ADVANCES IN MATHEMATICS 80, 39–77 (1990)

# Classical Invariant Theory and the Equivalence Problem for Particle Lagrangians. I. Binary Forms

PETER J. OLVER*

*School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455*

The problem of equivalence of binary forms under linear changes of variables is shown to be a special case of the problem of equivalence of particle Lagrangians under the pseudogroup of transformations of both the independent and dependent variables. The latter problem has a complete solution based on the equivalence method of Cartan. There are two particular rational covariants of any binary form which are related by a "universal function." The main result is that two binary forms are equivalent if and only if their universal functions are identical. Construction of the universal function from the syzygies of the covariants, and explicit reconstruction of the form from its universal function are also discussed. New results on the symmetries of forms, and necessary and sufficient conditions for the equivalence of a form to a monomial, or to a sum of two *n*th powers are consequences of this result. Finally, we employ some syzygies due to Stroh to relate our result to a theorem of Clebsch on the equivalence of binary forms. 

## 1. INTRODUCTION

The general *equivalence problem* is to determine when two geometric objects expressed in local coordinates are really both the same object under an appropriate change of coordinates. Cartan [2] developed a general algorithm, based on the theory of differential forms, which provides a systematic procedure for determining the necessary and sufficient conditions for the equivalence of geometric objects. Although it is a powerful explicit computational method, it has only recently begun to be applied to problems of interest, including differential equations [7, 15, 16], symmetry groups [14], control theory [5], and problems in the calculus of variations [7, 17]. The full range of applications to mathematics, physics, and engineering will, I believe, be quite extensive, and remains to be fully explored. This paper will describe a new and unexpected application of the equivalence method to classical invariant theory and the theory of poly-

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



0001-8708/90 $7.50

nomials. Surprisingly, the application originates as an equivalence problem in the calculus of variations.

The particular equivalence problem to be treated is to determine when two first order scalar variational problems,

$$\mathscr{L}[u] = \int L\left(x, u, \frac{du}{dx}\right) dx \qquad \text{and} \qquad \tilde{\mathscr{L}}[u] = \int \tilde{L}\left(\tilde{x}, \tilde{u}, \frac{d\tilde{u}}{d\tilde{x}}\right) d\tilde{x},$$

can be transformed into each other by a change of variables of the form

$$\tilde{x} = \varphi(x, u), \qquad \tilde{u} = \psi(x, u).$$

Cartan, in [3], actually solved this problem as a special case of a more general equivalence problem arising in differential geometry. It constitutes only one of at least four possible mathematically interesting equivalence problems for particle Lagrangians, which are formulated and solved in [17]. In Section 2 we outline Cartan's solution to the problem when the Lagrangian depends only on the derivative $p = du/dx$, which is the case of interest for classical invariant theory. In this case there is precisely one invariant for the problem, denoted by $I$, and one nontrivial "derived invariant," denoted by $J$. These two invariants are related by a single "universal function," denoted by $F$, whereby $J = F(I)$. In essence, the fundamental result of Cartan is that two Lagrangians $L(p)$ and $\tilde{L}(\tilde{p})$ are equivalent if and only if their universal functions are identical: $F = \tilde{F}$. See Theorems 2.4 and 2.5 for the precise formulation of this result.

Classical invariant theory is concerned with the properties of homogeneous polynomials or *forms* which are unchanged by linear coordinate transformations. A large amount of effort has been devoted to the determination of complete systems of invariants and covariants, which are well understood for forms of low degree, but become far too complicated as soon as the degree becomes even moderately large. One key use of the invariants and covariants is to investigate the *equivalence problem for forms*, which is to determine when two homogeneous polynomials can be transformed into each other by an appropriate linear transformation. A solution to this problem has important applications, not only to invariant theory itself, but also to the determination of canonical forms for polynomials.

The heart of this paper appears in Section 4. The fundamental observation is that the equivalence problem for binary forms can be recast as a special case of the Lagrangian equivalence problem, and hence has a complete solution based on Cartan's results. As a consequence of the solution to the Lagrangian equivalence problem, we deduce the result that for any binary form there are two fundamental absolute rational covariants, denoted $I^*$ and $J^*$, which completely characterize the complex equivalence problem. Assuming that the form is not the $n$th power of a linear form, there are two distinct cases: first, if $I^*$ is constant, then $\tilde{I}^*$ must have the same constant value in order that the forms be complex-equivalent; otherwise, there is a *universal function* relating the two covariants: $J^* = F^*(I^*)$, and two such forms are complex-equivalent if and only if their universal functions are identical: $F^* = \tilde{F}^*$. The equivalence problem for real polynomials can also be solved by the same methods, and, surprisingly, the solution requires only one or two additional sign restrictions, depending on the sign of the Hessian of the form and, if the degree is even, the sign of the form itself. Given the well-known proliferation of covariants and invariants as the degree of the form increases, the fact that the equivalence problem is solved using only two rational covariants is quite striking. The universal function $F^*$ must store a lot of information! See Theorems 4.2, 4.3 for the details.

Actually, in the classical literature, Clebsch [4] gave a different solution to the complex-equivalence problem for binary forms, based on their absolute invariants and associated linear covariants. In Section 7, we show how our main theorem is related to Clebsch's original result by appealing to some classical results on syzygies of the covariants of a binary form due to Stroh [26, 27]. This provides an alternative mechanism for constructing the universal function.

Two further applications of the method give additional new results in classical invariant theory, which are described in Section 6. The first is an explicit necessary and sufficient condition for determining when a binary form is equivalent to a monomial, generalizing the well-known vanishing Hessian test for when a form is a perfect $n$th power of a linear form. Secondly, the necessary and sufficient condition for determining when a binary form is equivalent to a sum of two $n$th powers is also found. In the final section, we address the problem of how to reconstruct the Lagrangian or binary form from the knowledge of its universal function. It turns to rest on the solution to a single first order ordinary differential equation.

Although the present method gives a complete solution to the equivalence problem, it does not make much headway with the related canonical form problem, which is to determine a complete list of elementary canonical forms for binary forms of a given degree. However, as shown in Section 5, once the canonical forms are known, the main theorem provides a simple mechanism for determining which canonical form (either real or complex) a given binary form can be cast into.

The equivalence problem for ternary forms, or forms involving even more variables, can also be related to a Cartan equivalence problem for multi-particle Lagrangians [23]. This Lagrangian equivalence problem is much less well understood, although recent unpublished results of Bryant and Gardner make substantial progress towards the intrinsic solution [6].

The application to the equivalence problem for forms in three or more variables will be the subject of the second paper in this series.

## 2. The Equivalence Problem for Particle Lagrangians

Consider a first order scalar variational problem

$$\mathscr{L}[u]=\int L(x, u, p)\, dx. \tag{2.1}$$

Both real and complex-valued variational problems are of interest, and we will discuss both versions of the equivalence problem. However, it is technically easier to present the complex case first, and so we begin by assuming that the *Lagrangian* $L(x, u, p)$ is an analytic function of three complex variables $x$, $u$, and $p\equiv du/dx$, defined on a domain $\Omega\subset\mathbb{C}^3$. We say that two Lagrangians $L$ and $\tilde{L}$ are *equivalent* if there exists an analytic change of variables

$$\tilde{x}=\varphi(x, u), \qquad \tilde{u}=\psi(x, u), \tag{2.2}$$

mapping one to the other. The change in the derivative is a linear fractional transformation

$$\tilde{p}=\frac{ap+b}{cp+d}, \tag{2.3}$$

where the coefficients

$$a=\frac{\partial\psi}{\partial u}, \qquad b=\frac{\partial\psi}{\partial x}, \qquad c=\frac{\partial\varphi}{\partial u}, \qquad d=\frac{\partial\varphi}{\partial x}, \tag{2.4}$$

may depend on $x$ and $u$. Equivalent Lagrangians must be related by the basic formula

$$L(x, u, p)=(cp+d)\,\tilde{L}(\tilde{x}, \tilde{u}, \tilde{p}), \tag{2.5}$$

stemming from the identification of the two one-forms

$$L(x, u, p)\, dx=\tilde{L}(\tilde{x}, \tilde{u}, \tilde{p})\, d\tilde{x}.$$

(The proposed equivalence problem is a somewhat restricted version of the "true" Lagrangian equivalence problem, in which one has the additional freedom of adding in a divergence term; it is, however, more general than the fiber-preserving equivalence problem, in which the new independent variables depend only on the old independent variables: $\tilde{x}=\varphi(x)$. These problems can also be solved by the same methods [17] but are not immediately relevant to our construction.)

Cartan developed a powerful algorithm that will solve such equivalence problems, leading to explicit necessary and sufficient conditions for equivalence. A complete solution to the equivalence problem under consideration appears in [3]. For the reader's convenience we will outline the basic method, but in the simpler special case when the Lagrangian $L = L(p)$ is a function of the derivative alone, as this is the case that is applicable to classical invariant theory. The first step in Cartan's method is to reformulate the problem in terms of differential forms. We introduce the one-forms

$$\omega_1 = du - p\,dx, \qquad \omega_2 = L(p)\,dx, \qquad \omega_3 = dp,$$

which, provided the Lagrangian does not vanish, constitute a *coframe*, or pointwise basis for the cotangent space $T^*\Omega$. From now on, we impose the condition $L(p) \neq 0$ for all $p \in \Omega$, which, except for the completely trivial case $L \equiv 0$, can always be realized by suitably shrinking the domain $\Omega$. The first basis element is the *contact form*, which must be preserved (up to multiple) in order that the derivative $p$ transform correctly, as in (2.3). The second basis element is the integrand in our variational problem, and the third is included just to complete the coframe. Similarly, we introduce the corresponding coframe for the transformed Lagrangian $\tilde{L}$:

$$\tilde{\omega}_1 = d\tilde{u} - \tilde{p}\,d\tilde{x}, \qquad \tilde{\omega}_2 = \tilde{L}(\tilde{p})\,d\tilde{x}, \qquad \tilde{\omega}_3 = d\tilde{p},$$

defined on a corresponding domain $\tilde{\Omega} \subset \mathbb{C}^3$.

LEMMA 2.1. *Two nonvanishing Lagrangians $L$ and $\tilde{L}$ are equivalent if and only if there exist complex-valued functions $A, B, C, D, E$ on $\Omega$, with $A, E \neq 0$, and a diffeomorphism $\Phi: \Omega \to \tilde{\Omega}$, such that the pull-back $\Phi^*: T^*\tilde{\Omega} \to T^*\Omega$ transforms the coframes as follows:*

$$\Phi^*(\tilde{\omega}_1) = A\omega_1, \qquad \Phi^*(\tilde{\omega}_2) = B\omega_1 + \omega_2, \qquad \Phi^*(\tilde{\omega}_3) = C\omega_1 + D\omega_2 + E\omega_3. \tag{2.6}$$

This condition can be restated more symmetrically as follows. Let $G$ denote the complex Lie group consisting of all matrices of the form

$$\begin{pmatrix} A & 0 & 0 \\ B & 1 & 0 \\ C & D & E \end{pmatrix}, \qquad A, B, C, D, E \in \mathbb{C}, \; A, E \neq 0.$$

Introduce the "lifted" coframes

$$\theta_1 = A\omega_1, \qquad \theta_2 = B\omega_1 + \omega_2, \qquad \theta_3 = C\omega_1 + D\omega_2 + E\omega_3, \tag{2.7}$$

and

$$\tilde{\theta}_1 = \tilde{A}\tilde{\omega}_1, \qquad \tilde{\theta}_2 = \tilde{B}\tilde{\omega}_1 + \tilde{\omega}_2, \qquad \tilde{\theta}_3 = \tilde{C}\tilde{\omega}_1 + \tilde{D}\tilde{\omega}_2 + \tilde{E}\tilde{\omega}_3,$$

which are now differential one-forms living on $\Omega \times G$ and $\tilde{\Omega} \times G$, respectively. There is a natural left action of $G$ on these two spaces (principal bundles), given by $g \cdot (x, h) = (x, g \cdot h)$. The equivalence condition of Lemma 2.1 can be readily translated into the following simpler equivalence condition for the lifted coframes, cf. [2, 3].

