INVARIANTS OF POLYNOMIALS AND BINARY FORMS

SVANTE JANSON

Abstract. We survey various classical results on invariants of polynomials, or equivalently, of binary forms, focussing on explicit calculations for invariants of polynomials of degrees 2, 3, 4.

1. Introduction

The purpose of this survey is to collect various classical (mainly 19th century) results on invariants of polynomials, focussing on explicit formulas for invariants of polynomials of degrees 2, 3, 4. Invariants of polynomials are equivalent to invariants of binary forms, so we begin (Section 2) with a summary of definitions and some key result for these, mainly based on Schur [18]; some other books on invariants (which we only partly have consulted) are Dickson [6], Elliott [7], Glenn [8], Hilbert [10], Olver [16]. See these books for further results and proofs. Some formulas below have been calculated using `Maple`.

The theory is really simpler and more symmetric for binary forms, and the obvious correspondence between binary forms and polynomials (see Section 3) makes it in principle trivial to transfer the definitions and results to polynomials. Nevertheless, since polynomials are so common in other parts of mathematics, we find it interesting to perform this translation explicitly and to give detailed formulas for polynomials.

Remark 1.1. The formulas are purely algebraic and are valid for any ground field of characteristic 0, for example $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$.

The formulas give invariants also for fields of finite characteristic, at least as long as it does not divide any denominator (for degree ≤ 4, only characteristic 2 or 3 may have such problems), but there are also other invariants in finite characteristic. One example is the invariant [5]

$$a_0^2 a_2 + a_0\, a_2^2 + a_0\, a_1^2 + a_1^2 a_2 - a_0^3 - a_2^3 \tag{1.1}$$

of a quadratic polynomial $a_0 x^2 + a_1 x + a_2$ in $\mathbb{F}_3$. (Cf. Section 7, and note that (1.1) does not vanish for $f(x) = x^2$, unlike the discriminant Δ.) See further [6] and, for example, [19].

We ignore trivial complications with the invariant that is identically 0; for example, we may say that there is no invariant of some type, really meaning that there is no such invariant that is not identically zero. Similarly, we for

Date: 14 February, 2011.

simplicity may say that an invariant Φ is the only invariant of some type, really meaning this up to constant factors, i.e. that every such invariant is a multiple $c\Phi$ of Φ (in other words, the space of such invariants is 1-dimensional). Note also that constant factors in the definition of specific invariants usually are uninteresting, and different choices of such factors often are made in different references.

When giving examples of different notations in other papers and books, we use subscripts; for example, $A_{[18]}$ means A in [18].

We denote falling factorials by

$$(n)_k := n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!} = \binom{n}{k}k!\,. \tag{1.2}$$

2. Invariants of binary forms

We begin by collecting some definitions and general results. See e.g. Elliott [7], Hilbert [10], Kung and Rota [14] and Schur [18] for further details. (In particular, see [14] for the *umbral calculus*, which is a useful method to describe and study invariants and covariants, but which will not be used here.)

Warning. Note that the notation in these and many other references is different, since the forms there are written as $\sum_{i=0}^{n}\binom{n}{i}a_i x^{n-i}y^i$ instead of (2.2) below; i.e., $a_{i\,[7]} = a_{i\,[10]} = a_{i\,[18]} = \breve{a}_i$, where

$$\breve{a}_i := \frac{a_i}{\binom{n}{i}}. \tag{2.1}$$

The variables $\breve{a}_i$ are often more convenient for theoretical purposes, see e.g. Example 2.9 below and [18, Satz 2.18] or [14], and they are generally used in standard treatments, but for our purposes we prefer our a_i, and will only rarely use $\breve{a}_i$.

Remark 2.1. The definitions in this section extend to forms in any number $n \geq 2$ variables, but we will only consider the binary case. See [7], [10] and [18, I].

A homogeneous binary form of degree (order) n can be written as

$$f(\mathbf{x}) = f(x,y) = \sum_{i=0}^{n} a_i x^{n-i} y^i. \tag{2.2}$$

We write $\mathbf{x} := (x,y)$ and $\mathbf{a} := (a_0,\dots,a_n)$. (We regard these as row vectors.) We sometimes use instead the notation $(x_1,x_2) = (x,y)$. We further write $\partial_x = \partial_1 = \partial/\partial x = \partial/\partial x_1$ and $\partial_y = \partial_2 = \partial/\partial y = \partial/\partial x_2$, and note that, for $0 \leq i \leq n$,

$$a_i = a_i(f) = \frac{1}{(n-i)!}\partial_x^{n-i} f(0,1) = \frac{1}{(n-i)!\,i!}\partial_x^{n-i}\partial_y^i f, \tag{2.3}$$

and thus

$$\check{a}_i = \frac{1}{n!}\partial_x^{n-i}\partial_y^i f. \tag{2.4}$$

We use occasionally subscripts $\langle n\rangle$ to denote the degree of the considered forms or polynomials; for example $a_{i\,\langle n\rangle}$.

A 2×2 matrix $T = \left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$ acts on the variables (to the right) by $\mathbf{x}' = \mathbf{x}T$ and on forms (to the left) by

$$Tf(\mathbf{x}) := f(\mathbf{x}T) = f(\alpha x + \gamma y, \beta x + \delta y). \tag{2.5}$$

This gives an action of the general linear group $GL(2)$ on the set of all binary forms of degree n.

Definition 2.2. A (projective) *invariant* (of binary forms of a given degree n) is a homogeneous polynomial $\Phi(f)$ in the coefficients $\mathbf{a}$ such that

$$\Phi(Tf) = |T|^w\Phi(f) \tag{2.6}$$

for some number w and all f and $T\in GL(2)$. The number w is the *weight* (or *index*) of Φ. We denote the *degree* of Φ by ν. (We generally use ν for the degree and w for the weight, sometimes without comment; similarly we later use μ for the order of covariants and seminvariants. There are no standard notations; some examples of other notations are $i_{[7]} = i_{[8]} = g_{[10]} = r_{[18]} = \nu$ for the degree and $w_{[7]} = k_{[8]} = p_{[10]} = p_{[18]} = w$ for the weight. Further $p_{[7]} = m_{[8]} = n_{[10]} = k_{[18]} = n$ for the degree of the form and $\omega_{[8]} = m_{[18]} = \mu$ for the order, see below.)

The weight w is necessarily an integer. Taking $T = \lambda I$, which gives $Tf = \lambda^n f$, we see that

$$n\nu = 2w. \tag{2.7}$$

Hence $w\geq 0$, and $w>0$ except in the trivial case $\nu = 0$ when the invariant is a constant.

Remark 2.3. If Φ satisfies the more general equation $\Phi(Tf) = c_T\Phi(f)$ for some collection of numbers c_T, then necessarily $c_T = |T|^w$ for some w, so Φ is an invariant as defined above. Similarly, in definitions below, we may equivalently allow arbitrary factors c_T in (2.8), (2.14), (2.16), (2.31), (2.32); these necessarily have to have the given form $|T|^w$ or $\alpha^\mu|T|^w$ for some w and μ.

Remark 2.4. The identity (2.6) is a polynomial identity in the entries of T, and thus it extends to all 2×2 matrices T, also singular. Thus $\Phi(Tf) = 0$ whenever T is singular, except in the trivial case $\nu = w = 0$ when Φ is a constant. The same applies to similar formulas below.

Definition 2.5. Similarly, a *joint invariant* of several forms $f_1,\dots,f_\ell$, of degrees $n_1,\dots,n_\ell$, is a polynomial in the coefficients of $f_1,\dots,f_\ell$, homogeneous of degrees $\nu_1,\dots,\nu_\ell$, respectively, such that

$$\Phi(Tf_1,\dots,Tf_\ell) = |T|^w\Phi(f_1,\dots,f_\ell) \tag{2.8}$$

for some w, the *weight* of Φ, and all $f_1, \dots, f_\ell$ and $T \in GL(2)$.

In this case we have

$$n_1\nu_1 + \cdots + n_\ell\nu_\ell = 2w. \tag{2.9}$$

Again w is an integer with $w \geq 0$, and $w > 0$ except in the trivial case of a constant invariant.

Remark 2.6. The assumption that Φ is homogeneous separately in the coefficients of each f_j is no real restriction, since any invariant polynomial Q can be decomposed into homogeneous components which are invariant.

Example 2.7. The *apolar invariant* of two binary forms $f(x,y) = \sum_{i=0}^n a_i x^{n-i} y^i$ and $g(x,y) = \sum_{i=0}^n b_i x^{n-i} y^i$ of the same degree n is

$$\begin{aligned} A(f,g) &:= \sum_{i=0}^n (-1)^i i!\,(n-i)!\,a_i b_{n-i} = n! \sum_{i=0}^n (-1)^i \frac{a_i b_{n-i}}{\binom{n}{i}} \\ &= n! \sum_{i=0}^n (-1)^i \binom{n}{i} \check{a}_i \check{b}_{n-i} \\ &= f(\partial_y, -\partial_x) g(x,y) = g(-\partial_y, \partial_x) f(x,y). \end{aligned} \tag{2.10}$$

This is a joint invariant of f and g of degrees $\nu_1 = \nu_2 = 1$ and weight n. (Our definition differs from [18] by a factor $n!$: $A_{[18]}(f,g) = A(f,g)/n!$.) The apolar invariant is also called *transvectant*, see Example 2.15 below. Using (2.3), we also have

$$A(f,g) = \sum_{i=0}^n (-1)^i (\partial_x^{n-i} f \cdot \partial_x^i g)(0,1). \tag{2.11}$$

Note that $A(g,f) = (-1)^n A(f,g)$; hence the apolar invariant is symmetric in f and g if n is even, and antisymmetric if n is odd.

The apolar invariant is the only joint invariant with degrees $\nu_1 = \nu_2 = 1$ of two binary forms of the same degree, and there are no such invariants of binary forms of different degrees [18, Satz 2.6].

See Example 2.16 for a generalization.

Example 2.8. Taking $f = g$ in Example 2.7 we obtain the *apolar invariant* (or *transvectant*, see Example 2.15) of a single binary form

$$A(f,f) := \sum_{i=0}^n (-1)^i i!\,(n-i)!\,a_i a_{n-i}; \tag{2.12}$$

this is an invariant of degree $\nu = 2$ and weight $w = n$, for any even n. (Note that $A(f,f) = 0$ when n is odd.)

In fact this is the only invariant of degree 2; if n is odd there is thus no such invariant [18, Satz 2.5].

Example 2.9. If $n = 2q$ is even, then the *Hankel determinant*

$$\mathrm{Han}(f) = \left|\breve{a}_{i+j}\right|_{i,j=0}^{q}, \qquad \text{with} \quad \breve{a}_l = a_l/\tbinom{n}{l}, \tag{2.13}$$

is an invariant of degree $\nu = q+1 = n/2+1$ and, by (2.7), weight $w = q(q+1)$. The Hankel determinant is also called the *catalecticant*.

2.1. **Covariants.**

Definition 2.10. More generally, a (projective) *covariant* is a polynomial $\Psi(f;\mathbf{x}) = \Psi(\mathbf{a};\mathbf{x})$ in $\mathbf{x}$ and the coefficients $\mathbf{a}$ of f such that

(i) Ψ is homogeneous in $\mathbf{a}$ of some degree ν, the *degree* of Ψ;
(ii) Ψ is homogeneous in $\mathbf{x}$ of some degree μ, the *order* of Ψ;
(iii)
$$\Psi(Tf;\mathbf{x}T^{-1}) = |T|^w \Psi(f;\mathbf{x}) \tag{2.14}$$
for some integer w, the *weight* of Ψ, and all forms f (of degree n) and all $T \in GL(2)$.

Hence, an invariant is a covariant of order 0.

The relation (2.7) generalizes to

$$n\nu = m + 2w. \tag{2.15}$$

Definition 2.11. Similarly, a *joint covariant* of forms $f_1,\dots,f_\ell$ of degrees $n_1,\dots,n_\ell$ is a polynomial $\Psi(f_1,\dots,f_\ell;\mathbf{x}) = \Psi(\mathbf{a}_1,\dots,\mathbf{a}_\ell;\mathbf{x})$ in the coefficients $\mathbf{a}_j$ of f_j, $j = 1,\dots,\ell$, that is homogeneous in each $\mathbf{a}_j$ of degree ν_j, homogeneous in $\mathbf{x}$ of degree μ, the *order* of Ψ, and such that

$$\Psi(Tf_1,\dots,Tf_\ell;\mathbf{x}T^{-1}) = |T|^w \Psi(f_1,\dots,f_\ell;\mathbf{x}) \tag{2.16}$$

for some integer w, the *weight* of Ψ, and all forms $f_1,\dots,f_\ell$ and all $T \in GL(2)$.

