INVARIANT THEORY, EQUIVALENCE PROBLEMS, AND THE CALCULUS OF VARIATIONS *

PETER J. OLVER†

Abstract: This paper surveys some recent connections between classical invariant theory and the calculus of variations, stemming from the mathematical theory of elasticity. Particular problems to be treated include the equivalence problem for binary forms, covariants of biforms, canonical forms for quadratic variational problems, and the equivalence problem for particle Lagrangians. It is shown how these problems are interrelated, and results in one have direct applications to the other.

1. Introduction

My mathematical researches into elasticity and the calculus of variations over the past eight years have led to several surprising connections with classical invariant theory. My original motivation for pursuing classical invariant theory arose through a study of existence theorems for non-convex problems in the calculus of variations of interest in nonlinear elasticity. This theory, due to John Ball, relies on a complete classification of all null Lagrangians, which are differential polynomials whose Euler-Lagrange equations vanish identically, or, equivalently, can be written as a divergence. The basic classification tool is a transform, analogous to the Fourier transform from analysis, originally introduced by Gel'fand and Dikii, [8], in their study of the Korteweg-deVries equation, and developed by Shakiban, [31]. This reduced the original problem to a question about determinantal ideals, which had been answered in fairly recent work in commutative algebra; see Ball, Currie and Olver, [1]. I further noticed that, when the relevant functions involved were homogeneous polynomials, the transform coincided with the classical symbolic method of classical invariant theory, but had the advantage of being applicable even when the functions were not polynomials, leading to a "symbolic method" for the "invariant theory of analytic functions". Moreover, even in the classical case of polynomials, the transform provides a ready mechanism for determination of the expression for classical covariants and invariants in terms of partial derivatives of the form, a significant problem mentioned by Kung and Rota, [20]. However, I have already written a survey of these results and applications, [25], and, as I want to discuss several more recent connections between classical invariant theory and the calculus of variations, space limitations will preclude any further presentation of this range of ideas.

Following upon these results, in the summer of 1981, John Ball asked me whether there were any conservation laws for nonlinear elasticity beyond the classical conservation laws of energy, momentum, etc. I set out to answer his question, but soon realized that the linear theory was not in all that great shape. A detailed study, [22], of the equations of linear isotropic elasticity in both two and three dimensions revealed new classes of conservation laws (despite claims in the literature to the contrary). The next logical

* Research supported in part by NSF Grant DMS 86-02004.

† School of Mathematics, University of Minnesota, Minneapolis, MN, USA 55455

step was to extend these results to anisotropic elasticity, but here a direct attack on the equations in physical coordinates proved to be too complicated. I hit upon the idea of employing some a priori change of variables, which would have the effect of placing the general quadratic variational problem into a much simpler canonical form, similar to the well-known canonical forms for ordinary polynomials. It turned out that the full power of classical invariant theory, modified to incorporate the theory of "biforms", was required to effect the classification of canonical forms for quadratic Lagrangians, and hence of linear elastic media, [28]. Importance consequences of this approach included the determination of "canonical elastic moduli", reducing the number of constants required to characterize and simplify the behavior of anisotropic elastic materials, [27], and a complete classification of conservation laws for arbitrary anisotropic planar elastic media, in which it was shown that there are always two infinite families of new conservation laws which depend on two arbitrary analytic functions of two complex variables, [29].

I still had not tried to tackle Ball's original question for nonlinear elasticity, and therefore started searching for an appropriate tool that would handle the nonlinear case as effectively as classical invariant theory had taken care of the linear case. Contemporaneously, I learned of a powerful method introduced by Elie Cartan for answering general (nonlinear) equivalence problems. By definition, an *equivalence* problem is to determine when two given objects, e.g. two polynomials, two differential equations, or two variational integrals, can be mapped into each other by an appropriate change of variables. In 1908, through his pioneering study of Lie pseudogroups, [3], Cartan proposed a fundamentally algorithmic procedure, based on the rapidly developing subject of differential forms, which would completely solve general equivalence problems, leading to necessary and sufficient conditions for equivalence of two objects. In spirit, Cartan's method is very much like classical invariant theory, in that it leads to certain functions of the original objects which are invariants of the problem, and so must have the same values for any two equivalent objects. But, even more, Cartan's theory tells you which invariants are really important as far as the equivalence problem is concerned, and gives the necessary and sufficient conditions for equivalence in terms of a finite number of these invariants. Although Cartan's method is extremely powerful, and received further developments in the 1930's for solving several equivalence problems of interest in differential equations and differential geometry, it never did catch on in the mathematical community at large. Most of the recent work can trace its inspiration back to a paper of R. Gardner in mathematical control theory, [7]. In the last few years, the method has had a number of successful applications to differential equations, [14], [16], calculus of variations, [17], differential operators and molecular dynamics, [18], etc., all of which have pointed to its increasing importance. However, the method still awaits a real popularization in the applied mathematical community as a straightforward, algorithmic, computational tool that will provide explicit and effective necessary and sufficient conditions for the solution of many equivalence problems of current mathematical and applied interest. I am convinced that there are even more profound, practical applications of Cartan's method to both pure and applied mathematics, not to mention physics and engineering, in the offing.

I still have not applied Cartan's method to nonlinear elasticity, although the relevant equivalence problem is now under investigation with Niky Kamran. However, while learning the method, I came across the remarkable observation that the fundamental equivalence problem of classical invariant theory, namely that of determining when two binary forms can be mapped to each other by a linear transformation, could be recast as a special case of a Cartan equivalence problem for a particular type of one-dimensional variational integral, a problem Cartan himself had solved in the 1930's, [5]. Even more surprising is the fact that the Cartan solution of the Lagrangian equivalence problem has, by this connection, profound consequences for classical invariant theory. In particular, a new solution to the equivalence problem for binary forms, based on just three of the associated covariants, as well as new results on symmetry groups and equivalence to monomials and sums of n^{th} powers are direct results of this remarkable connection, cf. [26], [30].

This paper will give a brief overview of these connections between the calculus of variations and classical invariant theory; proofs and more detailed developments of results can be found in the cited literature. This survey begins with a general discussion of equivalence problems, illustrated by several examples from classical invariant theory and the calculus of variations and their inter-relationships. In section 3, we introduce and compare the basic concepts of invariants, covariants and other kinds of invariant objects which make their appearance in the solutions to the equivalence problems discussed, which are presented in section 4. Section 5 includes a discussion of symmetry groups and how they arise from Cartan's approach.

2. Equivalence Problems

The general equivalence problem is to determine when two geometric or algebraic objects are really the same object, re-expressed in different coordinate systems. Of course, there are two underlying questions that must be precisely answered before we can mathematically formulate an equivalence problem: 1. Exactly what do we mean by two objects being the "same"? 2. Which changes of coordinates are to be allowed? Once we have been more precise in the specification of our equivalence problem, we can begin the mathematical analysis of our problem. In this section we briefly present several different types of equivalence problems arising in classical invariant theory and the calculus of variations, and discuss their inter-relationships.

The equivalence problem for forms

By a *form*[1] of *degree n*, we mean a homogeneous polynomial function

$$f(x) = \sum a_I x^I, \tag{2.1}$$

defined for $\mathbf{x} = (x^1, \ldots, x^m)$ in $\mathbb{R}^m$ or $\mathbb{C}^m$, where the sum is over all multi-indices $I = (i_1, \ldots, i_m)$, with $|I| \equiv i_1 + \ldots + i_m = n$, and where $x^I \equiv (x^1)^{i_1} \cdot \ldots \cdot (x^m)^{i_m}$. In classical invariant theory, the appropriate changes of coordinates are provided by the general linear

[1] Here we already encounter the first in a series of conflicting terminologies which will plague our attempts to unite these two fields. In classical invariant theory, a form means a homogeneous polynomial; in Cartan's equivalence method, differential forms are the key objects of interest. Needless to say, these are very different kind of objects. To distinguish them, we will always use "form" for a homogeneous polynomial, whereas "differential form" will always have "differential" in front of it for emphasis.

group $GL(m)$ (meaning either $GL(m,\mathbb{R})$ or $GL(m,\mathbb{C})$), which acts on the variables via the standard linear representation $\mathbf{x} \to A \cdot \tilde{\mathbf{x}}, A \in GL(m)$. The induced action on the forms, which takes $f(\mathbf{x})$ to

$$\tilde{f}(\tilde{\mathbf{x}}) = f(A \cdot \tilde{\mathbf{x}}) = \sum \tilde{a}_I \tilde{x}^I, \tag{2.2}$$

induces an action of $GL(m)$ on the coefficients $\mathbf{a} = (a_I)$ of the form, whose explicit expression is easy to derive, but not overly helpful. Two forms f and $\tilde{f}$ are called (real or complex) *equivalent* if they can be transformed into each other by a suitable element of $GL(m)$, so the basic equivalence problem here is to determine whether or not two given forms can be mapped into each other by a suitable linear transformation.