LEMMA 2.2. *Two nonvanishing Lagrangians $L$ and $\tilde{L}$ are equivalent if and only if there is a diffeomorphism $\Psi: \Omega \times G \to \tilde{\Omega} \times G$, which commutes with the natural left action of $G$ on these spaces and maps the lifted coframes directly to each other:*

$$\Psi^*(\tilde{\theta}_1) = \theta_1, \qquad \Psi^*(\tilde{\theta}_2) = \theta_2, \qquad \Psi^*(\tilde{\theta}_3) = \theta_3.$$

Using this formulation of the Lagrangian equivalence problem, we are now able to use the fundamental Cartan algorithm to effect its solution. Since this equivalence problem is particularly simple, we have the luxury of working "parametrically" throughout, in the original spirit of Cartan, as opposed to employing the intrinsic approach favored by Gardner [5]. This will immediately lead us to the explicit expressions for the invariants.

The key to Cartan's method is the elementary fact that the exterior derivative is an intrinsic operation. If two differential forms are related by the pull-back map: $\Psi^*(\tilde{\theta}) = \theta$, then their exterior derivatives must also be related in the same way: $\Psi^*(d\tilde{\theta}) = d\theta$. We therefore begin by computing the differentials $d\theta_i$. The resulting *structure equations* are found to be of the form

$$\begin{aligned} d\theta_1 &= \alpha \wedge \theta_1 + \tau_{112}\theta_1 \wedge \theta_2 + \tau_{123}\theta_2 \wedge \theta_3, \\ d\theta_2 &= \beta \wedge \theta_1 + \tau_{212}\theta_1 \wedge \theta_2 + \tau_{223}\theta_2 \wedge \theta_3, \\ d\theta_3 &= \gamma \wedge \theta_1 + \delta \wedge \theta_2 + \varepsilon \wedge \theta_3 + \tau_{312}\theta_1 \wedge \theta_2 + \tau_{323}\theta_2 \wedge \theta_3. \end{aligned} \tag{2.8}$$

Here $\alpha, \beta, \gamma, \delta, \varepsilon$ form a basis for the left-invariant (Maurer-Cartan) one-forms on the Lie group $G$,

$$\begin{pmatrix} \alpha & 0 & 0 \\ \beta & 0 & 0 \\ \gamma & \delta & \varepsilon \end{pmatrix} = dg \cdot g^{-1} = \begin{pmatrix} dA & 0 & 0 \\ dB & 1 & 0 \\ dC & dD & dE \end{pmatrix} \begin{pmatrix} A & 0 & 0 \\ B & 1 & 0 \\ C & D & E \end{pmatrix}^{-1},$$

and the *torsion coefficients* $\tau_{ijk}$, some of which are given below, are determined by explicit computation. The forms $\tilde{\theta}_i$ have analogous expressions for their differentials.

The next step is to *absorb* as much of the torsion as possible. Comparing the two expressions for the differentials $d\tilde{\theta}_i$ and $d\theta_i$, we see that the Maurer-Cartan forms $\alpha$ and $\tilde{\alpha}$, etc., must agree under the pull-back $\Psi^*$ up to a differential form on the base, i.e.,

$$\Psi^*(\tilde{\alpha}) = \alpha + z_1\theta_1 + z_2\theta_2 + z_3\theta_3, \text{ etc.,}$$

where the $z_i$'s are as yet unspecified functions. The goal of the absorption process is to replace each one-form $\alpha, \beta, \ldots$ in the structure equations (2.8) by an expression of the form $\alpha + z_1\theta_1 + z_2\theta_2 + z_3\theta_3$, etc., where the $z_i$'s are chosen so as to make as many of the torsion coefficients vanish as possible. The remaining unabsorbable torsion components are then invariants of the equivalence problem, and must have the same values for the two lifted coframes. For instance, looking at the equation for $d\theta_1$, we see that we can replace $\alpha$ by $\hat{\alpha} = \alpha - \tau_{112}\theta_2$ in order to make the coefficient of $\theta_1 \wedge \theta_2$ vanish:

$$d\theta_1 = \hat{\alpha} \wedge \theta_1 + \tau_{123}\theta_2 \wedge \theta_3.$$

The torsion component

$$\tau_{123} = \frac{A}{EL} \tag{2.9}$$

is, however, unabsorbable, and an invariant of the problem, meaning that it must have the same value for both Lagrangians:

$$\frac{\tilde{A}}{\tilde{E}\tilde{L}} = \frac{A}{EL}.$$

Similarly, in the equation for $d\theta_2$ we can absorb all the torsion components except

$$\tau_{223} = \frac{B - L_p}{EL}, \tag{2.10}$$

which forms a second invariant. (Subscripts on $L$ denote partial derivatives.) In the equation for $d\theta_3$ we can absorb all the torsion components, so there are no further invariants at this stage.

Now, according to Cartan, since the unabsorbable torsion coefficients are invariants which depend on the group parameters, they can be normalized to any convenient constant value by fixing some of the relevant group parameters. This has the effect of reducing the dimension of the underlying Lie group $G$, and thus simplifying the equivalence problem. The ultimate goal (barring prolongation [2, 5, 15]) is to get rid of the group parameters entirely, and then read off the (absolute) invariants, which will be functions of the base variables $(x, u, p)$ alone. For example, the first invariant (2.9) can be normalized to 1 by setting

$$A = EL, \tag{2.11}$$

whereas the second (2.10) can be normalized to 0 by setting

$$B = L_p. \tag{2.12}$$

This has the effect of reducing $G$ to a three-parameter group, with $C$, $D$, $E$ the remaining group parameters. Note that at this stage, since we have normalized $B$, the resulting form

$$\theta_2 = L_p \omega_1 + \omega_2 = (L - pL_p)\, dx + L_p\, du$$

is an invariant differential form for the Lagrangian $L$. This is the well known Cartan form from the calculus of variations [8].

We now substitute the expressions (2.11), (2.12) for the group parameters $A$ and $B$ in the original lifted coframe (2.7). We then recalculate the differentials $d\theta_i$, and apply the absorption procedure once again. The new structure equations have the form

$$\begin{aligned} d\theta_1 &= \varepsilon \wedge \theta_1 + \tau_{112}\theta_1 \wedge \theta_2 + \theta_2 \wedge \theta_3, \\ d\theta_2 &= \tau_{212}\theta_1 \wedge \theta_2 + \tau_{213}\theta_1 \wedge \theta_3, \\ d\theta_3 &= \gamma \wedge \theta_1 + \delta \wedge \theta_2 + \varepsilon \wedge \theta_3 + \tau_{312}\theta_1 \wedge \theta_2 + \tau_{323}\theta_2 \wedge \theta_3. \end{aligned}$$

The second equation provides two further unabsorbable pieces of torsion:

$$\tau_{212} = \frac{DLL_{pp}}{E^2L^2} \qquad \text{and} \qquad \tau_{213} = -\frac{L_{pp}}{E^2L}.$$

The first of these can be normalized to 0 by setting the group parameter $D = 0$. As to the second, there are two different possible normalizations at this stage. The "trivial" case is when $L_{pp} \equiv 0$, and hence $L$ is an affine function of $p$, i.e., $L = ap + b$; since this variational problem is completely trivial, we leave this case aside for the remainder of the equivalence procedure. Otherwise, by possibly shrinking the domain $\Omega$, we assume that neither $L$ nor $L_{pp}$ vanish in $\Omega$, and thus we can normalize $\tau_{213}$ to $-1$ by setting

$$E = \sigma \frac{\sqrt{L_{pp}}}{\sqrt{L}}, \tag{2.13}$$

where $\sigma = \pm 1$. A subtle but important point is that, although we can specify the square root branches for both $\sqrt{L_{pp}}$ and $\sqrt{L}$ at this stage (these branches will remain fixed through the rest of the computation), we are not allowed to *a priori* specify the sign of $\sigma$ itself; see the subsequent discussion for more details on this point. We have now reduced $G$ to a one-parameter group, with $C$ the only remaining undetermined parameter.

In the third and final loop through the equivalence procedure, we substitute the expressions for $D$ and $E$ into the formulas for the lifted coframe (2.7), and recompute the differentials. We find the new structure equations have the form

$$\begin{aligned} d\theta_1 &= \tau_{112}\theta_1 \wedge \theta_2 + \tau_{113}\theta_1 \wedge \theta_3 + \theta_2 \wedge \theta_3, \\ d\theta_2 &= -\theta_1 \wedge \theta_3, \\ d\theta_3 &= \gamma \wedge \theta_1 + \tau_{312}\theta_1 \wedge \theta_2 + \tau_{323}\theta_2 \wedge \theta_3. \end{aligned}$$

There are three further unabsorbable pieces of torsion. The first is

$$\tau_{113} = -\tfrac{1}{2}\sigma I_0(p),$$

where

$$I_0(p) = \frac{LL_{ppp} + 3L_p L_{pp}}{\sqrt{L}\,(\sqrt{L_{pp}})^3}. \qquad (2.14)$$

The other two torsion coefficients happen to be identical,

$$\tau_{112} = \tau_{323} = \sigma \frac{C}{\sqrt{L}\sqrt{L_{pp}}},$$

and can both be normalized to 0 by setting the group parameter $C = 0$.

