j}}{k_j}e^{k_jx+\nu k_j^2t}\right)^{n_j}. \tag{3.3.10}$$

In both (3.3.9) and (3.3.10) $n = \sum n_j$.

The full nth order term is a sum of such contributions for fixed n:

$$w_n(x,t) = \frac{1}{n}\left(\frac{-1}{2\nu}\right)^{n-1}\frac{\partial}{\partial x}\left(\int \sum c_{k_j}e^{k_jx+\nu k_j^2t}dx\right)^n. \tag{3.3.11}$$

The answer is a sum over n of the w_n which gives:

$$u(x,t) = -2\nu\frac{\partial}{\partial x}\log\left(1 - \frac{1}{2\nu}\int \sum_{k_j} c_{k_j}e^{k_jx+\nu k_j^2t}dx\right). \tag{3.3.12}$$

This is the Hopf–Cole solution (3.1.5).

The final enhancement is to note that (3.3.12) is really

$$u(x,t) = -2\nu\frac{\partial}{\partial x}\log\left(1 - \frac{1}{2\nu}\int_{-\infty}^{x} w_1(y,t)dy\right) \tag{3.3.13}$$

and use for w_1 any solution of the linearised approximation that is properly convergent on the boundary.

It only remains to make the connection to the initial value problem . If as before $u(x,0) = F(x)$, then the corresponding initial value of w_1 is,

$$w_1(x,0) = \frac{-1}{2\nu}F(x)\exp\left(\frac{-1}{2\nu}\int_{-\infty}^{x} F(y)dy\right). \tag{3.3.14}$$

Then, using the solution of the heat equation (3.3.6), the result (3.3.13) gives a solution to the initial value problem. This result agrees with that in [2], where the similar situation for Thomas' equation is also reviewed.

So the perturbation method has reproduced a known solution, which is an encouraging sign of its capabilities. Furthermore, this example may provide some small clue to the underlying reasons for the success of this technique. The Hopf–Cole solution is usually rationalised by study of the Lie symmetries of the Burgers equation (see [7], or [31] for more detail). It would be interesting to look for a group–theoretical reason for this phenomenon. We defer further comment until later.

3.4 The KdV Equation

For the KdV equation, direct application of the Stokes procedure again gives a complicated answer. The recurrence is:

$$\left(u_k e^{k^3 t}\right)_t = 6 \sum_{i=1}^{k-1} (k-i) u_i u_{k-i}, \tag{3.4.1}$$

which generates the sequence:

$$\begin{aligned} u_1 &= c_1 e^{-t} \\ u_2 &= c_2 e^{-8t} + c_1^2 e^{-2t} \\ u_3 &= c_3 e^{-27t} + \frac{3}{4} c_1^3 e^{-3t} + c_1 c_2 e^{-9t} \\ &\cdots . \end{aligned} \tag{3.4.2}$$

Again we turn to the perturbative expansion. The linearised problem is,

$$w_{1_t} = -w_{1_{xxx}}, \tag{3.4.3}$$

which is solved by a Stokes expansion to give:

$$w_1(x,t) = \sum_{k_j} c_{k_j} e^{k_j x - k_j^3 t}. \tag{3.4.4}$$

The next level equation (in w_2) contains the nonlinear driver:

$$w_{2_t} + w_{2_{xxx}} = 6w_1w_{1_x}. \tag{3.4.5}$$

Again, a series of such calculations allows the structure of the successive terms to be inferred. The computations are automated by a program to be found in Appendix A. This time the guess is somewhat complicated. A general term in w_n appears to be:

$$\begin{aligned} &2(n_1k_1 + n_2k_2 + \ldots + n_mk_m)^2 \\ &\times \sum_{perms} \frac{1}{(k_a + k_b)(k_b + k_c)\ldots(k_z + k_a)} \prod_j \frac{1}{n_j} \left(\frac{c_{k_j}}{k_j} e^{k_j x - k_j^3 t} \right)^{n_j}. \end{aligned} \tag{3.4.6}$$

with $\sum n_j = n$ as before. The sum over "perms" means the sum over all permutations of the indices $a, b, \ldots, z$ which contain no cycles of period less than n, the only exception being $n = 2$. An example may make this clearer: the coefficient of terms like $c_{k_1}c_{k_2}c_{k_3}c_{k_4}$ contains a factor

$$2\frac{(k_1k_2 + k_1k_3 + k_1k_4 + k_2k_3 + k_2k_4 + k_3k_4)}{(k_1 + k_2)(k_1 + k_3)(k_1 + k_4)(k_2 + k_3)(k_2 + k_4)(k_3 + k_4)}, \tag{3.4.7}$$

which is written as the prescribed sum over permutations,

$$\begin{aligned} &\frac{1}{(k_1 + k_2)(k_2 + k_3)(k_3 + k_4)(k_4 + k_1)} \\ &+\frac{1}{(k_1 + k_3)(k_3 + k_4)(k_4 + k_2)(k_2 + k_1)} \\ &+\frac{1}{(k_1 + k_4)(k_4 + k_2)(k_2 + k_3)(k_3 + k_1)}. \end{aligned}$$

In general there are $\frac{1}{2}(n-1)!$ such permutations.

Protracted study of (3.4.6) may convince the reader that such terms can actually be summed into a single generating function, using a set of n dummy variables x_s:

$$w_n = \frac{2}{n}\frac{\partial^2}{\partial x^2} \int_{-\infty}^{\frac{x}{2}} \int_{-\infty}^{\frac{x}{2}} \cdots \int_{-\infty}^{\frac{x}{2}} \prod_{s=0}^{n-1} \left(\sum_{k_j} \frac{c_{k_j}}{k_j} e^{k_j(x_s + x_{s+1}) - k_j^3 t} \right) \prod dx_s. \tag{3.4.8}$$

The x_s cycle round so that $x_n \equiv x_0$ and the lower limits of integration force the k_j to have non-negative real part. It appears to be difficult to prove this result but it has been checked up to w_4. We will ultimately judge it by the results it yields.

As in the case of the Burgers equation, we now notice that the sum in (3.4.8) is closely related to the solution of the linearised problem (3.4.3). If we take any solution of (3.4.3), written $\frac{\partial}{\partial x}F(x,t)$, then (3.4.8) may be written as:

$$w_n = \frac{2}{n}\frac{\partial^2}{\partial x^2}\int_{-\infty}^{\frac{x}{2}}\int_{-\infty}^{\frac{x}{2}}\cdots\int_{-\infty}^{\frac{x}{2}}\prod_{s=0}^{n-1}F\left(\frac{x_s+x_{s+1}}{2},t\right)\prod dx_s. \tag{3.4.9}$$

The full solution is then formally:

$$u(x,t) = \sum_{n=1}^{\infty} w_n. \tag{3.4.10}$$

When this is expanded it reads

$$\begin{aligned} u(x,t) = 2\frac{\partial^2}{\partial x^2}\Bigg(&\int^x F(x_0,t)dx_0 + \frac{1}{2}\int^x\int^x F(\frac{x_0+x_1}{2},t)F\left(\frac{x_1+x_0}{2},t\right)dx_1dx_0 \\ &+ \frac{1}{3}\int^x\int^x\int^x F\left(\frac{x_0+x_1}{2},t\right)F\left(\frac{x_1+x_2}{2},t\right)F\left(\frac{x_2+x_0}{2},t\right)dx_0dx_1dx_2 \\ &+ \cdots\Bigg). \end{aligned} \tag{3.4.11}$$

Clearly, there is no easy connection with the initial value problem. However, when the general solution to the linear equation is employed, the boundary conditions need not be the initial value conditions. Hence the lack of lower limits in the integrals in (3.4.11).

The form (3.4.11) can be related to the solution of the KdV equation which is supplied by the Inverse Scattering method [26]. In that formalism, the solution is given by,

$$u(x,t) = -2\frac{\partial}{\partial x}K(x,x,t), \tag{3.4.12}$$

where $K(x,z,t)$ satisfies the Marchenko equation:

$$\Omega(x+z,t) + K(x,z,t) + \int_{-\infty}^{x} K(x,y,t)\Omega(y+z,t)dy = 0,\ x \geq z. \tag{3.4.13}$$

Here $\Omega(x,0)$ is calculated from the inverse scattering data. The solution to (3.4.13) is given formally by the so-called Neumann expansion:

$$\begin{aligned} K(x,x,t) = &-\Omega(2x,t) + \int_{-\infty}^{x}\Omega(x+y,t)\Omega(y+x,t)dy \\ &- \int_{-\infty}^{x}\int_{-\infty}^{x}\Omega(x+y_1,t)\Omega(y_1+y_2,t)\Omega(y_2+x,t)dy_1dy_2 \\ &+ \cdots. \end{aligned} \tag{3.4.14}$$

The reader will find a short description of inverse scattering theory in Appendix B.

To make the connection between the two answers, simply perform one of the differentiations in (3.4.11) and compare with (3.4.14). It is found that the two expressions are the same up to the identification $F(x,t) = -\Omega(2x,t)$.

From this theory, we can compute particular solutions. One simple example is a time independent solution, starting from $F(x,t) = 1$. The series (3.4.11) gives

$$u(x,t) = -\frac{\partial^2}{\partial x^2}\log(1-x) = \frac{1}{(1-x)^2} \tag{3.4.15}$$

which may be readily verified to solve the KdV equation.

Difficulties of integration for other forms of $F(x,t)$ have so far prevented the discovery of rational solutions like $-\frac{x}{6t}$ by this method. We might reasonably expect that they should also be encompassed by (3.4.11) since the corresponding solution $-\frac{x}{t}$ for the Burgers equation arises from a choice $\frac{1}{\sqrt{\nu t}}\exp(-\frac{x^2}{4\nu t})$ for the argument of the logarithm in (3.3.13).

The soliton solutions of (3.1.3) are easily recovered. This can be verified simply by expanding the known soliton solutions (e.g. in Hirota's bilinear form) and comparing with the straightforward Stokes expansion (3.4.2) to the required order. (See [26] for a survey of the various forms of solution.) The answer is pleasingly simple. For the N-soliton solution:

$$\begin{aligned} c_2 &= -4d_1^2, \\ c_4 &= -8d_2^2, \\ &\vdots \\ c_{2N} &= -4Nd_N^2, \end{aligned} \tag{3.4.16}$$

and

$$c_i = 0, \tag{3.4.17}$$

otherwise. The d_i are the arbitrary constants in the determinantal form of the N-

soliton solution:

$$u = -2\frac{\partial^2}{\partial x^2}\log\det\begin{pmatrix} 1+\frac{d_1^2}{2k_1}e^{2(k_1x+k_1^3t)} & \cdots & \frac{d_1d_N}{k_1+k_N}e^{(k_1+k_N)x+(k_1^3+k_N^3)t} \\ \vdots & \ddots & \vdots \\ \frac{d_1d_N}{k_1+k_N}e^{(k_1+k_N)x+(k_1^3+k_N^3)t} & \cdots & 1+\frac{d_N^2}{2k_N}e^{2(k_Nx+k_N^3)t)} \end{pmatrix} \tag{3.4.18}$$

These solutions (3.4.18) can also be found directly from the perturbative expansion (3.4.11) by some rather formal manipulations, of the kind sometimes performed in quantum field theory [32]. First, we set up some infinite–dimensional traceology. We recall the fact that, for a square matrix A,

$$\operatorname{tr} A^n = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_n} A_{i_1i_2}A_{i_2i_3}\cdots A_{i_ni_1}. \tag{3.4.19}$$

The cyclic property of the indices resembles the form of (3.4.8) and this can be exploited by formally defining a square "matrix" B such that,

$$B(s,s') = F\left(\frac{s+s'}{2},t\right), \tag{3.4.20}$$

where we require the indices s and s' labelling rows and columns to be continuous rather than discrete. The definition is still useful since the equivalent of matrix multiplication is well defined by:

$$(BC)(s,s') = \int^x F\left(\frac{s+s''}{2},t\right)G\left(\frac{s''+s'}{2},t\right)ds'', \tag{3.4.21}$$

if $C(s,s') = G\left(\frac{s+s'}{2},t\right)$. This makes it clear that the identity operator is a Dirac delta–functional, since it obeys the relations,

$$\begin{aligned}\int^x F\left(\frac{s+s''}{2},t\right)\delta(s''-s')ds'' &= F\left(\frac{s+s'}{2},t\right),\\ \int^x \delta(s-s'')\delta(s''-s')ds'' &= \delta(s-s').\end{aligned} \tag{3.4.22}$$

Addition is defined as pointwise real number addition of the associated continuous functions. All this allows definition of a trace as:

$$\operatorname{tr} B = \int^x B(s,s)ds. \tag{3.4.23}$$

Given these definitions, it is not hard to prove a result analagous to the statement (3.4.19). Then, exploiting this property and the linearity of trace, (3.4.11) can be written:

$$\begin{aligned} u(x,t) &= 2\frac{\partial^2}{\partial x^2}\ \mathrm{tr}\left(B + \frac{B^2}{2} + \frac{B^3}{3} + \ldots\right) \\ &= -2\frac{\partial^2}{\partial x^2}\,\mathrm{tr}\,\log\left(\mathbb{I} - B\right), \end{aligned} \tag{3.4.24}$$

the final summation following only if $|B| < 1$. The logarithm is defined by the series expansion. The identity (δ-functional) is denoted by $\mathbb{I}$. The penultimate step involves defining a determinant–like object by the usual formula,

$$\det(\mathbb{I} - G) = \exp\left(\,\mathrm{tr}\,\log(\mathbb{I} - G)\right), \tag{3.4.25}$$

for one of our continuous matrices G. This tells us that (3.4.24) becomes:

$$u(x,t) = -2\frac{\partial^2}{\partial x^2}\log\det\left(\mathbb{I} - B\right). \tag{3.4.26}$$

The soliton solution (3.4.18) is then only a matter of a suitable choice of F to make the determinant have the right form.

$$F(x,t) = -\sum_{i=1}^{N} d_i^2 \exp\left(2k_i x + 2k_i^3 t\right) \tag{3.4.27}$$

fits the bill.

This result echoes the work by Wadati and Sawada [30] who discovered the "trace" representation of the soliton solutions with a similar type of perturbative expansion. The formalism is somewhat different to the present discussion but is tailor–made for the requirements of the soliton solutions. Wadati and Sawada do not discuss much the application of the method to general solutions of the KdV equation, nor do they realise its relation to Stokes expansions or its more general applicability.

The full theory, including a very detailed comparison with inverse scattering techniques, appears in Rosales' work [2]. In particular, there is a simplified treatment of the perturbation expansion using the dispersion relation of the KdV and a connection to the Marchenko equation which is similar to that presented here. Additionally, the Bäcklund transformations of the KdV are discussed in the light of this method.

3.5 Conclusion

If we judge the efficacy of a method by its results, then surely this should be judged a success. For the NLW equation, it generates a straightforward power series solution dependent on the Cauchy data. The Hopf–Cole form of solution is reconstructed for the Burgers equation. For the KdV equation, the perturbative expansion reproduces and generalises the known results of both the direct and inverse scattering techniques. In [1] there are a couple of examples of elementary non-soliton solutions of the KdV equation derived by this method. In [2] the reader will find many further examples, including mKdV, sine–Gordon and nonlinear Schrödinger equations, as well as examples in higher dimensions. However, all the tractable examples appear to be equations that are already integrable (in either of the senses mentioned in the abstract), and the suspicion must be that this technique tells us nothing about equations that are nonintegrable.

Despite its obvious successes and the work in [2] there is still no real understanding of why or how this technique works, or whether it points to deeper truths about the integration of evolution equations. However, it is tempting to conjecture that the underlying reasons are to do with the Lie symmetries of these equations. There are at least two pieces of evidence that support this idea. First, as mentioned earlier, the Hopf–Cole transformation is believed to be a group–theoretic phenomenon, and it is reproduced here. Likewise, the solitons of the KdV equation are essentially a restricted class of similarity solutions which can be explained using methods similar to those outlined in Section 1.6 (although the multisoliton solutions require use of generalised symmetries).

Furthermore, if we think specifically of the technique presented here (based on (3.2.1)) it is known that Lie symmetries can explain the existence of separation ansatze for differential equations. The best known example is the wave equation whose solution can be expressed as a product of functions of one variable, thereby reducing the problem to a set of ordinary differential equations. See [11] for a discus-

sion of the solutions to the wave equation and the Hamilton–Jacobi equation. The main theorem for these applications relates the symmetry groups to a set of Killing tensors of the manifold in which the equation is defined. This set defines appropriate coordinates for separation to take place. The application of these ideas to the ansatz (3.2.1) is still unclear. For a start, there is no generally accepted definition of the term "separable". One common definition requires the existence of an ansatz for the solution which reduces the problem to a set of ODEs. It is not entirely clear whether the infinite set of ODEs which appear in this method satisfy this definition, but it would be tempting to speculate that this is so. Whether this separability can be explained by study of the symmetries or generalised symmetries of these equations remains unexplored, both here and in the work of Rosales.

Chapter 4

The Bateman and Born–Infeld Equations

4.1 Introduction

Before studying the properties of the Universal Equations, we will look in some detail at a pair of closely related equations in two dimensions whose properties will later be generalised to the multi-variable case. The first is the Bateman equation, which is nothing more than the simplest possible Universal equation [33], and the second is the Born–Infeld equation, which describes minimal surfaces in $(2+1)$-dimensional Minkowski space.

Bateman's equation arose in a 1929 paper on the hydrodynamics of rotating fluids [34] when he studied the properties of the equation:

$$\Delta_1 = \phi_x^2\phi_{tt} - 2\phi_x\phi_t\phi_{xt} + \phi_t^2\phi_{xx} = 0. \tag{4.1.1}$$

Here ϕ is a scalar field and subscripts denote partial derivatives. Many of the interesting properties of the Bateman equation are dealt with in [33], and are generalised to higher dimensions in that and subsequent papers [35, 36, 37]. We will restrict our attention to two aspects of (4.1.1):

- The equation (4.1.1) admits an infinite–dimensional Lie point–symmetry group,

which includes General Linear ($GL(2,\mathbb{R})$) and conformal transformations of the base manifold and diffeomorphisms of the scalar field.

- The Bateman equation has an infinite number of inequivalent Lagrangian formulations. Any smooth function homogeneous of degree one in the derivatives ϕ_x and ϕ_t, and with nonvanishing Hessian determinant, works as a Lagrangian for (4.1.1).

The symmetry generators of the Bateman equation were found by the standard procedure in Section 1.5 and are listed in Table 1.1.