In the particular case of binary forms, meaning $m = 2$, so $f = f(x, y)$ depends on $\mathbf{x} = (x, y) = (x^1, x^2)$, we can replace $\mathbf{x}$ by the projective coordinate $p = x/y$, and write

$$g(p) = f(p, 1) \tag{2.3}$$

for the corresponding inhomogeneous polynomial. Now, on the projective line, the corresponding group of coordinate changes consists of the linear fractional transformations

$$\tilde{p} = \frac{ap + b}{cp + d}, \tag{2.4}$$

where $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2)$. Two n^{th} degree polynomials g and $\tilde{g}$ are *equivalent* if

$$g(p) = (cp + d)^n \tilde{g}(\tilde{p}) = (cp + d)^n \tilde{g}\left(\frac{ap + b}{cp + d}\right), \tag{2.5}$$

for some $A \in GL(2)$. The equivalence problem for inhomogeneous polynomials is to determine when two given polynomials can be mapped into each other by such a linear fractional transformation.

The equivalence problem for biforms.

A second important type of algebraic equivalence problem is provided by generalizing the considerations of part 1 to what will be called "biforms". By definition, a *biform* of *bidegree* (m, n) is a polynomial function

$$Q(\mathbf{x}, \mathbf{u}) = \sum a_{IJ} x^I u^J, \tag{2.6}$$

depending on two vector variables $(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^p \times \mathbb{R}^q$(or in $\mathbb{C}^p \times \mathbb{C}^q$), which for, fixed $\mathbf{u}$, is a homogeneous polynomial of degree m in $\mathbf{x}$, and, for fixed $\mathbf{x}$, is a homogeneous polynomial of degree n in $\mathbf{u}$. The appropriate changes of variable are the linear transformations in the Cartesian product group $GL(p) \times GL(q)$; the transformation $\mathbf{x} \to A \cdot \tilde{\mathbf{x}}, \mathbf{u} \to B \cdot \tilde{\mathbf{u}}$ maps the biform $Q(\mathbf{x}, \mathbf{u})$ to the biform

$$\tilde{Q}(\tilde{\mathbf{x}}, \tilde{\mathbf{u}}) = Q(A \cdot \tilde{\mathbf{x}}, B \cdot \tilde{\mathbf{u}}) = \sum \tilde{a}_{IJ} \tilde{x}^I \tilde{u}^J.$$

As with forms, two biforms are *equivalent* if they can be mapped to each other by a suitable group element, and we have a similar, but less well investigated, type of equivalence problem. (However, see Turnbull, [34], for a discussion of bilinear forms, and Weitzenböck, [35], for even more general types of polynomials.)

3. *Equivalence problems for variational integrals.* We now consider some of the possible equivalence problems associated with a general problem from the calculus of variations. Consider the integral

$$\mathcal{L}[\mathbf{u}] = \int_\Omega L(\mathbf{x}, \mathbf{u}^{(n)})dx. \tag{2.7}$$

Here the domain of integration is an open subset $\Omega \subset \mathbb{R}^p$, and the Lagrangian L is a smooth function of the independent variables $\mathbf{x} \in \Omega$, the dependent variables $\mathbf{u} \in \mathbb{R}^q$, and their derivatives up to some order n, denoted by $\mathbf{u}^{(n)}$. The typical calculus of variations problem is to find extremals (i.e. minimizers or maximizers) $\mathbf{u} = \mathbf{f}(\mathbf{x})$ for the integral $\mathcal{L}[\mathbf{u}]$ subject to suitable boundary conditions. Here we are more concerned with the integral itself, rather than the specific minimizers. There are at least four different versions of the notion of "equivalence of variational problems", depending on the type of changes of variables allowed, and the precise form the equivalence is to take. First, there are two immediately obvious possible choices of allowable coordinate changes:

1) The *fiber-preserving transformations*, in which the new independent variables depend only on the old independent variables, so the transformations have the form

$$\tilde{\mathbf{x}} = \phi(\mathbf{x}), \qquad \tilde{\mathbf{u}} = \psi(\mathbf{x}, \mathbf{u}). \tag{2.8}$$

2) The general *point transformations*, in which an arbitrary change of independent and dependent variables is allowed, and the transformations have the form

$$\tilde{\mathbf{x}} = \phi(\mathbf{x}, \mathbf{u}), \qquad \tilde{\mathbf{u}} = \psi(\mathbf{x}, \mathbf{u}). \tag{2.9}$$

Other possible classes of coordinate transformations include contact transformations, linear transformations, volume-preserving transformations, etc., etc., but these two are sufficient for our purposes. Furthermore, there are two choices for deciding when two Lagrangians are equivalent:

a) *Standard Equivalence*: Here we require that the two variational problems agree on all possible functions $\mathbf{u} = \mathbf{f}(\mathbf{x})$. This implies that the two Lagrangians are related by the change of variables formula for multiple integrals:

$$L(\mathbf{x}, \mathbf{u}^{(n)}) = \tilde{L}(\tilde{\mathbf{x}}, \tilde{\mathbf{u}}^{(n)}) \cdot \det \mathbf{J}, \tag{2.10}$$

where $\mathbf{J} = (D_i \phi^j)$ is the Jacobian matrix of the transformation. (Here, we are treating $\mathcal{L}[\mathbf{u}]$ as an oriented integral; otherwise we should put an absolute value on the factor $\det \mathbf{J}$.)

b) *Divergence Equivalence:* Here we only require that the variational problems agree on extremals, or, equivalently, that the associated Euler-Lagrange equations are mapped directly to each other by the change of variables. A standard result, [23; Theorem 4.7], says that two Lagrangians have the same Euler-Lagrange equations if and only if they differ by a divergence, so the two Lagrangians must be related by the formula

$$L(\mathbf{x}, \mathbf{u}^{(n)}) = \tilde{L}(\tilde{\mathbf{x}}, \tilde{\mathbf{u}}^{(n)}) \cdot \det \mathbf{J} + \mathrm{Div}\mathbf{F}, \tag{2.11}$$

where $\mathbf{F}(\mathbf{x}, \mathbf{u}^{(m)}) = (F_1, \dots F_p)$ is an arbitrary p-tuple of functions of $\mathbf{x}, \mathbf{u}$ and derivatives of $\mathbf{u}$.

Combining the two notions of equivalence with each of the two classes of coordinate transformations, we are led to four different equivalence problems for Lagrangians, such as the standard point transformation equivalence problem, the divergence fiber-preserving equivalence problem, etc. Depending on the context, each of these problems is important, and warrants a solution. To date, however, only some of the simpler cases, e.g. $p = 1$ or 2, $q = 1$, $n = 1$ or 2, have been looked at, and only the simplest case $p = q = n = 1$ has been solved completely, cf. [17].

4. *The equivalence problem for particle Lagrangians.*

Now we specialize the general discussion on equivalence problems in the calculus of variations to present one particular equivalence problem in detail. A *particle Lagrangian*is one involving only one independent variable, x, and we specialize to first order particle Lagrangians in one dependent variable also, u. Let $p \equiv \frac{du}{dx}$ denote the derivative variable. The equivalence problem is to determine when two first order scalar variational problems

$$\mathcal{L}[u] = \int L(x, u, p) dx, \text{ and } \tilde{\mathcal{L}}[u] = \int \tilde{L}(\tilde{x}, \tilde{u}, \tilde{p}) d\tilde{x},$$

can be transformed into each other by a point transformation

$$\tilde{x} = \phi(x, u), \qquad \tilde{u} = \psi(x, u), \tag{2.12}$$

without any additional divergence terms. Let us see what this entails.