We have now eliminated all the group parameters, or, in the terminology of the equivalence method, have reduced the problem to an $\{e\}$-structure. We can read off the invariants for the problem from the structure equations, which we find by direct computation to have the form

$$\begin{aligned} d\theta_1 &= -\tfrac{1}{2}\sigma I_0\theta_1 \wedge \theta_3 + \theta_2 \wedge \theta_3, \\ d\theta_2 &= -\theta_1 \wedge \theta_3, \\ d\theta_3 &= 0, \end{aligned} \qquad (2.15)$$

where the invariant coframe for the Lagrangian is given explicitly by

$$\begin{aligned} \theta_1 &= \sigma\sqrt{L}\sqrt{L_{pp}}\,(du - p\,dx), \\ \theta_2 &= (L - pL_p)\,dx + L_p\,du, \\ \theta_3 &= \sigma\frac{\sqrt{L_{pp}}}{\sqrt{L}}\,dp. \end{aligned} \qquad (2.16)$$

There is only one nonconstant invariant for this problem, namely, $\sigma I_0(p)$, which must have the same value for the two equivalent Lagrangians:

$$\sigma I_0(p) = \tilde{\sigma}\tilde{I}_0(\tilde{p}). \tag{2.17}$$

In other words, if $L$ and $\tilde{L}$ are equivalent Lagrangians, then the corresponding functions $I_0(p)$ and $\tilde{I}_0(\tilde{p})$ agree up to sign. The ambiguity in the sign of $\sigma I_0$ is, however, unavoidable. For instance, under the elementary transformation

$$\tilde{x} = x, \qquad \tilde{u} = -u, \tag{2.18}$$

the function $I_0(p)$ gets mapped to $-\tilde{I}_0(-\tilde{p})$ so the sign *can* change. However, by composing any change of variables with the orientation-reversing map (2.18), we can always change the sign of $I_0$ if required, so the ambiguity is of an inessential kind. We can avoid constantly referring to this ambiguity by using the rational differential function

$$I(p) = I_0(p)^2 = \frac{(LL_{ppp} + 3L_p L_{pp})^2}{LL_{pp}^3} \tag{2.19}$$

as our fundamental invariant from now on, so that the invariant equation (2.17) becomes

$$I(p) = \tilde{I}(\tilde{p}). \tag{2.20}$$

If $F(x, u, p)$ is any function on $\Omega$, its *covariant derivatives* with respect to the invariant coframe (2.16) are the three functions $F_{,1}$, $F_{,2}$, $F_{,3}$ defined by the formula

$$dF = \sum_{j=1}^{3} F_{,j}\theta_j. \tag{2.21}$$

Clearly the covariant derivatives of any invariant are also invariant, so in particular, the *derived invariants* $\sigma I_{0,1}$, $\sigma I_{0,2}$, $\sigma I_{0,3}$ are also invariants. Since $I_0$ depends only on $p$, the only one of these which does not automatically vanish is

$$\sigma I_{0,3} \equiv J = \frac{\sqrt{L}}{\sqrt{L_{pp}}} \frac{\partial I_0}{\partial p}.$$

Even though $I_0$ involves a square root,

$$J(p) = \frac{2L^2 L_{pp} L_{pppp} - 2LL_p L_{pp} L_{ppp} + 6LL_{pp}^3 - 3L_p^2 L_{pp}^2 - 3L^2 L_{ppp}^2}{2LL_{pp}^3} \tag{2.22}$$

is a single-valued rational function of the derivatives of $L$. Now, when $F=\sigma I_0$ (2.21) reduces to

$$d(\sigma I_0)=J\theta_3. \tag{2.23}$$

The key to the complete solution to any equivalence problem is the relationship between the basic invariants and their derived invariants. If the fundamental invariant $I$ is not constant, we can express the derived invariant $J$, given by (2.22), in terms of $I$.

DEFINITION 2.3. Let $L(p)$ be a Lagrangian depending only on $p$. Define the fundamental invariant $I$ and its derived invariant $J$ as in (2.19), (2.22). If $I$ is not constant, define the *universal function* $F$ so that

$$J=F(I). \tag{2.24}$$

If $I$ is constant, so $J=0$, then we set $F\equiv 0$.

We are explicitly allowing the possibility that the universal function $F$ is multiply-valued, and, indeed, this turns out to be the case for many Lagrangians of interest. In other words, what we are really doing is to view the functions $(I(p), J(p))$ as parametrizing a complex curve (the "universal curve") in $\mathbb{C}^2$, which we may identify with the graph of the universal function $F$.

The final result of the Cartan equivalence algorithm is the following theorem providing complete necessary and sufficient conditions for the solution to the Lagrangian equivalence problem. Its proof rests on Cartan's solution to the equivalence problem for $\{e\}$-structures and the complete integrability of the final structure equations.

THEOREM 2.4. *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 either*

(a) *the invariants $I$ and $\tilde{I}$ are both constant and the same: $I=\tilde{I}$, or*

(b) *the invariants $I$ and $\tilde{I}$ are both not constant, and the universal functions are identical: $\tilde{F}=F$.*

*Proof.* We can reconstruct the change of variables $\Phi\colon \Omega\to\tilde{\Omega}$, as given by (2.2), (2.3), by proving the complete integrability of the system

$$\Phi^*(\tilde{\theta}_1)=\theta_1, \qquad \Phi^*(\tilde{\theta}_2)=\theta_2, \qquad \Phi^*(\tilde{\theta}_3)=\theta_3, \tag{2.25}$$

where $\theta_1$, $\theta_2$, $\theta_3$, are the invariant coframe given by (2.16). The invariant coframes will match up provided the invariant equation (2.17) is satisfied; the system (2.25) will be integrable provided the derived invariants also match up, which will be assured by the identification of the two universal functions. In other words, the transformation (2.3) from $p$ to $\tilde{p}$ will be determined as a solution to the pair of equations

$$I(p) = \tilde{I}(\tilde{p}), \qquad J(p) = \tilde{J}(\tilde{p}). \tag{2.26}$$

The first of these equations can, except for a discrete collection of singular points, always be solved for complex $p$ and $\tilde{p}$, provided we analytically continue the Lagrangians $L$ and $\tilde{L}$, and hence the invariants $I$ and $\tilde{I}$. (Now we are required to also admit multiply valued Lagrangians!) Thus we have a solution to (2.17) for some choice of signs $\sigma$ and $\tilde{\sigma}$. The second equation in (2.26), in view of (2.24) and condition (b) of the theorem, just imposes the additional requirement that in our solution $p$ and $\tilde{p}$ map to the same branch of the universal function; again, analytic continuation shows that there is still always a complex analytic solution $\tilde{p} = \pi(p)$ to both equations in (2.26). Finally, we need to make sure that the induced signs $\sigma$ and $\tilde{\sigma}$ are consistent with the change of variables (2.25). However, according to (2.23), we have

$$\Phi^*\{\tilde{J}\tilde{\theta}_3\} = \Phi^*\{d[\tilde{\sigma}\tilde{I}_0(\tilde{p})]\} = d[\sigma I_0] = J\theta_3.$$

Thus, by the second equation in (2.26) the change of variables has the right "orientability" to be consistent with the third equation of (2.25). The theorem now follows directly from the general result of Cartan, cf. [3; 25, Theorem 4.1, p. 344; 15; 17].

Turning to the equivalence problem for real analytic Lagrangians under real change of variables, we find the basic procedure to be essentially the same, but with a few important distinctions. The underlying Lie group is the same, but the group parameters are now restricted to be real. The first loop through Cartan's algorithm proceeds as before, resulting in the same normalizations for the group parameters $A$ and $B$. In the second loop, we still set $D = 0$, but there are now *three* distinct possible normalizations for the torsion coefficient $\tau_{213}$, depending on the sign of the ratio $L_{pp}/L$; these lead to three distinct branches of the real equivalence problem. As before, the trivial case is when $L_{pp} \equiv 0$. The two remaining branches depend on whether $L_{pp}$ has the same or opposite sign to $L$ in $\Omega$. Thus we can normalize $\tau_{213}$ to either $-1$ or $+1$ by setting

$$E = \sigma \sqrt{\frac{|L_{pp}|}{|L|}}.$$

The third loop through the equivalence procedure leads to the final normalization $C = 0$. The structure equations now take the form

$$d\theta_1 = -\sigma I_0 \theta_1 \wedge \theta_3 + \theta_2 \wedge \theta_3,$$
$$d\theta_2 = \varepsilon \theta_1 \wedge \theta_3,$$
$$d\theta_3 = 0,$$

where $\varepsilon = -\text{sign}(L_{pp}/L)$. As before, $I(p) = \pm I_0(p)^2$ is the only nonconstant invariant for this problem, and must have the same value for the two equivalent Lagrangians. We deduce the analogous theorem solving the equivalence problem for real Lagrangians.

THEOREM 2.5. *Let $L(p)$ and $\tilde{L}(\tilde{p})$ be two nonvanishing real analytic Lagrangians which are not affine functions of p. Then L and $\tilde{L}$ are equivalent under a real analytic change of variables if and only if*

(a) *the universal functions relating these invariants are identical: $\tilde{F} = F$, and*

(b) *the invariant equation* (2.26) *has a real solution branch $\tilde{p} = \pi(p)$, on which the ratios $L_{pp}/L$ and $\tilde{L}_{\tilde{p}\tilde{p}}/\tilde{L}$ have the same sign.*

In particular, (2.26) requires that if the invariant $I$ determined from $L$ is constant, then $\tilde{I}$ must have the same constant value. The sign and solvability restrictions in (b) are essential; see the examples in Section 5 and [17]. Indeed, using analytic continuation, complex-equivalence is a purely local property, and the conditions of Theorem 2.4 reflect this. The same cannot be said of real equivalence, since there can be distinct, real-inequivalent forms of the same complex function.

## 3. INVARIANT THEORY OF BINARY FORMS

By a *binary form* of *degree n*, we mean a homogeneous polynomial function

$$f(x, y) = \sum_{i=0}^{n} \binom{n}{i} a_i x^i y^{n-i}, \tag{3.1}$$

defined on $\mathbb{R}^2$ or $\mathbb{C}^2$. The general linear group $GL(2)$ (meaning either $GL(2, \mathbb{R})$ or $GL(2, \mathbb{C})$) acts on the variables via the standard linear representation

$$\tilde{x} = ax + by, \qquad \tilde{y} = cx + dy, \qquad ad - bc \neq 0,$$

and hence acts on the coefficients $a_i$ of $f$ according to the transformation

$$\tilde{f}(\tilde{x}, \tilde{y}) = f(x, y).$$

A *covariant* of a binary form is a function $C(a_0, ..., a_n, x, y)$ which, except for a determinantal factor, does not change under this action of $GL(2)$:

$$C(a_0, ..., a_n, x, y) = (ad - bc)^g \tilde{C}(\tilde{a}_0, ..., a_n, \tilde{x}, \tilde{y}).$$

The number $g$ is called the *weight* of the covariant; if $g = 0$, we call $C$ an *absolute covariant*. If a covariant $C$ does not depend on $x$ or $y$ it is called an *invariant*. (In contrast, the invariants of the previous section would correspond to absolute covariants in the present terminology.) The symbolic method of classical invariant theory [4, 9, 11, 19] provides a ready means for constructing polynomial covariants and invariants for binary forms. 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; see Section 7). For example, the binary sextic ($n = 6$) has 26 independent polynomial covariants. Indeed, a complete system of covariants has been constructed only for forms of fairly low degree. (Although Gordan's method [9, 11] is constructive, it has only been carried through rigorously for binary forms of degrees $n \leqslant 6$ and $n = 8$, cf. [20, Section 4.1; 21, p. 100; 28].)

One of the principal goals of classical invariant theory is the classification of binary forms. This has direct implications for the problem of determining canonical forms, and the geometrical properties of the covariants themselves. Two forms $f$ and $\bar{f}$ are called (real or complex) *equivalent* if they can be transformed into each other by a suitable element of $GL(2)$. The goal then is to characterize equivalence classes of forms by suitable invariants or covariants, cf. [4]. 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 $i^3 - 27j^2$ which provides the key to the classification of quartic polynomials [11, p. 292].