The second assertion can actually be generalised somewhat by introducing an arbitrary dependence on ϕ into the Lagrangian. Suppose $\mathcal{L}(\phi, \phi_x, \phi_t)$ is a candidate Lagrangian. By Euler's theorem on homogeneous functions:

$$\phi_x \frac{\partial \mathcal{L}}{\partial \phi_x} + \phi_t \frac{\partial \mathcal{L}}{\partial \phi_t} = \mathcal{L}. \tag{4.1.2}$$

It follows easily that:

$$\begin{aligned} \phi_x \frac{\partial^2 \mathcal{L}}{\partial \phi_x^2} + \phi_t \frac{\partial^2 \mathcal{L}}{\partial \phi_x \partial \phi_t} &= 0, \\ \phi_x \frac{\partial^2 \mathcal{L}}{\partial \phi_x \partial \phi_t} + \phi_t \frac{\partial^2 \mathcal{L}}{\partial \phi_t^2} &= 0 \\ \phi_x \frac{\partial^2 \mathcal{L}}{\partial \phi \partial \phi_x} + \phi_t \frac{\partial^2 \mathcal{L}}{\partial \phi \partial \phi_t} &= \frac{\partial \mathcal{L}}{\partial \phi}. \end{aligned} \tag{4.1.3}$$

Now, if we expand the Euler–Lagrange expression,

$$\mathcal{EL} = \frac{\partial \mathcal{L}}{\partial \phi} - D_x\left(\frac{\partial \mathcal{L}}{\partial \phi_x}\right) - D_t\left(\frac{\partial \mathcal{L}}{\partial \phi_t}\right), \tag{4.1.4}$$

and use the homogeneity properties (4.1.3) the equation of motion becomes:

$$\mathcal{EL} = \frac{1}{\phi_x \phi_t} \frac{\partial^2 \mathcal{L}}{\partial \phi_x \partial \phi_t} \left(\phi_x^2 \phi_{tt} - 2\phi_x \phi_t \phi_{xt} + \phi_t^2 \phi_{xx}\right). \tag{4.1.5}$$

This is equivalent to (4.1.1) on extremals of $\mathcal{L}$. The coefficient multiplying the Bateman expression in (4.1.5) is a *variational integrating factor*, which is in agreement with the analysis of Anderson and Duchamp [38]. Notice that the condition on the

Hessian is replaced by the requirement that the second derivative of $\mathcal{L}$ appearing in (4.1.5) is nonzero.

Notice that if the Lagrangian is independent of ϕ, the Euler–Lagrange equation is always in the form of a conservation law, and that there are an infinite number of such laws. The question of how to construct Hamiltonian conservation laws that are properly in involution will be addressed in this chapter.

The Born–Infeld equation is a simplified, $(1+1)$-dimensional version of an equally venerable theory [39]. In physical coordinates it is [40],

$$\phi_x^2\phi_{tt} - 2\phi_x\phi_t\phi_{xt} + \phi_t^2\phi_{xx} = \phi_{tt} - \phi_{xx}. \tag{4.1.6}$$

We will mainly use the light–cone version:

$$\Delta_2 = \phi_x^2\phi_{tt} - (1 + 2\phi_x\phi_t)\,\phi_{xt} - \phi_t^2\phi_{xx} = 0. \tag{4.1.7}$$

The point symmetries of this equation are much more restricted — there is no infinite–dimensional component, and the symmetry generators just span the Poincaré symmetry of the aforementioned Minkowski space. Furthermore, (4.1.7) is derived from an essentially unique Lagrangian:

$$\mathcal{L}_{BI} = \sqrt{1 + 4\phi_x\phi_t}. \tag{4.1.8}$$

The two equations look very similar, and if we notice that (4.1.7) can be rescaled to give,

$$\Delta_2' = \phi_x^2\phi_{tt} - (\lambda + 2\phi_x\phi_t)\,\phi_{xt} - \phi_t^2\phi_{xx} = 0, \tag{4.1.9}$$

and set $\lambda = 0$, we recover (4.1.1). As we might expect, (4.1.1) and (4.1.7) share many properties.

In particular, they can both be solved using linearising transformations. For example, in common with other quasilinear second order equations [41], the Bateman equation can be linearised by a Legendre transform [37]. With the change of variables,

$$\begin{aligned} \xi &= \phi_x \\ \eta &= \phi_t \\ \omega(\xi,\eta) &= x\xi + y\eta - \frac{1}{2}\phi, \end{aligned} \tag{4.1.10}$$

(4.1.1) is equivalent to:

$$\xi^2\omega_{\xi\xi} + 2\xi\eta\omega_{\xi\eta} + \eta^2\omega_{\eta\eta} = 0. \tag{4.1.11}$$

This linearised equation supplies an implicit solution for (4.1.1) [37]. Any function homogeneous of degree zero or one in ξ, η is a solution of (4.1.11), so we write:

$$\omega = f\left(\frac{\xi}{\eta}\right) + (\xi + \eta)g\left(\frac{\xi}{\eta}\right), \tag{4.1.12}$$

where f and g are arbitrary functions. Differentiating this with respect to ξ and η yields the pair of equations:

$$\begin{aligned}\omega_\xi &= x = \frac{1}{\eta}(f' + (\xi + \eta)g') + g,\\ \omega_\eta &= t = -\frac{\xi}{\eta^2}(f' + (\xi + \eta)g') + g.\end{aligned} \tag{4.1.13}$$

From this we deduce that,

$$\omega + \phi = x\xi + t\eta = (\xi + \eta)g, \tag{4.1.14}$$

whence,

$$\phi = -f\left(\frac{\xi}{\eta}\right). \tag{4.1.15}$$

A particularly neat method of solution is to notice that (4.1.1) follows from the pair of evolution equations:

$$\begin{aligned}\phi_t &= \psi\phi_x\\ \psi_t &= \psi\psi_x.\end{aligned} \tag{4.1.16}$$

Note that the second of these is just the NLW equation. We can adapt the Taylor expansion technique outlined in Section 3.2 to this case. From the pair (4.1.16), we can prove that

$$\frac{\partial^n}{\partial t^n}\phi = \frac{\partial^{n-1}}{\partial x^{n-1}}\left(\psi^n\phi_x\right). \tag{4.1.17}$$

Now, using the initial data $\phi(x,0)$ and the first derivatives, the solution at a later time is given by a Taylor expansion:

$$\begin{aligned}\phi(x,t) &= \sum_{k=0}^{\infty}\frac{t^k}{k!}\frac{\partial^k}{\partial t^k}\phi(x,t)\Big|_{t=0}\\ &= \phi(x,0) + \sum_{k=1}^{\infty}\frac{t^k}{k!}\frac{\partial^{k-1}}{\partial x^{k-1}}\left(\psi(x,t)^k\phi_x(x,t)\right)\Big|_{t=0}.\end{aligned} \tag{4.1.18}$$

The last step uses (4.1.17). This is an explicit, calculable, power series solution to the Cauchy initial value problem for (4.1.1).

The Born-Infeld equation can also be linearised by the Legendre transform. The basic method can be found in [41] applied to the Euclidean equation of minimal surfaces. Alternatively, in [3] it is shown that (4.1.9) can be written as a first order equation in a similar manner to the Bateman equation with the help of the two independent roots u_1, u_2 of the quadratic equation for the characteristics [25]:

$$\phi_x^2 u^2 - (\lambda + 2\phi_x\phi_t)\, u + \phi_t^2 = 0. \tag{4.1.19}$$

The roots of this equation are

$$u_{\substack{1\\2}} = \frac{\lambda + 2\phi_x\phi_t \pm \sqrt{\lambda^2 + 4\lambda\phi_x\phi_t}}{2\left(\phi_x\right)^2} \tag{4.1.20}$$

Equation (4.1.9) can then be written in either of two forms,

$$\begin{aligned} \frac{\partial u_1}{\partial t} &= u_2 \frac{\partial u_1}{\partial x}, \\ \frac{\partial u_2}{\partial t} &= u_1 \frac{\partial u_2}{\partial x}. \end{aligned} \tag{4.1.21}$$

These equations possess an infinite number of conservation laws. It is straightforward to verify that:

$$\begin{aligned} \frac{\partial}{\partial t}\left(u_1 + u_2\right) &= \frac{\partial}{\partial x}\left(u_1 u_2\right); \\ \frac{\partial}{\partial t}\left(u_1^2 + u_1 u_2 + u_2^2\right) &= \frac{\partial}{\partial x}\left(u_1 u_2 \left(u_1 + u_2\right)\right); \\ \frac{\partial}{\partial t}\left(u_1^3 + u_1^2 u_2 + u_1 u_2^2 + u_2^3\right) &= \frac{\partial}{\partial x}\left(u_1 u_2 \left(u_1^2 + u_1 u_2 + u_2^2\right)\right) \end{aligned} \tag{4.1.22}$$

The general conservation law is,

$$\frac{\partial}{\partial t} S_k = \frac{\partial}{\partial x}\left(u_1 u_2 S_{k-1}\right), \tag{4.1.23}$$

where S_k denotes the symmetric polynomial of kth degree in u_1 and u_2. The proof is an uncomplicated induction.

The system (4.1.21) can then be linearised by a hodograph transformation to provide an implicit solution; the roles of dependent and independent variables are interchanged to give,

$$\begin{aligned} \frac{\partial x}{\partial u_2} &= -u_2 \frac{\partial t}{\partial u_2}, \\ \frac{\partial x}{\partial u_1} &= -u_1 \frac{\partial t}{\partial u_1}, \end{aligned} \tag{4.1.24}$$

with the solution,

$$\begin{aligned} x &= f(u_1) - u_1 f'(u_1) + g(u_2) - u_2 g'(u_2), \\ t &= f'(u_1) + g'(u_2), \end{aligned} \tag{4.1.25}$$

where f and g are arbitrary functions and a prime denotes differentiation with respect to the argument. Note that this is still some way off a solution for ϕ, which requires a solution of the above equations for ϕ_x and ϕ_t which may then in principle be integrated.

The complete solution may be obtained in principle by the inversion of the equations for x, t in terms of u_1, u_2, and the subsequent integration of the equations:

$$\begin{aligned} \frac{\partial \phi}{\partial x} &= \frac{1}{\sqrt{u_1} - \sqrt{u_2}}; \\ \frac{\partial \phi}{\partial t} &= \frac{\sqrt{u_1 u_2}}{\sqrt{u_1} - \sqrt{u_2}}. \end{aligned} \tag{4.1.26}$$

The consistency condition which guarantees integrability of these equations is simply the Born–Infeld equation itself. So specific choices of f and g will furnish explicit solutions of (4.1.7).

We can demonstrate a simple example of a solution generated by the results (4.1.25) and (4.1.26). A particularly convenient choice of f and g in the system (4.1.25) is $f(u_1) = u_1^2$ and $g(u_2) = -u_2^2$. This gives explicit expressions for u_1 and u_2 in terms of x and t:

$$\begin{aligned} u_1 &= -\frac{(4x + t^2)}{4t} \\ u_2 &= -\frac{(4x - t^2)}{4t}. \end{aligned} \tag{4.1.27}$$

These can be substituted back into the formulae (4.1.26) which can then be integrated to give an explicit solution:

$$\phi = -\frac{\sqrt{2}}{3}((\frac{4x-t^2}{-2t})^{\frac{3}{2}} + (\frac{4x+t^2}{-2t})^{\frac{3}{2}}) + \text{constant}. \tag{4.1.28}$$

In this chapter, we will concentrate on two further aspects of these equations, namely their generalised symmetries and the Hamiltonian structures associated with each. The aim will be to investigate two possible definitions of their integrability: linearisability and the existence of biHamiltonian structure and commuting symmetries. The generalised symmetries are calculated in the next section, and it is explained how these lead to linearisability of both equations. The Hamiltonian structures and conserved quantities of the Bateman equation are calculated in Section 4.3 in a formalism already applied to the Born–Infeld case [22]. Much of this material appears in [4].

4.2 Generalised Symmetries and Linearisability

To find the generalised (or "Lie–Bäcklund") symmetries of either of our two equations, we use the standard algorithm described in Section 1.7. Olver [7] and Stephani [11] provide further information. The calculation is essentially the same in both cases so we will restrict our attention to (4.1.9) and view the Bateman as the $\lambda = 0$ limit.

To begin with, we propose a symmetry generator $\mathbf{v}_Q$ in evolutionary form,

$$\mathbf{v}_Q = Q\frac{\partial}{\partial\phi}, \tag{4.2.1}$$

where the characteristic Q can depend on x, t, ϕ, and the derivatives of ϕ up to some specified order. Denoting the equation of interest by Δ, the symmetry condition is written:

$$\text{pr}\,\mathbf{v}_Q(\Delta) = 0, \tag{4.2.2}$$

on solutions of $\Delta = 0$, or more explicitly,

$$\frac{\partial\Delta}{\partial\phi_x}(D_xQ[\phi]) + \frac{\partial\Delta}{\partial\phi_t}(D_tQ[\phi]) + \frac{\partial\Delta}{\partial\phi_{xx}}(D_{xx}Q[\phi]) + \frac{\partial\Delta}{\partial\phi_{xt}}(D_{xt}Q[\phi]) + \frac{\partial\Delta}{\partial\phi_{tt}}(D_{tt}Q[\phi]) = 0. \tag{4.2.3}$$

The easiest choice of characteristic is one that depends on derivatives up to first order. Then (4.2.3) involves at most third order derivatives. By successively comparing coefficients of these derivatives, we find a set of conditions on $Q[\phi]$. The following results were found with the help of MAPLE.

First, the coefficients of the third order derivatives vanish on solutions of (4.1.1). To dispose of the coefficients of the quadratic second order derivative terms, Q must satisfy,

$$\phi_x^2 \frac{\partial^2 Q}{\partial \phi_x^2} + (\lambda + 2\phi_x \phi_t) \frac{\partial^2 Q}{\partial \phi_x \partial \phi_t} + \phi_t^2 \frac{\partial^2 Q}{\partial \phi_t^2} = 0. \tag{4.2.4}$$

This equation is essentially the same as the linearised form of the Born–Infeld equation and as such its solution is obtained implicitly by the method of characteristics [41, 25]. Its solution in general is not important — we only need to know that Q depends on a solution to a linear equation. However in the Bateman case ($\lambda = 0$), the resulting differential operator acting on Q factorises, and we find that any Q homogeneous of degree zero or one in the first derivatives of ϕ satisfies (4.2.4).

So the first order generalised symmetries of both equations depend on solutions to a linear differential equation. According to a theorem of Kumei and Bluman [31], this condition guarantees that there exists a transformation bijectively mapping (4.1.9) to a linear expression. The basic idea is that all linear equations admit a *superposition generator* dependent on an arbitrary solution of the equation and the linearisation finds coordinates in which the symmetry satisfying (4.2.4) acts as this superposition for the resultant linear problem.

The basic Kumei–Bluman theorem says that there is a one–to–one contact mapping,

$$x^i \to X^i = X^i[\phi]$$
$$\phi \to \omega = \omega[\phi],$$

from a given nonlinear differential equation $\Delta[\phi] = 0$ to a linear equation, if and only if Δ admits a symmetry generator of the form,

$$\mathbf{v}_Q = \sigma[\phi] Q(X[\phi]) \partial_\phi, \tag{4.2.5}$$

where $Q(X[\phi])$ is the solution to a linear differential equation, such as (4.2.4). (The notation $[\phi]$ indicates dependence on x, ϕ and first derivatives of ϕ.) This generator acts as the superposition generator for the associated linear equation.

The associated conditions on the new variables X^i and ω are written in terms of the *Lagrange bracket* defined for two functions $A[\phi]$ and $B[\phi]$ by,

$$[A,B]_L = (\partial_{\phi_i} A)(\partial_{x_i} B + \phi_i \partial_\phi B) - (\partial_{\phi_i} B)(\partial_{x_i} A + \phi_i \partial_\phi A). \tag{4.2.6}$$

We require that,

$$\begin{aligned} \left[X^i, X^j\right]_L &= 0 \\ \left[X^i, \omega\right]_L &= 0, \end{aligned} \tag{4.2.7}$$

for the transformation to satisfy the necessary contact conditions.

If there exists such a generator, then there exists a contact transformation bijectively mapping the p.d.e. to a linear equation. Comparing (4.2.1) with (4.2.5), we find that we can take $\sigma = k$ (a constant) and we can identify $X^i = \phi_i$. On imposing (4.2.7) we find that the X^i commute as required and that the second condition is satisfied by a transformation of the form (4.1.10).

Returning to the symmetry calculation, it is difficult to proceed to calculate further general constraints on Q without a convenient general solution for (4.2.4), which is awkward in the case $\lambda \neq 0$. However, for the Bateman equation, if we decide that Q is homogeneous of degree one in ϕ_x and ϕ_t then the coefficients of those terms linear in the second derivatives vanish identically on solutions. On the other hand, if the degree of homogeneity is zero, Q must obey:

$$\phi_x \frac{\partial Q}{\partial t} - \phi_t \frac{\partial Q}{\partial x} = 0. \tag{4.2.8}$$

This is identically solved by,

$$Q = F\left(x + t\frac{\phi_t}{\phi_x}\right), \tag{4.2.9}$$

where F is smooth and real valued. Furthermore, if Q is of the form $Q = F(\frac{\phi_t}{\phi_x})$, (4.2.8) becomes the Bateman equation, and thus disappears on solutions. Note that

it is automatically homogeneous of degree zero in ϕ_x and ϕ_t and so respects the conditions already derived. An arbitrary dependence on ϕ can also be incorporated into the solution.

The terms independent of second and third derivatives are also dealt with using the characteristic (4.2.9). Once again, an arbitrary dependence on ϕ is allowed. This "diffeomorphism invariance" of the Bateman equation has already been found among the point symmetries of the equation mentioned earlier.

It is possible to find simple examples of invariant solutions of the Bateman equation using the results above and the techniques of Section 1.6. Considering the case when Q is homogeneous of degree one in first derivatives, let us look at the particular example $Q_1 = x\phi_x + t\phi_t$. This is the evolutionary equivalent of the point symmetry,

$$\mathbf{v} = x\partial_x + t\partial_t, \tag{4.2.10}$$

which is a linear combination of two of the symmetries in Table 1.1. It is easy to verify that any function of x/t is invariant under this transformation. Substituting into the Bateman equation, the resulting reduced equation vanishes identically. Thus, any function ϕ dependent on x/t satisfies (4.1.1).

A symmetry of the type (4.2.9) is rather more awkward to deal with, as it is not the evolutionary form of any point symmetry. However, if we look at the simplest example, $Q_2 = x + t\frac{\phi_t}{\phi_x}$, the action of this symmetry reduces to that of Q_1 above, and so the reduction is the same.

Although the diffeomorphism symmetry is not variational (in the sense defined in (2.3.6)), it is instructive to study its effect on a Lagrangian $\mathcal{L}(\phi_x, \phi_t)$ of the Bateman equation. Writing the infinitesimal diffeomorphism generator as,

$$\mathbf{v}_D = f(\phi)\partial_\phi, \tag{4.2.11}$$

where f is some arbitrary smooth function of ϕ, we have:

$$\mathrm{pr}\,\mathbf{v}_D(\mathcal{L}) = f'(\phi)\mathcal{L} = \mathcal{L}'. \tag{4.2.12}$$

It is then easy to show that the associated Euler–Lagrange equation is,

$$\mathcal{E}(\mathcal{L}') = -f'(\phi)\mathcal{E}(\mathcal{L}) = 0. \tag{4.2.13}$$

If $f'(\phi)$ is nonzero, it acts as a variational integrating factor of the type mentioned in the previous section, and (4.2.13) reduces to the Bateman equation.

To summarise, the main lesson of this section is that the linearisability of both the Bateman and Born–Infeld equations is explained by the existence of a first order generalised symmetry whose characteristic solves (4.2.4). We will aim to generalise this result to arbitrary dimensions in the following chapter. In addition, we have explicitly derived two possible forms for the characteristic of a first order generalised symmetry of the Bateman equation: any function homogeneous of degree one in first derivatives of ϕ will do; otherwise any function of the form (4.2.9) can be used.