According to the chain rule, if $\tilde{x}, \tilde{u}$ are related to x, u according to (2.12), the change in the derivative p is given by a linear fractional transformation:

$$\tilde{p} = \frac{ap + b}{cp + d}, \tag{2.13}$$

where

$$a = \frac{\partial \psi}{\partial u}, \quad b = \frac{\partial \psi}{\partial x}, \quad c = \frac{\partial \phi}{\partial u}, \quad d = \frac{\partial \phi}{\partial x}. \tag{2.14}$$

Specializing the general transformation rule (2.10), we deduce that equivalent Lagrangians must be related by the basic change of variables formula

$$L(x, u, p) = (cp + d)\tilde{L}(\tilde{x}, \tilde{u}, \tilde{p}) \tag{2.15}$$

under (2.12), (2.13). Thus the original problem from the calculus of variations can be recast as a problem of determining when two functions of three variables are related by the formula (2.15) for some transformation of the form (2.12), (2.13). This is the basic problem solved by Cartan in [5].

A remarkable observation is that the equivalence condition (2.15) for particle Lagrangians and the equivalence condition (2.5) for binary forms are essentially the same! Indeed, if, given a nonhomogeneous polynomial $g(p)$ of degree n, we define the "Lagrangian"

$$L(p) = \sqrt[n]{g(p)} \tag{2.16}$$

then the relevant transformation rules are identical, and so the equivalence problem (2.5) for the polynomial $g(p)$ under the linear fractional transformation (2.4) is the *same* as the equivalence problem for the (x, u)-independent Lagrangian (2.16) under the transformation (2.12), (2.13), (2.15). Therefore, any solution to the Lagrangian equivalence problem immediately induces a solution to the equivalence problem for binary forms. This observation can be extended to forms in more variables, connecting the equivalence problem for forms with an equivalence problem for multi-particle Lagrangians, cf. [26].

5. *Equivalence problems for quadratic Lagrangians.*

Another important special class of equivalence problems from the calculus of variations comes from specializing the general considerations of part 3 to the special case of the divergence equivalence of quadratic variational problems

$$\mathcal{L}[\mathbf{u}] = \int \sum a^{\alpha\beta}_{IJ} \frac{\partial^n u^\alpha}{\partial x^I} \frac{\partial^n u^\beta}{\partial x^J} dx. \tag{2.17}$$

For simplicity, we assume that the coefficients $a^{\alpha\beta}_{IJ}$ are constants. These problems are motivated by the applications to linear elasticity, but they also arise in many other contexts. Mathematically, the relevant subclass of changes of variables that preseves the quadratic form of the Lagrangian are the linear change of variables $\mathbf{x} \to A \cdot \mathbf{x}, \mathbf{u} \to B \cdot \mathbf{u}$, already presented in our discussion of biforms. We define the symbol of a quadratic Lagrangian to be the biform

$$Q(\mathbf{x}, \mathbf{u}) = \sum a^{\alpha\beta}_{IJ} \, \mathbf{x}^I \mathbf{x}^J u^\alpha u^\beta, \tag{2.18}$$

bidegree $(2n, 2)$. It is not hard to show that, since we can add arbitrary divergences to our Lagrangian, the symbol is well-defined, independently of any quadratic divergence which might be added in. Moreover, except for an extra determinantal factor $\det A$ in (2.11), which can always be effectively eliminated by rescaling, and the replacement of A by A^{-1}, the transformation rules for quadratic Lagrangian and those for their symbols are exactly the same. Therefore, we deduce the important fact that quadratic Lagrangians are divergence equivalent under a linear change of variables if and only if their symbols are equivalent as biforms. Thus, this equivalence problem reduces to the previous algebraic equivalence problem.

3. *Invariants.*

Of fundamental importance to the solution of any equivalence problem are certain functions, the invariants, whose values do not change under the change of variables apposite to the problem. These can be appropriately defined for all of the equivalence problems considered above. Additional invariant quantities are provided by relative invariants or covariants, whose values change only by some multiplicative factor, and invariant differential forms, which are the key to Cartan's approach. The terminology here is slightly complicated by the different use of the term "invariant" in the different subject areas. In classical invariant theory, the invariants are distinguished from the more general covariants by the fact that they do not involve the variables $\mathbf{x}$. Furthermore, both invariants and covariants are really relatively invariant functions, as they do change by some multiplicative factor under the action of the general linear group. What we will be calling invariants would be known as absolute invariants in the classical terminology. In Cartan's approach, the distinction between invariants and covariants blurs, and they are all called invariants for simplicity. Thus, the invariants of Cartan's approach would be labelled absolute covariants in the classical invariant theory approach, while the invariants and covariants of classical invariant theory are really relative invariants according to the Cartan terminology. (I hope that this doesn't cause undue confusion for the reader!) Absolute invariants for two equivalent objects must agree identically, while relative invariants only need agree up to a factor. However, the vanishing of a relative invariant is an invariant condition, that often carries important geometric information about the object.

There are two basic methods for constructing invariants. In classical invariant theory, the powerful symbolic method provides a ready means of constructing all the (relative) polynomial invariants and covariants of a form. Hilbert's Basis Theorem says that there are a finite number of polynomially independent covariants for a form of a given degree, but the precise number of independent covariants increases rapidly with the degree n of the form (although this is partially mitigated by the presence of many polynomial syzygies). Indeed, a complete system of covariants has been constructed only for binary forms of degrees $2, 3, 4, 5, 6$, and 8. Despite the constructive methods used to generate the covariants themselves, it is by no means clear which covariants play the crucial role in the equivalence problem. For example, in the case of a binary quartic, there are two important invariants, denoted by i and j, but it is the strange combination $i^3 - 27j^2$ (the discriminant) which provides the key to the classification of quartic polynomials, [12; page 292]. Symbolic techniques can also be applied to construct covariants of biforms, [35], [24], and there is an analogous version of Hilbert's Basis Theorem. However, explicit Hilbert bases for even the simplest biforms are not known.

Cartan's approach provides an alternative method for constructing invariants, this time as functions of the partial derivatives of the basic object. Moreover, Cartan's algorithm automatically determines which invariants are important for the equivalence problem and readily gives necessary and sufficient conditions for equivalence based on the fundamental invariants. (We also note that Maschke, [21], developed a symbolic method for use in equivalence problems in Riemannian geometry. See Tresse, [33], for a treatment of relative invariants of differential equations.) We will see how the two approaches bear on each other in the equivalence problem considered. One conclusion will be that, as far as the equivalence problem is concerned, Cartan's approach is certainly the more powerful of the two.

Consider first the equivalence problem for forms. A classical *covariant* of weight w is a function $J(\mathbf{a}, \mathbf{x})$ depending on the coefficients $\mathbf{a} = (a_I)$ of the form and the variables $\mathbf{x}$, which, except for a determinantal factor, does not change under the action of $GL(n)$ given in (2.1), (2.2):

$$J(\tilde{\mathbf{a}}, \tilde{\mathbf{x}}) = (\det A)^w \cdot J(\mathbf{a}, \mathbf{x}). \tag{3.1}$$

If $w = 0$, we call J an *absolute covariant*; these are the invariants in Cartan's terminology. If a covariant J does not depend on $\mathbf{x}$ it is called an *invariant*, although here we will use the term *relative invariant* (unless $w = 0$).

In the case of a binary form $f(x, y)$, we list some of the most important classical covariants. First, the *Hessian*

$$H = (f, f)^{(2)} = \frac{2}{n^2(n-1)^2}(f_{xx}f_{yy} - f_{xy}^2), \tag{3.2}$$

is a covariant of weight 2 and degree $2n - 4$. The Jacobians

$$\begin{aligned} T = (f, H) &= \frac{1}{2n(n-2)}(f_x H_y - f_y H_x), \\ U = (H, T) &= \frac{1}{6(n-2)^2}(H_x T_y - H_y T_x), \end{aligned} \tag{3.3}$$

are covariants of weight 3 and 6 and degrees $3n - 6$ and $5n - 12$ respectively. In these formulae, the notation $(f, g)^{(k)}$ denotes the k^{th} transvectant of f and g. A classical result states that the k^{th} transvectant of any two covariants is again a covariant, and, moreover, all the polynomial covariants can be constructed using successive transvection, [9], [12].