We exhibit some of the elementary covariants of a binary form.[1] First the *Hessian* of a form $f$ of degree $n$, which is

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

[1] The normalization will be that given in Grace and Young [9] which is not the same as that of Gurevich, [11], or Kung and Rota [19]. Unfortunately, there are many different normalizations used in the literature, sometimes even within the same book. For instance the Hessian, as defined on page 3 of Grace and Young, does *not* agree with that used in chapter 6!

is a covariant of weight 2 and has degree $2n-4$. (The notation $(f, g)^{(k)}$ denotes the $k$th transvectant of $f$ and $g$, cf. [9].) Since the Jacobian of any two covariants is also a covariant, the polynomial

$$T=(f, H)=\frac{1}{2n(n-2)}(f_x H_y - f_y H_x), \tag{3.3}$$

is a covariant of weight 3 and degree $3n-6$. We will also need the covariant

$$U=(H, T)=\frac{1}{6(n-2)^2}(H_x T_y - H_y T_x), \tag{3.4}$$

which is of weight 6 and degree $5n-12$. Many more covariants can of course be constructed, see Section 7 for further important examples.

## 4. Equivalence of Binary Forms

We now proceed to connect the considerations of the preceding two sections. We begin by working over the complex numbers to avoid the extra sign and solvability restrictions that arise in the real domain. Let $f(x, y)$ be a complex-valued binary form of degree $n$. Introduce the homogeneous coordinate $p=x/y$, and write

$$g(p)=f(p, 1) \tag{4.1}$$

for the corresponding inhomogeneous polynomial. Note that the action of $GL(2)=GL(2, \mathbb{C})$ on the homogeneous coordinate $p$ reduces to the same linear fractional transformations as given in (2.3), with corresponding action

$$g(p)=(cp+d)^n\, \tilde{g}(\tilde{p})=(cp+d)^n\, \tilde{g}\left(\frac{ap+b}{cp+d}\right) \tag{4.2}$$

on the associated polynomials. Comparing (4.2) with the Lagrangian equivalence condition (2.5) we see that they will agree provided we define the "Lagrangian"

$$L(p)=\sqrt[n]{g(p)}. \tag{4.3}$$

We can either treat $L$ as a multiply-valued Lagrangian, or can choose any convenient branch of the $n$th root in (4.3); note that the different branches only differ by an $n$th root of unity, which can be taken care of by a simple rescaling of the Lagrangian. We avoid the branch points by restricting to a domain where $g$ does not vanish, which is also required for the application of the equivalence method to the Lagrangian $L$. The key observation is that the equivalence problem for the polynomial $g(p)$ under the linear fractional transformations (4.2) is the *same* as the equivalence problem for the $(x, u)$-independent Lagrangian $L(p)$ under the transformations (2.5). (The reader might object that the coefficients $a, b, c, d$ in (2.4) could depend on $x$ and $u$, but we can always "freeze" their values at any convenient point $(x_0, u_0)$ without affecting the equivalence of the Lagrangians.) Therefore, we obtain an immediate solution to the equivalence problem for the polynomial $g$, and hence for the original binary form $f$, by invoking Theorem 2.4 for the corresponding "Lagrangian" $L$.

We now translate Theorem 2.4 into the language of classical invariant theory by first evaluating the invariants $I$ and $J$ directly in terms of known covariants of the binary form $f$. In each of the covariants presented in Section 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 then proves the following:

LEMMA 4.1. *Let $g(p)$ be a polynomial of degree $n \geqslant 2$. Let $L(p)$ be defined by* (4.3). *Then we have the following identities among the invariants $I$ and $J$ of the Lagrangian $L$ and the covariants $H$, $T$, $U$ of the form $g$:*

$$L_{pp} = \frac{n-1}{2} L^{1-2n} H, \tag{4.4}$$

$$I = \frac{8(n-2)^2}{n^3(n-1)} \frac{T^2}{H^3}, \tag{4.5}$$

$$J = -\frac{12(n-2)^2}{n-1} \frac{gU}{H^3}. \tag{4.6}$$

(*The latter two identities require that the Hessian $H$ not vanish identically.*)

*Proof.* If $L$ is given by (4.3), then we easily find that

$$L_{pp} = \left[\frac{1}{n} g g_{pp} + \frac{1-n}{n^2} g_p^2\right] g^{1-2n}. \tag{4.7}$$

On the other hand, a homogeneous polynomial of degree $n$ can be reconstructed from its inhomogeneous form (4.1) by the elementary formula

$$f(x, y) = y^n g\left(\frac{x}{y}\right).$$

Differentiating, we find

$$f_{xx} = y^{n-2} g_{pp},$$
$$f_{xy} = (n-1)\, y^{n-2} g_p - x y^{n-3} g_{pp},$$
$$f_{yy} = n(n-1)\, y^{n-2} g - 2(n-1)\, x y^{n-3} g_p + x^2 y^{n-4} g_{pp}.$$

Substituting into (4.7), replacing $x/y$ by $p$ and comparing with (3.2) proves the first identity (4.4). The other two identities are proved similarly by further differentiations.

*Remark.* Formula (4.4) gives an elementary proof of the classical result that a binary form $f$ is the $n$th power of a linear form if and only if its Hessian vanishes identically, cf. [19, Proposition 5.3; 24, (3.3.14)].

Thus, except for inessential multiples, we can identify the Lagrangian invariant $I$ with the absolute rational covariant $T^2/H^3$, and the first derived Lagrangian invariant $J$ with the absolute rational covariant $gU/H^3$. Theorem 2.4 immediately implies the following solution to the equivalence problem of complex-valued binary forms.

THEOREM 4.2. *Let $f(x, y)$ be a binary form of degree $n$. Let $H$, $T$, $U$ be the covariants defined by* (3.2), (3.3), (3.4). *Suppose that the Hessian $H$ is not identically* 0, *so $f$ is not the* $n$th *power of a linear form. Define the fundamental rational covariants*

$$I^* = \frac{T^2}{H^3}, \qquad J^* = \frac{fU}{H^3}, \tag{4.8}$$

*which are both covariants of weight* 0 *and degree* 0. *If $I^*$ is not constant,* (*so $J^*$ does not vanish identically*), *define the* universal function $F^*$ *so that*

$$J^* = F^*(I^*). \tag{4.9}$$

*Two binary forms $f$ and $\tilde{f}$ are equivalent under the general linear group $GL(2, \mathbb{C})$ if and only if either*

(a) *the covariants $I^*$ and $\tilde{I}^*$ have the same constant values: $I^* = \tilde{I}^*$, or*

(b) *$I^*$ and $\tilde{I}^*$ are not constant, and the universal functions $F^*$ and $\tilde{F}^*$ are identical: $F^* = \tilde{F}^*$.*

Therefore a complete solution to the complex equivalence problem for binary forms depends on merely *two* absolute rational covariants: $I^*$ and $J^*$! As discussed in Section 2, the universal function can be multiply-valued, so one really should view it as defining a universal (rational) curve in $\mathbb{C}^2$. If the invariants $I^*$ and $\tilde{I}^*$ are not constants, and $f$ and $\tilde{f}$ have identical universal functions (curves), then one can explicitly determine all the transformations mapping $f$ to $\tilde{f}$ by solving the equation

$$I^*(p) = \tilde{I}^*(\tilde{p}), \qquad J^*(p) = \tilde{J}^*(\tilde{p}). \tag{4.10}$$

Of course, the second of these two equations merely serves to delineate the appropriate branch of $F$, and so rule out spurious solutions to the first equation which map between different branches. This is important, since the equation $I^* = \tilde{I}^*$, or, equivalently,

$$T(p)^2\, \tilde{H}(\tilde{p})^3 = \tilde{T}(\tilde{p})^2\, H(p)^3 \tag{4.11}$$

is, in general, a polynomial equation of degree $6n - 12$ for $p$ in terms of $\tilde{p}$; as such, many of its roots will be spurious, leading to the wrong branch of the universal function.

The corresponding classification problem for real polynomials works in exactly the same way, except we now need to worry about the signs of the form and its Hessian, and the solvability assumption of Theorem 2.5.

THEOREM 4.3. *Let $f(x, y)$ be a real binary form of degree $n$, which is not the $n$th power of a linear form, and let $g(p) = f(p, 1)$ be the corresponding inhomogeneous polynomial. Let $I^*$ and $J^*$ denote the rational covariants* (4.8), *and define the* universal function *$F^*$ as in* (4.9). (*If $I^*$ is constant, then $F^* = 0$.*) *Two real binary forms $f$ and $\tilde{f}$ are equivalent under the general linear group $GL(2, \mathbb{R})$ if and only if*

(a) *the universal functions $F^*$ and $\tilde{F}^*$ are identical: $F^* = \tilde{F}^*$, and*

(b) *the invariant equation* (4.10) *has a real solution branch $\tilde{p} = \pi(p)$, on which*

$$\operatorname{sign} H(p) = \operatorname{sign} \tilde{H}(\tilde{p}), \tag{4.12}$$

*and, if the degree of the form is even,*

$$\operatorname{sign} g(p) = \operatorname{sign} \tilde{g}(\tilde{p}). \tag{4.13}$$

*Proof.* As in the complex case, we begin by defining the Lagrangian $L(p) = \sqrt[n]{|g(p)|}$, but where we now take the absolute value before computing the $n$th root. Thus $L$ is always positive. Defining $\tilde{L}$ in analogously, we now invoke Theorem 2.5, which requires that the Hessians $H$ and $\tilde{H}$ have the same sign. Thus the two Lagrangians $L$ and $\tilde{L}$ are equivalent if and only if

$$|f(x, y)| = |\tilde{f}(\tilde{x}, \tilde{y})|.$$

If $f$ has odd degree, then this implies that $f$ and $\tilde{f}$ are equivalent, since we can always change the sign of $f$ by replacing $(x, y)$ by $(-x, -y)$. However, if the degree of $f$ is even, then the sign of $f$ (or $g$) is an invariant quantity, so we need an additional condition that $f$ and $\tilde{f}$ have the same sign. This explains the additional sign restriction (4.13) in this case.

Note that since the transformation (2.3) from $p$ to $\tilde{p}$ is a linear fractional transformation, the number of intervals on the real projective line where $H$ (and, in the even degree case, $g$ also) is positive or negative is an invariant of the form. For instance, if $H$ is positive definite, or has two simple real zeros, so is $\tilde{H}$. Note that there is still a possibility that two forms whose Hessians have the same number of sign changes and who have the same universal functions might still not be equivalent, since the invariant equation (4.10) might have no solutions which also satisfy the sign restrictions (4.12), (4.13).

## 5. Cubics and Quartics

In this section, we illustrate the general results of Section 4 by treating the equivalence problem for binary cubics and quartics in some detail, using the well known canonical forms for these polynomials. We begin by looking at the binary cubic

$$f(x, y) = ax^3 + 3bx^2y + 3cxy^2 + dy^3. \tag{5.1}$$

The universal function $F^*$ can be derived directly from the fundamental covariants of $f$ as follows. The Hilbert basis for the covariants of a cubic is provided by the form $f$ itself, the covariants $H$ and $T$, as in (3.2), (3.3), and the discriminant

$$\Delta = (H, H)^{(2)} = \tfrac{1}{18}(H_{xx}H_{yy} - H_{xy}^2),$$

which is an invariant. According to [9, p. 96], the fundamental syzygy among the irreducible covariants of the binary cubic is

$$-2T^2 = H^3 + \Delta f^2. \tag{5.2}$$

Moreover, according to [9, p. 97] (although the sign is wrong in this exercise), the reducible covariant $U$ can be written as

$$U = (H, T) = \tfrac{1}{2}\Delta f,$$

hence

$$-2T^2 = H^3 + 2fU. \tag{5.3}$$

Therefore

$$J^* = -I^* - \tfrac{1}{2}, \tag{5.4}$$

and the universal function $F^*(I^*) = -I^* - \frac{1}{2}$ is, in all cases, a linear function of $I^*$. Thus, for cubic forms, the universal function is always single-valued; this is no longer true for quartic forms. According to Theorem 4.2, the only features which will distinguish complex canonical forms of a binary cubic are whether the Hessian vanishes identically, and whether the invariant $I^*$ is a constant, i.e., whether or not $T^2$ is a constant multiple of $H^3$. Moreover, (5.3) shows that this latter possibility occurs if and only if the discriminant $\Delta$ vanishes, and the cubic has a double root. There are accordingly four distinct complex canonical forms for binary cubics:

(a) If $I^*$ is not constant, then $f$ has three simple roots, and is equivalent to $x^3 + y^3$.