4.3 Hamiltonian Analysis

The Hamiltonian and multi-Hamiltonian descriptions of the Born–Infeld equation have already been studied by Nutku and various collaborators [42, 22] and the aim here is to perform a similar analysis of the Bateman equation. We need to find an evolution equation (or system of evolution equations) in the form (2.5.3), with a suitable Hamiltonian and Poisson operator. We require that this system is equivalent to (4.1.1) under some suitable change of variables.

Ideally, we would like to be able to find two such structures for our evolution equations. By Magri's theorem [21], this would demonstrate the complete integrability of the equation and allow construction of the Lax pair representation and the associated conservation laws, provided some compatibility criteria were met.

The results of [42, 22] can be adapted to the present case simply by using the $\lambda = 0$ limit of the Born–Infeld Lagrangian (4.1.8). This would essentially repeat the work in [22] but, as we have seen, the set of admissible Lagrangians for (4.1.1) is far larger than this. We can discuss Hamiltonian formulations corresponding to a

large class of these non-standard Bateman Lagrangians and in general we will find behaviour quite different from the Born–Infeld case.

As an alternative, we might consider the pair (4.1.16) as a candidate for our Hamiltonian description, but if we view it as a system of evolution equations, then there is no viable Hamiltonian structure which works. (The Jacobi identity fails to close for any first–order skew operator which reproduces (4.1.16)). The second of the equations is Hamiltonian in its own right, but in order to adopt it we would have to use its partner as the relevant change of variables. Motivated as we are by the higher dimensional equations, it is hard to see how this would this would generalise to more than two independent variables.

Another route might be to use the Cauchy–Kowalevsky form of the equation of motion. Choosing t as the evolution parameter, this is:

$$\phi_{tt} = 2\left(\frac{\phi_t}{\phi_x}\right)\phi_{xt} - \left(\frac{\phi_t}{\phi_x}\right)^2 \phi_{xx}. \tag{4.3.1}$$

Now, changing variables so that,

$$\begin{aligned} r &= \phi_t \\ s &= \phi_x, \end{aligned} \tag{4.3.2}$$

we get the evolution equations:

$$\begin{aligned} r_t &= D_x\left(\frac{r^2}{s}\right) \\ s_t &= r_x. \end{aligned} \tag{4.3.3}$$

Again these fail to be Hamiltonian. Once more, the Jacobi identity fails to close for any suitable operator and candidate Hamiltonian which reproduce (4.3.3).

The most direct approach is suggested by the fact mentioned earlier, that if the Bateman Lagrangian has no ϕ dependence, (4.1.1) can be expressed as a conservation law:

$$D_x\left(\frac{\partial \mathcal{L}}{\partial \phi_x}\right) + D_t\left(\frac{\partial \mathcal{L}}{\partial \phi_t}\right) = 0. \tag{4.3.4}$$

The biHamiltonian structure of equations expressible in the form of a conservation law was demonstrated by Nutku [43]. As mentioned above, there are an infinite number of ways of expressing the Bateman equation as a divergence–free current, and there is a biHamiltonian structure associated with each.

Following [43], the divergence (4.3.4) may be rewritten as the smoothness condition for a function ψ.

$$\begin{aligned} \psi_x &= -\frac{\partial \mathcal{L}}{\partial \phi_t} \\ \psi_t &= \frac{\partial \mathcal{L}}{\partial \phi_x}. \end{aligned} \tag{4.3.5}$$

In accordance with the standard Hamiltonian technique, the field ϕ is mapped to a pair of variables u, v, where

$$\begin{aligned} u &= \phi_x, \\ v &= \psi_x. \end{aligned} \tag{4.3.6}$$

The t–derivatives can then be expressed in terms of these new variables:

$$\begin{aligned} \phi_t &= U(u,v), \\ \psi_t &= V(u,v). \end{aligned} \tag{4.3.7}$$

Consequently, the equation of motion and the smoothness condition for ϕ can be summarised by the condition of closure for the pair of exact 1-forms,

$$\begin{aligned} d\phi &= u\ dx + U\ dt, \\ d\psi &= v\ dx + V\ dt, \end{aligned} \tag{4.3.8}$$

namely that:

$$\begin{aligned} u_t - D_x(U) &= 0 \\ v_t - D_x(V) &= 0. \end{aligned} \tag{4.3.9}$$

Given our assumptions about the nature of the Lagrangian, we can deduce the form of U and V. From (4.1.3), we know that the derivatives of $\mathcal{L}$ with respect to ϕ_x

and ϕ_t are homogeneous of degree zero. So, let us write (4.3.5) as,

$$\begin{aligned} \psi_x &= F\left(\frac{\phi_t}{\phi_x}\right), \\ \psi_t &= G\left(\frac{\phi_t}{\phi_x}\right), \end{aligned} \tag{4.3.10}$$

where F and G are smooth, invertible functions. Solving the first of these relations for ϕ_t tells us that,

$$U(u,v) = u\ F^{-1}(v), \tag{4.3.11}$$

and substitution into the second,

$$V(u,v) = G\left(F^{-1}(v)\right). \tag{4.3.12}$$

At this point, we attempt to impose the Hamiltonian structure. A nice result about this type of equation [43] is that it admits the canonical, "flat" Poisson structure defined by the structure matrix,

$$\mathcal{J}_1 = \begin{pmatrix} 0 & D_x \\ D_x & 0 \end{pmatrix}, \tag{4.3.13}$$

and a Hamiltonian $\mathcal{H}_1$ dependent only on u and v, subject to the conditions,

$$\frac{\partial \mathcal{H}_1}{\partial u} = -V, \ \frac{\partial \mathcal{H}_1}{\partial v} = -U, \tag{4.3.14}$$

which in turn impose compatibility criteria on U and V.

In this example, the compatibility requires that:

$$G'\left(F^{-1}(v)\right)\frac{\partial}{\partial v}\left(F^{-1}(v)\right) = F^{-1}(v). \tag{4.3.15}$$

Bearing this in mind, we end up with a general Hamiltonian framework for the Bateman equation. The Hamilton equations are easily seen to reduce to the form,

$$\begin{aligned} u_t &= D_x\left(uF^{-1}(v)\right), \\ v_t &= F^{-1}(v)\ v_x. \end{aligned} \tag{4.3.16}$$

The structure matrix is $\mathcal{J}_1$ and the Hamiltonian is:

$$\mathcal{H}_1 = u\ G\left(F^{-1}(v)\right). \tag{4.3.17}$$

There is also a biHamiltonian structure associated with these equations. For this "conserved current" type of system there is always a conserved quantity of the form [43],

$$\mathcal{H}_0 = uv. \tag{4.3.18}$$

Following the standard program for biHamiltonian systems [7], we postulate this as a conserved Hamiltonian which reproduces the equations of motion when used with a second structure matrix $\mathcal{J}_2$:

$$\mathcal{J}_1 \delta\mathcal{H}_1 = \mathcal{J}_2 \delta\mathcal{H}_0. \tag{4.3.19}$$

Provided this second structure is "compatible" with the first, meaning that any linear combination of the two is also a Poisson structure, this construction will give rise to a (hopefully infinite) sequence of conserved quantities $\mathcal{H}_k$ such that,

$$\mathcal{J}_1 \delta\mathcal{H}_k = \mathcal{J}_2 \delta\mathcal{H}_{k-1}, \tag{4.3.20}$$

which can be generated by the recursion operator:

$$\mathcal{R} = \mathcal{J}_2 \mathcal{J}_1^{-1}. \tag{4.3.21}$$

We know that the general form of $\mathcal{J}_2$, as the most general first order, skew–adjoint, matrix differential operator is,

$$\mathcal{J}_2 = \begin{pmatrix} 2mD_x + m_x & 2pD_x + (p+q)_x \\ 2pD_x + (p-q)_x & 2nD_x + n_x \end{pmatrix}, \tag{4.3.22}$$

where m, n, p, q are functions of u and v. By using the condition (4.3.19) and demanding that it reproduces (4.3.16), restrictions on the form of $\mathcal{J}_2$ can be found. Nutku [43] proves that these restrictions also ensure the closure of the Jacobi identity. The reader will find REDUCE code in Appendix C to assist in verification of this fact.

It is difficult to write down a convenient general form for the most general structure so derived, but a useful simplifying assumption is to take $n = 0$. Then we find a possible candidate for the second structure given by,

$$m = \frac{k}{v^2},$$

$$
\begin{aligned}
n &= 0, \\
p &= \frac{1}{2}F^{-1}(v), \\
q &= \frac{1}{2}F^{-1}(v) + k',
\end{aligned}
\tag{4.3.23}
$$

where k and k' are constants. It is straightforward to ascertain that this structure is compatible with the $\mathcal{J}_1$: the combination $\mathcal{J}_1 + \mathcal{J}_2$ is skew adjoint and satisfies the Jacobi identity. So $\mathcal{J}_2$ can be used to find a second conserved Hamiltonian and thence a whole hierarchy of conserved quantities.

The associated recursion operator is the pseudo-differential operator,

$$
\mathcal{R} = \mathcal{J}_2\mathcal{J}_1^{-1} = \begin{pmatrix} F^{-1}(v) + \frac{\partial}{\partial v}(F^{-1})\, v_x\, D_x^{-1} & 2\frac{k}{v^2} - 2\frac{k}{v^3}v_x\, D_x^{-1} \\ 0 & F^{-1}(v) \end{pmatrix}, \tag{4.3.24}
$$

using the values of p, q, and m in (4.3.23).

We recall from Chapter 1 that a system of evolution equations of the form (1.7.11) which admits a recursion operator has a Lax representation:

$$
L_t = [\mathcal{A}, L]\,. \tag{4.3.25}
$$

The L is the recursion operator (4.3.24) and $\mathcal{A}$ is the Fréchet derivative:

$$
\mathcal{A} = \mathsf{D}_{\mathbf{K}} = \left(\frac{d}{d\varepsilon}\mathbf{K}[\mathbf{u}_\varepsilon]\right)\Big|_{\varepsilon=0}, \tag{4.3.26}
$$

which, in our example, is:

$$
\mathcal{A} = \begin{pmatrix} D_x\ (F^{-1}(v)\ .\ ) & D_x\ \left(u\frac{\partial}{\partial v}F^{-1}(v)\ .\ \right) \\ 0 & D_x\ (F^{-1}(v)\ .\ ) \end{pmatrix}. \tag{4.3.27}
$$

The argument of the $\mathcal{A}$ should be inserted as indicated by the dots. Substitution of (4.3.27) and (4.3.24) into (4.3.25) reproduces the Hamilton equations (4.3.16).

Now, using the recurrence (4.3.20), we can start to calculate conserved quantities. The first conserved Hamiltonian is found by considering:

$$
\mathcal{J}_2\delta\mathcal{H}_1 = \mathcal{J}_1\delta\mathcal{H}_2. \tag{4.3.28}
$$

It turns out to be:

$$
\mathcal{H}_2 = \int^v \left\{ u\left(F^{-1}(w)\right)^2 + \frac{2k}{w^2}G(F^{-1}(w)) \right\} dw. \tag{4.3.29}
$$

Continuing the process, we find the third and fourth Hamiltonians:

$$\mathcal{H}_3 = \int^v \left\{ u(F^{-1}(w))^3 + \frac{2k}{w^2} F^{-1}(w) G(F^{-1}(w)) + \frac{2k}{w^2} \int^w (F^{-1}(w'))^2 dw' \right\} dw;$$

$$\mathcal{H}_4 = \int^v \left\{ u(F^{-1}(w))^4 + \frac{2k}{w^2} F^{-1}(w) \int^w (F^{-1}(w'))^2 dw' \right.$$

$$\left. + \frac{2k}{w^2} (F^{-1}(w))^2 G(F^{-1}(w)) + \frac{2k}{w^2} \int^w (F^{-1}(w'))^3 dw' \right\} dw. \tag{4.3.30}$$

The subsequent members of the hierarchy become rapidly more complicated, but it is not hard to deduce from the emergent pattern that the infinite sequence of conserved Hamiltonians can be generated from:

$$\mathcal{H}_k = \int^v \left\{ \frac{2k}{w^2} \int^w (F^{-1}(w'))^{k-1} dw' + F^{-1}(w) \delta_v \mathcal{H}_{k-1} \right\} dw. \tag{4.3.31}$$

It is worth noting that the conserved-current type of equation found by Nutku admits a third independent Hamiltonian structure, compatible with the first two, with $\mathcal{H}_2$ as the relevant Hamiltonian density. Also, while the sequence of conserved quantities in principle continues *ad infinitum*, there may be certain particular cases when the sequence terminates or repeats itself after a finite number of steps. See [22] for an example of this behaviour arising in the biHamiltonian hierarchies of the Born–Infeld equation.

As a particular example of this construction, let us examine the case where the Lagrangian corresponds to the $\lambda = 0$ limit of the Born–Infeld Lagrangian:

$$\mathcal{L} = \sqrt{\phi_x^2 - \phi_t^2}, \tag{4.3.32}$$

leading to the Euler-Lagrange equation:

$$\frac{\partial}{\partial x} \left(\frac{\phi_x}{\sqrt{\phi_x^2 - \phi_t^2}} \right) - \frac{\partial}{\partial t} \left(\frac{\phi_t}{\sqrt{\phi_x^2 - \phi_t^2}} \right) = 0. \tag{4.3.33}$$

Following the prescription described in Equations (4.3.6...4.3.9), the correct variables to use are

$$u = \phi_x \text{ and } v = \frac{\phi_t}{\sqrt{\phi_x^2 - \phi_t^2}}.$$

In these variables (4.1.1) becomes:

$$u_t = \frac{\partial}{\partial x}\left(\frac{uv}{\sqrt{1+v^2}}\right)$$
$$v_t = \frac{\partial}{\partial x}\left(\sqrt{1+v^2}\right). \tag{4.3.34}$$

The first equation expresses the smoothness condition for ϕ, the second contains the dynamics of ϕ.

Now we express (4.3.34) as a biHamiltonian system. The Hamiltonian density for the first structure is straightforward:

$$H_1(\mathbf{u}) = u\sqrt{1+v^2}, \tag{4.3.35}$$

and the Poisson operator is $\mathcal{J}_1$, the canonical one for infinite-dimensional systems.

Now, as above we have to postulate an ansatz for the skew-adjoint operator $\mathcal{J}_1$ and then find when the ansatz reproduces (4.3.34) and satisfies the Jacobi identity. The interesting aspect of the example we have chosen is that this calculation has already been done for the space–time version of the Born–Infeld equation (4.1.6) [42, 22]. The similarity of (4.1.1) and (4.1.6) means that the argument in [22] is readily adaptable to the present problem. Note that (4.1.6) can be rescaled to give,

$$\phi_x^2\phi_{tt} - 2\phi_x\phi_t\phi_{xt} + \phi_t^2\phi_{xx} = \lambda(\phi_{tt} - \phi_{xx}). \tag{4.3.36}$$

So we can use the results of [22] directly and set $\lambda = 0$. Then we can confirm that the result has all the right properties. Rather than use the second Hamiltonian structure discussed above, let us use the third Hamiltonian structure found in [22], for no reason other than that the verification of the Jacobi identity is marginally less laborious.

The result is the following. Apart from the Hamiltonian structure defined by (4.3.35) and (4.3.13), the evolution equations (4.3.34) admit the additional Hamiltonian structure defined by the Hamiltonian density

$$H_1(\mathbf{u}) = \frac{uv^2}{2\sqrt{1+v^2}}, \tag{4.3.37}$$

and the operator

$$\mathcal{J}^{BI} = \begin{pmatrix} -2uvD_x - (uv)_x & (1+v^2)D_x + 3vv_x \\ (1+v^2)D_x - vv_x & 0 \end{pmatrix}. \tag{4.3.38}$$

It is straightforward to prove that the above operator and density reproduce (4.3.34) when inserted in (2.5.3). To prove that they are Hamiltonian requires proving skew-adjointness of $\mathcal{J}^{BI}$, and the Jacobi identity for the bracket. To evaluate $\mathcal{J}^{BI*}$, we can evaluate the adjoint of each entry in $\mathcal{J}^{BI}$ in the sense of (2.5.4) and then transpose the matrix. This is straightforward, and confirms the skew-adjointness of $\mathcal{J}^{BI}$. To verify the Jacobi identity, the most convenient formalism is that using functional multivectors as described in Section 2.5 and Olver [7].

Consider a pair of univectors $\theta = (\rho, \sigma)^T$, as constructed in Chapter 2, in which ρ and σ are functional 1-forms. Then the Jacobi identity for the operator $\mathcal{J}^{BI}$ corresponds to the equation:

$$\mathrm{pr}\,\mathbf{v}_{\mathcal{J}^{BI}\theta}(\Phi) = 0, \tag{4.3.39}$$

where

$$\Phi = \frac{1}{2}\int (\theta \wedge \mathcal{J}^{BI}\theta)dx, \tag{4.3.40}$$

and

$$\mathrm{pr}\,\mathbf{v}_{\mathcal{J}^{BI}\theta} = \sum_{\alpha,k} D_k((\mathcal{J}^{BI}\theta)^\alpha)\frac{\partial}{\partial u^\alpha_k} \tag{4.3.41}$$

is the prolongation of the vector field corresponding to $\mathcal{J}^{BI}\theta$.

To show (4.3.39), we first expand (4.3.40):

$$\begin{aligned}\Phi &= \frac{1}{2}\int (\rho,\sigma)\begin{pmatrix} -2uvD_x - (uv)_x & (1+v^2)D_x + 3vv_x \\ (1+v^2)D_x - vv_x & 0 \end{pmatrix}\begin{pmatrix}\rho\\ \sigma\end{pmatrix} \\ &= \frac{1}{2}\int -2uv(\rho\wedge\rho_x) + (1+v^2)(\rho\wedge\sigma_x) + 4vv_x(\rho\wedge\sigma) \\ &\quad + (1+v^2)(\sigma\wedge\rho_x)dx. \end{aligned} \tag{4.3.42}$$

Now evaluate (4.3.39):

$$\begin{aligned}&\mathrm{pr}\,\mathbf{v}_{\mathcal{J}^{BI}\theta}(\Phi) \\ &= \mathrm{pr}\,\mathbf{v}_{\mathcal{J}^{BI}\theta}\left(\frac{1}{2}\int \left[-2uv(\rho\wedge\rho_x) + (1+v^2)(\rho\wedge\sigma_x) + 4vv_x(\rho\wedge\sigma)\right.\right. \\ &\quad \left.\left. +(1+v^2)(\sigma\wedge\rho_x)\right]dx\right) \\ &= \frac{1}{2}\int \left[(-2uv\rho_x - (uv)_x\rho(1+v^2)\sigma_x + 3vv_x\sigma)\wedge(-2v(\rho\wedge\rho_x))\right. \\ &\quad \left. +((1+v^2)\rho_x - vv_x\rho)\wedge(-2u(\rho\wedge\rho_x) + 2v(\rho\wedge\sigma_x) + 4v_x(\rho\wedge\sigma) + 2v(\sigma\wedge\rho_x))\right.\end{aligned}$$

$$
\begin{aligned}
&+(2vv_x\rho_x + (1+v^2)\rho_{xx} - (vv_x)_x\rho - vv_x\rho_x) \wedge (4v(\rho \wedge \sigma))\Big]\ dx \\
&= \frac{1}{2}\int \Big[4v(1+v^2)(\rho_x \wedge \rho \wedge \sigma_x) + 4v(1+v^2)(\rho_{xx} \wedge \rho \wedge \sigma) \\
&\quad +4v_x(1+3v^2)(\rho_x \wedge \rho \wedge \sigma)\Big]\ dx \\
&= 0. \qquad (4.3.43)
\end{aligned}
$$

The last step follows because the integrand is a total derivative. This completes the proof.