Turning to the equivalence problem for biforms, we define a *covariant of biweight* (v, w) of a biform (2.6) to be a polynomial function $J(\mathbf{a}, \mathbf{x}, \mathbf{u})$, depending on the coefficients $\mathbf{a} = (a_{IJ})$ and the variables $\mathbf{x}, \mathbf{u}$, which satisfies:

$$J(\tilde{\mathbf{a}}, \tilde{\mathbf{x}}, \tilde{\mathbf{u}}) = (\det A)^v (\det B)^w J(\mathbf{a}, \mathbf{x}, \mathbf{u}), \quad \mathbf{A} = (A, B) \in GL(p) \times GL(q).$$

As with forms, *a relative invariant* is just a covariant which does not depend on the variables $\mathbf{x}$ or $\mathbf{u}$. If $v = w = 0$, then we have an absolute covariant.

For example, consider the first nontrivial case of a biform, a binary biquadratic

$$\begin{aligned} Q(\mathbf{x}, \mathbf{u}) = &a_{11}^{11}x^2u^2 + 2a_{12}^{11}xyu^2 + a_{22}^{11}y^2u^2 + 2a_{11}^{12}x^2uv + 4a_{12}^{12}xyuv + \\ &+ 2a_{22}^{12}y^2uv + a_{11}^{22}x^2v^2 + 2a_{12}^{22}xyv^2 + a_{22}^{22}y^2v^2. \end{aligned} \tag{3.5}$$

Here $\mathbf{x} = (x^1, x^2) = (x, y)$ and $\mathbf{u} = (u^1, u^2) = (u, v)$, and we have incorporated some multi-nomial coefficients to conform with the references. Since Q is a quadratic function of $\mathbf{x}$for each fixed $\mathbf{u}$, it is not hard to see that the discriminant

$$\Delta_{\mathbf{x}}(\mathbf{u}) = \frac{1}{4}(Q_{xx}Q_{yy} - Q_{xy}^2) \tag{3.6}$$

is a covariant of biweight (2,0). Similarly, the $\mathbf{u}$-discriminant

$$\Delta_{\mathbf{u}}(\mathbf{x}) = \frac{1}{4}(Q_{uu}Q_{vv} - Q_{uv}^2) \tag{3.7}$$

is a covariant of biweight $(0,2)$. These discriminants have the usual properties enjoyed by the discriminant of an ordinary quadratic polynomial; for instance, $\Delta_{\mathbf{x}}(\mathbf{u}_0) = 0$ implies that $Q(\mathbf{x}, \mathbf{u}_0)$ is a perfect square, etc. There is a mixed biquadratic covariant of biweight (2,2), which has the explicit formula

$$C_2 = \frac{1}{4}(Q_{xu}Q_{yv} - Q_{xv}Q_{yu}). \tag{3.8}$$

The simplest relative invariant of Q is the quadratic expression

$$I_2(\mathbf{a}) = 2a_{11}^{11}a_{22}^{22} - 4a_{12}^{11}a_{12}^{22} + 2a_{22}^{11}a_{11}^{22} - 4a_{11}^{12}a_{22}^{12} + 4(a_{12}^{12})^2, \tag{3.9}$$

and has biweight $(2,2)$. There is a single cubic invariant, namely

$$I_3(\mathbf{a}) = a_{11}^{11}a_{12}^{12}a_{22}^{22} - a_{11}^{11}a_{22}^{12}a_{12}^{22} - a_{12}^{11}a_{11}^{12}a_{22}^{22} + a_{12}^{11}a_{22}^{12}a_{11}^{22} + a_{22}^{11}a_{11}^{12}a_{12}^{22} - a_{22}^{11}a_{12}^{12}a_{11}^{22}. \tag{3.10}$$

Further invariants and covariants can be found by applying the technique of *composition* of covariants. If Q is any (bi)form, and J is a polynomial covariant for Q, then we can regard J itself as a (bi)form, whose coefficients are certain polynomial combinations of the coefficients of Q. Any covariant K, which depends directly on the coefficients of the new form J, is then, by composition, a covariant of the original biform Q, and is denoted by $K \circ J$. Thus, since the discriminant $\Delta_{\mathbf{u}}(\mathbf{x})$ of a binary biquadratic form is a binary quartic form in the variables $\mathbf{x} = (x, y)$, all the standard covariants of the binary quartic yield, under composition, covariants of the original biquadratic Q. Thus we have the Hessian of the $\mathbf{u}$-discriminant

$$H_{\mathbf{u}}(\mathbf{x}) = H \circ \Delta_{\mathbf{u}} = (\Delta_{\mathbf{u}}, \Delta_{\mathbf{u}})^{(2)}, \tag{3.11}$$

which is again a binary quartic in $\mathbf{x}$, and a covariant of biweight $(2,4)$, as well as the two relative invariants

$$i_{\mathbf{u}} = i \circ \Delta_{\mathbf{u}} = (\Delta_{\mathbf{u}}, \Delta_{\mathbf{u}})^{(4)} \quad \text{and} \quad j_{\mathbf{u}} = j \circ \Delta_{\mathbf{u}} = (\Delta_{\mathbf{u}}, H_{\mathbf{u}})^{(4)}, \tag{3.12}$$

which have biweights (4,4) and (6,6) respectively. Similarly, the Hessian of the $\mathbf{x}$-discriminant

$$H_{\mathbf{x}}(\mathbf{u}) = H \circ \Delta_{\mathbf{x}} = (\Delta_{\mathbf{x}}, \Delta_{\mathbf{x}})^{(2)}, \tag{3.13}$$

has biweight (4,2), and the two relative invariants

$$i_{\mathbf{x}} = i \circ \Delta_{\mathbf{x}} = (\Delta_{\mathbf{x}}, \Delta_{\mathbf{x}})^{(4)} \text{ and } j_{\mathbf{x}} = j \circ \Delta_{\mathbf{x}} = (\Delta_{\mathbf{x}}, H_{\mathbf{x}})^{(4)}, \tag{3.14}$$

have biweights (4,4) and (6,6), respectively. The two Hessians $H_{\mathbf{u}}$ and $H_{\mathbf{x}}$ are easily seen to be different quartic polynomials in general (even if one identifies the variables $\mathbf{x}$ and $\mathbf{u}$). Remarkably, the i and j invariants of the two discriminants are the same invariants of the original biquadratic polynomial Q.

Theorem 1, [24] Let Q be a binary biquadratic form. Let $\Delta_{\mathbf{x}}(\mathbf{u})$ and $\Delta_{\mathbf{u}}(\mathbf{x})$ be the two discriminants, which are quartic forms in $\mathbf{u}$ and $\mathbf{x}$ respectively. Then the invariants of these two quartic forms are the same:

$$i_{\mathbf{x}} = i \circ \Delta_{\mathbf{x}} = i_{\mathbf{u}} = i \circ \Delta_{\mathbf{u}}, \qquad j_{\mathbf{x}} = j \circ \Delta_{\mathbf{x}} = j_{\mathbf{u}} = j \circ \Delta_{\mathbf{u}}.$$

The structure of the roots of the two discriminants is an important invariant of the biquadratic (3.5), and provides the key to the determination of canonical forms and the solution to the equivalence problem. (See below.) For instance, if $\Delta_{\mathbf{x}}(\mathbf{u})$ has two double roots in one coordinate system, then it has two double roots in every coordinate system. Since the discriminant of a quartic, whose vanishing indicates the presence of repeated roots, is given by $i^3 - 27j^2$, cf. [12; page 293], Theorem 1 immediately implies the following interesting interconnection between the root structures of the two discriminants.

Corollary 2 The two discriminants $\Delta_{\mathbf{x}}(\mathbf{u})$ and $\Delta_{\mathbf{u}}(\mathbf{x})$ of a binary biquadratic either both have all simple roots, or both have repeated roots.

Note that it is *not* asserted that $\Delta_{\mathbf{x}}(\mathbf{u})$ and $\Delta_{\mathbf{u}}(\mathbf{x})$ have identical root multiplicities! For example, the biquadratic form $Q = x^2u^2 + xyv^2$ has $\mathbf{u}$-discriminant $\Delta_{\mathbf{u}} = -4x^3y$, which has a triple root at 0 and a simple root at ∞, whereas the $\mathbf{x}$-discriminant $\Delta_{\mathbf{x}} = v^4$ has a quadruple root at ∞. We also note that since the ratio i^3/j^2 essentially determines the cross ratio of the four roots of the quartic, [9], [12], the cross ratios of the roots of the two discriminants $\Delta_{\mathbf{x}}(\mathbf{u})$ and $\Delta_{\mathbf{u}}(\mathbf{x})$ must be the same.