We now have

$$n_1\nu_1 + \dots + n_\ell\nu_\ell = m + 2w. \tag{2.17}$$

Example 2.12. The form $f(\mathbf{x})$ itself is a covariant of degree $\nu = 1$, order $\mu = n$ and weight $w = 0$.

Example 2.13. The *Hessian covariant*

$$H(f) = H(f;\mathbf{x}) := \left|\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{1\le i,j\le 2} \tag{2.18}$$

is a covariant of degree $\nu = 2$, order $\mu = 2(n-2)$ and weight $w = 2$. (Other notation: $\mathcal{H}_{[10]} = (n(n-1))^{-2}H$.)

Example 2.14. The *Jacobian determinant*

$$J(f_1,f_2) = J(f_1,f_2;\mathbf{x}) := \left|\frac{\partial f_i}{\partial x_j}\right|_{1\le i,j\le 2} \tag{2.19}$$

is a joint covariant of degrees $\nu_1 = \nu_2 = 1$, order $n_1 + n_2 - 2$ and weight $w = 1$. Note that J is antisymmetric; $J(g,f) = -J(f,g)$, and $J(f,f) = 0$.

Example 2.15. The *k:th transvectant* $\{f,g\}_k$ is a joint covariant of two forms f and g of arbitrary degrees n_1 and n_2, defined by

$$\begin{aligned}\{f,g\}_k &= \Bigl(\frac{\partial}{\partial x_1}\frac{\partial}{\partial y_2} - \frac{\partial}{\partial x_2}\frac{\partial}{\partial y_1}\Bigr)^k f(\mathbf{x})g(\mathbf{y})\Big|_{\mathbf{y}=\mathbf{x}} \\ &= \sum_{i=0}^{k}(-1)^i\binom{k}{i}\frac{\partial^k f}{\partial x_1^{k-i}\partial x_2^i}\frac{\partial^k g}{\partial x_1^i\partial x_2^{k-i}}. \end{aligned} \tag{2.20}$$

Here $k \geq 0$ is an arbitrary positive integer, but it is easy to see that $\{f,g\}_k = 0$ unless $k \leq \min(n_1,n_2)$. (Trivially $\{f,g\}_0 = fg$.) It is easy to see that $\{f,g\}_k$ is a joint covariant of degrees $\nu_1 = \nu_2 = 1$, order n_1+n_2-2k and weight $w = k$. (Other notations: $\{f,g\}_k = (n_1)_k(n_2)_k(f,g)_{k\,[1]} = (f,g)^k_{[8]} = (n_1)_k(n_2)_k(f,g)_{k\,[10]} = (f,g)^{(k)}_{[16]} = [f,g]^k_{[17]}$.)

Furthermore, $\{f,g\}_k = (-1)^k\{g,f\}_k$, so $\{f,g\}_k$ is symmetric if k is even and anti-symmetric if k is odd. In particular, $\{f,f\}_k = 0$ for odd k, but for even $k \leq n$, $\{f,f\}_k$ is a non-trivial covariant of degree 2, order $2n-2k$ and weight k. (Other notations: $f_{k\,[10]} := \frac{1}{2}(f,f)_{k\,[10]} = \frac{1}{2}(n)_k^{-2}\{f,f\}_k$.)

The first transvectant is the Jacobian covariant in Example 2.14:

$$\{f,g\}_1 = J(f,g). \tag{2.21}$$

The second transvectant $\{f,f\}_2$ is (twice) the Hessian covariant in Example 2.13:

$$\{f,f\}_2 = 2H(f). \tag{2.22}$$

Furthermore, in the case $n_1 = n_2 = n = k$, $\{f,g\}_n$ is of order 0, i.e., an invariant. In this case, by a binomial expansion in (2.20), (2.4) and (2.10),

$$\begin{aligned}\{f,g\}_n &= \sum_{i=0}^{n}\binom{n}{i}(-1)^i\Bigl(\frac{\partial}{\partial x_1}\frac{\partial}{\partial y_2}\Bigr)^{n-i}\Bigl(\frac{\partial}{\partial x_2}\frac{\partial}{\partial y_1}\Bigr)^i f(\mathbf{x})g(\mathbf{y}) \\ &= \sum_{i=0}^{n}\binom{n}{i}(-1)^i n!\,\check{a}_i\,n!\,\check{b}_{n-i} \\ &= n!\,A(f,g). \end{aligned}\tag{2.23}$$

Hence the apolar invariant equals (apart from a factor $1/n!$) the nth transvectant $\{f,g\}_n$.

For relations between transvectants, and interpretations in terms of representations of SL_2, see Abdesselam and Chipalkatti [1].

Example 2.16. As shown in (2.23), the transvectant $\{f,g\}_n$ of two binary forms of equal degree n is (apart from a constant factor) their apolar invariant. More generally, if f and g are binary forms of degrees n and m with $n \geq m \geq 0$, the *apolar covariant* $\{f,g\}$ is defined as the highest non-trivial transvectant (i.e., the mth transvectant), which by (2.20) and a short calculation can be expressed as

$$\{f,g\} := \{f,g\}_m = m!\,g(-\partial_2,\partial_1)f(x_1,x_2). \tag{2.24}$$

(Hence, if $m = n$, $\{f, g\} = n!\,A(f, g)$, so the apolar covariant then reduces to the apolar invariant in Example 2.7, except for the trivial but inconvenient factor $n!$.)

By Example 2.15, $\{f, g\}$ is a joint covariant of degrees $\nu_1 = \nu_2 = 1$, order $\mu = n - m$ and weight $w = m$.

Note the asymmetry in the definition; we assume $n \geq m$.

Example 2.17. The *kth Gundelfinger covariant* $G_k(f)$, for $k = 0, 1, \dots$, is the $(k + 1) \times (k + 1)$ determinant

$$G_k(f) := \left| \frac{\partial^{2k} f(x, y)}{\partial x^{2k-i-j} \partial y^{i+j}} \right|_{0 \leq i,j \leq k}; \tag{2.25}$$

this is a covariant of degree $\nu = k + 1$, order $\mu = (k + 1)(n - 2k)$ and weight $w = k(k + 1)$, see [9] and [13]. Note that $G_0(f) = f$ and $G_1(f) = H(f)$, the Hessian covariant; further $G_k(f) = 0$ if $k > n/2$. If n is even and $k = n/2$, then, by (2.4) and (2.13),

$$G_{n/2}(f) = \left| n!\, \breve{a}_{i+j} \right|_{i,j=0}^{n/2} = n!^{n/2+1} \operatorname{Han}(f), \tag{2.26}$$

a constant times the Hankel determinant (catalecticant) in Example 2.9.

A covariant Ψ of order μ can be written $\Psi(f; \mathbf{x}) = \Phi(f) x_1^\mu + \dots$; we call the coefficient $\Phi(f)$ the *source* or *leading coefficient* of Ψ. (And similarly for joint covariants.) The source of Ψ is thus given by

$$\Phi(f) := \Psi(f; 1, 0); \tag{2.27}$$

equivalently,

$$\Phi(f) := \tfrac{1}{\mu!} \partial_1^\mu \Psi(f; \mathbf{x}). \tag{2.28}$$

Conversely, by (2.14), Ψ can be recovered by

$$\Psi(f; x, y) = x^{-2w} \Phi\big(T_{x,y}^{(1)} f\big) = x^{-w} \Phi\big(T_{x,y}^{(2)} f\big) = x^\mu \Phi\big(T_{x,y}^{(3)} f\big), \tag{2.29}$$

where

$$T_{x,y}^{(1)} := \begin{pmatrix} x & y \\ 0 & x \end{pmatrix}, \qquad T_{x,y}^{(2)} := \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}, \qquad T_{x,y}^{(3)} := \begin{pmatrix} 1 & y/x \\ 0 & 1 \end{pmatrix}. \tag{2.30}$$

2.2. **Seminvariants.**

Definition 2.18. A *seminvariant* (of binary forms of degree n) is a homogeneous polynomial $\Phi(f)$ in the coefficients $\mathbf{a}$ such that

$$\Phi(Tf) = \alpha^\mu |T|^w \Phi(f) \tag{2.31}$$

for some $\mu, w \geq 0$ and all f and T of the form $\left(\begin{smallmatrix} \alpha & 0 \\ \gamma & \delta \end{smallmatrix}\right)$. The number μ is the *order* and w is the *weight* of Φ. We denote the *degree* of Φ by ν.

Definition 2.19. Similarly, a *joint seminvariant* of several forms $f_1, \dots, f_\ell$, of degrees $n_1, \dots, n_\ell$, is a polynomial $\Phi(f_1, \dots, f_\ell)$ in the coefficients of $f_1, \dots, f_\ell$, homogeneous of degrees $\nu_1, \dots, \nu_\ell$, respectively, such that

$$\Phi(Tf_1, \dots, Tf_\ell) = \alpha^\mu |T|^w \Phi(f_1, \dots, f_\ell). \tag{2.32}$$

for some $\mu, w \geq 0$, the *order* and *weight* of Φ, and all $f_1, \dots, f_\ell$ and $T = \left(\begin{smallmatrix} \alpha & 0 \\ \gamma & \delta \end{smallmatrix}\right)$.

In other words, a (joint) seminvariant is an invariant for the subgroup of $GL(2)$ given by $\left\{\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right) \in GL(2) : \beta = 0\right\} = \left\{\left(\begin{smallmatrix} \alpha & 0 \\ \gamma & \delta \end{smallmatrix}\right) : \alpha\delta \neq 0\right\}$.

We still have (2.15) and (2.17), respectively. In fact, these are equivalent to invariance for all $T = \lambda I = \left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda \end{smallmatrix}\right)$. Consequently, if (2.15) or (2.17) holds, it is enough that (2.31) or (2.32) holds for T of the form $\left(\begin{smallmatrix} \alpha & 0 \\ \gamma & 1 \end{smallmatrix}\right)$; these are the transformations $(x, y) \mapsto (\alpha x + \gamma y, y)$, which form a group $A(1)$ obviously isomorphic to the group of affine maps $x \mapsto \alpha x + \gamma$ in one dimension.

Furthermore,we say that a coefficient a_i has *weight* i, and more generally that a monomial $a_0^{k_0} a_1^{k_1} a_2^{k_2} \cdots$ has weight $k_1 + 2k_2 + \cdots$. A polynomial in $\mathbf{a} = (a_0, \dots, a_n)$ is *isobaric* if all its terms has the same weight, and then this is said to be the weight of the polynomial. It is easily seen that the invariance (2.31) or (2.32) holds for all T of the form $\left(\begin{smallmatrix} 1 & 0 \\ 0 & \delta \end{smallmatrix}\right)$ if and only if Φ is isobaric of weight w. (In this case, (2.31)=(2.6) and (2.32)=(2.8).) Consequently, the invariance (2.31) or (2.32) holds for all diagonal matrices T if and only if Φ is homogeneous and isobaric and (2.15) or (2.17) holds. This leads to the following characterization, see [18, §II.2].

Theorem 2.20. *The following are equivalent for a polynomial Φ in the coefficients of one or several binary forms.*

(i) *Φ is a (joint) seminvariant*
(ii) *Φ is homogeneous and invariant for $A(1)$.*
(iii) *Φ is homogeneous and isobaric and invariant for all T of the form $\left(\begin{smallmatrix} 1 & 0 \\ t & 1 \end{smallmatrix}\right)$, i.e., translations $(x, y) \mapsto (x + t, y)$. (For such T, the invariance is simply $\Phi(Tf) = \Phi(f)$ or $\Phi(Tf_1, \dots, Tf_\ell) = \Phi(f_1, \dots, f_\ell)$.)*
(iv) *Φ is homogeneous and isobaric and satisfies the* Cayley–Aronhold differential equation

$$\Omega(\Phi) := \sum_{i=1}^{n} (n - i + 1) a_{i-1} \frac{\partial \Phi}{\partial a_i} = 0; \tag{2.33}$$

for a joint seminvariant $\Phi(\mathbf{a}_1, \dots, \mathbf{a}_\ell)$, with $\mathbf{a}_j = (a_{1,j}, \dots, a_{n_j,j})$, the equation takes the form

$$\Omega(\Phi) := \sum_{j=1}^{\ell} \sum_{i=1}^{n_j} (n_j - i + 1) a_{i-1,j} \frac{\partial \Phi}{\partial a_{i,j}} = 0, \tag{2.34}$$

where thus $\Omega = \sum_{j=1}^{\ell} \Omega_j$.