(b) If $I^*$ is constant, then $I^* = -\frac{1}{2}$; $f$ has a double root, and is equivalent to $x^2 y$.

(c) If $H \equiv 0$, then $f$ is a perfect cube, equivalent to $x^3$.

(d) The trivial case $f \equiv 0$.

The classification of real cubics requires an analysis of the sign of the Hessian $H$. Cases (b), (c), and (d) are unchanged. (In case (b) $H$ is necessarily negative definite since $H^3 = -2T^2$.) Case (a) splits into two subcases:

(a1) If the Hessian $H$ is negative semi-definite, then $f$ has three simple real roots, and is equivalent to $x^3 - xy^2$.

(a2) If the Hessian $H$ is indefinite, then $f$ has two complex conjugate roots, and is equivalent to $x^3 + y^3$.

This completes the classification problem for binary cubics. One further note: not every sign possibility is realized; the Hessian of a real binary cubic can never be positive definite. This fact serves to complicate the search for real canonical forms as it is never clear from the outset which types are possible.

Turning to the binary quartic

$$f(x, y) = ax^4 + 4bx^3y + 6cx^2y^2 + 4dxy^3 + ey^4, \tag{5.5}$$

we first note that there are two important invariants

$$i = (f, f)^{(4)} = 2ae - 8bd + 6c^2,$$

and

$$j = (H, f)^{(4)} = 6 \det \begin{vmatrix} a & b & c \\ b & c & d \\ c & d & e \end{vmatrix},$$

cf. [9, Section 89]. Let $H$, $T$, $U$ be as in (3.2), (3.3), (3.4). To evaluate the universal function, we use the fundamental syzygy

$$T^2 = -\tfrac{1}{2}H^3 + \tfrac{1}{4}if^2H - \tfrac{1}{6}jf^3, \tag{5.6}$$

and the identity

$$U = \tfrac{1}{6}jf^2 - \tfrac{1}{6}ifH, \tag{5.7}$$

cf. [9, pp. 98, 99]. We introduce the rational covariant $s = f/H$, which will be treated as a parameter in the ensuing equations. In terms of $s$, the absolute covariants $I^*$ and $J^*$ have the parametric form

$$I^* = -\tfrac{1}{2} + \tfrac{1}{4}is^2 - \tfrac{1}{6}js^3, \qquad J^* = \tfrac{1}{6}js^3 - \tfrac{1}{6}is^2.$$

These two equations give a simple parametrization of the universal curve. Eliminating $s$, we see that the universal function $J^* = F^*(I^*)$ appears as the implicit solution to the cubic equation

$$6j^2(2I^* + 2J^* + 1)^3 = i^3(2I^* + 3J^* + 1)^2. \tag{5.8}$$

Thus, the absolute invariant $j^2/i^3$ plays the crucial role in the equivalence problem, and completely determines the universal function $F^*$. In particular, the universal function is a single-valued linear function if either $i = 0$, so $F^*(I^*) = -I^* - \frac{1}{2}$, or $j = 0$, so $F^*(I^*) = -\frac{2}{3}I^* - \frac{1}{3}$; see [11, Exercise 25.6] for the geometric interpretation of these two conditions. In all other cases, the universal function is a multiply-valued function.

Over $\mathbb{C}$ there are six canonical forms for binary quartics, [11, p. 292]. They are distinguished by our invariants as follows:

| | | |
|---|---|---|
| (I) | $x^4 + 6\mu x^2y^2 + y^4$, $\mu \neq \pm\frac{1}{3}$, | $I^*$ not constant, |
| (II) | $x^2y^2 + y^4$, | $I^*$ not constant, |
| (III) | $x^2y^2$, | $I^* = 0$, |
| (IV) | $x^3y$, | $I^* = -\frac{1}{2}$, |
| (V) | $x^4$, | $H \equiv 0$, |
| (VI) | 0, | $f \equiv 0$. |

Note that in case I

$$i = 2 + 6\mu^2, \qquad j = 6\mu(1 - \mu^2),$$

hence the absolute invariant

$$\frac{j^2}{i^3} = \frac{\mu^2(1-\mu^2)}{8(1+3\mu^2)^3}$$

attains the same value in six different cases

$$\pm\mu, \qquad \pm\frac{1-\mu}{1+3\mu}, \qquad \pm\frac{1+\mu}{1-3\mu}.$$

These correspond to the transformations

$$(x, y) \to (\sqrt{-1}\,x, y),$$
$$(x, y) \to \frac{(x+y, x-y)}{\sqrt[4]{2+6\mu}},$$
$$(x, y) \to \frac{(\sqrt{-1}\,x+y, \sqrt{-1}\,x-y)}{\sqrt[4]{2-6\mu}},$$

mapping one quartic of type I to another quartic of type I. Thus the universal function, as determined by (5.8), will distinguish between genuinely inequivalent quartics of types I and II.

A corresponding classification of real binary quartics appears in [11, Exercises 25.13, 25.14]. For brevity, we just look at the analogues of case I. It is useful to introduce a more detailed classification than that in Gurevich:

| | | | |
|---|---|---|---|
| (Ia) | $x^4 + 6\mu x^2y^2 + y^4$, | $\mu < -\frac{1}{3}$, | $0_4$, $-$ |
| (Ib) | $x^4 + 6\mu x^2y^2 + y^4$, | $-\frac{1}{3} < \mu < 0$ or $\mu > 1$, | $+$, $0_4$ |
| (Ic) | $x^4 + 6\mu x^2y^2 + y^4$, | $0 \leqslant \mu < 1$, $\mu \neq \frac{1}{3}$, | $+$, $+$ |
| (Id) | $-(x^4 + 6\mu x^2y^2 + y^4)$, | $\mu < -\frac{1}{3}$, | $0_4$, $-$ |
| (Ie) | $-(x^4 + 6\mu x^2y^2 + y^4)$, | $-\frac{1}{3} < \mu < 0$ or $\mu > 1$, | $-$, $0_4$ |
| (If) | $-(x^4 + 6\mu x^2y^2 + y^4)$, | $0 \leqslant \mu < 1$, $\mu \neq \frac{1}{3}$, | $-$, $+$ |
| (Ig) | $x^4 + 6\mu x^2y^2 - y^4$, | $\mu \neq \pm\frac{1}{3}$ | $0_2$, $0_2$. |

The last column describes the pattern of signs for the function $g(p) = f(p, 1)$ and the Hessian $H(p)$, described at the end of Section 4. Here $+$ means positive (semi-)definite, $-$ means negative (semi-)definite,

$0_2$ means indefinite with two changes in sign on the projective line (i.e., the corresponding polynomial $f$ or $H$ has two simple real roots), and $0_4$ means indefinite with four changes in sign (i.e., four simple real roots). Thus, for example, in case Ia, $g(p)$ has four real roots, whereas $H(p)$ is negative definite.

Apparently, only cases Ia and Id are not distinguished by the sign pattern. However, there is a good reason for this. Given a quartic of type Ia, the transformation

$$(x, y) \to \frac{(x+y, x-y)}{\sqrt[4]{-2-6\mu}},$$

will transform it into a quartic of type Id, with

$$\tilde{\mu} = \frac{1-\mu}{1+3\mu}.$$

Thus Ia and Id really represent different canonical forms for the same quartic polynomial, and only one of the two is really necessary to make a complete list of canonical forms for real quartics.

## 6. Symmetries and Elementary Forms

Cartan's equivalence method gives us more, as it also determines the symmetry group of the Lagrangian. This fact can be exploited to prove some further results in classical invariant theory on the characterization of various types of elementary forms for polynomials.

Definition 6.1. Let $L(x, u, p)$ be a Lagrangian. Then a transformation (2.2) is called a *symmetry* of $L$ if the transformed Lagrangian $\tilde{L}$ is the same as $L$, i.e.,

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

The *symmetry group* of a Lagrangian $L$ is the group of all symmetry transformations.

In general, if an equivalence problem which reduces to an $\{e\}$-structure in dimension $n$ has $k$ functionally independent invariants, then the problem admits an $n-k$ dimensional symmetry group, cf. [2, 15, 17, 25]. (This is in addition to possible discrete symmetries of the problem.) In our case, $n=3$, and $k=0$ or 1, depending on whether or not the invariant $I(p)$ is a constant. Therefore we conclude:

THEOREM 6.2. *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 has 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.*

(The affine case does not follow from the results of Section 2, but is easily verified by direct computation.)

DEFINITION 6.3. Let $f(x, y)$ be a binary form. The *symmetry group* of $f$ is the subgroup consisting of all matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL(2)$ such that

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

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 (4.3) differ only by the translation group in $(x, u)$. Therefore, Theorem 6.2 immediately implies the following theorem on symmetries of binary forms.

THEOREM 6.4. *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 $\pm x^n$. The other two cases follow directly from Theorem 6.2.

In the case of constant invariant $I^*$, Theorem 6.4 provides an immediate test for determining whether a given form is equivalent to a monomial.

THEOREM 6.5. *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 $H^3$.*

*Proof.* Let $A \in GL(2, \mathbb{C})$ be the generator of the one-parameter symmetry group $e^{tA}$ of the form. We make a linear change of variables that places $A$ into Jordan canonical form $J$. We will show that the same change of variables changes $f$ into a monomial $\bar{f} = cx^i y^{n-i}$, and a simple complex rescaling will eliminate the coefficient $c$. Note that $e^{tJ}$ remains a symmetry group of the form $\tilde{f}$. If $A$ is diagonalizable, so $J = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$, then $e^{tJ}$ acts on the coefficients of $\tilde{f}$ by the induced scaling transformation

$$\tilde{a}_i \to e^{(\alpha i + \beta(n-i))t} \tilde{a}_i.$$

Clearly, this one-parameter family of transformations leaves the coefficient $\tilde{a}_i$ unchanged if and only if either $\tilde{a}_i = 0$ or $\alpha/\beta = 1 - i/n$. Thus, only one of the coefficients of $\tilde{f}$ can be nonzero, and the result follows. The only other case is when $A$ is not diagonalizable, so $J = \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix}$. It is not difficult to deduce that the one-parameter group $e^{tJ}$ can never be a symmetry group of a nonzero form $\tilde{f}$. This completes the proof.

In fact, we find by direct computation that $f$ is complex-equivalent to $x^k y^{n-k}$ if and only if

$$I^* = -\frac{(n-1)(n-2k)^2}{2(n-2)^2 k(n-k)}. \tag{6.1}$$

In particular, $I^*$ is always a negative rational number (if constant)!

In the case of real equivalence, there is an additional possibility since a polynomial might be complex-equivalent to a monomial, but not real-equivalent. For example, the quadratic form $x^2 + y^2$ is complex-equivalent to $xy$, but certainly not real-equivalent to any monomial. However, this is essentially the only counter-example.