To comply with the conditions of Magri's theorem, we need to prove that the two Hamiltonian structures are compatible, i.e. the operator $\mathcal{J}_1 + \mathcal{J}^{BI}$ is also a Hamiltonian operator. Since the resulting structure is very similar to $\mathcal{J}^{BI}$, the proofs of skew-adjointness and closure of the Jacobi identity are very similar to those above, and it is easily shown that the sum of the two structures is Hamiltonian. Therefore we have a *bona fide* bi-Hamiltonian system in the sense of Magri.

Following the standard prescription, we can construct a recursion operator for the system. As usual in these cases, this is

$$
\mathcal{R} = \mathcal{J}_1\mathcal{J}_0^{-1} = \begin{pmatrix} (1+v^2) + 3vv_xD_x^{-1} & -2uv - (uv)_xD_x^{-1} \\ 0 & (1+v^2) - vv_xD_x^{-1} \end{pmatrix}. \qquad (4.3.44)
$$

Finding a Lax-type representation for the equation, like (4.3.25) is the next step. The recursion operator (4.3.44) is L. Its partner is the Fréchet derivative of the right-hand side of (4.3.34),

$$
\mathcal{A} = \begin{pmatrix} D_x\left(v(1+v^2)^{-\frac{1}{2}}.\right) & D_x\left(u(1+v^2)^{-\frac{3}{2}}.\right) \\ 0 & D_x\left(v(1+v^2)^{-\frac{1}{2}}.\right) \end{pmatrix}, \qquad (4.3.45)
$$

where the dot indicates that the derivation acts on the argument of the operator also. It is now straightforward to verify that (4.3.25) reproduces the evolution equations (4.3.34) using the operators (4.3.44) and (4.3.45).

The interesting feature of this example is that it is one of the few cases where the recurrence fails to produce an independent sequence of conserved quantities. It is shown in [22] that the sequence of Hamiltonians can terminate after three iterations, or repeat itself, depending on the version of the equation considered. These properties remain true in this case.

To conclude, we have seen that the Bateman equation is equivalent to a large class of biHamiltonian evolution equations, and as such its integrability (in the sense of possession of an infinity of commuting generalised symmetries) is established. The construction is illustrated with an example. Comparing with the previous section, we see that the classification of the "integrability" of the Bateman equation can be defined either in terms of its linearisability, which is due to the existence of certain generalised symmetries of the equation, or by this biHamiltonian property. Building on the earlier work by Fairlie *et al* [33], we can include an arbitrary dependence on ϕ in the variational formulation of Equation (4.1.1).

Chapter 5

Properties of Universal Field Equations

5.1 Introduction

The Universal Field models proposed in [33, 35, 36, 37] are a class of field theories with a wide variety of attractive features:

- they may be formulated in an arbitrary number of space–time dimensions;
- they are either diffeomorphism or reparametrisation invariant in the dependent variables;
- they are Lorentz and Euclidean invariant (and in fact $GL(m, \mathbb{R})$ invariant) in the independent variables, and there is an additional linear invariance in the dependent variables;
- they are derived from an infinite number of inequivalent Lagrangians.

Furthermore, one of the scalar theories is a direct generalisation of the Bateman equation and is linearisable by a Legendre transformation in the manner described in Chapter 4 [37]. We will therefore describe it as "integrable" in accordance with our two-pronged definition of the term in Chapter 4.

The business at hand is to describe and explain the strange variational properties of the scalar Universal equations. We shall eventually arrive at a rather more general formulation of the model than is found in [33]. Consider a scalar field ϕ dependent on m space(-time) coordinates. In the jet-bundle language of Chapter 1, it will be sufficient to consider a trivial bundle $(\mathbb{R} \times \mathbb{R}^m, \pi, \mathbb{R}^m)$, and its space of 2-jets, $\mathrm{J}^2\pi$.

The variational derivation of the Universal equations is based on the idea of the *generic Euler hierarchy*, described in [36]. Consider a Lagrangian $\mathcal{L}_0 : \mathrm{J}^1\pi \to \mathbb{R}$ which only depends on the first derivatives of ϕ. Compute the variation of this and construct a new function,

$$\mathcal{L}_1 = F_1 \mathcal{E}\mathcal{L}_0, \tag{5.1.1}$$

where F_1 is some real-valued function dependent only on the first derivatives of ϕ. Then consider $\mathcal{L}_1$ to be the Lagrangian for a new equation, $\mathcal{E}\mathcal{L}_1$. Continue the process to arrive at:

$$\mathcal{L}_k = F_k \mathcal{E}\mathcal{L}_{k-1}. \tag{5.1.2}$$

This sequence terminates at the $(m+1)$th stage: $\mathcal{E}\mathcal{L}_m = 0$. Furthermore, at the mth stage the resulting Euler–Lagrange expression factorises, and one of the factors is "universal" — independent of the details of the initial and intermediate Lagrangians. On setting this Euler–Lagrange form to zero, we arrive at an equation of motion equivalent to the Monge–Ampère expression,

$$\Delta_{M-A} = \det H = 0, \tag{5.1.3}$$

where H is the Hessian matrix of the dependent variable ϕ: $H_{ij} = \phi_{ij}$.

Another scalar Universal equation of interest is referred to in [36] as the *generalised Bateman equation*. It arises as a special case of this construction. The idea is to use an initial Lagrangian $\mathcal{L}_0$ which is homogeneous of degree one in the first derivatives of ϕ, and to restrict all the F_k to have this property too. In addition, it is assumed that $\det H \neq 0$. Performing the iterative variational procedure leads to the cessation of the hierarchy a stage earlier than described above: $\mathcal{E}\mathcal{L}_{m-1} = 0$. Remarkably, the penultimate Euler–Lagrange expression $\mathcal{E}\mathcal{L}_{m-2}$ is again a product, and again one of

the its factors is "universal" — independent of the details of the initial and all the intermediate Lagrangians. Setting this expression to zero gives the Universal Field Equation of [33] and sequels.

The equation is written,

$$\Delta = \det \begin{pmatrix} 0 & \phi_{x_1} & \cdots & \phi_{x_m} \\ \phi_{x_1} & \phi_{x_1 x_1} & \cdots & \phi_{x_1 x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{x_m} & \phi_{x_m x_1} & \cdots & \phi_{x_m x_m} \end{pmatrix} = 0, \tag{5.1.4}$$

or in components,

$$\Delta = \varepsilon_{i_1 \ldots i_m} \varepsilon_{j_1 \ldots j_m} \phi_{i_1} \phi_{j_1} \phi_{i_2 j_2} \ldots \phi_{i_m j_m} = 0. \tag{5.1.5}$$

The component form makes it easy to verify that Δ can also be expressed as:

$$\Delta = \operatorname{tr}\left(G H^{\dagger}\right), \tag{5.1.6}$$

where the matrix G has components

$$G_{ij} = \phi_i \phi_j \tag{5.1.7}$$

and

$$H^{\dagger} = \operatorname{adj}(H) \tag{5.1.8}$$

is the classical adjoint matrix of H. Alternatively, we could define a new matrix U such that (5.1.6) is equivalent to the equation:

$$\Delta = \operatorname{tr}(UH). \tag{5.1.9}$$

The explicit form of U is easily deduced from (5.1.5). It is:

$$U = \varepsilon_{i_1 \ldots i_m} \varepsilon_{j_1 \ldots j_m} \phi_{i_1} \phi_{j_1} \phi_{i_2 j_2} \ldots \phi_{i_{m-1} j_{m-1}}. \tag{5.1.10}$$

We can write another form for the Universal equation [3] which allows us to demonstrate the existence of an infinite number of conservation laws. Define $m-1$ new variables, $u^i = \frac{\phi_i}{\phi_1}$ (for $i = 2, \ldots, m$) and form a matrix (which we will denote C) with entries:

$$C_{ij} = u^i_{j+1} - u^j u^i_1. \tag{5.1.11}$$

As above, the subscripts represent partial derivatives. Explicit calculation shows that the determinant of C is equivalent to Δ as a source form.

As an example, consider the case where $m = 3$. Define $u = \frac{\phi_2}{\phi_1}$, $v = \frac{\phi_3}{\phi_1}$. Then (5.1.4) is equivalent to,

$$\begin{vmatrix} u_2 - uu_1 & u_3 - vu_1 \\ v_2 - uv_1 & v_3 - vv_1 \end{vmatrix} = 0, \tag{5.1.12}$$

or more explicitly, as.

$$u_2v_3 - u_3v_2 - u\left(u_1v_3 - v_1u_3\right) - v\left(u_2v_1 - u_1v_2\right) = 0. \tag{5.1.13}$$

There are an infinite number of conservation laws of the form,

$$\begin{aligned}
&\frac{\partial}{\partial x_1}\left[\frac{u^r}{r}\frac{\partial}{\partial x_2}\left(\frac{v^{s+1}}{s+1}\right) - \left(\frac{v^{s+1}}{s+1}\right)\frac{\partial}{\partial x_2}\left(\frac{u^r}{r}\right) + \frac{v^s}{s}\frac{\partial}{\partial x_3}\left(\frac{u^{r+1}}{r+1}\right)\right. \\
&\left. - \left(\frac{u^{r+1}}{r+1}\right)\frac{\partial}{\partial x_3}\left(\frac{v^s}{s}\right)\right] \\
&+ \frac{\partial}{\partial x_2}\left[\left(\frac{u^r}{r}\right)\frac{\partial}{\partial x_3}\left(\frac{v^s}{s}\right) - \left(\frac{v^s}{s}\right)\frac{\partial}{\partial x_3}\left(\frac{u^r}{r}\right) - \left(\frac{u^r}{r}\right)\frac{\partial}{\partial x_1}\left(\frac{v^{s+1}}{s+1}\right)\right. \\
&\left. + \left(\frac{v^{s+1}}{s+1}\right)\frac{\partial}{\partial x_1}\left(\frac{u^r}{r}\right)\right] \\
&+ \frac{\partial}{\partial x_3}\left[\left(\frac{v^s}{s}\right)\frac{\partial}{\partial x_2}\left(\frac{u^r}{r}\right) - \left(\frac{u^r}{r}\right)\frac{\partial}{\partial x_2}\left(\frac{v^s}{s}\right) - \left(\frac{v^s}{s}\right)\frac{\partial}{\partial x_1}\left(\frac{u^{r+1}}{r+1}\right)\right. \\
&\left. + \left(\frac{u^{r+1}}{r+1}\right)\frac{\partial}{\partial x_1}\left(\frac{v^s}{s}\right)\right] = 0,
\end{aligned} \tag{5.1.14}$$

where r and s are positive integers. A subsidiary condition is that the matrix in (5.1.12) is symmetric.

Both equations find their main application in differential geometry. The Monge–Ampère equation describes surfaces with zero Gauss curvature whereas the Bateman–type equation can be used to describe developable surfaces [44]. There is no known physical application of (5.1.4), although A.N. Leznov has demonstrated a connection to the ADHM ansatz for Self–Dual Monopoles [45].

This chapter is concerned mainly with the generalised symmetries and variational features of the model. For the generalised Bateman case, we will find that the first order generalised symmetries are basically those of a linear equation and this provides

the linearisation process described in [37]. In addition, the symmetries of both kinds of Universal equation will provide us with clues to the nature of the variational properties of both the generic and Bateman constructions. We will then use the techniques described in Chapter 2 to provide substantially simplified and generalised proofs of the main results on the both hierarchies. Also, we will try to see if the Born–Infeld equation in two dimensions has an m-dimensional generalisation which shares the linearisability property of the Universal equation. The variational and symmetry properties of this equation are discussed in the light of the analysis of the Universal equations.

We will also briefly consider the problem of a Hamiltonian structure for the Universal equations, mainly using the formal covariant formalism due to De Donder, Weyl, Cartan and others. We will discuss the difficulties in finding a biHamiltonian structure similar to that which we derived for the Bateman equation.

Basic references (apart from the original papers [33, 35, 36, 37]) are Over [7], the review articles by Anderson [15, 17] and his recent work with Pohjanpelto [16, 46]. Another view of the universal equations and their variational symmetry properties is given by Grigore [47, 48].

Elements of this chapter appeared in [3] and [5].

5.2 Generalised Symmetries

In order to understand the generic and Bateman hierarchies a little better (with a view to generalisation), it will be helpful to know more about the generalised symmetries of the various equations that appear in them. To that end we will carry out a detailed analysis of each of each equation in the manner described in Section 1.7. We will look for first order generalised symmetries — in other words symmetries whose evolutionary characteristic depends on the x_i, ϕ and the first derivatives of ϕ.

We begin with the equations of the generic hierarchy. The first equation in the

sequence is:

$$\begin{aligned}\Delta_1^G &= \mathcal{E}\mathcal{L}_0 \\ &= \frac{\partial \mathcal{L}_0}{\partial \phi} - \frac{\partial^2 \mathcal{L}_0}{\partial \phi \phi_i}\phi_i - \frac{\partial^2 \mathcal{L}_0}{\partial \phi_i \partial \phi_j}\phi_{ij} = 0. \end{aligned} \tag{5.2.1}$$

Using the standard procedure, we seek a generalised symmetry in the evolutionary form (1.7.3). The symmetry condition is:

$$\mathrm{prv}_Q(\Delta_1^G) = \frac{\partial \Delta_1^G}{\partial \phi} Q + \frac{\partial \Delta_1^G}{\partial \phi_i} D_i Q + \frac{\partial \Delta_1^G}{\partial \phi_{ij}} D_{ij} Q = 0, \tag{5.2.2}$$

where Q is the characteristic, on solutions of Δ_1^G. The total derivatives D_i and D_{ij} have the usual expansions:

$$\begin{aligned} D_i F &= \frac{\partial F}{\partial x_i} + \frac{\partial F}{\partial \phi_J}\phi_{Ji}, \\ D_{ij} F &= \frac{\partial^2 F}{\partial x_i \partial x_j} + \frac{\partial^2 F}{\partial x_i \partial \phi_K}\phi_{Kj} + \frac{\partial F}{\partial \phi_J}\phi_{Jij} + \frac{\partial^2 F}{\partial x_j \partial \phi_J}\phi_{Ji} + \frac{\partial^2 F}{\partial \phi_J \partial \phi_K}\phi_{Ji}\phi_{Kj}, \end{aligned} \tag{5.2.3}$$

where the multi-indices J, K will actually only have length zero or one in the case we are considering. After expanding the total derivatives, we find that the coefficients of the third derivatives vanish identically, given that the first prolongation of Δ_1^G must vanish. This property passes on to the subsequent equations in the sequence. There is little else we can say without further information about the initial Lagrangian.

However, we can repeat the work in Section 4.2 for the higher-dimensional universal equations. Again, the Kumei-Bluman theorem will furnish an explanation for the linearisability of the generalised Bateman case. Furthermore, we will use the generalised symmetries of the Monge-Ampère equation to motivate the discussion in the next section concerning the termination of both hierarchies.

We will begin with the equation (5.1.3). It turns out that the first order generalised symmetries of this equation span a rather large infinity of possibilities. The symmetry equation is quite straightforward:

$$\mathrm{prv}_Q(\Delta_{M-A}) = \frac{\partial \Delta_{M-A}}{\partial \phi_{ij}} D_{ij} Q[\phi] = \frac{1}{m} H_{ij}^{\dagger} D_{ij} Q[\phi] = 0. \tag{5.2.4}$$

On expanding the total derivative, we find that the coefficients of the third derivatives of ϕ vanish on solutions of (5.1.3) (after taking into account the first prolongation of

the equation of motion) and the first symmetry constraint comes from the term of order $(m+1)$ in the second derivatives, which gives:

$$H^{\dagger}_{ij}\frac{\partial^2 Q}{\partial\phi_k\partial\varphi_l}\phi_{ik}\varphi_{jl}=0. \tag{5.2.5}$$

on solutions. Writing,

$$[H_Q]_{kl}=\frac{\partial^2 Q}{\partial\phi_k\partial\phi_l}, \tag{5.2.6}$$

we find that (5.2.5) becomes:

$$\text{tr}\left(H^{\dagger}HH_QH\right)=\det H\ \text{tr}\left(H_QH\right)=0, \tag{5.2.7}$$

and this is identically true on solutions of $\Delta_{M-A}=0$. This pattern continues for the terms of order m in second derivatives, and we find that a Q with arbitrary dependence on the first derivatives is a generalised symmetry for (5.1.3).

This fact provides us with an important clue to the nature of the variational sequence. If these generalised symmetries are variational, then any function of the first derivatives is the characteristic of a conservation law for Δ_{M-A}. In that case, the next Lagrangian $F_m\Delta_{M-A}$ is a total divergence (see (2.3.11)), and the Euler variation annihilates it, terminating the hierarchy. This must be the case, given the converse of Noether's theorem. We will formalise and extend this idea in Subsection 5.3.3.

The next task is to study the effect of introducing homogeneity into the picture. We begin with a minor diversion to review a few basic formulae. We will make extensive use of Euler's famous theorem on homogeneous functions — namely that, if $\mathcal{F}(\mathbf{x})$ is homogeneous of degree α in its arguments $\mathbf{x}=(x_1,...,x_n)$, then:

$$\sum_{i=1}^{n}x_i\frac{\partial\mathcal{F}}{\partial x_i}=\alpha\mathcal{F}. \tag{5.2.8}$$

Of course, this amounts to the action of a scaling vector field on the function $\mathcal{F}$.

A couple of easy results follow from this. First, we can compute the scaling properties of derivatives of $\mathcal{F}$, namely that any kth order derivative of $\mathcal{F}$ is homogeneous of degree $\alpha-k$:

$$(\alpha-k)\partial_K\mathcal{F}=x_j\partial_{Kj}\mathcal{F}. \tag{5.2.9}$$

The proof is just an easy induction from differentiation of the formula (5.2.8). Also we can deduce that an Euler variation $\mathcal{EF}(\phi_i)$ is an expression homogeneous of degree $\alpha - 2$ in first derivatives and degree one in second derivatives. If the Lagrangian depends on the field ϕ, the homogeneity is lost unless $\alpha = 1$. To summarise:

$$(\alpha - 2)\mathcal{EF}(\phi_i) = \phi_j \frac{\partial}{\partial \phi_j}(\mathcal{EF}). \tag{5.2.10}$$

All of this generalises easily to the Lagrangians $\mathcal{L}_k$ of the Bateman hierarchy. Remembering the definition of such Lagrangians and starting from $\mathcal{L}_0$ homogeneous of degree one in first derivatives, it is not hard to see that $\mathcal{L}_k$ is *(a)* homogeneous of degree $1 - k$ in the first derivatives and *(b)* homogeneous of degree k in the second derivatives. It then follows that.

$$\mathcal{L}_k = \phi_I \frac{\partial \mathcal{L}_k}{\partial \phi_I}, \tag{5.2.11}$$

where I is a multi-index of length one or two. It is clear from this construction that each Δ_k in either the generic or Bateman hierarchy is homogeneous of degree k in the second derivatives of ϕ. In the generalisation which we are about to discuss in Section 5.3. where the $\mathcal{L}_k$ can have arbitrary ϕ dependence, the property (5.2.11) cannot be regarded as describing the prolonged action of a scaling symmetry on $\mathcal{L}_k$. The vector field $\phi_I \partial \phi_I$ is just a particular section of $TJ^2\pi$.