In (iii), it suffices to consider the special $T = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, i.e., $(x, y) \mapsto (x+y, y)$.

Remark 2.21. Using $\breve{a}_i := a_i / \binom{n}{i}$ as in [18], (2.33) becomes

$$\Omega(\Phi) = \sum_{i=1}^{n} i \breve{a}_{i-1} \frac{\partial \Phi}{\partial \breve{a}_i} = 0. \tag{2.35}$$

Remark 2.22. There is also a dual differential operator

$$\Omega^*(\Phi) := \sum_{i=0}^{n-1}(i+1)a_{i+1}\frac{\partial\Phi}{\partial a_i} = \sum_{i=0}^{n-1}(n-i)\breve{a}_{i+1}\frac{\partial\Phi}{\partial\breve{a}_i}, \tag{2.36}$$

and similarly for joint seminvariants with $\Omega^* := \sum_j \Omega_j^*$. The differential equation

$$\Omega^*(\Phi) = 0 \tag{2.37}$$

holds for *invariants* Φ, but not for other seminvariants. In fact, (2.37) is a necessary and sufficient condition for a seminvariant Φ to be an invariant [18, Sätze 2.1–2.2]. (Other notations: $\mathcal{D}_{[18]} = \Omega$, $\Delta_{[18]} = \Omega^*$.)

Obviously, an invariant is a seminvariant. Moreover, there is an important correspondence between covariants and seminvariants.

Theorem 2.23. *For any n, there is a one-to-one correspondence between covariants Ψ and seminvariants Φ, such that Φ is the source of Ψ; see* (2.27)*–*(2.29)*.*

More generally, for any $n_1, \dots, n_\ell$, there is a one-to-one correspondence between joint covariants and joint seminvariants given by taking the source (leading coefficient).

The degrees, order and weight are preserved by this correspondence.

Remark 2.24. Another way to recover the covariant Ψ from its source Φ is by the formula [18, pp. 56–58]

$$\Psi = \sum_{j=0}^{\mu}\frac{(\Omega^*)^j(\Phi)}{j!}x^{\mu-j}, \tag{2.38}$$

where μ is the order. Since further $(\Omega^*)^{\mu+1}\Phi = 0$, the sum can also be written $\sum_{j=0}^{\infty}(\Omega^*)^j(\Phi)\,x^{\mu-j}/j!$.

Remark 2.25. We have defined the weight of a covariant so that it equals the weight of its source. It is easy to see, arguing as for Theorem 2.20, that if we give x weight 1 and y weight 0, then a covariant is isobaric, with each term of weight $w+\mu$. (Some references, e.g. [8], call our w the *index* of the covariant, and call $w+\mu$ the *weight*, but we do not make this definition. Note that if we instead give x weight 0 and y weight -1, then the covariant is isobaric with weight w.)

Example 2.26. The source of $f(\mathbf{x})$, i.e., the seminvariant corresponding to $f(\mathbf{x})$, see Example 2.12, is a_0. This has degree 1, order n, weight 0.

Example 2.27. The Hessian seminvariant H_0 is the source of the Hessian covariant H in Example 2.13. It is, by a simple calculation,

$$H_0(f) := 2n(n-1)a_0a_2 - (n-1)^2a_1^2 = n^2(n-1)^2(\breve{a}_0\,\breve{a}_2 - \breve{a}_1^2). \tag{2.39}$$

H_0 has degree 2, order $2n-4$ and weight 2.

Example 2.28. The Jacobian joint seminvariant of two binary forms $f(x) = \sum_{i=0}^{n_1} a_i x^{n_1-i} y^i$ and $g(x) = \sum_{i=0}^{n_2} b_i x^{n_2-i} y^i$, corresponding to the Jacobian joint covariant in Example 2.14, is, by a simple calculation,

$$n_1 a_0 b_1 - n_2 b_0 a_1 = n_1 n_2 (\check{a}_0 \check{b}_1 - \check{b}_0 \check{a}_1). \tag{2.40}$$

This has degrees $\nu_1 = \nu_2 = 1$, order $n_1 + n_2 - 2$ and weight 1.

Example 2.29. The source g_k of the kth Gundelfinger covariant in Example 2.17 is, by (2.25), (2.27) and (2.4),

$$\begin{aligned} g_k(f) = G_k(f; 1, 0) &= \left| \frac{\partial^{2k} f}{\partial x^{2k-i-j} \partial y^{i+j}}(1,0) \right|_{0 \le i,j \le k} \\ &= \left| \frac{1}{(n-2k)!} \frac{\partial^n f}{\partial x^{n-i-j} \partial y^{i+j}} \right|_{0 \le i,j \le k} \\ &= \bigl((n)_{2k}\bigr)^{k+1} \bigl| \check{a}_{i+j} \bigr|_{i,j=0}^{k}. \end{aligned} \tag{2.41}$$

This is a seminvariant of degree $\nu = k + 1$, order $\mu = (k + 1)(n - 2k)$ and weight $w = k(k + 1)$ by Example 2.17.

Cf. the special cases in Example 2.27 ($k = 1$) and Example 2.9 ($n = 2k$).

Example 2.30. The source $\tau_k(f, g)$ of the transvectant $\{f, g\}_k$ is given by, using (2.28), (2.20), (2.23), (2.11), (2.3) and (2.1),

$$\begin{aligned} \tau_k(f,g) &= \frac{1}{(n_1+n_2-2k)!} \partial_1^{n_1+n_2-2k} \Bigl(\Bigl(\frac{\partial}{\partial x_1} \frac{\partial}{\partial y_2} - \frac{\partial}{\partial x_2} \frac{\partial}{\partial y_1} \Bigr)^k f(\mathbf{x}) g(\mathbf{y}) \Big|_{\mathbf{y}=\mathbf{x}} \Bigr) \\ &= \frac{1}{(n_1+n_2-2k)!} \Bigl(\Bigl(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial y_1} \Bigr)^{n_1+n_2-2k} \Bigl(\frac{\partial}{\partial x_1} \frac{\partial}{\partial y_2} - \frac{\partial}{\partial x_2} \frac{\partial}{\partial y_1} \Bigr)^k f(\mathbf{x}) g(\mathbf{y}) \Big|_{\mathbf{y}=\mathbf{x}} \Bigr) \\ &= \frac{1}{(n_1+n_2-2k)!} \Bigl(\frac{\partial}{\partial x_1} \frac{\partial}{\partial y_2} - \frac{\partial}{\partial x_2} \frac{\partial}{\partial y_1} \Bigr)^k \Bigl(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial y_1} \Bigr)^{n_1+n_2-2k} f(\mathbf{x}) g(\mathbf{y}) \Big|_{\mathbf{y}=\mathbf{x}} \\ &= \frac{1}{(n_1-k)!\,(n_2-k)!} \Bigl(\frac{\partial}{\partial x_1} \frac{\partial}{\partial y_2} - \frac{\partial}{\partial x_2} \frac{\partial}{\partial y_1} \Bigr)^k \partial_1^{n_1-k} f(\mathbf{x}) \partial_1^{n_2-k} g(\mathbf{y}) \Big|_{\mathbf{y}=\mathbf{x}} \\ &= \frac{1}{(n_1-k)!\,(n_2-k)!} \bigl\{ \partial_1^{n_1-k} f,\, \partial_1^{n_2-k} g \bigr\}_k \\ &= \frac{k!}{(n_1-k)!\,(n_2-k)!} A_{\langle k \rangle} \bigl(\partial_1^{n_1-k} f,\, \partial_1^{n_2-k} g \bigr) \\ &= \frac{k!}{(n_1-k)!\,(n_2-k)!} \sum_{i=0}^{k} (-1)^i \bigl(\partial_1^{n_1-i} f \cdot \partial_1^{n_2-k+i} g \bigr)(0,1) \\ &= \frac{k!}{(n_1-k)!\,(n_2-k)!} \sum_{i=0}^{k} (-1)^i (n_1-i)!\, a_i\, (n_2-k+i)!\, b_{k-i} \\ &= \frac{n_1!\, n_2!}{(n_1-k)!\,(n_2-k)!} \sum_{i=0}^{k} (-1)^i \binom{k}{i} \check{a}_i\, \check{b}_{k-i}. \end{aligned} \tag{2.42}$$

Note from Example 2.15 that $\tau_1(f,g)$ is the Jacobian seminvariant in Example 2.28, and $\tau_2(f,f) = 2\,H_0(f)$, the Hessian seminvariant in Example 2.27, while if $n_1 = n_2 = n$, then $\tau_n(f,g) = \{f,g\}_n$ is $n!\,A(f,g)$, the apolar invariant in Example 2.7. Further, the special case $n_2 = k$ yields the source of the apolar covariant in Example 2.16.

The group $GL(2)$ is generated by the subgroup $A(1)$ and the reflection $\rho = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ which interchanges x and y. Hence, Φ is invariant if and only it is invariant under both $A(1)$ and ρ, i.e., if and only if it is a seminvariant that is invariant under ρ. We have $\rho(x,y) = (y,x)$ and thus, by (2.5), $\rho f(x,y) = f(y,x)$. We denote ρf by $f^\dagger$. It follows from (2.2) that if f has coefficients $\mathbf{a} = (a_0,\dots,a_n)$ as in (2.2), then $f^\dagger$ has coefficients

$$\mathbf{a}^\dagger := (a_n,\dots,a_0). \tag{2.43}$$

This leads to the following companion to Theorem 2.20, which can be used together with Theorem 2.20 to find convenient criteria for invariants.

Theorem 2.31. *The following are equivalent for a polynomial Φ in the coefficients of one or several binary forms.*

(i) Φ *is a (joint) invariant.*
(ii) Φ *is a (joint) seminvariant and* $\Phi(\mathbf{a}^\dagger) = (-1)^w\Phi(\mathbf{a})$ *or* $\Phi(\mathbf{a}_1^\dagger,\dots,\mathbf{a}_\ell^\dagger) = (-1)^w\Phi(\mathbf{a}_1,\dots,\mathbf{a}_\ell)$.
(iii) Φ *is a (joint) seminvariant of order* $\mu = 0$.
(iv) Φ *is a (joint) seminvariant and* $n\nu = 2w$ *or* $n_1\nu_1 + \dots + n_\ell\nu_\ell = 2w$ *for the degree(s) and the weight (i.e.,* (2.7) *or* (2.9) *holds).*

Proof. (i) $\iff$ (ii) by the discussion above.
(i) $\iff$ (iii) by the correspondence in Theorem 2.23 and the fact that an invariant is a covariant of order 0 and conversely.
(iii) $\iff$ (iv) by (2.15) and (2.17). □

It is obvious that we can take linear combinations of (joint) invariants, covariants or seminvariants with the same degrees, weights and orders. Furthermore, a product of (joint) invariants, covariants or seminvariants is always another invariant, covariant or seminvariant, with degrees, weights and orders in the factors added. Consequently, an isobaric polynomial in invariants is another invariant; the same is true for covariants and seminvariants provided the result also is homogeneous in the coefficients $a_0,\dots,a_n$.

We say that a set $\mathcal{B}$ of invariants (etc.) is a *basis* if every invariant (etc.), of forms of the given degree(s), is a (necessarily isobaric) polynomial in elements of $\mathcal{B}$. (Less formally, one also says that the invariants in $\mathcal{B}$ are all invariants, thus really meaning that every invariant is a polynomial of invariants in $\mathcal{B}$.) It is proved by Gordan (and more generally by Hilbert), that for any n, there exists a finite basis of the invariants (covariants or seminvariants).

We have also the following.

Theorem 2.32. *A covariant of a sequence of covariants* $\Psi_1(\mathbf{a}_1, \dots, \mathbf{a}_l; \mathbf{x})$*,* $\Psi_2(\mathbf{a}_1, \dots, \mathbf{a}_l; \mathbf{x})$*, … is itself a covariant.*

Example 2.33. As said in Example 2.12, the form $f(\mathbf{x})$ itself is a covariant of degree 1, order n and weight 0. Thus f^2 is a covariant of degree 2, order $2n$ and weight 0. Hence, see Example 2.8, the apolar invariant $A(f^2, f^2)$ is an invariant (note that $2n$ is even); it is easily seen that this invariant has degree 4 and weight $2n$, cf. (2.7). (It is shown in [18, p. 42] that $A(f^2, f^2)$ does not vanish identically for any $n \geq 2$.)