THEOREM 6.6. *A binary form $f$ is real-equivalent either to a monomial $\pm x^k y^{n-k}$, or to $(x^2 + y^2)^k$, $n = 2k$, if and only if the covariant $T^2$ is a constant multiple of $H^3$.*

This can be proved from the real canonical form for the generator of the symmetry group, or, more simply, by looking at which complex monomials can be transformed into real polynomials.

In the case when the invariant $I^*$ is not constant, the discrete symmetry group of the binary form can be determined directly from the invariant equation (4.10) (taking the two forms to be the same).

THEOREM 6.7. *Let $f$ be a binary form which is not complex-equivalent to a monomial. Then the (necessarily discrete) symmetry group of $f$ consists of all transformations $\tilde{p} = \pi(p)$ which solve the equations*

$$I^*(\tilde{p}) = I^*(p), \qquad J^*(\tilde{p}) = J^*(p). \tag{6.2}$$

*(As before, the second equation merely makes sure we are in the correct branch of the universal curve of $f$.)*

There is also a similar result for the real symmetry group of a real binary form, the only difference is that the symmetry transformations must preserve the sign of the Hessian, and, in the case of even degree, the form itself.

Another simple result, which follows even more directly from the main theorem is the characterization of when a form is equivalent to the sum of two $n$th powers.

THEOREM 6.8. *A binary form of degree $n \geqslant 3$, which is not equivalent to a monomial, i.e., $I^*$ is not constant, is complex-equivalent to a sum of two* $n$th *powers, i.e., to $x^n + y^n$, if and only if its universal function takes the form*

$$J^* = -\frac{n}{3n-6}\left(I^* + \frac{1}{2}\right), \tag{6.3}$$

*or, equivalently, the covariants $f$, $H$, $T$, $U$ are related by the equation*

$$(3n-6)\,fU + nT^2 + \tfrac{1}{2}nH^3 = 0. \tag{6.4}$$

*Proof.* It is a simple matter to check that the particular form $x^n + y^n$ has universal function given by (6.3); indeed, an elementary computation finds

$$H = 2x^{n-2}y^{n-2},$$

$$T = x^{n-3}y^{n-3}(x^n - y^n),$$

$$U = -\frac{n}{3n-6}x^{2n-6}y^{2n-6}(x^n + y^n),$$

hence

$$I^* = \frac{x^n}{8y^n} - \frac{1}{4} + \frac{y^n}{8x^n},$$

$$J^* = -\frac{n}{3n-6}\left(\frac{x^n}{8y^n} + \frac{1}{4} + \frac{y^n}{8x^n}\right),$$

from which (6.3) follows. The result now follows directly from Theorem 4.2.

According to Sylvester's theorem, [18; 19, Theorem 5.1], binary forms of odd degree are essentially classified according to the number of $n$th powers they can be written as the sum of. Theorem 6.8 gives a simple check for the case of two $n$th powers; it would be interesting to see how the universal function distinguishes between sums of three or more $n$th powers, although this appears to be much more difficult as it is no longer single-valued.

## 7. UNIVERSAL SYZYGIES AND UNIVERSAL FUNCTIONS

For forms of degree higher than 4, the same method of employing syzygies among the covariants can be used to generate implicit relations for the universal function, but the intervening expressions are more complicated. The classification of syzygies has been most deeply studied by Stroh [26, 27], and we begin by quoting some of his results.

Throughout this section, $f$ will be a complex binary form of degree $n \geqslant 4$. Let

$$m = \left[\frac{n}{2}\right], \qquad m' = \left[\frac{n-1}{2}\right],$$

so that if $n = 2m$ is even, then $m' = m - 1$, whereas if $n = 2m + 1$ is odd, then $m' = m$. In either case, $m + m' = n - 1$. We order the covariants of $f$ by their weight. Up to multiple, the only covariant of weight 1 is the form $f$ itself. There are $m$ independent covariants of weight 2, namely the transvectants

$$S_k \equiv (f, f)^{(2k)}, \qquad k = 1, ..., m. \tag{7.1}$$

Note that $S_k$ has degree $2n - 4k$; also $S_1 = H$ is the Hessian, and, if $n = 2m$ is even, $S_m$ is an invariant of $f$. There are many possible covariants of weight 3; for our purposes, the most important of these are the Jacobians of the covariants of weight 2 with $f$, which we denote by

$$T_k \equiv (S_k, f) = ((f, f)^{(2k)}, f), \qquad k = 1, ..., m'. \tag{7.2}$$

(If $n = 2m$ is even, then $T_m = 0$, since $S_m$ is an invariant.) Note that $T_k$ has degree $3n - 4k - 2$; in particular $T_1 = T$ is the covariant used in our fundamental Theorem 4.2. We begin by proving the important result that, by utilizing certain basic syzygies, every covariant of a higher degree binary form can be rewritten in terms of the $n$ fundamental covariants $f$, $S_1, ..., S_m, T_1, ..., T_{m'}$. This result is a direct consequence of some syzygy theorems due to Stroh, which we now state.

THEOREM 7.1. *Let $C$ be a covariant which is of weight $w \geqslant 4$ in the coefficients of the binary form $f$ of degree $n \geqslant 4$. Then there is a syzygy of the form*

$$f^2 C = \Phi(D_{k_1}, ..., D_{k_l}),$$

*where $\Phi$ is a polynomial function of certain covariants, labelled $D_{k_v}$, each of which has weight strictly less than $w$.*

If the covariant happens to be the Jacobian of a pair of simpler covariants, then there is an alternative version of this result.

THEOREM 7.2. *Let $C$ be a covariant of weight $w \geqslant 4$, which can be written as the Jacobian of two simpler covariants $C'$, $C''$, each of weight at least 2, i.e.,*

$$C = (C', C'').$$

*Then there is a syzygy of the form*

$$fC = \Phi(D_{k_1}, ..., D_{k_l}),$$

*where $\Phi$ is a polynomial function of covariants each of weight strictly less than $w$.*

Such a result does not hold for the covariants of weight three (or less). However, there is a modification which allows us to express any covariants of weight 3 in terms of the fundamental covariants $f, S_1, ..., S_m, T_1, ..., T_{m'}$.

THEOREM 7.3. *Let $C$ be a covariant which is of weight $w = 3$ in the coefficients of the binary form $f$. Then there is a syzygy of the form*

$$f^3 C = \Phi(f, S_1, ..., S_m, T_1, ..., T_{m'}),$$

*where $\Phi$ is a polynomial in the fundamental covariants* (7.1), (7.2). (*If $C$ coincides with one of the covariants $T_k$, then this syzygy is the trivial identity $f^3 T_k = f^3 T_k$.*)

The proof of these results can be found in [26]. As a consequence of Stroh's syzygy theorems, and an obvious induction on the weight of a covariant, we find that any covariant or invariant of an $n$th order binary form ($n \geqslant 4$) can be re-expressed as a function of the $n$ fundamental covariants $f, S_1, ..., S_m, T_1, ..., T_{m'}$.

COROLLARY 7.4. *Let $f$ be an binary form of degree $n \geqslant 4$. Let $C$ be any polynomial covariant of $f$. Then there is a polynomial syzygy of the form*

$$f^k C = \Phi(f, S_1, ..., S_m, T_1, ..., T_{m'}), \tag{7.3}$$

*for some integer $k \geqslant 0$.*

See [10, (3.3)] for a related, but slightly different kind of result.

For our purposes, it is slightly better to introduce a "homogenized" version of these syzygies. We introduce the fundamental absolute covariants

$$s_k = \frac{f^{2k-2} S_k}{H^k}, \qquad k = 2, ..., m,$$
$$t_j = \frac{f^{2j-2} T_j}{H^{j+(1/2)}}, \qquad j = 1, ..., m'. \tag{7.4}$$

Since the exponents are fractional, these functions are algebraic rational functions. Nevertheless, by suitably modifying the definition of covariant, we easily see that these are all covariants of weight 0 and degree 0. In particular, we have

$$I^* = \frac{T^2}{H^3} = t_1^2. \tag{7.5}$$

If $C$ is any covariant of weight $w$ and degree $d$, then the function $f^{-w}(f^{-2}H)^{(nw-d)/4}\,C$ is a covariant of weight 0 and degree 0. Thus, if we divide (7.3) through by $f^k$, a simple homogeneity argument shows that we can rewrite the syzygy in the form

$$C = f^w \left(\frac{H}{f^2}\right)^{(nw-d)/4} \cdot \quad \tilde{\Phi}(s_2, \ldots, s_m, t_1, \ldots, t_{m'}), \tag{7.6}$$

where

$$\tilde{\Phi}(s_2, \ldots, s_m, t_1, \ldots, t_{m'}) = \Phi(1, 1, s_2, \ldots, s_m, t_1, \ldots, t_{m'}).$$

As a specific example, consider the covariant $U = (H, T)$. Here Theorem 7.2 is applicable. A relatively straightforward calculation using the symbolic method of classical invariant theory, [9, 11, 19], shows that the basic syzygy (7.3) takes the following form:

$$fU = -\frac{1}{2} H^3 - T^2 + \frac{1}{6}\frac{n-3}{n-2} f^2 S_2 H, \tag{7.7}$$

where $S_2 = (f, f)^{(4)}$. (It would take us too far afield to present the details of the symbolic method required in the calculation here, but the reader can reconstruct it from the given references.) Note that if $f$ is a binary quartic, then eliminating the invariant $j$ from the syzygies (5.6), (5.7) reproduces (7.7). If we divide (7.7) by $H^3$, we deduce the fundamental relation

$$J^* = -\frac{1}{2} - I^* + \frac{1}{6}\frac{n-3}{n-2} s_2. \tag{7.8}$$

Thus (7.5), (7.8) provide two equations linking our absolute rational covariants $I^*$, $J^*$ with 2 of the algebraic covariants (7.4), namely, $s_2$ and $t_1$. To eliminate both of these "parameters" and thereby obtain the explicit formula for the universal function $F^*$, we require yet one more relation among the algebraic covariants. The way to accomplish this is to employ the syzygies (7.3) for the invariants of the form. If there are enough invariants, then we will (in principle) be able to deduce the universal function by eliminating all the parameters $s_2, \ldots, s_m, t_1, \ldots, t_{m'}$, and thereby obtaining a single universal relation connecting the covariants $I^*$, $J^*$. An important theorem of Hilbert [13; 24, (3.4.9)], will guarantee that there are enough independent invariants for our purposes.

THEOREM 7.5. *Let $f$ be a binary form of degree $n \geq 3$. Then the ring of invariants of $f$ has transcendence degree $n-2$ over $\mathbb{C}$.*

In particular, Hilbert's theorem guarantees the existence of $n-2$ independent invariants, which we denote by $I_1, ..., I_{n-2}$. Then according to (7.3) there are syzygies of the form

$$I_\nu = \Psi_\nu(s_1, s_2, ..., s_m, t_1, ..., t_{m'}), \qquad \nu = 1, ..., n-2, \tag{7.9}$$

where

$$\Psi_\nu(s_1, s_2, ..., s_m, t_1, ..., t_{m'}) \equiv s_1^{w_\nu/4} \tilde{\Phi}_\nu(s_2, ..., s_m, t_1, ..., t_{m'}),$$

$w_\nu$ denotes the degree of $I_\nu$, and $s_1$ denotes the additional rational covariant

$$s_1 = \frac{H^n}{f^{2n-4}},$$

which has weight 4 and degree 0. Clearly, if we can eliminate the $n-1$ parameters $s_1, s_2, ..., s_m, t_1, ..., t_{m'}$ from the $n$ equations (7.5), (7.8), (7.9), then we will be left with a single equation relating the covariants $I^*$ and $J^*$, which will give us the universal function $F^*$. In this way, we will directly relate the invariants of a binary form to the universal function.