There are at least three possible ways to prove Theorem 1. One is to explicitly write out the invariants $i_{\mathbf{u}}, j_{\mathbf{u}}, i_{\mathbf{x}}$ and $j_{\mathbf{x}}$, and compare terms. This was the original version of the proof, and was effected on an Apollo computer using the symbolic manipulation language SMP. The explicit formula for $j_{\mathbf{u}}$ runs to two entire printed pages! A second approach is to write out the formulas for the invariants in terms of the partial derivatives of the biquadratic form. The final approach is to work entirely symbolically; this last proof is the easiest for hand computation, and can be found in [24].

We now turn to a discussion of the invariants for Lagrangian equivalence problems. For specificity we consider example 4 of section 2 - the standard equivalence problem for a particle Lagrangian under point transformations. We assume that we are at a point where neither L nor the second derivative L_{pp} vanishes. (In particular, we are excluding the elementary affine Lagrangians $a(x,u)p + b(x,u)$.) The simplest invariant of this problem is the rational differential function

$$I = \frac{(LL_{ppp} + 3L_pL_{pp})^2}{LL_{pp}^3}. \tag{3.15}$$

The next most complicated one is

$$J(p) = \frac{2L^2L_{pp}L_{pppp} - 2LL_pL_{pp}L_{ppp} + 6LL_{pp}^3 - 3L_p^2L_{pp}^2 - 3L^2L_{ppp}^2}{2LL_{pp}^3}. \tag{3.16}$$

Below we shall see how both of these formulas are found using the Cartan method.

The numerator and denominator of these two absolute invariants are relative invariants. Indeed, according to (2.15),

$$\tilde{L} = (cp+d)^{-1}L,$$

hence, differentiating using the chain rule and (2.14), we find

$$\tilde{L}_{\tilde{p}\tilde{p}} = (ad-bc)^{-2}(cp+d)^3L,$$
$$\tilde{L}\tilde{L}_{\tilde{p}\tilde{p}\tilde{p}} + 3\tilde{L}_{\tilde{p}}\tilde{L}_{\tilde{p}\tilde{p}} = (ad-bc)^{-3}(cp+d)^5(LL_{ppp} + 3L_pL_{pp}).$$

This proves the invariance of I directly, the corresponding result for J follows after one further differentiation.

So far we have been discussing just scalar-valued invariants. We can consider vector-valued invariants (or even more general quantities) when we have an induced action of the relevant coordinate changes on some other vector space. The most important class of such invariants are the invariant differential forms, where the exterior powers of the cotangent space have the standard induced action under coordinate changes. In the Cartan equivalence method, one begins by encoding the original equivalence problem into a problem involving the mapping of non-invariant differential one-forms on the manifold. For the standard particle Lagrangian equivalence problem under point transformations we are led to introduce the one-forms

$$\omega_1 = L(x,u,p)dx, \qquad \omega_2 = du - p\ dx, \tag{3.17}$$

the first of which is essentially the integrand, and the second of which is known as the *contact form*, which must be preserved (up to multiple) in order that the derivative $p$transform correctly, as in (2.13). We introduce the corresponding one-forms for the transformed Lagrangian $\tilde{L}$:

$$\tilde{\omega}_1 = \tilde{L}(\tilde{x},\tilde{u},\tilde{p})d\tilde{x}, \qquad \tilde{\omega}_2 = d\tilde{u} - \tilde{p}\ d\tilde{x}.$$

Then we have the following reformulation of our basic equivalence problem.

Lemma 3 Two nonvanishing Lagrangians L and $\tilde{L}$ are equivalent if and only if there exist functions $A(x,u,p)$, $B(x,u,p)$, with $B \neq 0$, and a diffeomorphism $(\tilde{x},\tilde{u},\tilde{p}) = \Phi(x,u,p)$ such that the one-forms are related according to

$$\Phi^*(\tilde{\omega}_1) = \omega_1 + A\omega_2, \qquad \Phi^*(\tilde{\omega}_2) = B\omega_2. \tag{3.18}$$

under the pull-back map Φ^*.

Now that we have recast the original equivalence problem into an equivalence problem involving differential forms, we are ready to apply the Cartan algorithm. Unfortunately, space considerations will preclude a discussion of the details of the algorithm, and we refer the interested reader to the references [4], [7], [15], [17] for the details. Suffice it to say that the method is completely algorithmic, to the extent that it can be programmed onto a symbolic manipulation computer package, of which several exist. Barring complications, the final result of the Cartan method is to produce a list of invariant one-forms, which can then be used to produce scalar invariants and the complete solution to the equivalence problem.

For instance, in the Lagrangian equivalence problem, if the Lagrangian does not depend on x or u, the invariant forms resulting from Cartan's method are

$$\begin{aligned}
\theta_1 &= \pm\sqrt{|LL_{pp}|}(du - p\ dx),\\
\theta_2 &= (L - pL_p)dx + L_p du,\\
\theta_3 &= \pm\sqrt{\left|\frac{L_{pp}}{L}\right|}dp.
\end{aligned} \tag{3.19}$$

(If L does depend on x and u, there are much more complicated expressions for the invariant forms, [17].) As the reader can check directly, as long as $L \neq 0$, $L_{pp} \neq 0$,(again we exclude the trivial affine Lagrangians), these differential forms constitute an invariant "coframe", or basis for the cotangent space at each point of of the (x, u, p)-space, meaning that under the point transformations (2.12), (2.13), with the Lagrangians matching up as in (2.15), the corresponding differential forms satisfy the invariance conditions

$$\tilde{\theta}_i = \theta_i, \qquad i = 1, 2, 3. \tag{3.20}$$

(The ambiguous $\pm$ signs in the coframe are unavoidable and stem from the ambiguity of the square root.) The differential form θ_2 is the well-known Cartan form from the calculus of variations, also known as Hilbert's invariant integral.

Once we have determined an invariant coframe, there is a straightforward method for producing invariant scalar-valued functions, such as those in (3.15), (3.16). It relies on the fact that the exterior derivative operation is invariant under coordinate changes, so we can differentiate the invariant coframe elements to determine new invariant forms. We can evaluate each $d\theta_i$ directly, and rewrite the resulting two-forms in terms of wedge products of the invariant coframe:

$$d\theta_i = \sum C^i_{jk} \theta_j \wedge \theta_k. \tag{3.21}$$

These are known as the *structure equations* for our problem. It is readily seen using (3.20) that the so-called *torsion coefficients* C^i_{jk} must all be scalar invariants of the problem, i.e. $\tilde{C}^i_{jk} = C^i_{jk}$. For the Lagrangian equivalence problem, with $L = L(p)$, the structure equations take the explicit form

$$\begin{aligned} \theta_1 &= -\frac{1}{2} I_0 \theta_1 \wedge \theta_3 + \theta_2 \wedge \theta_3, \\ d\theta_2 &= \mp \theta_1 \wedge \theta_3, \\ d\theta_3 &= 0. \end{aligned} \tag{3.22}$$

Hence, the only nonconstant torsion coefficient is the invariant

$$I_0 = \pm \frac{L L_{ppp} + 3 L_p L_{pp}}{\sqrt{|L| |L_{pp}|^3}}.$$

If we square I_0 to eliminate the ambiguous sign, we recover the earlier invariant (3.15).

Further scalar invariants are found by re-expressing the exterior derivatives of the invariants appearing in the structure equations in terms of the invariant coframe, leading to the "derived invariants". In our case, since L only depends on p, we find

$$dI_0 = J \theta_3,$$

where the derived invariant

$$J = \pm \sqrt{\left| \frac{L}{L_{pp}} \right|} \, \frac{\partial I_0}{\partial p}$$

is the same invariant given in (3.16). This process can lead to higher and higher order derived invariants; for instance the equation $dJ = K\theta_3$ leads to the second order derived invariant $K = \pm\sqrt{|L/L_{pp}|}\cdot J_p$, etc., etc. For the full Lagrangian equivalence problem, there are not one but three fundamental invariants appearing in the corresponding structure equations, and a host of interesting derived invariants, cf. [17].