Example 2.34. The Hessian covariant $H(f; \mathbf{x})$ in Example 2.13 has degree 2 and order $2n - 4$; hence the apolar invariant $A(H(f; \mathbf{x}), H(f; \mathbf{x}))$ is an invariant of degree 4 and, by (2.7), weight $2n$. (It is shown in [18, p. 43] that $A(H(f), H(f))$ does not vanish identically for any $n \geq 2$.)

2.3. **Rational invariants.** By definition, invariants etc. are required to be polynomials in the coefficients. A few times we will consider a minor extension.

Definition 2.35. A *rational invariant* (joint invariant, covariant, etc.) is a rational function of the coefficients that has the invariance property (2.6) (etc.).

It is easily seen that a rational invariant is the same as a quotient of two invariants (etc.) [18, Satz 1.4]. Note that a rational invariant may be infinite or undefined for certain values of the coefficients.

We will in the sequel sometimes use rational seminvariants of the form $a_0^{-k}\Phi$, where Φ is a seminvariant. Another interesting case is the following.

Definition 2.36. An *absolute invariant* is a rational invariant with weight $w = 0$; it is thus a rational function of the coefficients that satisfies

$$\Phi(Tf) = \Phi(f) \tag{2.44}$$

for all $T \in GL(2)$.

By (2.6), there are no non-trivial invariants that are absolute invariants; we have to consider rational invariants here. Any absolute invariant is the quotient Φ_1/Φ_2 of two invariants of the same weight, and thus the also the same degree; conversely, any such quotient is an absolute invariant. One example is given in Example 9.2.

2.4. **Dimensions.** The set of all covariants of degree ν and weight w of binary forms of a given degree n is a linear space. We let $N(n, \nu, w)$ be its dimension, i.e., the number of linearly independent covariants of this degree and weight. By Theorem 2.23, $N(n, \nu, w)$ is also the dimension of the linear space of seminvariants of degree ν and weight w. (Theorem 2.23 yields an isomorphism between the two linear spaces.) Note that we get the invariants of degree ν by taking $w = n\nu/2$ (provided this is an integer), see (2.7) and (2.15).

The number $N(n,\nu,w)$ can be computed as follows by a formula by Cayley (the first complete proof was given by Sylvester), see [18, Sätze 2.21–2.22].

Let $\begin{bmatrix} n \\ k \end{bmatrix}_q$ be the *Gaussian polynomial* defined by

$$\begin{bmatrix} n \\ k \end{bmatrix}_q := \frac{\prod_{i=1}^n (1-q^i)}{\prod_{i=1}^k (1-q^i) \prod_{j=1}^{n-k} (1-q^j)} = \frac{\prod_{i=n-k+1}^n (1-q^i)}{\prod_{i=1}^k (1-q^i)}. \tag{2.45}$$

(See further e.g. Andrews [2].) We let $[q^w]P(q)$ denote the coefficient of q^w in a polynomial $P(q)$.

Theorem 2.37. *If $2w \le n\nu$, then*

$$\begin{aligned} N(n,\nu,w) &= [q^w]\left((1-q)\begin{bmatrix} n+\nu \\ n \end{bmatrix}_q\right) = [q^w]\left((1-q)\begin{bmatrix} n+\nu \\ \nu \end{bmatrix}_q\right) \\ &= [q^w]\frac{\prod_{i=\nu+1}^{\nu+n}(1-q^i)}{\prod_{i=2}^{n}(1-q^i)} = [q^w]\frac{\prod_{i=n+1}^{n+\nu}(1-q^i)}{\prod_{i=2}^{\nu}(1-q^i)} \\ &= [q^w]\begin{bmatrix} n+\nu \\ n \end{bmatrix}_q - [q^{w-1}]\begin{bmatrix} n+\nu \\ n \end{bmatrix}_q. \end{aligned} \tag{2.46}$$

If $2w > n\nu$, then $N(n,\nu,w) = 0$.

It follows that if we fix n and w, $N(n,\nu,w)$ is the same for all $\nu \ge w$, and is given by a simple generating function; see also Remark 3.20 below.

Corollary 2.38. *If $n \ge 2$ and $\nu \ge w$, then*

$$N(n,\nu,w) = [q^w]\prod_{i=2}^{n}(1-q^i)^{-1}. \tag{2.47}$$

Proof. The factors $1-q^i$ with $i \ge \nu+1 > w$ do not affect $[q^w]$. □

3. Invariants of polynomials

We may identify the binary form $\widetilde{f}(x,y) = \sum_{i=0}^n a_i x^{n-i} y^i$ and the polynomial $f(x) = \sum_{i=0}^n a_i x^{n-i}$; this gives a one-to-one correspondence between binary forms of degree n and polynomials of degree (at most) n described by $f(x) = \widetilde{f}(x,1)$ and, conversely, $\widetilde{f}(x,y) = y^n f(x/y)$. We let $\mathcal{P}_n$ denote the set of all such polynomials $\sum_{i=0}^n a_i x^{n-i}$.

Remark 3.1. When n is given, we will say "polynomial of degree n" for any polynomial $\sum_{i=0}^n a_i x^{n-i}$, even if $a_0 = 0$. (As just said, this gives a correspondence with binary forms of degree n.) This thus includes polynomials of lower degrees. See further Subsection 3.2.

A transform $T = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ acts on binary forms by (2.5); this transfer to the action

$$Tf(x) = (\beta x + \delta)^n f\Bigl(\frac{\alpha x + \gamma}{\beta x + \delta}\Bigr) \tag{3.1}$$

on polynomials.

We define a (projective) invariant or seminvariant of a polynomial f (of some given degree) as an invariant or seminvariant of the corresponding binary form $\widetilde{f}$; similarly, a (projective) covariant is a polynomial $\Psi(\mathbf{a};x)$ of some degree μ (or less) in x such that the corresponding binary form $\widetilde{\Psi}(\mathbf{a};x,y)$ of degree μ is a covariant of the form $\widetilde{f}$; these definitions extend to joint invariants etc. in the obvious way.

Thus, a polynomial $\Phi(f)$ in the coefficients of a polynomial $f(x) = \sum_{i=0}^n a_i x^{n-i}$ is an invariant if

$$\Phi\Bigl((\beta x+\delta)^n f\bigl(\frac{\alpha x+\gamma}{\beta x+\delta}\bigr)\Bigr) = (\alpha\delta-\beta\gamma)^w \Phi\bigl(f(x,y)\bigr) \tag{3.2}$$

for all f and $\left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$.

Similarly, by Theorem 2.20(ii), a polynomial $\Phi(f)$ in the coefficients of a polynomial $f(x) = \sum_{i=0}^n a_i x^{n-i}$ is a (projective) seminvariant if it is homogeneous and invariant for $A(1)$, i.e.

$$\Phi\bigl(f(\alpha x+\gamma)\bigr) = \alpha^{\mu+w}\Phi\bigl(f(x,y)\bigr) \tag{3.3}$$

for all f and (α,γ). In other words, a seminvariant is the same as an *affine invariant* for polynomials. (However, we continue to use the traditional term seminvariant).

The same applies with obvious modifications to joint invariants and seminvariants.

Example 3.2. If f is a polynomial of degree n and $\widetilde{f}$ the corresponding binary form, then $x\frac{\partial\widetilde{f}}{\partial x} + y\frac{\partial\widetilde{f}}{\partial y} = n\widetilde{f}$. It follows after some calculations that, for $y=1$,

$$\begin{vmatrix} \frac{\partial^2\widetilde{f}}{\partial x^2} & \frac{\partial^2\widetilde{f}}{\partial x\partial y}\\ \frac{\partial^2\widetilde{f}}{\partial x\partial y} & \frac{\partial^2\widetilde{f}}{\partial y^2}\end{vmatrix} = \begin{vmatrix} \frac{\partial^2\widetilde{f}}{\partial x^2} & (n-1)\frac{\partial\widetilde{f}}{\partial x}\\ (n-1)\frac{\partial\widetilde{f}}{\partial x} & n(n-1)\widetilde{f}\end{vmatrix} = n(n-1)\widetilde{f}\frac{\partial^2\widetilde{f}}{\partial x^2} - (n-1)^2\Bigl(\frac{\partial\widetilde{f}}{\partial x}\Bigr)^2.$$

Hence the Hessian covariant, see Example 2.13, of a polynomial f of degree n is the polynomial

$$H(f;x) := n(n-1)f(x)f''(x) - (n-1)^2(f'(x))^2. \tag{3.4}$$

The source of $H(f)$ is a seminvariant $H_0(f)$ of degree 2, order $2n-4$ and weight 2; by Example 2.27, it is given by

$$H_0(f) = 2n(n-1)a_0a_2 - (n-1)^2a_1^2. \tag{3.5}$$

Example 3.3. Similar calculation show that the Jacobian joint covariant of two polynomials f and g of degrees n_1 and n_2 is given by

$$J(f,g) = n_2f'g - n_1fg'. \tag{3.6}$$

In particular, it follows from (3.4) and (3.6) that

$$H(f) = (n-1)J_{\langle n-1,n\rangle}(f',f). \tag{3.7}$$

Example 3.4. The calculations in Example 3.2 generalize to the Gundelfinger covariants in Example 2.17 and show that the kth Gundelfinger covariant of a polynomial f of degree n is the determinant

$$G_k(f;x) = \left| \frac{(n-2k+i+j)!}{(n-2k)!} f^{(2k-i-j)}(x) \right|_{0\le i,j\le k} = \left| \frac{(n-i-j)!}{(n-2k)!} f^{(i+j)}(x) \right|_{0\le i,j\le k}. \tag{3.8}$$

Theorems 2.20 and 2.31 translate to criteria for (joint) invariants and seminvariants of polynomials. For example, we have the following.

Theorem 3.5. *A polynomial Φ in the coefficients of one or several polynomials is a (joint) seminvariant if and only if Φ is homogeneous and isobaric and invariant for all translations $x \mapsto x + x_0$, i.e., $\Phi(f(x+x_0)) = \Phi(f(x))$ or $\Phi(f_1(x+x_0),\dots,f_\ell(x+x_0)) = \Phi(f_1,\dots,f_\ell)$.*

Theorem 3.6. *The following are equivalent for a polynomial Φ in the coefficients of one or several binary forms.*

(i) *Φ is a (joint) invariant.*
(ii) *Φ is a (joint) seminvariant and $\Phi(\mathbf{a}^\dagger) = \Phi(\mathbf{a})$ or $\Phi(\mathbf{a}_1^\dagger,\dots,\mathbf{a}_\ell^\dagger) = \Phi(\mathbf{a}_1,\dots,\mathbf{a}_\ell)$.*
(iii) *Φ is a (joint) seminvariant of order $\mu = 0$.*
(iv) *Φ is a (joint) seminvariant and $n\nu = 2w$ or $n_1\nu_1 + \dots + n_\ell\nu_\ell = 2w$ for the degree(s) and the weight (i.e., (2.7) or (2.9) holds).*

Here $\mathbf{a}^\dagger$ is given by (2.43); if $\mathbf{a}$ are the coefficients of f, then these are the coefficients of the reflected polynomial $f^\dagger(x) := x^n f(1/x)$.

3.1. **Derivatives.** The derivative $f'(x)$ is *not* a (projective) covariant. (If it were, it would be of order $\mu = n-1$ and its source would be na_0; however, na_0 is a seminvariant of order n, not $n-1$, so (2.15) would not hold.) Nevertheless, it is, as well as higher derivaties $f^{(j)}$, obviously invariant under translations (and affine maps), and Theorem 3.5 implies the following, together with the obvious generalization to joint seminvariants.

Theorem 3.7. *If $\Phi_{\langle m\rangle}$ is a seminvariant of polynomials of degree $m \le n$, then $\Phi_{\langle m\rangle}(f^{(n-m)})$ is a seminvariant of polynomials of degree n. If $\Phi_{\langle m\rangle}$ has degree ν, weight w and order μ, then $\Phi(f^{(n-m)})$ has the same degree ν and weight w, while the order is increased to $\mu + (n-m)\nu$.*

(This theorem is equivalent to [18, Satz 2.18]; note that the form given there requires using the variables $\breve{a}_i$.)

Proof. The formula for the order follows from (2.15). □

Remark 3.8. In particular, even if Φ is an invariant ($\mu = 0$), $\Phi(f^{(n-m)})$ is not; it is only a seminvariant since its order is $(n-m)\nu > 0$. It follows that the covariant corresponding to $\Phi(f^{(n-m)})$ can not be obtained immediately from the covariant corresponding to Φ.