As a final point, we can combine the last result with (5.2.9) to show that for $\mathcal{L}_0$ homogeneous of degree one in first derivatives,

$$\phi_i \frac{\partial^2 \mathcal{L}_0}{\partial \phi_i \partial \phi_j} = 0, \tag{5.2.12}$$

which in turn implies the singularity condition:

$$\det\left(\frac{\partial^2 \mathcal{L}_0}{\partial \phi_i \partial \phi_j}\right) = 0, \tag{5.2.13}$$

since ϕ is arbitrary. A good account of the classical field theory of systems with singular Lagrangians is found in [49].

Returning to the symmetry properties, the homogeneity of the Lagrangian may be described by the action of the dilation field:

$$\mathbf{v}_\phi = \phi \frac{\partial}{\partial \phi}, \tag{5.2.14}$$

assuming no explicit dependence of the Lagrangian on ϕ. Given that the k-th Lagrangian in the hierarchy, $\mathcal{L}_k$, is homogeneous to some degree in each of its arguments, we can write,

$$\text{pr } \mathbf{v}_\phi(\mathcal{L}_k) = \alpha \mathcal{L}_k, \tag{5.2.15}$$

where α is the sum of the degrees of homogeneity of ϕ, and its derivatives. Thus the dilation field is not a *variational* symmetry of the problem. Nonetheless, using the basic identity,

$$\mathcal{E}(\text{pr } \mathbf{v}_\phi(\mathcal{L}_k)) = \text{pr } \mathbf{v}_\phi(\mathcal{E}\mathcal{L}_k) + \mathsf{D}_\phi^*(\mathcal{E}\mathcal{L}_k), \tag{5.2.16}$$

where D_ϕ^* is the adjoint of the Fréchet derivative with respect to ϕ, it is straightforward to verify that:

$$\text{pr } \mathbf{v}_\phi(\mathcal{E}\mathcal{L}_k) = (\alpha - 1)\mathcal{E}\mathcal{L}_k. \tag{5.2.17}$$

Hence, the dilation field (5.2.14) is a symmetry of the Euler–Lagrange equations for $\mathcal{L}_k$ when $\alpha = 1$.

It then remains to test the behaviour of $\mathbf{v}_\phi$ on the next member of the series. Take $\alpha = 1$ and

$$\mathcal{L}_{k+1} = F_{k+1}\mathcal{E}\mathcal{L}_k \tag{5.2.18}$$

where

$$\text{pr } \mathbf{v}_\phi(F_{k+1}) = F_{k+1}. \tag{5.2.19}$$

Then, using the derivation property of the Lie derivative, we have easily that,

$$\text{pr } \mathbf{v}_\phi(\mathcal{L}_{k+1}) = \mathcal{L}_{k+1}, \tag{5.2.20}$$

and so the analysis is unchanged: the dilations form a symmetry of the resulting Euler–Lagrange equations. So all the equations of the hierarchy share the property of total homogeneity of degree one.

If we turn our attention specifically to the symmetry algebra of the Universal equation (5.1.4), it turns out to be a little more restrictive than in the generic case. The determining equation is,

$$\frac{\partial \Delta}{\partial \phi_i}(D_i Q[\phi]) + \frac{\partial \Delta}{\partial \phi_{(ij)}}(D_{ij} Q[\phi]) = 0, \tag{5.2.21}$$

on solutions of $\Delta = 0$. Again, we will only look for first order generalised symmetries. Our experience with the Bateman equation suggests that the symmetries dependent on first derivatives are responsible for the linearisability property.

Given this assumption, (5.2.21) expands to become,

$$\varepsilon_{i_1 \ldots i_m} \varepsilon_{j_1 \ldots j_m} \phi_{i_1} \phi_{j_1} \phi_{i_2 j_2} \cdots \phi_{i_m j_m} + (m-1) U_{kl} \delta_{(ij)(kl)} D_{ij} Q = 0, \tag{5.2.22}$$

where,

$$U_{kl} = \varepsilon_{i_1 \ldots i_{m-1} k} \varepsilon_{j_1 \ldots j_{m-1} l} \phi_{i_1} \phi_{j_1} \phi_{i_2 j_2} \cdots \phi_{i_{m-1} j_{m-1}} \tag{5.2.23}$$

is the sort of matrix appearing in (5.1.9).

Following the usual algorithm, we must set the left side of (5.2.22) to zero order-by-order in the derivatives, taking into account the equation $\Delta = 0$. Using the first prolongation of Δ, we find that third order derivatives automatically vanish, and so the first task is to find the condition for the vanishing of terms of order m in the second derivatives.

Extracting the relevant term from (5.2.22), we define a matrix S such that:

$$\begin{aligned} S_{kl} \frac{\partial^2 Q}{\partial \phi_k \partial \phi_l} &= (m-1) H_{ki} U_{ij} H_{jl} \frac{\partial^2 Q}{\partial \phi_k \partial \phi_l} \\ &= 0, \end{aligned} \tag{5.2.24}$$

on solutions of $\Delta = 0$. To simplify this, we use the form (5.1.9) to rewrite Δ using the cyclicity and linearity of trace and the properties of the classical adjoint:

$$\Delta = \frac{1}{m} \frac{1}{\det H} \operatorname{tr} \left(H U H H^{\dagger} \right), \tag{5.2.25}$$

and then, by associativity,

$$\Delta = \frac{1}{m(m-1)} \frac{1}{\det H} \operatorname{tr} \left(S H^{\dagger} \right). \tag{5.2.26}$$

On comparison with (5.1.6) we find that $S = m(m-1) \det H \; G$ and the "on-shell" symmetry condition is satisfied by any S =(some scalar factor)$\times \det H \; G$. This produces a symmetry condition,

$$\phi_i \phi_j \frac{\partial^2 Q}{\partial \phi_i \partial \phi_j} = 0, \tag{5.2.27}$$

which is a straightforward generalisation of (4.2.4), and is satisfied by any Q homogeneous of degree zero or one in the first derivatives of ϕ.

As with the Bateman, this condition guarantees linearisability of (5.1.4) via the Kumei–Bluman theorem [31]. The analysis is essentially the same as in Section 4.2. The form of the symmetry condition leads us to use new coordinates $X^i = \phi_i$ and Φ for the independent and dependent variables respectively. As before the preservation conditions for the contact ideal are that,

$$\begin{aligned} \left[X^i, X^j\right]_L &= 0, \\ \left[X^i, \Phi\right]_L &= 0. \end{aligned} \tag{5.2.28}$$

where the Lagrange bracket is defined as in (4.2.6). Once again, the X^i automatically commute with one another, and the commutation of Φ with X is guaranteed if Φ is the Legendre transform of ϕ:

$$\Phi(X) = X^i \phi_i - \frac{1}{2}\phi. \tag{5.2.29}$$

The linearised equation is then:

$$X^i X^j \frac{\partial^2 \Phi}{\partial X^i \partial X^j} = 0. \tag{5.2.30}$$

This is the result presented in [37].

The rationalisation for the termination of the generic hierarchy applies here, too. The fact that any function homogeneous of weight one in first derivatives is a symmetry means the next Lagrangian must be a divergence, and so the hierarchy terminates.

Having disposed of the term with m second derivatives, we need to equate the term with $m-1$ second derivatives to zero. The relevant equation can be written as:

$$H^{\dagger}_{ij} \frac{\partial Q}{\partial x_i} \phi_j + 2(m-1) U_{kl} \delta_{(ij)(kl)} \left(\frac{\partial^2 Q}{\partial x_i \partial \phi_n} \phi_{jn} + \frac{\partial^2 Q}{\partial \phi \partial \phi_n} \phi_i \phi_{jn} \right) = 0. \tag{5.2.31}$$

This can be solved by a characteristic of the form,

$$Q = g(\phi) F(\eta_i), \tag{5.2.32}$$

where g is an arbitrary smooth, real-valued function, $\eta_i = x_i\phi_i$ (no sum) and F is a smooth, real-valued function which respects the homogeneity properties that we have decided on for Q. Notice that the diffeomorphism symmetry of the equation is included in this solution.

Finally, to get rid of the remaining terms, we use (5.2.32) and arrive eventually at the condition:

$$\varepsilon_{i_1\ldots i\ldots i_m}\varepsilon_{j_1\ldots j\ldots j_m}\phi_{i_1}\phi_{j_1}\phi_{i_2j_2}\ldots\phi_i\phi_j\ldots\phi_{i_mj_m}\,(gF''+2g'F'+g''F)=0. \qquad (5.2.33)$$

This is identically true due to the antisymmetry of the ε symbol.

To summarise, the conclusions of this section are that:

- the Monge–Ampère equation (5.1.3) has a large class of first order generalised symmetries — any function of the first derivatives of ϕ is the characteristic of a generalised symmetry;
- any function homogeneous of degree zero or one in the first derivatives of ϕ is the characteristic of a generalised symmetry of (5.1.4);
- since this fact follows from a linear equation, the Bateman–type Universal equations are linearisable.

5.3 Lagrangian Properties

Our mission is to try to understand and generalise the generic and Bateman hierarchies described in Section 5.1 from the point of view of the variational theory developed in Chapter 2. In particular, we wish to know under what circumstances we can introduce dependence on the field ϕ into the initial Lagrangian and the multiplier functions F_k. We will find that only in certain circumstances does the Euler hierarchy define a "universal" theory. In addition, we will think about what the iterative procedure means in terms of the formal variational theory of Section 2.4, incorporating the information we have from Section 5.2. We will find that the iterated equations

of motion are related to determining equations for variational symmetries of higher members of the hierarchy.

5.3.1 Generic Hierarchy

We begin by analysing the Euler hierarchy in the light of our knowledge of the variational calculus presented in Chapter 2. Initially, we will only assume that the initial Lagrangian depends on the field and its first derivatives, $\mathcal{L}_0 : \mathrm{J}^1\pi \to \mathbb{R}$, and make no assumptions about any symmetries of the resulting action.

From $\mathcal{L}_0$, we derive an Euler–Lagrange form:

$$\Delta_1 = \mathsf{D}^*_{\mathcal{L}_0}(1) = \mathcal{E}\mathcal{L}_0. \tag{5.3.1}$$

From this we construct the next Lagrangian.

$$\mathcal{L}_1 = F_1\Delta_1, \tag{5.3.2}$$

where we assume that F_1 depends only on ϕ and its first derivatives. The next source form follows readily using the product formula (2.2.11):

$$\Delta_2 = \mathcal{E}\mathcal{L}_1 = \mathsf{D}^*_{\Delta_1}(F_1) + \mathsf{D}^*_{F_1}(\Delta_1). \tag{5.3.3}$$

Now, by the Helmholtz condition (see (2.4.2)), the first term is just the Fréchet derivative:

$$\begin{aligned} \mathsf{D}^*_{\Delta_1}(F_1) &= \mathsf{D}_{\Delta_1}(F_1) \\ &= \frac{\partial \Delta_1}{\partial \phi_J} D_J(F_1) \end{aligned} \tag{5.3.4}$$

The second term on the right–hand side of (5.3.3) just turns out to be,

$$\mathsf{D}^*_{F_1}(\Delta_1) = \Delta_1 \frac{\partial F_1}{\partial \phi} - D_i\left(\frac{\partial F_1}{\partial \phi_i}\Delta_1\right), \tag{5.3.5}$$

since F_1 only depends on ϕ and its first derivatives.

Putting (5.3.4) and (5.3.5) into (5.3.3), a short calculation tells us that:

$$\begin{aligned} \Delta_2 = \Delta_1 \mathcal{E}F_1 &+ \frac{\partial \Delta_1}{\partial \phi}\left(F_1 - \frac{\partial F_1}{\partial \phi_i}\phi_i\right) + \frac{\partial \Delta_1}{\partial \phi_i}\frac{\partial F_1}{\partial \phi}\phi_i \\ &+ \frac{\partial \Delta_1}{\partial \phi_{ij}}\left(D_i\left(\frac{\partial F_1}{\partial \phi}\phi_j\right) + D_i\left(\frac{\partial F_1}{\partial \phi_k}\right)\phi_{jk}\right). \end{aligned} \tag{5.3.6}$$

The important point to notice is that this construction guarantees that there are no derivatives of order higher than two in the resulting equation of motion. This means that when the process is continued, the analysis is essentially unchanged. We are led to the iterative formula for the $(k+1)$th Euler–Lagrange form:

$$\Delta_{k+1} = \Delta_k \mathcal{E} F_k + \frac{\partial \Delta_k}{\partial \phi}\left(F_k - \frac{\partial F_k}{\partial \phi_i}\phi_i\right) + \frac{\partial \Delta_k}{\partial \phi_i}\frac{\partial F_k}{\partial \phi}\phi_i + \frac{\partial \Delta_k}{\partial \phi_{ij}}\left(D_i\left(\frac{\partial F_k}{\partial \phi}\phi_j\right) + D_i\left(\frac{\partial F_k}{\partial \phi_l}\right)\phi_{jl}\right). \tag{5.3.7}$$

There is no such recursive definition if F_k depends on second or higher derivatives.

We are interested in sequences generated by this kind of recursion which terminate after a finite number of iterations. The expression (5.3.7) simplifies greatly if we restrict attention to F_k that are *(1)* independent of ϕ, and *(2)* homogeneous of degree one in the first derivatives. Then we find:

$$\Delta_{k+1} = \Delta_k \mathcal{E} F_k + \frac{\partial \Delta_k}{\partial \phi_{ij}} D_i\left(\frac{\partial F_k}{\partial \phi_l}\right)\phi_{jl}, \tag{5.3.8}$$

or more symmetrically,

$$\Delta_{k+1} = \left(\frac{\partial \Delta_k}{\partial \phi_{kl}}\phi_{ik}\phi_{jl} - \Delta_k \phi_{ij}\right)\frac{\partial^2 F_k}{\partial \phi_i \partial \phi_j}. \tag{5.3.9}$$

This is precisely the recurrence found by Fairlie and Govaerts ([36]) in their treatment of the generic hierarchy. They found that if $\mathcal{L}_0$ and all the F_k were independent of ϕ then the sequence terminated in a Monge–Ampère equation. This sequence only produces such a result if it is assumed that the initial Lagrangian is independent of ϕ in which case the condition *(2)* can be relaxed anyway, so there is no new information. Otherwise, (5.3.8) fails to reproduce the generic hierarchy because the successive source forms Δ_k have inhomogeneous terms due to the derivatives with respect to ϕ.

So far, all attempts to generalise the construction of the generic hierarchy by relaxing the assumption of ϕ-independence of the functions F_k have failed. Sample computer calculations carried out on MAPLE show that such constructions do not terminate in universal, exact source forms at either the $(m-1)$th or mth stages.

We proceed to generate the equations of the generic hierarchy under the restriction that the $\mathcal{L}_0$ and F_k are ϕ-independent. We know that the first equation (5.3.1) has the explicit form,

$$\Delta_1 = -\frac{\partial^2 \mathcal{L}_0}{\partial \phi_i \partial \phi_j}\phi_{ij}, \tag{5.3.10}$$

which can be written:

$$\Delta_1 = -\operatorname{tr}(HM_0), \tag{5.3.11}$$

where M_0 is the Hessian matrix of $\mathcal{L}_0$ with respect to its dependence on first derivatives of ϕ. It is then a straightforward matter to apply the recursion relation (5.3.9) to this starting term. Using the notation,

$$\begin{aligned} [M_k]_{ij} &= \frac{\partial^2 F_k}{\partial \phi_i \partial \phi_j}, \\ P_k &= HM_k, \end{aligned} \tag{5.3.12}$$

the first few terms of the hierarchy are:

$$\begin{aligned}
\Delta_1 &= -\operatorname{tr}(P_0), \\
\Delta_2 &= \operatorname{tr}(P_0)\operatorname{tr}(P_1) - tr(P_0P_1), \\
\Delta_3 &= -\operatorname{tr}(P_2(P_0P_1 + P_1P_0)) + \operatorname{tr}(P_0)\operatorname{tr}(P_1P_2) + \operatorname{tr}(P_1)\operatorname{tr}(P_0P_2) \\
&\quad + \operatorname{tr}(P_2)\operatorname{tr}(P_0P_1) - \operatorname{tr}(P_0)\operatorname{tr}(P_1)\operatorname{tr}(P_2), \\
\Delta_4 &= -\operatorname{tr}(P_3(P_1P_2P_0 + P_0P_1P_2 + P_2P_0P_1 + P_0P_2P_1 + P_1P_0P_2 + P_2P_1P_0)) \\
&\quad + \operatorname{tr}(P_1P_2)\operatorname{tr}(P_0P_3) + \operatorname{tr}(P_2P_3)\operatorname{tr}(P_0P_1) + \operatorname{tr}(P_1P_3)\operatorname{tr}(P_0P_2) \\
&\quad + \operatorname{tr}(P_0)(\operatorname{tr}(P_1P_2P_3) + \operatorname{tr}(P_3P_2P_1)) + \operatorname{tr}(P_1)(\operatorname{tr}(P_0P_2P_3) + \operatorname{tr}(P_3P_2P_0)) \\
&\quad + \operatorname{tr}(P_2)(\operatorname{tr}(P_0P_1P_3) + \operatorname{tr}(P_3P_1P_0)) - \operatorname{tr}(P_0)\operatorname{tr}(P_1)\operatorname{tr}(P_2P_3) \\
&\quad - \operatorname{tr}(P_0)\operatorname{tr}(P_2)\operatorname{tr}(P_1P_3) - \operatorname{tr}(P_1)\operatorname{tr}(P_2)\operatorname{tr}(P_0P_3) \\
&\quad - \operatorname{tr}(P_3)\Delta_3 \\
&\quad \vdots
\end{aligned} \tag{5.3.13}$$

The subsequent members of the hierarchy become combinatorially more complicated.

We can use the recursion relation as the basis for a plausibility argument that demonstrates the vanishing of the Δ_k at a certain stage, and hence when the F_k

is the characteristic of a conservation law. We wish to know when Δ_k vanishes for *any* possible choices of F_k. (Of course, we are only interested in those cases when the Hessian of F_k is non-vanishing.) This amounts to solving the matrix differential equation,

$$H\Gamma^k H = \Delta_k H, \tag{5.3.14}$$

where,

$$\Gamma^k_{ij} = \frac{\partial \Delta_k}{\partial \phi_{ij}}. \tag{5.3.15}$$

The equation (5.3.14) is solved by any Δ_k of the form $\Delta =$ (some factor) $\times \det H$, where "some factor" is independent of the second derivatives of ϕ. We know from the variational calculation that the second derivatives only enter the source forms such that Δ_k is homogeneous of degree k in the second derivatives. Therefore, this solution can only work at the mth stage.