4. Solution of Equivalence Problems

The invariant quantities play a fundamental role in the resolution of any equivalence problem. Basically, one tries to characterize the equivalence of two objects in terms of the associated invariants. The Cartan approach is especially efficacious in this regards, in that it readily identifies which invariants are of fundamental importance for the equivalence problem, and, moreover, through the powerful Cartan-Kähler existence theorem for exterior systems of differential equations, provides *necessary and sufficient conditions* for equivalence based on these invariants. Roughly, the key to the complete solution of the equivalence problem is the functional relationship between the invariants which appear in the structure equations and their corresponding derived invariants, as discussed above. (Here I am glossing over several complications, including structures of higher order, and structures that require prolongation.) These functional relations, which lead to the concept of a determining function for the equivalence problem, are the principal objects of interest. It is best if we illustrate this with a particular example - the standard particle Lagrangian equivalence problem in the special case when the Lagrangian only depends on the derivative variable p. In this case, there is one fundamental invariant appearing in the structure equations (3.22), namely $I_0(p)$, or, better, its square $I(p)$. If I happens to be constant, then its value must remain unchanged under the point transformations (2.12), and so must have the same value for both Lagrangians. Otherwise, if $I(p)$ is not constant, we express the derived invariant $J(p)$, cf. (3.16), in terms of I, leading to an equation of the form $J = F(I)$. The scalar function F is called the *determining function* for our equivalence problem, since it effectively determines the equivalence class of a given Lagrangian. Since F may well turn out to be a multiply-valued function, it is better to view the invariants I and J as parametrizing a curve in $\mathbb{C}^2$, which we may identify with the "graph" of the determining function F.

Definition 4 Let $L(p)$ be a complex-analytic Lagrangian depending only on the derivative variable p. The *universal curve* corresponding to L is the complex curve

$$\mathcal{C} \equiv \{(I(p), J(p)) : p \in \mathbb{C}\} \subset \mathbb{C}^2. \tag{4.1}$$

(If I is constant, so $J = 0$, then $\mathcal{C}$ reduces to a single point.)

The universal curve is an invariant for the Lagrangian, so that two equivalent Lagrangians have identical universal curves. The Cartan method shows that, moreover, barring the trivial affine Lagrangians, the universal curve provides the complete necessary and sufficient conditions for the solution to the Lagrangian equivalence problem.

Theorem 5 Let $L(p)$ and $\tilde{L}(\tilde{p})$ be two complex analytic Lagrangians which are not affine functions of p. Then L and $\tilde{L}$ are equivalent under a complex analytic change of variables if and only if their universal curves are identical: $\mathcal{C} = \tilde{\mathcal{C}}$.

In particular, if curve degenerates to a point, the invariants I and $\tilde{I}$ are both constant, and they must the same: $I = \tilde{I}$. For a real-valued Lagrangian, the theorem is essentially the same, except that one must add in the additional condition that the sign of the second derivatives L_{pp} must match that of $\tilde{L}_{\tilde{p}\tilde{p}}$ in order that the resulting change of variables be real, cf. [30].

According to the remarks in section 2.4, the equivalence problem for binary forms is a special case of the general Lagrangian equivalence problem, when the Lagrangian is the n^{th} root of a polynomial of degree n. We can therefore translate Theorem 5 into the language of classical invariant theory by evaluating the invariants I and J directly in terms of known covariants of the binary form f. In each of the covariants (3.2), (3.3), we can replace x and y by the homogeneous coordinate p to find corresponding covariants of the polynomial $g(p)$; we use the same symbols for these covariants. A simple exercise in differentiation using the formula $f(x, y) = y^n g(x/y)$ will prove the following formula:

$$L_{pp} = \frac{n-1}{2} L^{1-2n} H. \tag{4.2}$$

Note that (4.2) provides a simple proof of the classical fact that a binary form is the n^{th} power of a linear form if and only if its Hessian is identically 0. Indeed, if $H \equiv 0$, then the Lagrangian L must be an affine function of p, i.e. $L = ap + b$, which implies that $g(p) = (ap + b)^n$.

Assume that this is not the case, i.e. H does not vanish identically. Then further differentiations prove that

$$I = \frac{8(n-2)^2}{n^3(n-1)} \frac{T^2}{H^3}, \qquad J = -\frac{12(n-2)^2}{n-1} \frac{gU}{H^3}. \tag{4.3}$$

Discarding the inessential constants, we deduce the following solution to the equivalence problem for binary forms.

Theorem 6 Let $f(x, y)$ be a binary form of degree n. Let H, T, U be the covariants defined by (3.2), (3.3). Suppose that the Hessian H is not identically 0, so f is not the nth power of a linear form. Define the fundamental absolute rational covariants

$$I^* = \frac{T^2}{H^3}, \qquad J^* = \frac{fU}{H^3}, \tag{4.4}$$

which are both covariants of weight 0 and degree 0. The functions I^* and J^* parametrize a rational curve $\mathcal{C}^*$ in the projective plane $\mathbb{CP}^2$, called the*universal curve*associated with the binary form f. (If I^* is constant, the curve reduces to a single point.) Then two binary forms f and $\tilde{f}$ are equivalent under the general linear group $GL(2, \mathbb{C})$if and only if their universal curves are identical: $\mathcal{C}^* = \tilde{\mathcal{C}}^*$.

Therefore a complete solution to the complex equivalence problem for binary forms depends on merely two absolute rational covariants $- I^*$ and J^*! Buchberger, [2], has developed a computationally effective method based on the idea of a Gröbner basis which can be used to eliminate the parameter p from the definition of the universal curve, and thereby give an implicit formula for the curve. This could provide a computationally effective method to find the universal curve associated with a binary form and explicitly solve the equivalence problem in a form amenable to symbolic computation. Clebsch, [6], gives another solution to the equivalence problem for binary forms based on the absolute invariants of the forms, assuming the existence of suitable linear or quadratic covariants. However, his approach is not applicable to all forms, whereas the Cartan approach is valid in all cases. In [30], the relationship between the Clebsch and Cartan approaches is explained, and the special role of the null forms made clear. When the universal curve does not degenerate to a single point, the linear fractional transformations (2.13) themselves mapping equivalent Lagrangians, or equivalent binary forms, to each other can also be explicitly determined using the universal curve. Let $\mathbf{I}^* = (I^*, J^*) : \mathbb{CP}^1 \to \mathcal{C}^*$ denote the map parametrizing the universal curve.

Theorem 7 Let f and $\tilde{f}$ be equivalent binary forms, so that the universal curves are the same $\mathcal{C}^* = \tilde{\mathcal{C}}^*$. If the curve does not degenerate to a single point, then the implicit equation

$$\tilde{\mathbf{I}}^*(\tilde{p}) = \mathbf{I}^*(p), \tag{4.5}$$

which has a discrete set of solutions, determines all the linear fractional transformations mapping f to $\tilde{f}$.

In other words, we can explicitly determine all the linear fractional transformations mapping f to $\tilde{f}$ by solving the equations

$$I^*(p) = \tilde{I}^*(\tilde{p}), \qquad J^*(p) = \tilde{J}^*(\tilde{p}).$$

Of course, the second of these two equations merely serves to delineate the appropriate branch of the universal curve, and so rule out spurious solutions to the first equation which map between different branches.

Another (related) approach to the solution of an equivalence problem is to solve the more difficult canonical form problem, which is to find a complete list of simple, canonical forms for the given type of objects. The invariants and covariants will then determine which of the canonical forms a given object is equivalent to. This appears to be difficult to do for general Cartan equivalence problems, since there are uncountably many different equivalence classes. However, for many binary forms, canonical forms are well known, cf. [12]. In the case of biforms, the solution to the equivalence problem for binary biquadratics presented in [28] depends on the enumeration of all possible canonical forms. These are listed in table 1. It can be shown that every binary biquadratic is complex-equivalent to exactly one of the 20 classes of canonical forms. (It is not quite uniquely equivalent to one of the canonical forms, since there are certain discrete automorphisms taking one of the canonical forms to another in the same class for the first four classes.). The real equivalence problem has the same sort of classification, except there are various subclasses given by placing $\pm$ signs in front of the squared terms x^2u^2, etc. The different canonical forms are basically distinguished by the root structure of the associated discriminants $\Delta_{\mathbf{x}}$, $\Delta_{\mathbf{u}}$ discussed above. Thus the first class corresponds to those biforms whose discriminants have four simple complex roots. (Note that according to Corollary 2, if one discriminant has all simple roots, so does the other.) Some of the classes need more sophisticated invariants to distinguish them. For instance, both case 13 and case 17 have discriminants with two pairs of double roots, but they are distinguished by the fact that the invariant I_3, cf. (3.10), vanishes for case 17, but is nonzero for case 13. See [28] for the full details of this classification.