However, there are some nontrivial technical problems to be overcome. Let $k = (k_1, ..., k_{n-2})$ be a point in $\mathbb{C}^{n-2}$, and define the polynomial ideal $\mathcal{I}_k$ to be that generated by the polynomials

$$\Psi_\nu(s_1, s_2, ..., s_m, t_1, ..., t_{m'}) - k_\nu, \qquad \nu = 1, ..., n-2.$$

To each binary form $f$ we associate the ideal $\mathcal{I}_f \equiv \mathcal{I}_k$ in the ring $\mathbb{C}[s_1, ..., s_m, t_1, ..., t_{m'}]$ of polynomials in the $n$ variables $s_1, ..., s_m, t_1, ..., t_{m'}$, called the *determining ideal* of $f$, in which the parameter values $k_\nu = I_\nu$ are given by the invariants of $f$. The question of whether we can eliminate all the parameters from Eqs. (7.5), (7.8), (7.9) to deduce the form of the universal curve is intimately related to the structure of the ideal $\mathcal{I}_f$. In particular, if $\mathcal{I}_f$ is irreducible, and of dimension 1, then its set of zeros determines a curve in $\mathbb{C}^n$. We can use this curve to parametrize the covariants $I^*$, $J^*$ according to (7.5), (7.8). Eliminating the parameter leaves us with an implicit formulation of the universal function in terms of the invariants of the form, analogous to our construction for binary quartics in Section 5. However, in view of the following considerations, we do not expect the ideal $\mathcal{I}_f$ to always have dimension 1.

A binary form all of whose invariants vanish is called a *null-form*. Hilbert [13] was the first to emphasize the importance of null forms in the general theory. A basic theorem is the following characterization of null forms cf. [13, p. 301; 20, Section 4.1; 21, p. 110].

THEOREM 7.6. *Let $f(x, y)$ be a binary form of degree $n$. Then all the invariants of $f$ vanish if and only if $f$ has a root of multiplicity strictly greater than $n/2$.*

Let $f$ be a null form of degree $n$. If we translate the root of multiplicity $m > n/2$ to 0, we can then perform a further normalization to place $f$ into the canonical form

$$f(x, y) = x^m(y^j + b_{j-2}x^2y^{j-2} + \cdots + b_0x^j), \qquad m > j,$$

where $b_{j-2} = 1$ or 0, and the degree of $f$ is $n = m + j$. Therefore, for $n \leqslant 6$ there is a discrete set of inequivalent null forms, (including most of the monomials), whereas if $n \geqslant 7$, there is an $[(n-5)/2]$ parameter family of "generic" null forms of degree $n$, together with additional lower dimensional subfamilies of "special" null forms. Note that this immediately implies that the ideal $\mathcal{I}_0$, which corresponds to any null form, cannot be either irreducible or one-dimensional for $n \geqslant 7$, as otherwise Eqs. (7.5), (7.8), (7.9) would lead to a single universal curve, which would be insufficient to distinguish between the parametrized families of inequivalent null forms, and would thereby violate Theorem 4.2.

If the form $f$ is *not* a null form, then at least one of the invariants $I_1, \ldots, I_{n-2}$ does not vanish. Suppose $I_1 \neq 0$. Then we can immediately eliminate the parameter $s_1$ from the syzygies (7.9) by introducing the absolute invariants

$$A_\nu = \frac{I_\nu^{w_1}}{I_1^{w_\nu}}, \qquad \nu = 2, \ldots, n-2. \tag{7.10}$$

If, in addition, we know that the determining ideal $\mathcal{I}_f$ is irreducible and of dimension 1, then elimination of the parameters $s_1, \ldots, s_m, t_1, \ldots, t_{m'}$, from (7.5), (7.8), (7.9), will lead to an explicit formula for the universal function involving only the absolute invariants (7.10). Thus in this case, *the universal function is uniquely determined by the absolute invariants of the form.*

THEOREM 7.7. *Suppose $f$ and $\bar{f}$ are binary forms of degree $n \geqslant 4$, neither of which is complex equivalent to a monomial. (Thus neither Hessian vanishes identically, and the invariant $I^*$ is not constant in either case.) Assume that the determining ideal $\mathcal{I}_f$ is irreducible and of dimension 1. Then $f$ is complex equivalent to $\bar{f}$ if and only if $f$ and $\bar{f}$ have the same absolute invariants.*

I conjecture that the conditions on the ideal $\mathscr{I}_f$ hold when all the invariants $I_\nu \neq 0$ do not vanish, and, possibly, when any one invariant is nonzero; if true, only the null forms would not be covered by Theorem 7.7, and we would have the result that two stable forms are equivalent if and only if their absolute invariants are the same. However, I have no proof of this statement, and must leave it in conjectural form. If it is true, then Theorem 7.7 would be a generalization of a result of Clebsch [4, p. 365; 13, Section 10] which says that two binary forms of odd degree are equivalent if and only if all their absolute invariants are the same, *and* each form has two independent linear covariants $\alpha x + \beta y$, $\alpha' x + \beta' y$, meaning the determinant $\alpha\beta' - \alpha'\beta$ does not vanish. There is also a similar result [4, p. 421; 13, Section 10] for forms of even degree which have three suitably independent quadratic covariants. Clebsch's conditions imply stability, but, as the example $x^4 y + xy^4$ shows, the converse does not hold.

EXAMPLE 7.8. In the case of a quintic form, there are 23 fundamental covariants in the Hilbert basis, which Stroh takes to be

$$\begin{array}{lll}
H = (f, f)^{(2)}, & i = (f, f)^{(4)}, & T = (f, H), \\
q = (f, i), & j = -(i, f)^{(2)}, & m = (i, H) \\
h = -(i, H)^{(2)} - \frac{3}{10} i^2, & A = (i, i)^{(2)}, & r = (j, H), \\
\varepsilon = (j, i), & \alpha = -(j, i)^{(2)}, & \eta = (h, i), \\
\tau = -(h, i)^{(2)}, & n = (j, h), & \beta = (i, \alpha), \\
\vartheta = -(i, \tau), & B = (i, \tau)^{(2)}, & Q = (j, \tau), \\
\gamma = (\tau, \alpha), & C = (\beta, \alpha), & \delta = -(\vartheta, \alpha), \\
R = (\beta, \gamma). & &
\end{array}$$

In particular,

$$S_1 = H, \qquad S_2 = i, \qquad T_1 = T, \qquad T_2 = q.$$

The invariants are $A$, $B$, $C$, $R$ of which only the first three are required to determine the universal function. (There is a single syzygy among the four invariants.) The absolute invariants are

$$\frac{B}{A^2}, \qquad \frac{C}{A^3}.$$

According to Stroh [26, pp. 105–106] (see also Hammond, [12]) the 18 fundamental syzygies of the quintic stemming from Theorems 7.1, 7.2, 7.3 are

$$f^3j=\tfrac{3}{2}f^2iH-3H^3-6T^2,$$
$$fm=iT-Hq,$$
$$f^2h=H^2i+fHj-\tfrac{1}{2}f^2i^2+2qT,$$
$$f^2A=-2fij-i^2H-2q^2,$$
$$fr=Hm+jT,$$
$$f\varepsilon=im+jq,$$
$$f\alpha=j^2-\tfrac{1}{2}i^3-HA-2ih,$$
$$f\eta=qh-ir,$$
$$f\tau=2jh-H\alpha-\tfrac{1}{3}i^2j,$$
$$fn=T\alpha-2H\eta-fi\varepsilon,$$
$$f\beta=2i\eta-q\alpha,$$
$$f\vartheta=in-q\tau,$$
$$fB=h\alpha-\tfrac{1}{2}i^2\alpha-\tfrac{1}{3}Aij-2j\tau,$$
$$fQ=jn+\tau m,$$
$$f\gamma=2\tau\eta+\alpha n,$$
$$fC=2\beta\eta-\alpha^2j+\alpha Ah+\tfrac{3}{2}i\alpha\tau,$$
$$f\delta=2\vartheta\eta-\tfrac{5}{2}j\alpha\tau-Bf\alpha+h\alpha^2,$$
$$fR=Ah\gamma+Bh\beta-j\alpha\gamma+\tfrac{3}{2}\beta\tau^2+\tfrac{3}{2}i\tau\gamma.$$

It is a simple matter to determine the three syzygies (7.9) from these equations, which for a binary quintic take the form

$$A=s_1\tilde{\Phi}_1(s_2,t_1,t_2),\qquad B=s_1^2\tilde{\Phi}_2(s_2,t_1,t_2),\qquad C=s_1^3\tilde{\Phi}_3(s_2,t_1,t_2),$$

where

$$s_1=\frac{H^5}{f^6},\qquad s_2=\frac{f^2i}{H^2},\qquad t_1=\frac{T}{H^{3/2}},\qquad t_2=\frac{f^2q}{H^{5/2}}.$$

The simplest of these three functions is

$$A=s_1(6s_2-4s_2^3+12t_1^2-2t_2^2).$$

Unfortunately, the expressions for $\tilde{\Phi}_2$, $\tilde{\Phi}_3$ are extremely complicated, so an explicit determination of the universal function is not possible in general. (This may be related to the insolubility of the quintic by radicals.)

Leaving aside the monomials, we find that in the case of a binary quintic, Clebsch's theorem does not apply in the canonical cases

$$x^5 + y^5, \qquad x^4 y + x y^4, \qquad x^3(x^2 + y^2).$$

The reader can check that, based on Clebsch's calculations [4, Section 93], these forms are distinguished from each other by their absolute invariants, even though they do not fall under the realm covered by his theorem. The third form is (up to equivalence) the only null quintic which is not equivalent to a monomial. Presumably, this implies that for a binary quintic, for all $k$, including 0, the ideal $\mathscr{I}_k$ satisfies our hypotheses of being irreducible and one-dimensional; however, I do not have a complete proof of this. It also appears that reducibility problems for these ideals will arise only for forms of degree 7 or more.

## 8. Integration of the Universal Equation

The final topic to be discussed is how (in principle) to reconstruct a Lagrangian or form from its universal function $F$. Note first that since the invariants $I$ and $J$ are rational functions of the derivatives of the Lagrangian $L$, the universal relation

$$J = F(I) \tag{8.1}$$

is a fourth order ordinary differential equation for $L(p)$, whose general solution is a four-parameter family of complex-equivalent Lagrangians. (Real-equivalence must still take into account the sign and solvability restrictions of Theorem 2.5.) *Moreover*, the universal equation retains the general linear group $GL(2)$ as a symmetry group, since if $L(p)$ is a solution, so is any equivalent Lagrangian $\tilde{L}$. Therefore, Lie's theory of ordinary differential equations enables us to reduce this fourth order ordinary differential equation to a first order ordinary differential equation, cf. [22, Example 2.59]. (If the general linear group $GL(2)$ were a solvable Lie group, we could integrate (8.1) by quadratures.) We begin by restating the symmetry condition:

LEMMA 8.1. *The universal equation is invariant with respect to the four-parameter Lie group with infinitesimal generators*

$$\mathbf{v}_1 = \frac{\partial}{\partial p}, \qquad \mathbf{v}_2 = p\frac{\partial}{\partial p}, \qquad \mathbf{v}_3 = p^2\frac{\partial}{\partial p} + pL\frac{\partial}{\partial L}, \qquad \mathbf{v}_4 = L\frac{\partial}{\partial L}.$$

*Proof.* This is just the infinitesimal version of the statement that if $L(p)$ is a solution, so is $\tilde{L}(\tilde{p})$, as given by (2.5). Alternatively, this can be checked directly by noting that $I$ and $J$, as given by (2.19), (2.22), form a complete set of functionally independent differential invariants for the prolonged group action $pr^{(4)}GL(2, \mathbb{R})$, cf. [22].