Recall that we use subscripts $\langle n\rangle$ (on coefficients or seminvariants) to denote the degree of the considered polynomials. We have

$$a_{i\,\langle n-1\rangle}(f') = (n-i)a_i(f), \qquad \breve{a}_{i\,\langle n-1\rangle}(f') = n\breve{a}_i(f), \tag{3.9}$$

and, more generally,

$$a_{i\,\langle m\rangle}(f^{(n-m)}) = (n-i)_{n-m}\, a_i(f), \quad \breve{a}_{i\,\langle m\rangle}(f^{(n-m)}) = (n)_{n-m}\, \breve{a}_i(f). \tag{3.10}$$

Example 3.9. Applying the Hessian seminvariant for degree $n-1$ to f', we obtain by (2.39)

$$\begin{aligned} H_{0\,\langle n-1\rangle}(f') &= 2(n-1)(n-2)na_0(n-2)a_2 - (n-2)^2((n-1)a_1)^2 \\ &= (n-2)^2\, H_0(f), \end{aligned} \tag{3.11}$$

so, apart from a constant factor, we obtain the Hessian covariant for degree n.

Example 3.10. Similarly, for the Jacobian joint seminvariant in Example 2.28,

$$J_{\langle n_1-1,n_2\rangle}(f',g) = (n_1-1)J(f,g). \tag{3.12}$$

Example 3.11. For the kth Gundelfinger seminvariant we obtain by (2.41) and (3.9), generalizing (3.11),

$$\begin{aligned} g_{k\,\langle n-1\rangle}(f') &= \bigl((n-1)_{2k}\bigr)^{k+1} \bigl|\breve{a}_{i+j\,\langle n-1\rangle}(f')\bigr|_{i,j=0}^{k} \\ &= \bigl((n-1)_{2k}\bigr)^{k+1} \bigl|n\breve{a}_{i+j}(f)\bigr|_{i,j=0}^{k} = \bigl(n(n-1)_{2k}\bigr)^{k+1} \bigl|\breve{a}_{i+j}(f)\bigr|_{i,j=0}^{k} \\ &= (n-2k)^{k+1} g_k(f). \end{aligned} \tag{3.13}$$

Example 3.12. The kth transvectant seminvariant $\tau_k(f,g)$ is by (2.42) obtained by applying the apolar invariant to suitable derivatives:

$$\tau_{k\,\langle n_1,n_2\rangle}(f,g) = \frac{k!}{(n_1-k)!\,(n_2-k)!} A_{\langle k\rangle}\bigl(f^{(n_1-k)},\, g^{(n_2-k)}\bigr). \tag{3.14}$$

As a consequence,

$$\tau_{k\,\langle n_1-1,n_2\rangle}(f',g) = (n_1-k)\tau_{k\,\langle n_1,n_2\rangle}(f,g) \tag{3.15}$$

and, more generally,

$$\tau_{k\,\langle n_1-\ell_1,n_2-\ell_2\rangle}(f^{(\ell_1)},g^{(\ell_2)}) = (n_1-k)_{\ell_1}(n_2-k)_{\ell_2}\tau_{k\,\langle n_1,n_2\rangle}(f,g), \tag{3.16}$$

so, apart from a constant factor, we obtain the kth transvectant seminvariant of the original fumctions. Note the special cases in Example 3.9 ($k=2$, $\ell_1=\ell_2=1$) and Example 3.10 ($k=1$, $\ell_1=1$, $\ell_2=0$).

3.2. **Restriction to lower degree.** Let Φ be a seminvariant of polynomials of degree (at most) n. If $m < n$, then $\mathcal{P}_m \subset \mathcal{P}_n$, so every polynomial of degree m can be regarded as a polynomial $\sum_{i=0}^n a_i x^{n-i}$ with $a_0 = \cdots = a_{n-m-1} = 0$; thus, $\Phi(f)$ is defined for every such polynomial. (See Remark 3.1.)

Note that we write a polynomial of degree $m < n$ as $\sum_{j=0}^m a_{j\,\langle m\rangle} x^{m-j} = \sum_{i=0}^n a_{i\,\langle n\rangle} x^{n-i}$, and thus

$$a_{i\,\langle n\rangle} = \begin{cases} a_{i-(n-m)\,\langle m\rangle}, & \text{if } i \geq n-m \\ 0. & \text{if } i < n-m. \end{cases} \tag{3.17}$$

We denote the restriction of a seminvariant Φ to polynomials of degree m by $\Phi|_{\langle m\rangle}$.

Theorem 3.13. *A seminvariant Φ of polynomials of degree n is also a seminvariant of polynomials of any given lower degree $n-j$. If Φ has degree ν, weight w and order μ, then its restriction $\Phi|_{\langle n-j\rangle}$ has degree ν, weight $w - jn$ and order $\mu + jn$.*

Proof. It is an immediate consequence of Theorem 3.5 that $\Phi|_{\langle n-j\rangle}$ is a seminvariant. The degree is the same, but the weight of each a_i is decreased by j by (3.17), and thus the new weight is $w - j\nu$. The new order is by (2.15) given by

$$(n-j)\nu - 2(w - j\nu) = n - 2w + j\nu = m + j\nu. \qquad \square$$

In particular, $\Phi|_{\langle n-j\rangle}$ has order $\mu + j\nu \geq \nu > 0$ for any seminvariant Φ and any $j > 0$; hence, a non-trivial restriction is never an invariant, even if Φ is an invariant.

Example 3.14. The restriction $H_0|_{\langle n-1\rangle}$ of the Hessian seminvariant in Example 3.2 is, recalling $a_0 = 0$,

$$-(n-1)^2 a_{1\,\langle n\rangle}^2 = -(n-1)^2 a_{0\,\langle n-1\rangle}^2. \tag{3.18}$$

This has degree 2, weight 0 and order $2n-2$, in agreement with Theorem 3.13.

Example 3.15. Combining Theorem 3.13 and Theorem 3.7, we see that if Φ is a seminvariant of polynomials of degree n, then so is $\Phi(f') = \Phi|_{\langle n-1\rangle}(f')$, and more generally $\Phi(f^{(j)})$ for every $j \geq 1$. If Φ has degree ν, weight w and order $\mu = n\nu - 2w$, then $\Phi(f')$ is a seminvariant with degree ν, weight $w - \nu$ and order $\mu + 2\nu$.

3.3. **Reduced form.** The *reduced form* of a polynomial $f(x) = \sum_{i=0}^n a_i x^{n-i}$ of degree n is the polynomial

$$\widehat{f}(x) = \sum_{i=0}^n \widehat{a}_i x^{n-i} := f\Big(x - \frac{a_1}{na_0}\Big); \tag{3.19}$$

note that $\widehat{a}_0 := a_0$ and $\widehat{a}_1 = 0$. The reduced form is thus the unique translation $f(x - x_0)$ of f with vanishing coefficient for the second highest degree x^{n-1}. Explicitly, by (3.19) and binomial expansions,

$$\widehat{a}_i = \sum_{j=0}^{i} a_j \binom{n-j}{n-i} \Bigl(-\frac{a_1}{n\,a_0}\Bigr)^{i-j}. \tag{3.20}$$

The coefficient $\widehat{a}_i$ is a_0 times a polynomial of degree i in $(a_j/a_0)_{j=1}^n$, and thus $a_0^{i-1}\widehat{a}_i$ is a polynomial in $a_0, \dots, a_n$. This polynomial is homogeneous of degree i and, as is easily checked, isobaric with weight i. Furthermore, the reduced form is the same for all translations $f(x-x_0)$, so its coefficients are translation invariant.

Theorem 3.16. *The coefficients $\widehat{a}_i$ of the reduced form of f are rational seminvariants; more precisely $\widehat{a}_i$ is a seminvariant divided by a_0^{i-1}. The seminvariant $a_0^{i-1}\widehat{a}_i$ has degree and weight $\nu = w = i$ and thus order $\mu = (n-2)i$.*

Example 3.17. The constant term

$$\widehat{a}_n = \widehat{f}(0) = f\Bigl(-\frac{a_1}{n\,a_0}\Bigr) = \sum_{j=0}^{n} a_j \Bigl(-\frac{a_1}{n\,a_0}\Bigr)^{n-j} \tag{3.21}$$

is a real seminvariant and $a_0^{n-1}\widehat{a}_n = a_0^{n-1} f(-a_1/na_0)$ is a seminvariant of degree and weight n and order $n(n-2)$.

Note that every coefficient $\widehat{a}_i$ can be obtained as the constant term of a derivative $\widehat{f}^{(n-i)} = \widehat{f^{(n-i)}}$, cf. Theorem 3.7.

Example 3.18. The first non-trivial reduced coefficient is

$$\widehat{a}_2 = a_0 \binom{n}{2} \Bigl(\frac{a_1}{n\,a_0}\Bigr)^2 - a_1(n-1)\frac{a_1}{n\,a_0} + a_2 = a_2 - \frac{(n-1)\,a_1^2}{2n\,a_0} = \frac{H_0}{2n(n-1)\,a_0}, \tag{3.22}$$

see (2.39). The seminvariant $a_0\widehat{a}_2$ is thus a constant times the Hessian seminvariant H_0.

Every homogeneous and isobaric polynomial in $\widehat{a}_2, \dots, \widehat{a}_n$ times a power a_0^s is a rational seminvariant, and a seminvariant if the exponent s is large enough. Conversely, every seminvariant Φ is translation invariant, and thus $\Phi(f) = \Phi(\widehat{f})$; hence every seminvariant is a polynomial in a_0 and $\widehat{a}_2, \dots, \widehat{a}_n$. Up to powers of a_0, every seminvariant is thus a polynomial in the seminvariants $a_0^{i-1}\widehat{a}_i$. However, these seminvariants do not form a basis (when $n \geq 3$), since we may need need negative powers of a_0 in the representation. For example, for $n = 3$, by (8.3)–(8.5),

$$\Delta = -4a_0\,\widehat{a}_2^3 - 27a_0^2\,\widehat{a}_3^2 = \frac{-4(a_0\widehat{a}_2^2)^3 - 27(a_0^2\widehat{a}_3)^2}{a_0^2}. \tag{3.23}$$

In general, we have the following theorem.

Theorem 3.19. *If Φ is a seminvariant with degree ν and weight w, then*

$$\Phi = a_0^{\nu-w} G\big((a_0^{i-1}\widehat{a}_i)_{i=2}^n\big) \tag{3.24}$$

for some isobaric polynomial G of weight w. Consequently, Φ is a polynomial in a_0 and $a_0\widehat{a}_2, \dots, a_0^{n-1}\widehat{a}_n$ if and only if $\nu \geq w$.

If $\nu \geq w$, then (3.24) gives a one-to-one correspondence between seminvariants with degree ν and weight w, and isobaric polynomials $G\big((a_0^{i-1}\widehat{a}_i)_{i=2}^n\big)$ of weight w.

Proof. Each term $a_0^{i-1}\widehat{a}_i$ has the same degree and weight, and thus so has every (isobaric) polynomial $G\big((a_0^{i-1}\widehat{a}_i)_{i=2}^n\big)$ in them, while a_0 has degree 1 and weight 0. Hence an isobaric term $a_0^s G\big((a_0^{i-1}\widehat{a}_i)_{i=2}^n\big)$ has weight v and degree $s + v$ for some w, and thus we must have $v = w$ and $s = \nu - v = \nu - w$. □

By Example 4.2 below, the discriminant Δ has $\nu = 2(n-1)$ and $w = n(n-1)$, so $\nu - w = -(n-2)(n-1) < 0$ for any $n \geq 3$, and then Δ is not a polynomial in a_0 and $a_0^{i-1}\widehat{a}_i$.

Remark 3.20. In the case $\nu \geq w$, we see again that the dimension $N(n, \nu, w)$ is independent of ν as long as $\nu \geq w$. Moreover, $N(n, \nu, w)$ then equals the number of isobaric monomials of weight w in $(\widehat{a}_i)_{i=2}^n$; this number has the generating function $\prod_{i=2}^n (1 - q^i)^{-1}$, which yields another proof of Corollary 2.38.

4. Invariants and roots

Let the polynomial f of degree n have roots $\xi_1, \dots, \xi_n$ (possibly in some extension of the ground field). Then, as is well-known,

$$a_i = (-1)^i a_0 e_i(\xi_1, \dots, \xi_n), \tag{4.1}$$

where e_i is the i:th symmetric polynomial; note that e_i has degree i. If $\Phi(f)$ is a seminvariant, we can thus write $\Phi(f)$ as a polynomial $\Phi^*(\xi_1, \dots, \xi_n; a_0)$.