Table 1: Canonical Forms for Binary Biquadratics.

1.	$x^2u^2 + y^2v^2 + \alpha(x^2v^2 + y^2u^2) + 2\beta xyuv,$	$(1+\alpha^2-\beta^2)^2 \neq 4\alpha^2, \alpha \neq 0,$
2.	$x^2u^2 + y^2v^2 + \alpha(x^2v^2 + y^2u^2) + 2\beta xyuv,$	$(1+\alpha^2-\beta^2)^2 = 4\alpha^2, \alpha \neq 0,$
3.	$x^2u^2 + y^2v^2 + 2\beta xyuv,$	$\beta^2 \neq 1,$
4.	$x^2u^2 + y^2v^2 + 2xyuv,$	
5.	$x^2u^2 + y^2v^2 + y^2u^2 + 2\beta xyuv,$	$\beta^2 \neq 1,$
6.	$x^2u^2 + y^2v^2 + y^2u^2 + 2xyuv,$	
7.	$x^2u^2 - x^2v^2 + xyu^2 + xyuv^2,$	
8.	$x^2u^2 - y^2u^2 + x^2uv + y^2uv,$	
9.	$x^2u^2 + y^2v^2 + y^2uv + 2xyuv,$	
10.	$x^2u^2 + y^2uv,$	
11.	$x^2u^2 + xyuv^2,$	
12.	$x^2uv + xyu^2,$	
13.	$x^2u^2 + y^2u^2 + xyuv,$	
14.	$x^2u^2 + x^2v^2 + xyuv,$	
15.	$x^2u^2 + y^2u^2,$	
16.	$x^2u^2 + x^2v^2,$	
17.	$x^2u^2 + xyuv,$	
18.	$x^2u^2,$	
19.	$xyuv,$	
20.	$0.$	

Each of these canonical biforms corresponds to a canonical form for a first order planar quadratic Lagrangian, i.e. $p = q = 2, n = 1$, in (2.17). In linear elasticity, the physically important Legendre-Hadamard strong ellipticity condition, [10], [27], requires that the symbol of the Lagrangian be a positive definite real biform, i.e. $Q(\mathbf{x}, \mathbf{u}) > 0$ for $\mathbf{x}$, $\mathbf{u} \neq \mathbf{0}$. Therefore, the only equivalence classes of interest there are the first two, case 1 corresponding to an anisotropic elastic material, case 2 to an isotropic material. The corresponding Lagrangian takes the form

$$u_x^2 + v_y^2 + \alpha(u_y^2 + v_x^2) + 2\beta u_x v_y.$$

This particular Lagrangian can be shown to be just a rescaled version of the stored energy function for an orthotropic elastic material, which, in three dimensions, is an anisotropic material with three reflectional planes of material symmetry, cf. [10]. (A block of wood is a good example of such a material.) The constants α and β are called the canonical elastic moduli of the material. Thus, the invariant theory produces the surprising new result that every linear, planar elastic material is equivalent to an orthotropic material, determined by just two canonical elastic moduli (as opposed to the standard six moduli for a planar anisotropic material described in any text book on linear elasticity). Complex variables methods, well-known for orthotropic materials, but considerably more cumbersome for more general anisotropic materials, can now be readily employed for the analysis of physical problems.

5. Symmetry Groups

For any equivalence problem, the symmetry group of an individual object is the group of "self-equivalences" - i.e. the group consisting of all allowable changes of variables which leave the given object unchanged. For example, in the case of a binary form, the symmetry group is the subgroup consisting of all matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2)$ such that

$$f(ax + by, cx + dy) = f(x, y).$$

Similarly, for one of our Lagrangian equivalence problems, the symmetry group of a given Lagrangian is the group of all transformations (either fiber-preserving or point transformations, depending the allowed changes of variables) which map the Lagrangian to itself (either with or without a divergence, depending on the notion of equivalence used). For instance, if the Lagrangian is $L = p^2$, then the scaling group $(x, u) \to (\lambda x, \lambda^2 u)$, $\lambda > 0$, is a symmetry group for the fiber-preserving equivalence problem (and hence for any of the more complicated equivalence problems). The transformation group $(x, u) \to (x, u + \varepsilon x)$, $\varepsilon \in \mathbb{R}$, maps L to $(p + \varepsilon)^2 = p^2 + 2\varepsilon p + \varepsilon^2$, and so is not a symmetry group of the standard Lagrangian equivalence problem; however, the two extra terms can be written as a divergence, $2\varepsilon p + \varepsilon^2 = D(2\varepsilon u + \varepsilon^2 x)$, and hence this group *is* a symmetry group of the divergence equivalence problem. Similarly, in the case of the quadratic Lagrangian or biforms, we can talk about linear symmetries. For instance, the "isotropic" case 2 of table 1 is distinguished from the more generic "anisotropic" case 1 by the fact that it possesses an additional one-parameter rotational symmetry group beyond the obvious scaling symmetry common to all quadratic Lagrangians. Thus, two-dimensional isotropic elastic materials are distinguished from more general anisotropic materials by this additional one-parameter symmetry group.

One of the great strengths of the Cartan method is that it immediately provides the dimension of the symmetry group of any given object.

Theorem 8 Suppose we have solved an equivalence problem on an n dimensional space by constructing an invariant coframe. Let r be the rank of the invariant coframe, meaning the number of functionally independent invariants appearing among the torsion coefficients C^i_{jk} and their derived invariants. Then the symmetry group of the problem is an $n - r$ dimensional Lie group.

(Here we have phrased the theorem so as to ignore additional complications which can arise when there are infinite pseudo-groups of symmetries, or when the problem must be "prolonged".) See Hsu and Kamran, [14], for an application of this result to the study of symmetry groups of ordinary differential equations. Here we use Theorem 8 to provide a complete determination of the possible symmetry groups for our Lagrangian equivalence problem.

Theorem 9 Let $L(p)$ be a Lagrangian which depends only on p. Then the two-parameter translation group $(x, u) \to (x + \delta, u + \varepsilon)$, $\delta, \varepsilon \in \mathbb{C}$, is always a symmetry group. If L is an affine function of p, then L possesses an infinite-dimensional Lie pseudogroup of symmetries depending on two arbitrary functions. If L is not an affine function of p, and the invariant I is constant, then L admits an additional one-parameter group of symmetries. If the invariant I is not constant, then the symmetry group of L is generated by only the translation group and, possibly, discrete symmetries.

Indeed, the rank of the invariant coframe (3.19) is one, unless the invariant I is constant, in which case there are no non-constant invariants for the problem, and the rank is zero. (The affine case does not follow from the equivalence method results as presented above, since we explicitly excluded this case from consideration, but is easily verified by direct computation.) It is easy to see that, except in the case when the Lagrangian is an affine function of p, the symmetry groups of a binary form f and the corresponding Lagrangian (2.16) differ only by the translation group in (x, u). Therefore, Theorem 8 immediately implies the following theorem on symmetries of binary forms.

Theorem 10 Let $f(x, y)$ be a binary form of degree n.
i) If $H \equiv 0$, then f admits a two-parameter group of symmetries.
ii) If $H \not\equiv 0$, and I^* is constant, then f admits a one-parameter group of symmetries.
iii) If $H \not\equiv 0$, and I^* is not constant, then f admits at most a discrete symmetry group.

(Case *i*) is proved by direct computation, using the fact that f is the n^{th} power of a linear form, and hence equivalent to $\pm x^n$.)

In the case when the invariant I^* is not constant, so the universal curve $\mathcal{C}^*$ is really a curve, we can combine Theorems 7 and 10 to determine the cardinality of the discrete symmetry group of a binary form.

Theorem 11 Let f be a binary form with non-constant invariant I^*. Let $d$denote the covering degree of the universal curve $\mathbf{I}^* : \mathbb{CP}^1 \to \mathcal{C}^*$, i.e. the number of points in the inverse image $\mathbf{I}^{*-1}\{z\}$ of a generic point $z \in \mathcal{C}^*$. Then the symmetry group of f is a finite group consisting of d elements.