A proof of these results was presented in [36] for the case when the initial and intermediate Lagrangians are independent of ϕ. We will attempt to demonstrate these properties in a slightly different fashion using some results from linear algebra.

Let us assume for the moment that the functions F_k are all identical to $\mathcal{L}_0$, so $P_k = P$ and $M^k = M$ for all k. We can understand the structure of these equations a little more by recalling the *Cayley–Hamilton theorem.* It is well–known that an square $(m \times m)$ matrix A is a root of its own characteristic polynomial. In other words, given a polynomial $\sigma(\lambda)$ in some scalar λ defined by,

$$\sigma(\lambda) = \det(\lambda \mathbb{I} - A), \tag{5.3.16}$$

and of the form:

$$\sigma(\lambda) = \lambda^m + b_{m-1}\lambda^{m-1} + b_{m-2}\lambda^{m-2} + \ldots + b_1\lambda + b_0, \tag{5.3.17}$$

then the Cayley–Hamilton theorem [50] states that:

$$\sigma(A) = A^m + b_{m-1}A^{m-1} + b_{m-2}A^{m-2} + \ldots + b_1 A + b_0 \mathbb{I} = 0. \tag{5.3.18}$$

(The symbol $\mathbb{I}$ denotes the unit matrix.) Furthermore, it is not difficult to prove that $b_{m-1} = -\operatorname{tr}(A)$ and $b_0 = \det(-A)$.

This construction can be generalised[1] as follows. Consider the recursion relation,

$$T_{k+1} = AT_k - \frac{\mathrm{tr}\,(AT_k)\mathbb{I}}{k+1},\ T_0 = \mathbb{I}, \tag{5.3.19}$$

for an $(m \times m)$ matrix A. This generates a sequence of matrix polynomials, and from an induction on the $m = 2$ case it can be shown that the Cayley–Hamilton theorem is encapsulated by the statement $T_m = 0$. The most useful by–product of this treatment is a sequence of expansions for the determinants E_m of the matrix A in terms of its trace for all values of m. We know that the coefficient b_0 of the identity matrix in (5.3.18) is, up to sign, $\det(A)$. So we have:

$$\begin{aligned}
E_2 &= \frac{1}{2}\left((\,\mathrm{tr}\,(A))^2 - \,\mathrm{tr}\,(A^2)\right), \\
E_3 &= \frac{1}{6}\left((\,\mathrm{tr}\,(A))^3 - 3\,\mathrm{tr}\,(A)\,\mathrm{tr}\,(A^2) + 2\,\mathrm{tr}\,(A^3)\right), \\
E_4 &= \frac{1}{4!}\left((\,\mathrm{tr}\,(A))^4 + 8\,\mathrm{tr}\,(A)\,\mathrm{tr}\,(A^3) + 3\,\mathrm{tr}\,(A^2)\,\mathrm{tr}\,(A^2) - 6\,\mathrm{tr}\,(A^4)\right. \\
&\quad \left. -6\,\mathrm{tr}\,(A^2)(\,\mathrm{tr}\,(A))^2\right), \\
&\ \vdots
\end{aligned} \tag{5.3.20}$$

Furthermore, the $E_2, \ldots, E_{m-1}$ correspond (up to sign) to the coefficients $b_{m-2}, \ldots, b_1$ in the expansion (5.3.17). The sequence (5.3.20) can also be generated by a recurrence relation,

$$E_{k+1} = \frac{1}{k+1}(\,\mathrm{tr}\,(A)\,E_k - \mathcal{P}E_k), \tag{5.3.21}$$

where the linear operator $\mathcal{P}$ is a derivation that acts on traces by:

$$\mathcal{P}\,\mathrm{tr}\left(A^k\right) = k\,\mathrm{tr}\left(A^{k+1}\right). \tag{5.3.22}$$

In the case when the P_i are identical to one another, the sequence (5.3.20) and the hierarchy (5.3.13) are essentially the same (modulo overall sign and a combinatorial factor). We can demonstrate this by defining the action of an operator $\mathcal{P}'$ on a scalar Ξ constructed from the matrix $P = P_i$ by:

$$\mathcal{P}'\,\Xi = \mathrm{tr}\,(H\Gamma' P). \tag{5.3.23}$$

[1] I am grateful to D.B. Fairlie for pointing this out to me.

The symbol Γ' is a matrix defined by:

$$\Gamma'_{ij} = \frac{\partial \Xi}{\partial \phi_{ij}}. \tag{5.3.24}$$

Then it is easy to check that $\mathcal{P}'$ has the property which defines $\mathcal{P}$, namely that,

$$\mathcal{P}' \operatorname{tr}\left(P^k\right) = k \operatorname{tr}\left(P^{k+1}\right), \tag{5.3.25}$$

for any integer k. So the recursion relation (5.3.9) becomes:

$$\Delta_{k+1} = \mathcal{P}'\Delta_k - \Delta_k \operatorname{tr}(P), \tag{5.3.26}$$

which is corresponds up to an overall factor to (5.3.21). So the Euler hierarchy (5.3.13) is the same as the Cayley–Hamilton sequence (5.3.20).

In the case when $F_k = \mathcal{L}_0$, we can now state that the mth member of the generic hierarchy is.

$$\Delta_m = \det(P) = \det(M)\det(H), \tag{5.3.27}$$

which is equivalent to Δ_{M-A} when set to zero, and,

$$\Delta_{m+1} \equiv 0, \tag{5.3.28}$$

as described in the introduction.

For the complete formulation when the P_i are not necessarily the same, we must generalise this construction still further. This seems to be a difficult task, and has not been accomplished to the same degree of success as the "generalised determinant" argument in [36] to which it seems to be related. Nevertheless, we can make a few general remarks. Consider m distinct square $m \times m$ matrices $\{A_1, \ldots, A_m\}$. We recursively define the matrices,

$$T_{k+1}(A_1, \ldots, A_{k+1}) = A_{k+1}\mathcal{S}T_k(A_1, \ldots, A_k) - \frac{\operatorname{tr}\left(A_{k+1}\mathcal{S}T_k(A_1, \ldots, A_k)\right)}{(k+1)!}\mathbb{I},\ T_0 = \mathbb{I}, \tag{5.3.29}$$

where the operator $\mathcal{S}$ is defined by,

$$\mathcal{S}T_k(A_1, \ldots, A_k) = \sum_{\sigma \in C_k} T_k(A_{\sigma(1)}, \ldots, A_{\sigma(k)}), \tag{5.3.30}$$

and C_k denotes the kth cyclic group. It should be clear that all these objects reduce to the standard Cayley–Hamilton matrices when all the As are equal. The generalisation of the Cayley–Hamilton theorem is then the statement that,

$$\mathcal{S}T_m(A_1, \ldots . A_m) \equiv 0. \tag{5.3.31}$$

If we then define the generalised Cayley–Hamilton terms to be,

$$E'_{k+1}(A_1, \ldots, A_{k+1}) = \frac{\text{tr } (A_{k+1}\mathcal{S}T_k(A_1, \ldots, A_k))}{(k+1)!}, \tag{5.3.32}$$

then these quantities correspond to the general sequence of source forms (5.3.13). We can generate them directly by a recursion similar to (5.3.26):

$$\Delta_{k+1} = \mathcal{P}_{k+1}\Delta_k - \text{tr } (A_{k+1})\,\Delta_k,\ \Delta_0 = 1. \tag{5.3.33}$$

where the operator $\mathcal{P}_k$ is a linear derivation defined by its action on the trace of a product of matrices:

$$\mathcal{P}_k \,\text{tr } (A_{i_1} \ldots A_{i_k}) = \text{tr } (A_{k+1}\mathcal{S}(A_1 \ldots A_k)))\,. \tag{5.3.34}$$

By a similar argument to that in (5.3.23...5.3.26), we can show that the recursion relations (5.3.33) and (5.3.9) are essentially the same.

To complete the analysis, we need to prove a multiplication theorem for the E' in analogy to the result 5.3.27. It seems to be very difficult to find a widely applicable statement of this nature, but there seem to be two important special cases:

$$\begin{aligned} E'_m(A_1B, \ldots, A_mB) &= E'_m(BA_1, \ldots, BA_m), \\ &= \text{tr } (A_m\mathcal{S}T_{m-1}(A_1, \ldots, A_{m-1})B\mathcal{S}T_{m-1}(B, \ldots, B)) \end{aligned} \tag{5.3.35}$$

for the final equation of the generic hierarchy, and (for later reference),

$$\begin{aligned} E'_{m-1}(A_1B, \ldots, A_{m-1}B) &= E'_{m-1}(BA_1, \ldots, BA_{m-1}), \\ &= \text{tr } (\mathcal{S}T_{m-1}(A_1, \ldots, A_{m-1})\mathcal{S}T_{m-1}(B, \ldots, B)) \end{aligned} \tag{5.3.36}$$

for the penultimate equation.

Now, we know from the Cayley–Hamilton theorem $T_{m-1}(B,\ldots,B) = B^{\dagger}$, so we find,

$$E'_{m-1}(A_1B,\ldots,A_mB) = -\operatorname{tr}\left(\mathcal{S}T_{m-1}(A_1,\ldots,A_{m-1})B^{\dagger}\right). \tag{5.3.37}$$

Likewise, we know that $B\mathcal{S}T_{k-1}(B,\ldots,B) = -\det(-B)\mathbb{I}$, and so for the final generalised determinant we deduce,

$$E'_m(A_1B,\ldots,A_mB) = -\det(-B)E'_m(A_1,\ldots,A_m). \tag{5.3.38}$$

Thus, substituting $P_k = HM_k$ for the A_kB in the multiplication theorem (5.3.35), we see that Δ_m is equivalent to the universal equation Δ_{M-A}.

This completes the discussion of the generic case.

5.3.2 Bateman Hierarchy

We now assume that the initial Lagrangian and the F_k have arbitrary dependence on ϕ and are homogeneous of degree one in first derivatives of ϕ. This greatly civilises the shapes of both Δ_1 and Δ_2 from their original forms (5.3.1) and (5.3.6). If we use the properties (5.2.9), (5.2.10) and (5.2.11), we recalculate Δ_1 and Δ_2 to be,

$$\begin{aligned}\Delta_1 &= -\frac{\partial^2\mathcal{L}_0}{\partial\phi_i\phi_j}\phi_{ij},\\ \Delta_2 &= \left(\frac{\partial\Delta_1}{\partial\phi_{kl}}\phi_{ik}\phi_{jl} - \Delta_1\phi_{ij}\right)\frac{\partial^2F_1}{\partial\phi_i\partial\phi_j},\end{aligned} \tag{5.3.39}$$

from which we deduce precisely the same recurrence (5.3.9) without the restriction that the F_k need be independent of ϕ.

The recursive procedure now defines a set of equations identical in form to (5.3.13), and with all of the linear-algebraic properties of the generic equations, and we can apply all of the results derived from the Cayley–Hamilton theorem.

For example, in the simplified case where $F_k = \mathcal{L}_0$, equation (5.3.27) demonstrates the termination of the Bateman hierarchy as in the generic case. By (5.2.13), $\det(M)$ is identically zero for homogeneous Lagrangians and so Δ_m is also identically zero.

By the same arguments, (5.2.12) we know that $GM = 0$ identically, and this can be solved [33] by taking $M^\dagger =$ (some constant) $\times\, G$. Now, the $(m-1)$th Cayley–Hamilton matrix expression, T_{m-1}, is just $A^\dagger$, which in our case means:

$$\begin{aligned}\Delta_{m-1} &= \text{tr}\left(M^\dagger H^\dagger\right)\\ &= (\text{some constant}) \times \text{tr}\left(GH^\dagger\right),\end{aligned} \tag{5.3.40}$$

and this is recognised as giving the Universal equation in the form (5.1.6).

For the more general case, when $F_k \neq \mathcal{L}_0$, the multiplication theorem (5.3.36) tells us that:

$$\Delta_{m-1} = \text{tr}\left(\mathcal{S}T_{m-1}(M_0,\ldots,M_{m-2})H^\dagger\right). \tag{5.3.41}$$

Then we use a similar argument to that above. We know that $GM_k \equiv 0$ for all k and from (5.3.31) that,

$$\mathcal{S}T_m(M_0,\ldots,M_{m-1}) = \sum_{\sigma\in C_m} M_{\sigma(m-1)}\mathcal{S}T_{m-1}(M_{\sigma(0)},\ldots,M_{\sigma(m-2)}) = 0, \tag{5.3.42}$$

and so we can solve this equation by an assignment,

$$\mathcal{S}T_{m-1}(M_{\sigma(0)},\ldots,M_{\sigma(m-2)}) = (\text{some constant}) \times G, \tag{5.3.43}$$

for any cyclic permutation σ. Using this in (5.3.41) allows us to assign Δ_{m-1} as in (5.3.40).

5.3.3 Geometrical Interpretation

There is a nice interpretation of all this in terms of the theory of the Euler–Lagrange complex studied in Section 2.4. Consider a Lagrangian $\lambda_0 \in \Omega^{(n,0)}$. The corresponding variation is the source form $\Delta_1 d\phi \wedge \omega = \mathcal{E}(\lambda) \in \Omega^{n,1}$. Now consider taking a Lie derivative of Δ_1 with respect to the prolongation of some generalised evolutionary vector field X_{F_1}. For an arbitrary source form Δ, let us write the Cartan formula:

$$\mathsf{L}_{\text{pr}\,X_{F_1}}(\Delta\; d\phi\wedge\omega) = \delta_V\left(X_{F_1}\lrcorner\;\Delta\; d\phi\wedge\omega\right) + X_{F_1}\lrcorner\;\delta_V\left(\Delta\; d\phi\wedge\omega\right). \tag{5.3.44}$$

A very similar formula appears in [16] for the Lie derivative with respect to a point symmetry vector field. Each term in (5.3.44) has a simple interpretation in the calculus of variations. The first term on the right is an Euler variation, and so it vanishes if X_{F_1} is the characteristic of a conservation law. The δ_V in the third term is just a Helmholtz operator, so it vanishes if Δ is an exact Euler–Lagrange form. If we expand these terms, we find that (5.3.44) is identical to (2.3.8), so this Lie derivative vanishes identically if X_1 is a variational symmetry of Δ. Not surprisingly, this type of formula is central in the recent studies of the generalised Noether theorem by Anderson and Pohjanpelto [16, 46].

Returning to our example, (5.3.44) takes on a particularly simple form when applied to Δ_1. Since Δ_1 is an exact Euler–Lagrange form, the Helmholtz term vanishes and we are left with the form Δ_2 as defined in (5.3.3). This interpretation holds for all Δ_k, so we can view the Euler hierarchy as repeated application of the Lie derivative to successive source forms. So the recursive definition becomes:

$$\Delta_{k+1}\, d\phi \wedge \omega = \mathsf{L}_{\mathbf{prv}_{F_k}} (\Delta_k\, d\phi \wedge \omega)\,. \tag{5.3.45}$$

The "universal" theory rests on the observation that at a certain stage all the source forms so defined, in either the generic or Bateman hierarchy, are equivalent to one another and their flows defined by the Lie derivative are identically zero. In such a situation, the iteration vanishes identically, and this yields a product expansion of the type (5.3.3) which is equal to zero. Compare this to the condition (2.3.13) and we see that the F_k must be the characteristic of a conservation law. Explicitly, we deduce that any function of the first derivatives of ϕ is a variational symmetry of the Monge–Ampère equation and that any function of homogeneous of weight one in the ϕ_i but with arbitrary dependence on ϕ is a variational symmetry of the Bateman–type Universal equation (5.1.4). This is the result which we glimpsed in Section 5.2.

Changing emphasis, we can view the F_k in the generic hierarchy as unknown functions to be determined, and then we can interpret Δ_2 as the equation determining the variational symmetry algebra of the Euler–Lagrange form Δ_1. The third and subsequent source forms $\Delta_3, \ldots, \Delta_m$ are a set of nested equations determining vari-

ational symmetry algebras for their immediate predecessors. It would be interesting to know what information, if any, can be gleaned from these higher equations about the symmetries of the original equation Δ_1.

5.3.4 Conclusions

We have analysed the circumstances under which the Euler hierarchy method can be used with Lagrangians which have explicit dependence on the field ϕ. It was found that the generic hierarchy could not admit any dependence on the field and still terminate in a universal source form. However, the Bateman hierarchy is algebraically unchanged by the inclusion of ϕ dependence, and terminates in the standard Universal Field Equation.

Motivated by the generalised symmetries of the equations, an interpretation of the Euler iteration has been suggested which characterises the method as the successive application of a a generalised Lie derivative operator to the initial Lagrangian. The geometrical significance of such a procedure remains unclear.

5.4 Interlude: The Born–Infeld Equation Revisited

We saw in Section 4.2 that the linearisability properties of the Bateman and Born–Infeld equations were essentially facets of a common generalised symmetry. Unfortunately, there is no such correspondence between the Universal equation and the multi-dimensional Born–Infeld model studied in [51]. However, it was shown in [3] that there does exist a family of equations in several dimensions which are Lorentz invariant, which can be reduced to the Born–Infeld equation in two dimensions, and which have precisely the same linearisability properties as the Universal equations.

The idea is to start with the form (5.1.6). The crucial observation is this: under a Lorentz transformation generated by the matrix Λ, both G and $H^\dagger$ transform in the

same way,

$$G' = \Lambda^{-1} G \Lambda, \quad H^{\dagger\prime} = \Lambda^{-1} H^{\dagger} \Lambda. \tag{5.4.1}$$

From (5.1.6) it is easy to see that the Universal equation is invariant under Lorentz transformations. The argument in [3] is that this form can be exploited to produce a new family of equations which have the properties of Lorentz invariance and linearisability. G is essentially idempotent — $G^2 = \mathrm{tr}\,(G)G$ — and so the only other matrix available to us which transforms according to (5.4.1) is the metric tensor η_{ij}. Thus the following was postulated as a generalisation of (4.1.7),

$$\sum_{i,j} H^{\dagger}_{ij} \left(G_{ji} + f\left(\mathrm{tr}\,(\eta G)\right) \eta_{ji} \right) = 0, \tag{5.4.2}$$

where f is an arbitrary, smooth, real-valued function of the quadratic Lorentz invariant constructed from ϕ_i. (The possibility of an explicit dependence upon ϕ is excluded.) This is *not* the usual minimal-surface/p-brane equation which has not proved to be integrable [52, 53].

Therefore it seems that we can construct a new set of Lorentz-invariant models in four dimensions that are integrable, in the sense of being linearisable. The family of such models is very small: essentially there are only the Self-Dual Yang Mills and Self-Dual Einstein equations [54, 55] and their supersymmetric extensions, which are linearisable by twistor methods, the relativistic string equations [56, 57, 58], and the equations proposed here, which are directly related to linear equations through the Legendre transform.