Theorem 4.1. *A polynomial $\Phi^*(\xi_1, \dots, \xi_n; a_0)$ is a seminvariant of degree ν and weight w if and only if $\Phi^*(\xi_1, \dots, \xi_n; a_0) = a_0^\nu \varphi(\xi_1, \dots, \xi_n)$ where*

- (i) *φ is symmetric in $\xi_1, \dots, \xi_n$;*
- (ii) *φ is homogeneous of degree w in $\xi_1, \dots, \xi_n$;*
- (iii) *φ is translation invariant, i.e., $\varphi(\xi_1 - x_0, \dots, \xi_n - x_0) = \varphi(\xi_1, \dots, \xi_n)$. Equivalently, $\varphi(\xi_1, \dots, \xi_n)$ is a polynomial in the differences $\xi_j - \xi_n$.*
- (iv) *$\nu \geq \deg_{\xi_1}\big(\varphi(\xi_1, \dots, \xi_n)\big)$, the degree of ξ_1 in $\varphi(\xi_1, \dots, \xi_n)$.*

Furthermore, Φ^* is an invariant if and only the above holds and $n\nu = 2w$; in this case

$$(\xi_1 \cdots \xi_n)^\nu \varphi(\xi_1^{-1}, \dots, \xi_n^{-1}) = (-1)^w \varphi(\xi_1, \dots, \xi_n). \tag{4.2}$$

Proof. Recall that every symmetric polynomial is a polynomial in $e_1, \dots, e_n$. Then use Theorem 3.5 and (4.1) and note that each ξ_j has weight 1 by the fact that a_i has weight i. This might yield terms containing negative powers of a_0, and (iv) is necessary and sufficient for Φ to be a polynomial in $a_0, \dots, a_n$. The symmetry $\Phi(\mathbf{a}^\dagger) = (-1)^w \Phi(\mathbf{a})$ translates to (4.2). We omit the details. □

Example 4.2. The *discriminant* of f is

$$\Delta(f) := a_0^{2n-2}\Delta_0(f) = a_0^{2n-2} \prod_{1\le i<j\le n} (\xi_i - \xi_j)^2, \tag{4.3}$$

This is symmetric and has degree $w = n(n-1)$ in $\xi_1, \dots, \xi_n$. It follows from Theorem 4.1, since $\nu = 2(n-1)$, that the discriminant Δ is an invariant of degree $\nu = 2(n-1)$ and weight $w = n(n-1)$. (Other notation: $\Delta = D_{[18]}$.)

Example 4.3. The sum $\sum_{1\le i<j\le n}(\xi_i-\xi_j)^2$ satisfies (i)–(iii) in Theorem 4.1, and has degree 2 in ξ_1, so $a_0^2 \sum_{1\le i<j\le n}(\xi_i - \xi_j)^2$ is a seminvariant. We have, using (4.1),

$$\begin{aligned}\sum_{1\le i<j\le n} (\xi_i - \xi_j)^2 &= (n-1)\sum_{i=1}^n \xi_i^2 - 2\sum_{1\le i<j\le n} \xi_i\xi_j \\ &= (n-1)\Bigl(\sum_{i=1}^n \xi_i\Bigr)^2 - 2n\sum_{1\le i<j\le n} \xi_i\xi_j = (n-1)\Bigl(\frac{-a_1}{a_0}\Bigr)^2 - 2n\frac{a_2}{a_0},\end{aligned}$$

so

$$a_0^2 \sum_{1\le i<j\le n} (\xi_i - \xi_j)^2 = (n-1)a_1^2 - 2na_0a_2 = -\frac{1}{n-1}H_0(f), \tag{4.4}$$

where H_0 is the Hessian seminvariant in Examples 2.27 and 3.2.

Example 4.4. Let $\bar\xi := \frac{1}{n}\sum_{i=1}^n \xi_i = -a_1/na_0$. Then the roots of the reduced polynomial $\widehat{f}$ are $\xi_1 - \bar\xi, \dots, \xi_n - \bar\xi$. Any symmetric homogeneous polynomial in $\xi_1 - \bar\xi, \dots, \xi_n - \bar\xi$ satifies Theorem 4.1(i)–(iii), and multiplied by a suitable power of a_0, it is thus a seminvariant. Since any such polynomial can be written as an isobaric polynomial in $\widehat{a}_i/a_0$, this also follows by Theorem 3.16 or Theorem 3.19.

In particular, the elementary symmetric polynomials e_k yield the rational seminvariants

$$e_k(\xi_1 - \bar\xi, \dots, \xi_n - \bar\xi) = (-1)^k \widehat{a}_k/a_0. \tag{4.5}$$

Example 4.5. As another example of the construction in Example 4.4, consider the power sum

$$S_k := \sum_{i=1}^n (\xi_i - \bar\xi)^k \tag{4.6}$$

and the seminvariant $a_0^k S_k$. Note that $S_0 = n$ is a constant and $S_1 = 0$. Further,

$$S_2 = \sum_{i=1}^{n} \xi_i^2 - n\bar{\xi}^2 = \frac{1}{2n} \sum_{i,j=1}^{n} (\xi_i - \xi_j)^2, \tag{4.7}$$

so by (4.4),

$$a_0^2 S_2 = -\frac{1}{n(n-1)} H_0(f). \tag{4.8}$$

Furthermore, S_k can be expressed in $e_1, \dots, e_k$ by the standard generating function identity

$$\log\left(\sum_{k=0}^{\infty} e_k(-t)^k\right) = \sum_{i=1}^{n} \log(1 - t\xi_i) = -\sum_{k=1}^{\infty} S_k \frac{t^k}{k}. \tag{4.9}$$

which leads to the classical *Newton's identities* (with $e_0 = 1$ and $e_k = 0$ for $k > n$),

$$ke_k = \sum_{i=1}^{k} (-1)^{i-1} e_{k-i} S_i, \qquad k \geq 1. \tag{4.10}$$

In our situation, the arguments are $\xi_1 - \bar{\xi}, \dots, \xi_n - \bar{\xi}$; thus $S_1 = e_1 = 0$ and we have, for example,

$$S_2 = e_1^2 - 2e_2 = -2e_2, \tag{4.11}$$

$$S_3 = 3e_3, \tag{4.12}$$

$$S_4 = -4e_4 + 2e_2^2. \tag{4.13}$$

Thus, by (4.5) and Example 3.18, we obtain the seminvariants

$$a_0^2 S_2 = -2a_0^2 e_2 = -2a_0 \widehat{a}_2 = -\frac{1}{n(n-1)} H_0, \tag{4.14}$$

$$a_0^3 S_3 = 3a_0^3 e_3 = -3a_0^2 \widehat{a}_3, \tag{4.15}$$

$$a_0^4 S_4 = -4a_0^4 e_4 + 2(a_0^2 e_2)^2 = -4a_0^3 \widehat{a}_4 + 2(a_0 \widehat{a}_2)^2. \tag{4.16}$$

Note that $a_0^k S_k$ has degree and weight $\nu = w = k$ (see Theorem 4.1).

Example 4.6. Consider the random variable $X = \xi_Y$, where $Y \in \{1, \dots, n\}$ is a random index (with uniform distribution). Then X has mean $\mathbb{E}\, X = \bar{\xi}$ and centred moments

$$\mathbb{E}(X - \mathbb{E}\, X)^k = \mathbb{E}(X - \bar{\xi})^k = \tfrac{1}{n} S_k. \tag{4.17}$$

Thus, $a_0^k\, \mathbb{E}(X - \mathbb{E}\, X)^k$ equals n^{-1} times the seminvariant in Example 4.5.

Kung and Rota [14, §7.6] suggested studying the *cumulants* χ_k of X, $k \geq 2$. These are defined by the generating function

$$\exp\left(\sum_{k=1}^{\infty} \chi_k \frac{t^k}{k!}\right) = \mathbb{E}\, e^{tX} = \sum_{k=0}^{\infty} \mathbb{E}\, X^k \frac{t^k}{k!}, \tag{4.18}$$

and thus, since $\chi_1 = EX = \overline{\xi}$,

$$\exp\left(\sum_{k=2}^{\infty} \chi_k \frac{t^k}{k!}\right) = \mathbb{E}\, e^{t(X-\overline{\xi})} = \sum_{k=0}^{\infty} \frac{1}{n} S_k \frac{t^k}{k!}. \tag{4.19}$$

By expanding, we obtain the standard formulas for χ_k as a polynomial in $S_1, \dots, S_k$, or, using (4.9), in $e_1, \dots, e_k$. For example,

$$\chi_2 = \mathbb{E}(X - \mathbb{E}\,X)^2 = \frac{1}{n} S_2 = -\frac{2}{n} e_2, \tag{4.20}$$

$$\chi_3 = \mathbb{E}(X - \mathbb{E}\,X)^3 = \frac{1}{n} S_3 = \frac{3}{n} e_3, \tag{4.21}$$

$$\chi_4 = \mathbb{E}(X - \mathbb{E}\,X)^4 - 3\left(\mathbb{E}(X - \mathbb{E}\,X)^2\right)^2 = \frac{1}{n} S_4 - \frac{3}{n^2} S_2^2 = -\frac{4}{n} e_4 + \frac{2n-12}{n^2} e_2^2. \tag{4.22}$$

It follows from (4.19) that $a_0^k \chi_k$ is an isobaric polynomial in $a_0^j S_j$, $j = 2, \dots, k$, and thus a seminvariant, with degree and weight $\nu = w = k$. For example, by (4.20)–(4.22) and (4.14)–(4.16),

$$a_0^2 \chi_2 = -\frac{2}{n} a_0^2 e_2 = -\frac{2}{n} a_0 \widehat{a}_2 = -\frac{1}{n^2(n-1)} H_0, \tag{4.23}$$

$$a_0^3 \chi_3 = \frac{3}{n} a_0^3 e_3 = -\frac{3}{n} a_0^2 \widehat{a}_3, \tag{4.24}$$

$$a_0^4 \chi_4 = -\frac{4}{n} a_0^4 e_4 + \frac{2n-12}{n^2}\left(a_0^2 e_2\right)^2 = -\frac{4}{n} a_0^3 \widehat{a}_4 + \frac{2n-12}{n^2}\left(a_0 \widehat{a}_2\right)^2. \tag{4.25}$$

4.1. **The case $a_0 = 0$.** We have implicitly assumed $a_0 \neq 0$ above. If $a_0 = 0$, then f has degree at most $n-1$, and thus at most $n-1$ roots. We then adopt the projective view and regard ∞ as a root of multiplicity $n - \deg(f)$, so that f still has n roots (counted with multiplicity); these correspond (just as in the case $a_0 \neq 0$) to the zeros of the corresponding binary form $\sum_{i=0}^{n} a_i x^{n-i} y^i$ of degree n.

We can apply a limit argument to find the expression for a seminvariants in the roots of f in this case too.

Theorem 4.7. *Let Φ be a seminvariant of polynomials of degree n, and that $\Phi(f) = a_0^\nu \varphi(\xi_1, \dots, \xi_n)$ for some polynomial φ in the roots $\xi_1, \dots, \xi_n$ of f. Then the restriction to polynomials of degree $n-1$ is given by $\Phi|_{\langle n-1\rangle}(f) = a^\nu_{0\,\langle n-1\rangle} \varphi|_{\langle n-1\rangle}(\xi_1, \dots, \xi_{n-1})$, where $\varphi|_{\langle n-1\rangle}$ is obtained from φ by first replacing each monomial $\xi_1^{j_1} \cdots \xi_n^{j_n}$ by $\xi_1^{j_1} \cdots \xi_{n-1}^{j_{n-1}}$ if $j_n = \nu$ and by 0 otherwise, and then multiplying by $(-1)^\nu$.*

Note that $j_n \leq \nu$ for every term by Theorem 4.1(iv) (and symmetry).

In other words, we delete all terms in φ not containing a factor ξ_n^ν, and replace each factor ξ_n^ν by $(-1)^\nu$.