We now proceed to implement the basic Lie reduction method, using, in order, the one parameter groups generated by $\mathbf{v}_4$, $\mathbf{v}_1$, $\mathbf{v}_2$. (The order is motivated by the commutator table for the group.)

*Step* 1. To straighten out the scale group generated by $\mathbf{v}_4$, we replace $L$ by

$$M = \log L.$$

We find that in terms of $M$,

$$I = \frac{(M_{ppp} + 6M_p M_{pp} + 4M_p^3)^2}{(M_{pp} + M_p^2)^3}, \qquad J = \frac{\partial \sqrt{I}/\partial p}{\sqrt{M_{pp} + M_p^2}},$$

and so (8.1) reduces to a third order equation for $\tilde{M} = M_p$.

*Step* 2. To use the translational invariance, we let

$$y = \tilde{M}, \qquad Q = \tilde{M}_p,$$

be the new independent and dependent variable, in terms of which

$$I = \frac{(QQ_y + 6yQ + 4y^3)^2}{(Q + y^2)^3}, \qquad J = \frac{Q \cdot \partial \sqrt{I}/\partial y}{\sqrt{Q + y^2}},$$

and hence (8.1) is just a second order equation for $Q(y)$.

*Step* 3. Finally, we use the scaling invariance of the second order ordinary differential equation for $Q$, which is given by the one-parameter group $(y, Q) \mapsto (\lambda y, \lambda^2 Q)$. The invariants of the group action are

$$z = \frac{Q}{y^2}, \qquad R = \frac{Q_y}{y},$$

where we regard $R$ as a function of $z$. We find that in terms of $R$ and $z$,

$$I = \frac{(zR + 6z + 4)^2}{(z + 1)^3}, \qquad J = \frac{z(R - 2z)\, \partial \sqrt{I}/\partial z}{\sqrt{z + 1}}.$$

However, rather than work with the principal invariant $R$, it is much more convenient to take

$$\hat{R}=\sqrt{I}=\frac{zR+6z+4}{(z+1)^{3/2}}$$

as the new dependent variable, in terms of which

$$J=\{(z+1)\,\hat{R}-2\sqrt{z+1}\,(z+2)\}\,\hat{R}_z.$$

Therefore, we have reduced the universal equation for the form $Q$ to a (complicated) first order ordinary differential equation, namely,

$$\frac{d\hat{R}}{dz}=\frac{F(\hat{R})}{(z+1)\,\hat{R}-2\sqrt{z+1}\,(z+2)}, \tag{8.2}$$

in which $F$ is the universal function. Retracing our various reductions, we find that we have proved the following.

THEOREM 8.2. *Let $F$ be a given function. Let $\hat{R}(z)$ be the general solution to* (8.2). *Let*

$$Q(y)=\exp\left\{\int^y \frac{dz}{2z-z\hat{R}(z)}+c_1\right\} \quad \textit{and} \quad S(w)=\int^w \frac{dy}{Q(y)}.$$

*Then*

$$L(p)=c_3\exp\left\{\int^p S^{-1}(p+c_2)\,dp\right\}$$

*describes the complete four-parameter family of Lagrangians which lead to the prescribed universal function $F$.*

Unfortunately, the explicit implementation of this construction is usually too complicated to complete.

## 9. Further Questions

(1) The case when the universal function is single-valued is important. According to Theorems 6.6 and 6.8 it occurs when $f$ is equivalent to a monomial or to a sum of two $n$th powers. However, it also occurs for other special types of forms, for instance when the transvectant $(f,f)^{(4)}=0$ vanishes, cf. (5.8), (7.8). Precisely when is $F^*$ single-valued, and what are the geometric consequences?

(2) In general, the universal function $F^*$ is not explicit. However, since $F^*$ is analytic, we can compare two universal functions $F^*$ and $\tilde{F}^*$ by expanding them both in power series. (In the multiple-valued case, we must take care that we are comparing the same sheets of the associated universal rational curve.) The question is, given the degree of the form, how many terms in the power series for $F^*$ must be examined to ensure equivalence?

(3) The fundamental result has reduced the problem of the equivalence of binary forms to the problem of determining when their universal curves are the same. However, this reduced problem is still not entirely trivial: given two parametrized curves in the plane, how can you determine whether they are in fact the same curve? The Gröbner basis method of Buchberger [1, p. 75] appears to offer a practical, computational method for answering this question. It would be very interesting to try to implement Buchberger's algorithm in this context.

(4) Is there a purely invariant-theoretic proof of Theorem 4.2? Also, what is the geometric or algebraic significance of the associated variational problem, the invariant coframe and the universal function or universal curve?

(5) A theorem of Gundelfinger, cf. [18], gives generic conditions for a binary form of degree $n$ to be written as a sum of $k$ $n$th powers. It is unclear how Gundelfinger's result and the associated theory of apolarity is related to the result in Theorem 6.8.

(6) Extensions to ternary and higher dimensional forms can be made. Bryant and Gardner [6] have looked at the corresponding Lagrangian equivalence problem using intrinsic calculations; however, a complete solution, and the classical invariant theoretic consequences remain under active investigation.

## ACKNOWLEDGMENTS

It is a pleasure to thank Niky Kamram and Bill Shadwick for igniting my interest in the Cartan equivalence method, and for many valuable discussions and correspondence on the subject. I thank Gian-Carlo Rota for encouraging my forays into classical invariant theory, and Robert Gardner for useful remarks on the equivalence problem. Also, I thank Frank Grosshans for alerting me to some errors in an earlier version of this manuscript, and pointing out the key role that the null forms play in the further development of the theory.

## REFERENCES

1. B. BUCHBERGER, 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, pp. 59–87, Springer-Verlag, New York, 1988.

2. E. CARTAN, Les problèmes d'équivalence, *in* "Œuvres Complètes," Part II, Vol. 2, pp. 1311–1334, Gauthiers-Villars, Paris, 1952.
3. E. CARTAN, Sur un problème d'équivalence et la théorie des espaces métriques généralisés, *in* "Œuvres Complètes," Part. III, Vol. 2, pp. 1131–1153, Gauthiers-Villars, Paris, 1955.
4. A. CLEBSCH, "Theorie der Binären Algebraischen Formen," B. G. Teubner, Leipzig, 1872.
5. R. B. GARDNER, Differential geometric methods interfacing control theory, *in* "Differential Geometric Control Theory," (R. W. Brockett *et al.*, Eds.), pp. 117–180, Birkhauser, Boston, 1983.
6. R. B. GARDNER, personal communication.
7. R. B. GARDNER AND W. F. SHADWICK, Equivalence of one dimensional Lagrangian field theories in the plane I, *in* "Global Differential Geometry and Global Analysis," (D. Ferus *et al.*, Eds.), Lecture Notes in Math., Vol. 1156, pp. 154–179, Springer-Verlag, New York, 1985.
8. H. GOLDSCHMIDT AND S. STERNBERG, The Hamilton–Cartan formalism in the calculus of variations, *Ann. Inst. Fourier* **23** (1973), 203–269.
9. J. H. GRACE AND A. YOUNG, "The Algebra of Invariants," Cambridge Univ. Press, Cambridge, 1903.
10. F. D. GROSSHANS, Constructing invariant polynomials via Tschirnhaus transformations, *in* "Invariant Theory," (S. S. Koh, Ed.), Lecture Notes in Math, Vol. 1278, pp. 95–102, Springer-Verlag, New York, 1987.
11. G. B. GUREVICH, "Foundations of the Theory of Algebraic Invariants," P. Noordhoff Ltd., Groningen, Holland, 1964.
12. J. HAMMOND, Syzygy tables for the binary quintic, *Amer. J. Math.* **8** (1885), 19–25.
13. D. HILBERT, Über die vollen Invariantensysteme, *in* "Ges. Abh.," Vol. II, pp. 287–344, Springer-Verlag, Berlin, 1933.
14. L. HSU AND N. KAMRAN, Classification of second-order ordinary differential equations admitting Lie groups of fiber-preserving point symmetries, *Proc. London Math. Soc.* **58** (1989), 387–416.
15. N. KAMRAN, Contributions to the study of the equivalence problem of Elie Cartan and its applications to partial and ordinary differential equations, Mémoires Cl. Sci. Acad. Roy. Belgique, to appear.
16. N. KAMRAN, K. G. LAMB, AND W. F. SHADWICK, The local equivalence problem for $d^2y/dx^2 = F(x, y, dx/dy)$ and the Painlevé transcendants, *J. Differential Geom.* **22** (1985), 139–150.
17. N. KAMRAN AND P. J. OLVER, Equivalence problems for first order Lagrangians on the line, *J. Diff. Eq.* **80** (1989), 32–78.
18. J. P. S. KUNG, Canonical forms for binary forms of even degree, *in* "Invariant Theory," (S. S. Koh, Ed.), Lecture Notes in Math., Vol. 1278, pp. 52–61, Springer-Verlag, New York, 1987.
19. J. P. S. KUNG AND G.-C. ROTA, The invariant theory of binary forms, *Bull. Amer. Math. Soc.* **10** (1984), 27–85.
20. D. MUMFORD AND J. FOGARTY, "Geometric Invariant Theory," Springer-Verlag, New York, 1982.
21. P. E. NEWSTEAD, "Introduction to Moduli Problems and Orbit Spaces," Tata Institute of Fundamental Research Lectures on Mathematics, Vol. 51, Narosa, New Delhi, 1978.
22. P. J. OLVER, "Applications of Lie Groups to Differential Equations," Graduate Texts in Mathematics, Vol. 107, Springer-Verlag, New York, 1986.
23. P. J. OLVER, Classical invariant theory and the equivalence problem for particle Lagrangians, *Bull. Amer. Math. Soc.* **18** (1988), 21–26.
24. T. A. SPRINGER, "Invariant Theory," Lecture Notes in Math., No. 585, Springer-Verlag, New York, 1977.

25. S. STERNBERG, "Lectures on Differential Geometry," Prentice-Hall, Englewood Cliffs, NJ, 1964.
26. E. STROH, Ueber eine fundamentale Eigenschaft des Ueberschiebungsprocesses und deren Verwerthung in der Theorie der binären Formen, *Math. Ann.* **33** (1889), 61–107.
27. E. STROH, Die fundamentalen Syzyganten der binären Form sechster Ordnung, *Math. Ann.* **34** (1890), 306–318.
28. E. B. VINBERG, Effective invariant theory, *Amer. Math. Soc. Transl.* **137** (1987), 15–19.