Indeed, the symmetries will be determined by all solutions to the implicit equation

$$\mathbf{I}^*(\tilde{p}) = \mathbf{I}^*(p). \tag{5.1}$$

Moreover, since I^* and J^* are rational functions, the degree cannot be infinite, so we deduce that a binary form cannot have an infinite, discrete symmetry group.

A second consequence of Theorem 11 is the interesting result that any binary form with constant invariant I^* is equivalent to a monomial. The proof is elementary once the symmetry generator is placed in Jordan canonical form, cf. [30].

Theorem 12 A binary form f is complex-equivalent to a monomial, i.e. to $x^k y^{n-k}$, if and only if the covariant T^2 is a constant multiple of H^3, or, equivalently, its universal curve degenerates to a single point.

Another new result which follows directly from the Cartan approach is a complete characterization of those binary forms which can be written as the sum of two n^{th} powers.

Theorem 13 A binary form of degree $n \geq 3$ is complex-equivalent to a sum of two n^{th} powers, i.e. to $x^n + y^n$, if and only if the invariant I^* is not constant, and its universal curve is an affine subspace of $\mathbb{CP}^2$ of the explicit form

$$J^* = -\frac{n}{3n-6}(I^* + \frac{1}{2}). \tag{5.2}$$

This is equivalent to the condition that the covariants f, H, T, U are related by the equation

$$(3n-6)fU + nT^2 + \frac{1}{2}nH^3 = 0. \tag{5.3}$$

(Incidentally, there are other binary forms whose universal curves are also affine subspaces of $\mathbb{C}^2$, one example being a binary quartic whose roots are in equianharmonic ratio. An interesting open problem is to characterize all such binary forms.)

There is a classical theorem, due to Gundelfinger, cf. [11], [19], which gives an alternative generic test for determining how many binary forms of degree n can be written as the sum of k n^{th} powers. It would be interesting to find a relationship between Gundelfinger's result and the criterion in Theorem 13 in the case of two n^{th} powers. Another interesting line of investigation would be to see how the universal curve distinguishes between sums of three or more n^{th} powers, although this appears to be much more difficult as it is no longer determined by a single-valued function of I^*.

Acknowledgments

I would like to thank Dennis Stanton for his effort in organizing such a stimulating conference. I would also like to express my thanks to Gian-Carlo Rota for encouraging my forays into classical invariant theory. Also, I would like to thank Bernd Sturmfels for alerting me to the possibility of using Gröbner bases to explicitly determine the universal curve of a binary form.

References

[1] Ball, J.M., Currie, J.C. and Olver, P.J., Null Lagrangians, weak continuity, and variational problems of arbitrary order, *J. Func. Anal.* **41** (1981), 135–174

[2] Buchberger, B., Applications of Gröbner bases in non-linear computational geometry, in: *Scientific Software*, J.R. Rice, ed., IMA Volumes in Mathematics and its Applications, vol. 14, Springer-Verlag, New York, 1988

[3] Cartan, E., Les sous-groupes des groupes continus de transformations, *Oeuvres Complétes*, part. II, vol. 2, Gauthiers-Villars, Paris, 1953, pp. 719–856.

[4] Cartan, E., Les problèmes d'équivalence, *Oeuvres Complétes*, part. II, vol. 2, Gauthiers-Villars, Paris, 1952, pp. 1311–1334

[5] Cartan, E., Sur un problème d'équivalence et la théorie des espaces métriques généraléss, *Oeuvres Complètes*, part. III, vol. 2, Gauthiers-Villars, Paris, 1955, 1131–1153.

[6] Clebsch, A., *Theorie der Binären Algebraischen Formen*, B.G. Teubner, Leipzig, 1872

[7] Gardner, R.B., Differential geometric methods interfacing control theory, in: *Differential Geometric Control Theory*, R.W. Brockett et. al., eds., Birkhauser, Boston, 1983, pp. 117–180

[8] Gel'fand, I.M. and Dikii, L.A., Asymptotic behavior of the resolvent of Sturm-Liouville equations and the algebra of the Korteweg-deVries equation, *Russ. Math. Surveys* **30** (1975), 77-113

[9] Grace, J.H. and Young, A., *The Algebra of Invariants*, Cambridge Univ. Press, Cambridge, 1903

[10] Green, A.E. and Zerna, W., *Theoretical Elasticity*, The Clarendon Press, Oxford, 1954

[11] Gundelfinger, S., Zur Theorie der binären Formen, *J. Reine Angew. Math.* **100** (1886), 413-424

[12] Gurevich, G.B., *Foundations of the Theory of Algebraic Invariants*, P. Noordhoff Ltd., Groningen, Holland, 1964

[13] Hilbert, D., Über die vollen Invariantensysteme, in: *Ges. Abh. II*, Springer-Verlag, Berlin, 1933, 287-344

[14] Hsu, L. and Kamran, N., Classification of second-order ordinary differential equations admitting Lie groups of fiber-preserving symmetries, *Proc. London Math. Soc.*, to appear

[15] Kamran, N., Contributions to the study of the equivalence problem of Elie Cartan and its applications to partial and ordinary differential equations, preprint, 1988.

[16] Kamran, N., Lamb, K.G. and Shadwick, W.F., The local equivalence problem for $d^2y/dx^2 = F(x,y,dx/dy)$ and the Painlevé transcendants, *J. Diff. Geom.* **22** (1985), 139-150

[17] Kamran, N., and Olver, P.J., The equivalence problem for particle Lagrangians, preprint, 1988

[18] Kamran, N., and Olver, P.J., Equivalence of differential operators, preprint, 1988

[19] Kung, J.P.S., Canonical forms for binary forms of even degree, in: *Invariant Theory*, S.S. Koh, ed., Lecture Notes in Math, vol. 1278, Springer-Verlag, New York, 1987, pp. 52-61

[20] Kung, J.P.S. and Rota, G.-C., The invariant theory of binary forms, *Bull. Amer. Math. Soc.* **10** (1984), 27–85

[21] Maschke, H., A symbolic treatment of the theory of invariants of quadratic differential quantics of n variables, Trans. Amer. Math. Soc. **4** (1903), 445–469

[22] Olver, P.J., Conservation laws in elasticity. II. Linear homogeneous isotropic elastostatics, *Arch Rat. Mech. Anal.* **85** (1984), 131–160

[23] Olver, P.J., *Applications of Lie Groups to Differential Equations*, Graduate Texts in Mathematics, vol. 107, Springer-Verlag, New York, 1986

[24] Olver, P.J., Invariant theory of biforms, preprint, 1986

[25] Olver, P.J., Invariant theory and differential equations, in: *Invariant Theory*, S.S. Koh, ed., Lecture Notes in Mathematics, vol. 1278, Springer-Verlag, New York, 1987, pp. 62–80

[26] Olver, P.J., Classical invariant theory and the equivalence problem for particle Lagrangians, *Bull. Amer. Math. Soc.* **18** (1988), 21-26

[27] Olver, P.J., Canonical elastic moduli, *J. Elasticity* **19** (1988), 189-212

[28] Olver, P.J., The equivalence problem and canonical forms for quadratic Lagrangians, *Adv. Appl. Math.* **9** (1988), 226-257

[29] Olver, P.J., Conservation laws in elasticity. III. Planar anisotropic elastostatics, *Arch. Rat. Mech. Anal.*, to appear

[30] Olver, P.J., Classical invariant theory and the equivalence problem for particle Lagrangians. I. Binary forms, *Adv. in Math.*, to appear

[31] Shakiban, C., A resolution of the Euler operator II, *Math. Proc. Camb. Phil. Soc.* **89** (1981), 501-510

[32] Stroh, E., Ueber eine fundamentale Eigenschaft des Ueberschiebungsprocesses und deren Verwerthung in der Theorie der binären Formen, *Math. Ann.* **33** (1889), 61–107

[33] Tresse, M.A., *Détermination des Invariants Ponctuels de l'Èquation Différentielle Ordinaire du Second Ordre* $y'' = \omega(x, y, y')$, S. Hirzel, Leipzig, 1896

[34] Turnbull, H.W., The invariant theory of a general bilinear form, *Proc. London Math. Soc.* **33** (1932), 1–21

[35] Weitzenböck, R., *Invariantentheorie*, P. Noordhoff, Groningen, 1923