Proof. Fix $\xi_1, \dots, \xi_{n-1}$ and let $\xi_n = b/a_0$ for some b; now let $a_0 \to 0$. (This limit can be done in a purely formal way, for any field, and does not really assume any kind of continuity. All quantities below are polynomials in a_0, and we just substitute $a_0 = 0$ in them.) Then, for $1 \le i \le n$,

$$\begin{aligned} a_0 e_i(\xi_1, \dots, \xi_n) &= a_0 e_i(\xi_1, \dots, \xi_{n-1}) + a_0 \xi_n e_{i-1}(\xi_1, \dots, \xi_{n-1}) \\ &\to b e_{i-1}(\xi_1, \dots, \xi_{n-1}). \end{aligned}$$

Hence, comparing with (4.1), we see that the coefficients of the polynomial f with roots $\xi_1, \dots, \xi_n$ and leading coefficient a_0 tend to the coefficients of the polynomial f_1 of degree $n-1$ with roots $\xi_1, \dots, \xi_{n-1}$ and leading term $-bx^{n-1}$, i.e., leading coefficient $a_{0\,\langle n-1\rangle}(f_1) = -b$. Consequently, $\Phi(f) \to \Phi(f_1)$. The result follows by noting that, as $a_0 \to 0$, $a_0^\nu \xi_n^{j_n} \to 0$ if $j_n < \nu$, while $a_0^\nu \xi_n^\nu = b^\nu = (-a_{0\,\langle n-1\rangle}(f_1))^\nu$. □

Example 4.8. Applying Theorem 4.7 to the discriminant in (4.3) we find for a polynomial f of degree $n-1$

$$\Delta_{\langle n\rangle}(f) = a_0^{2n-2} \prod_{1\le i<j\le n-1} (\xi_i - \xi_j)^2 = a_0^2\, \Delta_{\langle n-1\rangle}(f). \tag{4.26}$$

If we repeat, we find that for any f of degree $n-2$ (or smaller), $\Delta_{\langle n\rangle}(f) = 0$, in accordance with our view that then f has a double root ∞.

4.2. **Joint invariants.** Theorem 4.1 extends to the case of several polynomials $f_1, \dots, f_\ell$. Let the polynomial f_j have degree n_j and roots $\xi_1^{(j)}, \dots, \xi_n^{(j)}$ (possibly in some extension of the ground field). Then, by (4.1),

$$a_i(f_j) = (-1)^i a_0(f_j) e_i(\xi_1^{(j)}, \dots, \xi_n^{(j)}), \tag{4.27}$$

and if $\Phi(f_1, \dots, f_\ell)$ is a joint seminvariant, we can write it as a polynomial $\Phi^*(\xi_1^{(j)}, \dots, \xi_n^{(j)}; a_0(f_1), \dots, a_0(f_\ell))$ in all roots and leading coefficients.

Theorem 4.9. *A polynomial $\Phi^*(\xi_1^{(1)}, \dots, \xi_n^{(\ell)}; a_0(f_1), \dots, a_0(f_\ell))$ is a joint seminvariant of $f_1, \dots, f_\ell$ with degrees $\nu_1, \dots, \nu_\ell$ and weight w if and only if $\Phi^*(\xi_1^{(1)}, \dots, \xi_n^{(\ell)}; a_0) = \prod_{j=1}^{\ell} a_0(f_j)^{\nu_j} \cdot \varphi(\xi_1^{(1)}, \dots, \xi_n^{(\ell)})$, where*

(i) *φ is symmetric in each set of roots $\xi_1^{(j)}, \dots, \xi_n^{(j)}$, $j = 1, \dots, \ell$;*
(ii) *φ is homogeneous of degree w in $\xi_1^{(1)}, \dots, \xi_n^{(\ell)}$;*
(iii) *φ is translation invariant, i.e., $\varphi(\xi_1^{(1)}-x_0, \dots, \xi_n^{(\ell)}-x_0) = \varphi(\xi_1^{(1)}, \dots, \xi_n^{(\ell)})$.*
(iv) *$\nu_j \ge \deg_{\xi_1^{(j)}}\big(\varphi(\xi_1^{(1)}, \dots, \xi_n^{(\ell)})\big)$, the degree of $\xi_1^{(j)}$ in $\varphi(\xi_1^{(1)}, \dots, \xi_n^{(\ell)})$.*

Furthermore, Φ^* is an invariant if and only the above holds and $n_1\nu_1 + \dots + n_\ell \nu_\ell = 2w$; in this case

$$\prod_{j=1}^{\ell} (\xi_1^{(j)} \cdots \xi_n^{(j)})^{\nu_j} \cdot \varphi\big((\xi_1^{(1)})^{-1}, \dots, (\xi_n^{(\ell)})^{-1}\big) = (-1)^w \varphi(\xi_1^{(1)}, \dots, \xi_n^{(\ell)}). \tag{4.28}$$

Proof. As for Theorem 4.1, with obvious modifications. □

Example 4.10. The *resultant* of two polynomials $f = \sum_{i=0}^{n} a_{n-i}x^i$ and $g = \sum_{j=0}^{m} b_{m-j}x^j$ of degrees n and m and with roots $\xi_1, \dots, \xi_n$ and $\eta_1, \dots, \eta_m$ is

$$R(f,g) := a_0^m b_0^n \prod_{i=1}^{n}\prod_{j=1}^{m}(\xi_i - \eta_j). \tag{4.29}$$

This is symmetric in $\xi_1^{(j)}, \dots, \xi_n^{(j)}$ and in $\eta_1, \dots, \eta_m$ and has degree $w = nm$ in $\xi_1, \dots, \xi_n, \eta_1, \dots, \eta_m$. Theorem 4.9 applies, with $n_1 = n$, $n_2 = m$, $\nu_1 = m$, $\nu_2 = n$, and thus $n_1\nu_1 + n_2\nu_2 = 2mn = 2w$; hence the resultant R is a joint invariant of degrees (m,n) and weight nm.

Note that $R(g,f) = (-1)^{mn}R(f,g)$ and the formulas, see e.g. [11],

$$R(f,g) = a_0^m \prod_{i=1}^{n} g(\xi_i) = (-1)^{mn} b_0^n \prod_{j=1}^{m} f(\eta_j). \tag{4.30}$$

Example 4.11. Consider the resultant of f and $H(f)$, where f is a polynomial of degree n. By (3.4), $H(f)(\xi_i) = -(n-1)^2(f'(\xi_i))^2$, and thus

$$\begin{aligned} R\bigl(f, H(f)\bigr) &= a_0^{2n-4}(-1)^n(n-1)^{2n}\prod_{i=1}^{n} f'(\xi_i)^2 = a_0^{-2}(-1)^n(n-1)^{2n}R(f,f')^2 \\ &= (-1)^n(n-1)^{2n}\Delta^2, \end{aligned} \tag{4.31}$$

since $\Delta = (-1)^{n(n-1)/2}a_0^{-1}R(f,f')$, see [11].

Example 4.12. Consider again two polynomials $f = \sum_{i=0}^{n} a_{n-i}x^i$ and $g = \sum_{j=0}^{m} b_{m-j}x^j$ of degrees n and m and with roots $\xi_1, \dots, \xi_n$ and $\eta_1, \dots, \eta_m$. The difference $m\sum_{i=1}^{n}\xi_i - n\sum_{j=1}^{m}\eta_j$ satisfies (i)–(iii) in Theorem 4.9, and has degree 1 in ξ_1 and η_1, so $a_0b_0\bigl(m\sum_{i=1}^{n}\xi_i - n\sum_{j=1}^{m}\eta_j\bigr)$ is a joint seminvariant. We have, using (4.1),

$$a_0b_0\Bigl(m\sum_{i=1}^{n}\xi_i - n\sum_{j=1}^{m}\eta_j\Bigr) = a_0b_0\Bigl(m\frac{-a_1}{a_0} - n\frac{-b_1}{b_0}\Bigr) = -ma_1b_0 + na_0b_1,$$

so this equals the Jacobian seminvariant in Examples 2.28 and 3.3.

4.3. **Covariants.** Similarly, a covariant can be written as a polynomial in x and the roots $\xi_1, \dots, \xi_n$ of f. The following theorem yields an explicit formula.

Theorem 4.13. *Let Ψ be a covariant of polynomials of degree n, and let Φ be its source. Suppose that $\Phi(f) = a_0^\nu\varphi(\xi_1, \dots, \xi_n)$ as in Theorem 4.1. Then*

$$\Psi(f;x) = a_0^\nu \prod_{i=1}^{n}(x-\xi_i)^\nu \cdot \varphi\Bigl(\frac{1}{x-\xi_1}, \dots, \frac{1}{x-\xi_n}\Bigr).$$

Proof. Let $\widetilde{f}$ be the binary form corresponding to f, and let $\widetilde{\Psi}(\widetilde{f})$ be the covariant corresponding to $\Psi(f)$; thus $\Psi(f;x) = \widetilde{\Psi}(\widetilde{f};x,1)$; further, let $\widetilde{\Phi}$ be the source of $\widetilde{\Psi}$.

The reflection $\rho(x,y) := (y,x)$ has matrix $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$ and determinant $|\rho| = -1$; thus (2.14) yields

$$\Psi(f;0) = \widetilde{\Psi}(\widetilde{f};0,1) = (-1)^w\widetilde{\Psi}(\rho\widetilde{f};1,0) = (-1)^w\widetilde{\Phi}(\rho\widetilde{f}) = (-1)^w\Phi(f^\dagger), \tag{4.32}$$

where $f^\dagger$ is the polynomial corresponding to $\rho\widetilde{f}(x,y) = \widetilde{f}(y,x) = \sum_{i=0}^n a_i y^{n-i}x^i$. We have

$$\begin{aligned} f^\dagger(x) &= \sum_{i=0}^n a_i x^i = x^n f(1/x) = a_0 x^n \prod_{i=1}^n \bigl(x^{-1} - \xi_i\bigr) = a_0 \prod_{i=1}^n (1 - x\xi_i) \\ &= a_0 \prod_{i=1}^n (-\xi_i) \prod_{i=1}^n (x - \xi_i^{-1}), \end{aligned} \tag{4.33}$$

with roots $\xi_1^{-1}, \dots, \xi_n^{-1}$. Consequently,

$$\Phi(f^\dagger) = \Bigl(a_0 \prod_{i=1}^n (-\xi_i)\Bigr)^\nu \varphi(\xi_1^{-1}, \dots, \xi_n^{-1}) \tag{4.34}$$

and thus by (4.32), since φ is homogeneous of degree w,

$$\Psi(f;0) = (-1)^w\Phi(f^\dagger) = a_0^\nu \prod_{i=1}^n (-\xi_i)^\nu \varphi\Bigl(\frac{1}{-\xi_1}, \dots, \frac{1}{-\xi_n}\Bigr). \tag{4.35}$$

We have $\Psi(f;x) = \Psi^*(\xi_1, \dots, \xi_n; x)$ for some polynomial Ψ^*, and

$$\Psi^*(\xi_1, \dots, \xi_n; x) = \Psi^*(\xi_1 - x, \dots, \xi_n - x; 0) \tag{4.36}$$

by translation invariance. The result follows by (4.36) and (4.35). □

Example 4.14. As a trivial example, the covariant f has source a_0, and Theorem 4.13 yields, with $\varphi = 1$, $f = a_0 \prod_{i=1}^n (x - \xi_i)$.

Example 4.15. The source of the Hessian covariant is, by (4.4),

$$H_0(f) = -(n-1)a_0^2 \sum_{1\le i<j\le n} (\xi_i - \xi_j)^2. \tag{4.37}$$

Hence Theorem 4.13 shows that the Hessian covariant is given by

$$\begin{aligned} H(f;x) &= -(n-1)a_0^2 \prod_{k=1}^n (x-\xi_k)^2 \sum_{1\le i<j\le n} \Bigl(\frac{1}{x-\xi_i} - \frac{1}{x-\xi_j}\Bigr)^2 \\ &= -(n-1)a_0^2 \sum_{1\le i<j\le n} (\xi_i - \xi_j)^2 \prod_{k\ne i,j} (x-\xi_k)^2. \end{aligned} \tag{4.38}$$

Note that by extracting the leading coefficients (the coefficients of x^{2n-4}), we recover (4.4).

The extension to joint covariants is straightforward; we leave the formulation to the reader and give only a simple example.

Example 4.16. The source of the Jacobian covariant $J(f,g)$ is by Example 4.12

$$a_0b_0\Bigl(m\sum_{i=1}^n \xi_i - n\sum_{j=1}^m \eta_j\Bigr) \tag{4.39}$$

and thus

$$J(f,g) = a_0b_0\prod_{i=1}^n(x-\xi_i)\prod_{j=1}^m(x-\eta_j)\left(m\sum_{i=1}^n(x-\xi_i)^{-1} - n\sum_{j=1}^m(x-\eta_j)^{-1}\right). \tag{4.40}$$

5. Some characterizations of vanishing invariants

In some cases, there are simple characterizations of vanishing invariants or covariants. For example, the following basic result is an immediate consequence of (4.3). For simplicity, we assume in this section that $a_0 \neq 0$, i.e., that the actual degree is n; the results immediately extend to the case $a_0 = 0$ by projective invariance (considering roots at infinity), see for example Example 4.8. (The results all have invariant formulations for binary forms.)