The multivariable version of the Legendre transform is a straightforward generalisation of (4.1.10). If we transform coordinates $\phi_i = \xi_i$, $i = 1, \ldots, m$ and introduce a function $w(\xi_i)$ defined by

$$\phi\left(x_1, x_2, \ldots, x_m\right) + w\left(\xi_1, \xi_2, \ldots, \xi_m\right) = x_1\xi_1 + x_2\xi_2 + \ldots, x_m\xi_m. \tag{5.4.3}$$

$$\xi_i = \frac{\partial \phi}{\partial x_i}, \quad x_i = \frac{\partial w}{\partial \xi_i}, \qquad \forall\, i. \tag{5.4.4}$$

The second derivatives ϕ_{ij} transform to derivatives of w according to the equation $\Phi W = \mathbb{I}$, where Φ and W are the Hessian matrices of ϕ and w respectively. Then,

assuming that Φ is invertible, and

$$\frac{\partial^2\phi}{\partial x_i\partial x_j} = \left(W^{-1}\right)_{ij}, \quad \frac{\partial^2 w}{\partial\xi_i\partial\xi_j} = \left(\Phi^{-1}\right)_{ij}. \tag{5.4.5}$$

Then (5.4.2) is transformed to,

$$\sum_{i,j}\left(\xi_i\xi_j + f\left(\sum\xi_k^2\right)\eta_{ji}\right)\frac{\partial^2 w}{\partial\xi_i\partial\xi_j} = 0, \tag{5.4.6}$$

a linear second order equation for w. Introducing the radial variable $\rho = \sqrt{\sum\xi_k^2}$ this equation takes the form

$$\left[\left(\rho^2 + f(\rho)\right)\frac{\partial^2}{\partial\rho^2} + \frac{mf(\rho)-1}{\rho}\frac{\partial}{\partial\rho} + \frac{f(\rho)}{2\rho^2}\sum_{i,j}\left(\xi_i\frac{\partial}{\partial\xi_j} - \xi_j\frac{\partial}{\partial\xi_i}\right)^2\right]w = 0. \tag{5.4.7}$$

Here we recognise a generalised total angular momentum operator,

$$\sum_{i<j}\left(\xi_i\frac{\partial}{\partial\xi_j} - \xi_j\frac{\partial}{\partial\xi_i}\right)^2,$$

whose eigenfunctions are just harmonic functions on the $m-1$ sphere, with eigenvalues $-k(k+m-2)$, k integral. Then solutions can be found by the method of of separation of variables as $w = \sum_k F_k(\rho)\times$(general harmonic of degree k), where $F_k(\rho)$ is a solution to the ordinary differential equation,

$$\left(\rho^2 + f(\rho)\right)\frac{d^2F_k}{d\rho^2} + \frac{mf(\rho)-1}{\rho}\frac{dF_k}{d\rho} - k(k+m-2)\frac{f(\rho)}{\rho^2}F_k = 0. \tag{5.4.8}$$

Given such a solution, the x variables can be recovered parametrically in terms of ξ_j from $x_i = \frac{\partial w}{\partial\xi_i}$. It remains to use the definition of w in terms of ϕ and eliminate the variables ξ_j and hence solve for ϕ.

This is not necessarily a practical proposition, as the following example may show. In 2+1 dimensions (5.4.2) takes the form,

$$\varepsilon_{ijk}\varepsilon_{\alpha\beta\gamma}\phi_{i\alpha}\phi_{j\beta}\left(\phi_\gamma\phi_k - f(\rho)\eta_{\gamma k}\right) = 0. \tag{5.4.9}$$

It seems that the most convenient example we can discuss is given by the choice $f(\rho) = -\rho^2$ rather than just a constant, λ. (Choosing a multiple of ρ^2 fails to remove

the cubic singularity that occurs with the first derivative.) Then, the ODE (5.4.8) becomes:

$$\frac{3\rho^2+1}{\rho}\frac{dF_k}{d\rho} - k\,(k+1)\,F_k = 0. \tag{5.4.10}$$

The solution F_k is:

$$F_k = F_{0_k}\left(3\rho^2+1\right)^{\frac{k(k+1)}{6}}, \tag{5.4.11}$$

which permits a general solution for ω. We found no invertible solution which provided an explicit expression for $\phi\,(x_1, x_2, x_3)$.

The construction of an infinite set of conservation laws for the equation (5.4.9) has also defied analysis as yet. It seems reasonable to expect that the procedure based on (5.1.12) should generalise to the deformed equation (5.4.9), but we have not been able to accomplish this.

Finally, let us briefly examine the Lagrangian properties of our generalised Born–Infeld model. It is shown in [3] that (5.4.2) arises as the penultimate element of an Euler hierarchy, in precisely the same manner as the Universal equation with the exception that the starting and intermediate Lagrangians are unique. Suppose we construct a generic Euler hierarchy based on the unique starting Lagrangian,

$$\mathcal{L}_0 = \sqrt{\phi_i\phi_j\eta_{ij} + \lambda}, \tag{5.4.12}$$

and with all the subsequent F_k equal to $\mathcal{L}_0$. Then, from the work in Subsection 5.3.1 we know that the penultimate equation in the hierarchy is given by,

$$\Delta_{m-1} = \mathrm{tr}\left(M^{\dagger}H^{\dagger}\right), \tag{5.4.13}$$

where M is the type of matrix defined in (5.3.12). In this case, the form of M is:

$$M_{ij} = \frac{2\,(\lambda + \mathrm{tr}\,(\eta G))\,\eta_{ij} - 2\eta_{ik}\eta_{jl}G_{kl}}{(\lambda + \mathrm{tr}\,(\eta G))^{\frac{3}{2}}}. \tag{5.4.14}$$

Studying particular examples motivates the following ansatz for the adjoint of M,

$$M^{\dagger}_{ij} = \frac{\alpha}{(\,\mathrm{tr}\,(\eta G) + \lambda)^{\beta}}\,(G_{ij} + \lambda\eta_{ij}), \tag{5.4.15}$$

where α and β are some constants which need not be determined. It is fairly straightforward matter (remembering the properties of G and η) to show that this ansatz is correct, by demonstrating that the product $MM^{\dagger}$ is a multiple of the identity. Substituting (5.4.15) back into (5.4.13) gives a source form that is clearly equivalent to (5.4.2) with the arbitrary function f given the constant value λ.

In the case where the dimension of space time is two, $\mathcal{L}_0$ is simply the Lagrangian for the standard Born–Infeld equation (4.1.7).

To summarise, (5.4.2) is constructed to be a linearisable, Lorentz–invariant equation dependent on an arbitrary number of space–time variables. Its linearisability has been explicitly demonstrated and the form of its solutions has been outlined. In addition, the equation seems to be derived from a variational principle that is part of an Euler hierarchy.

5.5 Hamiltonian Formulation

In this section we will briefly consider candidates for a Hamiltonian formulation of the Bateman–type Universal equations (5.1.4, 5.1.5, 5.1.6, 5.1.9) in an arbitrary number of dimensions. There are a number of difficulties associated with this task. The first is that the singularity of the initial Lagrangian may lead (in some cases) to gauge–type freedom of the resulting Hamiltonian structure. The second problem is that the linear point invariances of the equations mean that there is no preferred time coordinate; as well as Lorentz invariance, there is Euclidean invariance and so any splitting of the coordinates of the base manifold into "space" and "time" variables is utterly artificial. Therefore, we would like to have a covariant analogue of the standard Hamiltonian picture. Third, we have no convenient "conserved current" type of formalism like that used in Section 4.3 to elucidate the Hamiltonian structure of the Bateman equation, and so the relation to the Lagrangian formalism is not entirely clear.

It is also worth remarking that our ideas about biHamiltonian structure do not generalise easily to higher–dimensional theories (see [59] for comments), although

biHamiltonian structures have been postulated for the Self–Dual Yang–Mills [60] and Anti-Self-Dual Einstein [61] models. However, in those cases knowledge of zero–curvature formulations and other structures provide clues to the possible structure of the recursion operator and hence the Poisson structures. There is no such clue available to us here.

As it happens, finding a covariant Hamiltonian language helps resolve all three problems mentioned above, and allows us to bypass the question of biHamiltonian structure. Such a theory exists in a number of different forms due to people like Cartan, Weyl, De Donder and, more recently, G. Sardanashvily. In addition to being nicely covariant, it has the added advantage of defining a finite–dimensional analogue of phase–space, so we will not need the material in Chapter 2 on infinite–dimensional Hamiltonian systems, and an analogue of the Legendre transform of classical mechanics will provide us with the link to the Lagrangian picture. A comprehensive account of the various guises of the theory may be found in Sardanashvily's lectures [10]. The form we will use here will be referred to as the Hamilton–Cartan formalism, for which the seminal modern reference is [62], although we will rely on the much more readable papers by Shadwick [63, 64, 65, 66]. Other useful references are Dickey [67] and Kanatchikov [68].

We start by calculating the covariant canonical variables for all the equations of the Bateman hierarchy. The Hamilton–Cartan formalism assigns to each field a set of conjugate momenta. In our case, the momenta conjugate to ϕ are given by,

$$\pi_k^I = \frac{\delta \mathcal{L}_k}{\delta \phi_I}, \ |I| = 1, 2; \tag{5.5.1}$$

or, more concretely:

$$\pi_k^i = \frac{\partial \mathcal{L}_k}{\partial \phi_i} - \partial_j \frac{\partial \mathcal{L}_k}{\partial \phi_{ij}}, \tag{5.5.2}$$

$$\pi_k^{ij} = \frac{\partial \mathcal{L}_k}{\partial \phi_{ij}}. \tag{5.5.3}$$

As before, $\mathcal{L}_k$ is the Lagrangian of the $(k+1)$th equation of the hierarchy. In general, the conjugate momenta can carry a multi-index of length equal to the order of the

Lagrangian. In our case, all the Lagrangians only depend on second order derivatives at most.

Given these new variables we can ask if there is an analogue of the Legendre transform which maps the variational picture to the Hamiltonian one without degeneracy. In order to answer this we need to discuss the idea of regularity of the Lagrangian in this new formalism. The relevant definition, to be found in [66], is the following: we require all of the matrices,

$$\frac{\partial \pi_k^I}{\partial \phi_J},$$

where I is just a uni-index and J is a *tri*-index, and

$$\frac{\partial \pi_k^J}{\partial \phi_K}.$$

where J and K are both *bi*-indices, to have maximal rank. It is not difficult to see that, in this case, the condition,

$$\det\left(\frac{\partial^2 \mathcal{L}}{\partial \phi_A \partial \phi_B}\right) \neq 0, \tag{5.5.4}$$

for *bi*-indices A and B, is sufficient to ensure regularity and conversely.

It should be clear that not all of the Lagrangians of the hierarchy will be regular. We have already decided (Equation (5.2.13)) that the initial Lagrangian $\mathcal{L}_0$ is singular. Similarly, we know that $\mathcal{L}_1$ is homogeneous of degree one in the second derivatives and hence,

$$\phi_{ij} \frac{\partial^2 \mathcal{L}_1}{\partial \phi_{ij} \partial \phi_{kl}} = 0, \tag{5.5.5}$$

by Euler's theorem. From this homogeneous linear system, we can argue that the condition (5.5.4) fails to be satisfied.

We will leave aside those two cases for the moment and concentrate on the regular Lagrangians. Now, as in the usual variational formulation of mechanics, there is a Legendre transform. The generalised Hamilton–Cartan version gives:

$$\mathcal{H}_k = \phi_I \pi_k^I - \mathcal{L}_k. \tag{5.5.6}$$

Using the homogeneity of the Lagrangian and expanding the definitions of the momenta, this becomes:

$$\mathcal{H}_k = -\phi_i D_j \frac{\partial \mathcal{L}_k}{\partial \phi_{ij}} \tag{5.5.7}$$

$$= -\phi_i D_j \pi_k^{ij}. \tag{5.5.8}$$

This analysis applies to all the Lagrangians of the hierarchy.

Now, the substance of the Hamilton–Cartan approach is the following theorem, to be found in its full generality in [66]: *assuming the regularity condition (5.5.4) holds*, the Euler–Lagrange equations corresponding to $\mathcal{L}_k$ are equivalent to the covariant Hamiltonian equations.

$$\phi_i = \frac{\partial \mathcal{H}_k}{\partial \pi_k^i}, \qquad \text{(no sum over } k\text{)} \tag{5.5.9}$$

$$\phi_{ij} = \frac{\partial \mathcal{H}_k}{\partial \pi_k^{ij}}, \qquad \text{(no sum over } k\text{)} \tag{5.5.10}$$

$$D_i \pi_k^i = -\frac{\partial \mathcal{H}_k}{\partial \phi}, \tag{5.5.11}$$

$$D_i \pi_k^{ij} = -\frac{\partial \mathcal{H}_k}{\partial \phi_i}. \tag{5.5.12}$$

These equations were first written down in the 1930s by De Donder.

This picture differs in one important respect from the familiar Hamiltonian setup. There, each of the coordinates has assigned to it a single conjugate momentum. Here ϕ is equipped with no less than $\frac{1}{2}m(m+3)$ conjugate variables, which does not sit easily with the idea of a symplectic structure on phase space. Kanatchikov [68] suggests a resolution of this paradox, at least in the case of first order Lagrangians. His idea is that to each field y^a in a first order Lagrangian field theory, one should assign the $(m-1)$-form,

$$\pi_a = \pi_a^i \partial_i \lrcorner\, \omega, \tag{5.5.13}$$

where the π_a^i are the sort of variables defined in (5.5.1) and ω is the volume form on M. Then, the De Donder equations can be written,

$$dy^a = \frac{\partial H}{\partial \pi_a^i} dx^i,$$

$$d\pi_a = -\frac{\partial H\omega}{\partial y^a}. \tag{5.5.14}$$

Kanatchikov goes on to define what he calls a *polysymplectic form.* We will not elaborate on this remark, which is merely intended to point out that this apparent paradox may have a general resolution.

The question that arises naturally at this stage is the definition of a Poisson bracket. Again, there are as many answers to this problem as there are authors, but a common thread seems to be to make use of an analogue of the well known isomorphism between the Poisson bracket and the Lie bracket of Hamiltonian vector fields. We have determined a Lie algebra of infinitesimal symmetries of the Universal field equations. This induces a candidate Poisson bracket between the associated conserved currents. Shadwick [65] defines the Poisson bracket between equivalence classes of conserved currents as follows: given two forms $\xi, \zeta \in \Omega^{(n-1,0)}$ that are conserved currents in the sense of (2.4.18), the Poisson bracket between the associated equivalence classes of currents, denoted $[\xi]$ and $[\zeta]$ respectively, is defined by:

$$\{[\xi],[\zeta]\} = [X \lrcorner\, d\zeta], \tag{5.5.15}$$

where X is the vector whose characteristic is the characteristic of ξ. Kanatchikov's approach [68] results in a slightly more complicated algebraic structure, but the basic idea is the same.

From the work in Sections 5.3 and 5.2, we know that the Bateman Universal equation possesses an infinite number of conservation laws. It would be nice to be able to prove the complete integrability of the equation in the Hamiltonian sense by demonstrating the involution of these laws. The integrability of the Sine–Gordon equation has been demonstrated in this manner in [64]. Unfortunately, without an explicit expression for the currents themselves (not just their characteristics), this seems difficult. Nonetheless, it should be true in principle.

Finally, we end by remarking that the constrained case when the Lagrangian is singular is a little more complicated. Sardanashvily [10] argues that in this case the De Donder equations are not appropriate and that a somewhat different definition of

the Legendre transformation should be used. The application of his ideas to this case remains unexplored.

Chapter 6

Conclusions and Outlook

Let us briefly look at the main conclusions of this thesis, and the prospects for further research.

6.1 Conclusions

Throughout this work, we have been interested in the properties of partial differential equations that are integrable in at least one of two senses: they may be linearisable by a suitable change of variables, or possess an infinite sequence of commuting conserved quantities. We were interested in how these methods were connected to the symmetry algebras of the equations, and in any Lagrangian or Hamiltonian structure.

In Chapter 3, we looked at a series ansatz for the solution of some well-known p.d.e s that are integrable by one or other of the definitions above. The ansatz separated the equations into nested sequences of o.d.e s which we were able to solve and subsequently sum the solutions by *ad hoc* methods. It turns out that the greater part of this work is subsumed into that of Rosales [2], apart from the application to the nonlinear wave equation (3.1.1), which is almost the simplest case. The chapter includes some speculation about the connection to the symmetry properties of the equations.

In Chapter 4, we looked at two equations which are (1+1)-dimensional analogues

of equations studied in Chapter 5 — the Bateman and Born–Infeld equations. Here, we generalised the Lagrangian derivation of the Bateman equation to include arbitrary dependence on the field ϕ, and introduced a new argument to solve the Born–Infeld equation and demonstrate the existence of an infinity of conservation laws.

We went on to find a condition on the first order generalised symmetries of both equations, and used an argument due to Kumei and Bluman to demonstrate that this condition explains the linearisability of both equations. In the case of the Bateman equation, we were able to explictly solve this condition to find the form of all possible first–order generalised symmetries.

Finally, the biHamiltonian structure of the Bateman equation was explained using a formalism due to Nutku and collaborators. Fittingly, given our knowledge of its Lagrangian properties, we found that it has an infinite number of such structures. We calculated the associated conserved Hamiltonians.

Chapter Five attempted to generalise some of these results to arbitrary space–time dimension, and additionally to explain the Euler hierarchy process by which the Universal equations are derived. We began by finding the first order generalised symmetries of both the generic and Bateman–type Universal models. In the Monge–Ampère case, any function of the first derivatives is the characteristic of a generalised symmetry. In the Bateman case, the result is a straightforward generalisation of the work in Chapter 4: any function homogeneous of degree zero or one in the first derivatives characterises a generalised symmetry.

We used these symmetries to motivate our approach to the Euler hierarchy. We began by attempting to include explicit field dependence into the members of the hierarchy. It turns out that this is only possible in the Bateman case. In that case we further proved, using an extension of the Cayley–Hamilton theorem, that the resulting equations are algebraically very similar to those in the ϕ-independent case, and that they terminate in an identical universal equation.

Given our characterisation of the generalised symmetries and the Euler hierarchy, it is possible to interpret the iterative Euler process as the successive application of

Lie derivatives to source forms in the Euler–Lagrange complex mentioned in Chapter 2.

In order to connect with Chapter 4, we introduced a multi-dimensional analogue of the Born–Infeld equation that is linearisable by a Legendre transform. The equation is also Lorentz invariant. The linearisation was carried out explicitly, and the form of the solution was explained. It also turns out that the model may be derived from a generic Euler hierarchy.

The chapter concluded with some general speculation on the possible form that a Hamiltonian description of the equations might take.

6.2 Further Research

The outstanding problems arising from Chapter 3 have already been mentioned there. For what classes of equation is the ansatz (3.2.1) useful? Does it constitute a separable solution? If so, does it have an explanation in terms of the symmetries of the equation of interest? Furthermore, is it important or desirable for the equations to have some underlying Hamiltonian structure?

The Universal equations are still quite poorly understood. The priority must be to find more examples of such equations. The obvious starting point is the work already done by Fairlie and Govaerts on the multi-component field generalisations of the model. What are the generalised symmetries of such models? Do they imply linearisability or some other route to integrability? Can we introduce explicit field dependence, and does the Lie derivative interpretation hold for these equations? Even more generally, we might speculate about the existence of models formulated in variables of higher, or half-integral spin. Questions about the Hamiltonian integrability of these problems appear to be very involved and may have to await the introduction of a larger body of theory.

Given a larger stable of examples, we might profitably consider questions of principle, such as what the Lie derivative iterations tell us about the equations that they

yield. An even longer term goal could be the formulation of a quantum version of such models. Finally, there is still no known physical application of these ideas — finding one might transform the problem from a pretty but pointless exercise into one of interest to applied researchers.

Appendix A

MAPLE code to calculate Stokes' coefficients

```
#This calculates the coefficients up to arbitrary order for u_t=uu_x#

v[1]:=c[1];
u[1]:=v[1]*exp(a*t);
#THIS IS A SUB-PROCEDURE#
g:=proc(m)
          local p,q,r,s,y,b,j;
             p:=sum(i*b[i]*t^(i-1),i=1..m-1);
             q:=expand(sum(v[l]*(m-l)*v[m-l],l=1..m-1));
             for j from 1 by 1 to m-1 do
              assign(r[j]=coeff(p,t,j-1));
              assign(s[j]=coeff(q,t,j-1));
              assign(y[j]=solve(r[j]=s[j],b[j]));
             od;
             assign(v[m]=(c[m]+sum(y[h]*t^h,h=1..m-1)));
             assign(u[m]=v[m]*exp(m*a*t));
```

```
            print(u[m]);
          end;
#THIS PROCEDURE CALCULATES THE COEFFICIENTS U[j]#
f:=proc(n)
           local k;
           for k from 2 by 1 to n do
            g(k);
           od;
           end;
#END

#This is for the KdV equation#

u[2]:=c[2]*z^8;
h:=proc(nn)
     for mm from 2*nn+1 by 1 to 40 do
        assign(c[mm]=0);
     od;
     for kk from 1 by 2 to 2*nn-1 do
        assign(c[kk]=0);
     od;
     for kkk  from 3 by 2 to 40 do
        assign(u[kkk]=0);
     od;
     end;
g:=proc(m)
      local p,q,r,s,y,a,j;
       p:=sum(a[i]*z^i,i=1..m^3-1);
       q:=expand(6*sum(u[l]*(m-l)*u[m-l],l=1..m-1));
```

```
        for j from 1 by 1 to m^3-1 do
          assign(r[j]=coeff(p,z,j));
          assign(s[j]=coeff(q,z,j));
          assign(y[j]=solve(r[j]=s[j],a[j]));
        od;
        assign(u[m]=c[m]*z^(m^3)+sum((y[h]*z^h)/(m^3-h),h=1..m^3-1));
#        print(u[m]); #[OPTIONAL]
       end;
f:=proc(n,xx)
         local k;
         h(n);
         for k from 4 by 2 to xx do
           g(k);
         od;
         u1:=simplify(expand(sum(u[o]*b^o,o=1..xx)));
         end;
#END
```

Appendix B

The Inverse Scattering Method

The aim of the inverse scattering technique is to find solutions of nonlinear equations like the KdV (3.1.3). In each case, the line of attack is via the *Sturm–Liouville* or *Schrödinger* problem:

$$\psi_{xx} - (\lambda - u(x))\psi = 0. \tag{B.0.1}$$

Here, u is interpreted as a potential function which determines the shape of the functions ψ.

The construction of solutions to this equation (given a particular u) is a so–called *scattering problem*, and is a standard exercise in quantum mechanics, for example. The form of the solutions is well known. If the constant λ is greater than zero, then the ψ has an oscillatory behaviour described by a continuous spectrum of the form:

$$\hat{\psi}(x;k) \sim \begin{cases} e^{-ikx} + b(k)e^{ikx} & \text{as } x \to +\infty \\ a(k)e^{ikx} & \text{as } x \to -\infty \end{cases} \tag{B.0.2}$$

where $k = \sqrt{\lambda}$. On the other hand, if $\lambda < 0$, the spectrum is discrete and non-oscillatory:

$$\psi_n(x) \sim c_n exp(-\kappa_n x) \text{ as } x \to +\infty, \tag{B.0.3}$$

with $\kappa = \sqrt{-\lambda}$, and the index n ranging from one to some integer N.

Generally speaking, the cases of interest are those where u decays rapidly at positive and negative infinity, but need not be smooth. The ψ is usually assumed

to be smooth and, in the case of the discrete eigenfunctions, square integrable. This allows the c_n to be fixed by normalisation.

The crucial observation is that it is actually possible to reconstruct the function u from knowledge of the scattering data (B.0.2) and (B.0.3). To justify this, we will appeal to an analogy with the wave equation which can be viewed as a limiting case of our problem when u is small. The complete argument can be found in [26]. If we consider the two-dimensional wave equation for a field ϕ,

$$\phi_{xx} - \phi_{zz} = 0, \tag{B.0.4}$$

then we can Fourier transform this using.

$$\hat{\psi}(x;k) = \int_{-\infty}^{\infty} \phi(x,z)e^{ikz}dz. \tag{B.0.5}$$

The transformed problem is the eigenfunction equation:

$$\hat{\psi}_{xx} + k^2\hat{\psi} = 0. \tag{B.0.6}$$

If we wish to find solutions to this which behave like $\hat{\psi} \sim e^{ikx}$, we can do this by prescribing,

$$\phi(x,z) = \delta(x-z) + K(x,z), \tag{B.0.7}$$

where K is a solution of (B.0.4) which is identically zero if $z < x$. Then the inverse Fourier transform gives,

$$\hat{\psi}(x;k) = e^{ikx} + \int_x^{\infty} K(x,z)e^{ikz}, \tag{B.0.8}$$

and so $\hat{\psi}$ has the right behaviour.

Our problem is essentially this one with the embellishments that u is not negligible and so the eigenvalue problem is complicated, and our prescription for ψ is different. If we try a similar tack with this more involved situation [26], we end up with the following result: if we prescribe the behaviour (B.0.2,B.0.3), then we can find,

$$u(x) = -2\frac{d}{dx}K(x,x), \tag{B.0.9}$$

where K satisfies the *Marchenko* equation,

$$K(x,z) + \Omega(x+z) + \int_x^\infty K(x,y)\Omega(x+y)dy = 0, \tag{B.0.10}$$

and,

$$\Omega(X) = \sum_{n=1}^{N} c_n^2 \exp(-\kappa_n X) + \frac{1}{2\pi}\int_{-\infty}^{\infty} b(k)e^{ikX}dk. \tag{B.0.11}$$

What does this have to do with the KdV equation? Let us rewrite (3.1.3),

$$u_t = 6uu_x - u_{xxx}. \tag{B.0.12}$$

and perform the so–called *Miura transformation*,

$$u = v^2 + v_x. \tag{B.0.13}$$

For the moment, we regard time as merely a parameter. A couple of lines of algebra turn (B.0.12) into,

$$\left(2v + \frac{\partial}{\partial x}\right)(v_t - 6v^2 v_x + v_{xxx}) = 0. \tag{B.0.14}$$

We deduce that any solution of the *modified KdV equation* is also a solution of the KdV.

Now, the Miura map (B.0.13) is linearised by the Riccati–type substitution $v = \psi_x/\psi$ to give:

$$\psi_{xx} - u\psi = 0. \tag{B.0.15}$$

Then, using the Galilean boost symmetry calculated in (1.5.16), we can replace u with $u - \lambda$ (since the t-dependence is merely parametric) to give,

$$\psi_{xx} + (\lambda - u)\psi = 0, \tag{B.0.16}$$

our familiar eigenvalue problem.

The inverse scattering technique then proceeds as follows. We are given the initial profile of the solution $u(x,0)$. Then we:

- work out the scattering data for this initial profile using the standard techniques;
- evolve these data in time;

- use the inverse scattering results to find $u(x,t)$.

The method works because the second step is astonishingly simple. It turns out [26] that the c_n, $a(k)$ and $b(k)$ (which must all depend parametrically on t) change according to:

$$\begin{aligned} \frac{dc_n}{dt} &= 4\kappa_n^3 c_n, \\ \frac{da}{dt} &= 0, \\ \frac{db}{dt} &= 8ik^3 b. \end{aligned} \tag{B.0.17}$$

The time evolution follows easily.

To find the solitons of the KdV equation, the method requires only that a special initial profile be chosen. This is a sech^2 function of x, and it has the advantage of making the the reflection function $b(k)$ zero. This simplifies the solution of the Marchenko equation by the Neumann expansion (3.4.14).

Appendix C

REDUCE code for Jacobi identities

Program to calculate functional multivector corresponding to the Jacobi identity

```
comment set up dependence of u and v on x;
depend u,x$
depend v,x$

ux:=df(u,x)$
vx:=df(v,x)$

comment assign derivatives of the Hamiltonian;
depend h1,v$

m:=u*df(h1,v)$
n:=0$

depend h2,v$
```

```
p:=h2+1/2$
q:=p-h1$

comment define the chain rule for differentiation;
let {df(h2,x)=df(h2,v)*vx,df(h1,x)=df(h1,v)*vx}$

comment define the rules for functional 1-forms;
noncom f,g,fx,gx,fxx,gxx$
for all z let df(f(z),z)=fx(z)$
for all z let df(g(z),z)=gx(z)$
for all z let df(fx(z),z)=fxx(z)$
for all z let df(gx(z),z)=gxx(z)$
for all z let f(z)*f(z)=0$
for all z let g(z)*g(z)=0$
for all z let fx(z)*fx(z)=0$
for all z let gx(z)*gx(z)=0$
for all z let fxx(z)*fxx(z)=0$
for all z let gxx(z)*gxx(z)=0$
for all z let f(z)*g(z)=-g(z)*f(z)$
for all z let f(z)*fx(z)=-fx(z)*f(z)$
for all z let f(z)*gx(z)=-gx(z)*f(z)$
for all z let f(z)*fxx(z)=-fxx(z)*f(z)$
for all z let f(z)*gxx(z)=-gxx(z)*f(z)$
for all z let g(z)*fx(z)=-fx(z)*g(z)$
for all z let g(z)*gx(z)=-gx(z)*g(z)$
for all z let g(z)*fxx(z)=-fxx(z)*g(z)$
for all z let g(z)*gxx(z)=-gxx(z)*g(z)$
for all z let fx(z)*gx(z)=-gx(z)*fx(z)$
for all z let fx(z)*fxx(z)=-fxx(z)*fx(z)$
```

```
for all z let fx(z)*gxx(z)=-gxx(z)*fx(z)$
for all z let gx(z)*fxx(z)=-fxx(z)*gx(z)$
for all z let gx(z)*gxx(z)=-gxx(z)*gx(z)$
for all z let fxx(z)*gxx(z)=-gxx(z)*fxx(z)$

comment define the Poisson operator;
m1:=mat((m,p),(p,n))$
m2:=mat((df(m,x),df(p,x)+df(q,x)),(df(p,x)-df(q,x),df(n,x)))$

f1:=mat((fx(x)),(gx(x)))$
f2:=mat((f(x)),(g(x)))$

jt:=2*m1*f1+m2*f2$

comment calculate the functional multi-vector \phi;
phi:=f2(1,1)*jt(1,1)+f2(2,1)*jt(2,1)$

comment calculate the action of the prolonged Poisson vector-field on
\phi;
pr:=jt(1,1)*df(phi,u)+jt(2,1)*df(phi,v)+df(jt(1,1),x)*df(phi,ux)+
df(jt(2,1),x)*df(phi,vx);

end;
```

It remains to check whether this prolonged multivector is a total derivative or not.

Bibliography

[1] D.B. Fairlie and J.A. Mulvey. A perturbative solution to the KdV equation. Preprint DTP/93/39 (unpublished), *1993.*

[2] R.R. Rosales. *Stud. Appl. Math.* **59** 117–151, (1978).

[3] D.B. Fairlie and J.A. Mulvey. *J. Phys.* **A27** 1317–1324, *1994.*

[4] J.A. Mulvey. BiHamiltonian formulations of the Bateman equation. Preprint DTP/95/51, to be published in *Phys. Lett.* **A**, *1995.*

[5] J.A. Mulvey. Properties of the scalar Universal equations. Preprint DTP/95/53, submitted to *J. Phys.* **A**, *1995.*

[6] D.J. Saunders. *The Geometry of Jet Bundles.* Cambridge University Press, *1989.*

[7] P.J. Olver. *Application of Lie Groups to Differential Equations.* Springer–Verlag, 2nd edition, *1993.*

[8] R. Hermann. *Geometry, Physics, and Systems.* Marcel Dekker, *1973.*

[9] A. Echeverría, M.C. Muñoz Lecanda, and R-R. Narciso. Geometry of Lagrangian first-order classical field theories. Preprint dg-ga/9505004, *1995.*

[10] G. Sardanashvily. Five Lectures on Jet Manifold Theory. Preprint hep-th/9411089, *1994.*

[11] H. Stephani. *Differential Equations: Their Solution Using Symmetries.* Cambridge University Press, *1989.*

[12] P. Winternitz. . In L.A. Ibort and M.A. Rodriguez, editors, *Integrable systems, Quantum Groups, and Quantum Field Theories.* Kluwer, *1993.*

[13] W. Hereman. *Euromath Bull.* **2** 45–79, (1993).

[14] G. Landi, G. Marmo, and G. Vilasi. *J. Math. Phys.* **35** 808–815, *1994.*

[15] I.M. Anderson. *Contemp. Math.* **132** 51–73, *1992.*

[16] I.M. Anderson and J. Pohjanpelto. *Math. Ann.* **299** 191–222, *1994.*

[17] I.M. Anderson. *Archivum Mathematicum (Brno)* **24** 181–202, *1988.*

[18] D. Krupka. . In S. Benenti, M. Francaviglia, and A. Lichnerowicz, editors, *Proc. IUTUM-ISIMM Symposium in Modern developments in analytical mechanics.* Atti della Accademia delle Scienze di Torino, *1983.*

[19] J.E. Marsden and T.S. Ratiu. *Introduction to Mechanics and Symmetry.* Springer–Verlag, *1994.*

[20] P.J. Olver. *Math. Proc. Camb. Phil. Soc.* **88** 71–88, *1980.*

[21] G. Magri. *J. Math. Phys.* **19** 1156–1162, *1978.*

[22] M. Arik, F. Neyzi, Y. Nutku, P.J. Olver, and J. Verosky. *J. Math. Phys.* **30** 1338–1344, *1989.*

[23] D.B. Fairlie. Equations of hydrodynamic type. Durham Preprint DTP/93/31, *1993.*

[24] G.G. Stokes. *Camb. Trans.* **8** 441–473, *1847.*

[25] G.B. Whitham. *Linear and Nonlinear Waves.* Wiley, *1974.*

[26] P.G. Drazin and R.S. Johnson. *Solitons.* C.U.P., *1989.*

[27] A.M. Polyakov. *Nucl. Phys.* **B396** 367–385, *1992.*

[28] C.S Gardner, J.M. Greene, M.D. Kruskal, and R.M. Miura. *Phys. Rev. Lett.* **19** 1095–1097, *1967.*

[29] R. Hirota. *Phys. Rev. Lett.* **27** 1192–1194, *1971.*

[30] M. Wadati and K. Sawada. *J. Phys. Soc. Japan* **48** 312–318, *1980.*

[31] S. Kumei and G. Bluman. *SIAM J. Appl. Math.* **42** 1157–1173, *1982.*

[32] L.S. Brown. *Quantum Field Theory.* Cambridge University Press, *1992.*

[33] D.B. Fairlie, J. Govaerts, and A. Morozov. *Nucl. Phys.* **B373** 214–232, *1992.*

[34] H. Bateman. *Proc. Roy. Soc. London* **A125** 598–618, *1929.*

[35] D.B. Fairlie and J. Govaerts. *Phys. Lett.* **281B** 49–53, *1992.*

[36] D.B. Fairlie and J. Govaerts. *J. Math. Phys.* **33** 3543–3546, *1992.*

[37] D.B. Fairlie and J. Govaerts. *J. Phys.* **A26** 3339–3347, *1993.*

[38] I.M. Anderson and T.E Duchamp. *J. Diff. Eq.* **51** 1–47, *1984.*

[39] M. Born and A. Infeld. *Proc. Roy. Soc.* **A224** 425–450, *1934.*

[40] B.M. Barbashov and N.A. Chernikov. *Sov. Phys. JETP* **24** 93–102, *1966.*

[41] R Courant and D Hilbert. *Methods of Mathematical Physics*, volume II. New York: Interscience, *1962.*

[42] Y. Nutku and P.J. Olver. *J. Math. Phys.* **29** 1610–1619, *1988.*

[43] Y. Nutku. *J. Math. Phys.* **26** 1237–1242, *1985.*

[44] W. Klingenberg. *A Course in Differential Geometry.* Springer–Verlag, (1978).

[45] A.N. Leznov. Local solutions of the d'Alembert and double d'Alembert equations in three and four dimensions. Protvino Preprint, *1994.*

[46] I.M. Anderson and J. Pohjanpelto. *Math. Ann.* **301** 627–653, *1995.*

[47] D.R. Grigore. *J. Phys.* **A28** 2921–2937, *1995.*

[48] D.R. Grigore. *J. Phys.* **A28** L49–L57, *1995.*

[49] V.V. Nesterenko. *J. Phys.* **A22** 1673–1687, *1989.*

[50] M.L. Curtis. *Abstract Linear Algebra.* Springer–Verlag, *1990.*

[51] B.M. Barbashov and N.A. Chernikov. *Sov. Phys. JETP* **24** 437–442, *1967.*

[52] M. Bordemann and J. Hoppe. *Phys. Lett.* **317B** 315–320, *1993.*

[53] M. Bordemann and J. Hoppe. *Phys. Lett.* **325B** 359–365, *1994.*

[54] C.N. Yang. *Phys. Rev. Lett.* **38** 1377–1379, *1977.*

[55] R.S. Ward. *Phys. Lett.* **61A** 81–82, *1977.*

[56] W.T. Shaw. *Class. Quantum Grav.* **2** L113–L119, *1985.*

[57] D.B. Fairlie and C.A. Manogue. *Phys. Rev.* **D34** 1832–1834, *1986.*

[58] D.B. Fairlie and C.A. Manogue. *Phys. Rev.* **D36** 475–489, *1987.*

[59] I. Dorfman. *Dirac Structures and Integrability of Nonlinear Evolution Equations.* Wiley, *1993.*

[60] J. Schiff. Self–Dual Yang–Mills and the Hamiltonian structures of integrable systems. Preprint IASSNS-HEP 92/34, hep-th/9211070, *1992.*

[61] I.A.B. Strachan. *J. Math. Phys.* **36** 3566–3573, *1995.*

[62] H. Goldschmidt and S. Sternberg. *Ann. Inst. Fourier (Grenoble)* **23.1** 203–267, *1973.*

[63] W.F. Shadwick. *Rep. Math. Phys.* **18** 243–256, *1982.*

[64] W.F. Shadwick. *Lett. Math. Phys.* **4** 241–248, *1980.*

[65] W.F. Shadwick. *Lett. Math. Phys.* **5** 137–141, *1981.*

[66] W.F. Shadwick. *Lett. Math. Phys.* **6** 409–416, *1982.*

[67] L.A. Dickey. *Soliton Equations and Hamiltonian Systems.* World Scientific, *1991.*

[68] Igor V. Kanatchikov. Basic structures of the covariant canonical formalism for fields based on the De Donder–Weyl theory. Preprint hep-th/9410238, *1994.*