Theorem 5.1. *The discriminant $\Delta(f) = 0$ if and only if f has a double root (in some extension field).*

Equivalently, a binary form $f(x,y)$ has discriminant 0 if and only if it has a square factor $(ax+by)^2$.

Theorem 5.2. *The Hessian covariant $H(f) = 0$ if and only if f has a single root, i.e., $\xi_1 = \dots = \xi_n$; equivalently, $f(x) = a_0(x-\xi)^n$ for some a_0 and ξ.*

Equivalently, a binary form $f(x,y)$ has Hessian covariant $H(f) = 0$ if and only if it equals $c(ax+by)^n$ for some a,b,c.

Theorem 5.3. *The Jacobian joint covariant $J(f,g) = 0$ if and only if f and g have the same roots, and their multiplicities always are in the same proportion n_1/n_2; equivalently, $f(x) = a_0h(x)^{d_1}$ and $g(x) = b_0h(x)^{d_2}$ for some polynomial h and some integers $d_1,d_2 \geq 1$. In particular, if f and g have the same degree, then $J(f,g) = 0$ if and only if f and g are proportional.*

Proof. Suppose that ξ is a root of f or g, and let the multiplicities of the root be $k_1 \geq 0$ and $k_2 \geq 0$ (with $k_1 + k_2 > 0$). By a projective transformation we may assume that $\xi = 0$. Then $f(x) = ax^{k_1} + \dots$ and $g(x) = bx^{k_2} + \dots$ (showing the lowest degree terms only), with $a,b \neq 0$, and Example 3.3 shows that

$$J(f,g) = (n_2k_1 - n_1k_2)abx^{k_1+k_2-1} + \dots \tag{5.1}$$

Hence $J(f,g) = 0$ implies $n_2k_1 - n_1k_2 = 0$, and thus both k_1 and k_2 are nonzero and $k_1/k_2 = n_1/n_2$. This shows that f^{n_2} and g^{n_1} have the same roots, with the same multiplicities, and thus $f^{n_2} = cg^{n_1}$ for some c. The result follows by the unique factorization of polynomials into irreducible ones.

Conversely, if $f = a_0 h^{d_1}$ and $g = b_0 h^{d_2}$, then

$$J(f, g) = a_0 b_0 d_1 h^{d_1 - 1} d_2 h^{d_2 - 1} J(h, h) = 0. \qquad \square$$

Proof of Theorem 5.2. If $H(f) = 0$, then (3.7) yields $J(f', f) = 0$, and thus Theorem 5.3 yields $f' = ah^{d_1}$, $f = bh^{d_2}$ for some polynomial h and constants a, b, d_1 and d_2. Then $n-1 = d_1 \deg(h)$ and $n = d_2 \deg(h)$, and consequently $1 = (d_2 - d_1) \deg(h)$, which implies $\deg(h) = 1$ and $d_2 = n$.

The converse follows directly from (3.7) and Theorem 5.3. $\square$

Theorem 5.4 ([18, Satz 2.11])**.** *All non-constant invariants of a polynomial of degree n vanish if and only if the polynomial has a root of multiplicity $> n/2$. (This includes the case when the actual degree is $< n$; there is a root at ∞ of multiplicity more than $n/2$ when the degree is $< n/2$.)*

Note that seminvariants still may be non-zero.

Example 5.5. The seminvariant $a_0(f) = 0$ if and only if $\deg(f) \le n - 1$. If, for example, $f(x) = x^n$, then $a_0 \neq 0$ while all invariants vanish by Theorem 5.4.

Example 5.6. If $f(x) = x^{n-1}(x - b)$, with $a \neq 0$, then f has a root $\xi_1 = \cdots = \xi_{n-1} = 0$ of multiplicity $(n - 1)$ and a simple root $\xi_n = b$. If $n \ge 3$, then all invariants of f vanish by Theorem 5.4. However, the Hessian seminvariant is $-b^2 \neq 0$ by (4.4) (or by (2.39) and $a_1 = -b$, $a_2 = 0$).

Similarly, again by (4.4), the Hessian seminvariant is non-zero for any polynomial with all n roots real, unless all roots coincide.

Theorem 5.2 characterizes the polynomials that can be written as an nth power $c(x - \xi)^n$. There is a generalization (due to Gundelfinger [9], see also Kung [13]) to sums $\sum_{i=1}^{m} c_i(x - \xi_i)^n$ of a given number m of such powers; however, we also have to include limit cases corresponding to several coinciding ξ_i, and the precise statement is as follows.

Theorem 5.7. *The following are equivalent, for any polynomial f of degree n and $1 \le m \le n$.*

(i) $G_m(f) = 0$.

(ii) *f belongs to the closure $\overline{\mathcal{P}}_{n,m}$ of the set of polynomials $\mathcal{P}_{n,m} := \{\sum_{i=1}^{m} c_i(x - \xi_i)^n : c_i, \xi_i \in F\}$.*

(iii) *$f = \sum_{i=1}^{l} \sum_{j=0}^{m_i - 1} c_{ij}(x - \xi_i)^{n-j}$ for some $l \le m$, $m_i \ge 1$ with $\sum_{i=1}^{l} m_i = m$, $c_{ij} \in F$ and $\xi_i \in F^*$, for $i = 1, \dots, l$ and $j = 0, \dots, m_i - 1$.*

(iv) *There exists a polynomial g of degree (at most) m such that the apolar invariant $\{f, g\}_m = 0$.*

Remark 5.8. In (iii), we allow the possibility $\xi_i = \infty$; in this case we use the interpretation $(x - \infty)^{n-j} := x^j$ (which is natural from a projective perspective).

By "closure" in (ii), we mean in the ordinary topological sense (identifying a polynomial with its vector $\mathbf{a}$ of coefficients) if, for example, we consider the field of rational, real or complex numbers. In general, the closure can be interpreted algebraically, as the set of all $f = f_0$ for some family f_ε of polynomials, with coefficients that are polynomials in a parameter $\varepsilon \in F$ (or $\varepsilon \in \mathbb{Q}$), such that $f_\varepsilon \in \mathcal{P}_{n,m}$ for all $\varepsilon \neq 0$.

In particular, if $m > n/2$, then $\overline{\mathcal{P}}_{n,m} = \mathcal{P}_n$, i.e., every polynomial is in $\overline{\mathcal{P}}_{n,m}$, since then G_m vanishes identically on $\mathcal{P}_n$, see Example 2.17. If n is even, then $G_{n/2}$ is a multiple of the catalecticant $\operatorname{Han}(f)$, see Example 2.17, and thus we have the corollary:

Corollary 5.9. *If n is even then $\operatorname{Han}(f) = 0$ if and only if $f \in \overline{\mathcal{P}}_{n,n/2}$, i.e., if and only if f is as in Theorem 5.7*(iii) *with $m = n/2$.*

The relation between (iii) and (iv) in Theorem 5.7 can be made more precise as follows, see [14].

Theorem 5.10. *The following are equivalent, for a polynomial f of degree n and given $\xi_i \in F^*$ and $m_i \geq 1$, $i = 1, \dots, l$, with $m := \sum_{i=1}^{l} m_i \leq n$,*

(i) *$f = \sum_{i=1}^{l} \sum_{j=0}^{m_i-1} c_{ij}(x - \xi_i)^{n-j}$ for some $c_{ij} \in F$, $i = 1, \dots, l$ and $j = 0, \dots, m_i - 1$.*

(ii) *If $g = \prod_{i=1}^{l}(x - \xi_i)^{m_i}$, then the apolar invariant $\{f, g\}_m = 0$. (Note that g is a poynomial of degree m.)*

If $\xi_i = \infty$, we interpret $(x - \infty)^{n-j}$ as x^j in (i), as above, and $(x - \infty)^{m_j}$ as 1 in (ii).

6. Invariants of polynomials of degree 1

All seminvariants of a linear polynomial $a_0 x + a_1$ are of the form $c\, a_0^w$. In other words, $\{a_0\}$ is a basis for the seminvariants.

There are no invariants (except constants). (The discriminant is trivially 1.)

7. Invariants of polynomials of degree 2

We consider invariants etc. of a polynomial $f(x) = a_0x^2 + a_1x + a_2$ of degree 2 (a *quadratic* polynomial).

7.1. Invariants.

The discriminant is, as is well-known and easily verified,

$$\Delta(f) = a_1^2 - 4a_0a_2. \tag{7.1}$$

The discriminant is an invariant of degree $\nu = 2$ and weight $w = 2$.

The reduced form of f is

$$\widehat{f}(x) := f(x - a_1/2a_0) = a_0x^2 - \frac{a_1^2 - 4a_0a_2}{4a_0} = a_0x^2 - \frac{\Delta}{4a_0}; \tag{7.2}$$

hence the only non-trivial coefficient of the reduced form is $\widehat{a}_2 := -\Delta/4a_0$, so $a_0\widehat{a}_2$ is the invariant $-\Delta/4$ of degree $\nu = 2$ and weight $w = 2$, cf. Theorem 3.16.

The apolar invariant of f, see Example 2.8, is

$$A(f,f) = 4a_0a_2 - a_1^2 = -\Delta. \tag{7.3}$$

This is another invariant of degree and weight $\nu = w = 2$.

The Hankel determinant (catalecticant) of f, see Example 2.9, is

$$\operatorname{Han}(f) = \begin{vmatrix} \breve{a}_0 & \breve{a}_1 \\ \breve{a}_1 & \breve{a}_2 \end{vmatrix} = \begin{vmatrix} a_0 & \frac{1}{2}a_1 \\ \frac{1}{2}a_1 & a_2 \end{vmatrix} = a_0a_2 - \frac{1}{4}a_1^2 = -\frac{1}{4}\Delta. \tag{7.4}$$

Again, this is an invariant of degree $\nu = 2$ and weight $w = 2$.

The Hessian covariant is by Example 2.13 a covariant of order $2(n-2) = 0$ for $n = 2$, i.e., an invariant. Thus, $H_0 = H$. We have, using Example 3.2,

$$H(f) = H(f;x) = 4a_0(a_0x^2 + a_1x + a_2) - (2a_0x + a_1)^2 = 4a_0a_2 - a_1^2 = -\Delta. \tag{7.5}$$

Once again, this is an invariant of degree $\nu = 2$ and weight $w = 2$.

Of course, these invariants are multiples of each other. In fact, as said in Example 2.8 for any n, there is no other invariants of degree 2. Moreover, Δ is a basis for the invariants, i.e., every invariant is $c\Delta^\ell$ for some c and ℓ [18, Sätze 1.9 and 2.8].

7.2. **Seminvariants and covariants.** The leading coefficient a_0 is a seminvariant of degree 1 and weight 0. It is the source of the covariant $f(x)$ of degree 1, order 2 and weight 0, see Examples 2.12 and 2.26.

Theorem 7.1 ([18, Satz 2.16])**.** *The covariants f and Δ form a basis of the covariants; thus $\{a_0, \Delta\}$ is a basis of the seminvariants [18, Satz 2.16].*

Hence, the only seminvariant of degree ν and weight w (up to constant factors) is $a_0^{\nu-w}\Delta^{w/2}$ provided w is even and $\nu \geq w$; there are no seminvariants for other ν and w.

As a further example, the only non-trivial Gundelfinger covariant, see Example 2.17, is the invariant $G_1(f) = H(f) = -\Delta$; by (2.26) we also have $G_1(f) = 4\operatorname{Han}(f)$ in accordance with (7.4).

7.3. **Seminvariants of f'.** Since the only seminvariant of a linear function is the leading coefficient a_0, the only seminvariant of $f'(x) = 2a_0x + a_1$ is $2a_0$.

7.4. **The case $a_0 = 0$.** When $a_0 = 0$, i.e., considering the restriction to polynomials of degree 1, the essentially only non-trivial formula is $\Delta(a_1x + a_2) = a_1^2$, or, equivalently,

$$\Delta_{\langle 2\rangle}(a_{0\,\langle 1\rangle}x + a_{1\,\langle 1\rangle}) = a_{0\,\langle 1\rangle}^2, \tag{7.6}$$

cf. Example 4.8. In particular, for $f \in \mathcal{P}_2$,

$$\Delta_{\langle 2\rangle}(f') = 4a_0^2. \tag{7.7}$$

7.5. **Seminvariants and roots.** By Example 4.2,

$$\Delta = a_0^2(\xi_1 - \xi_2)^2. \tag{7.8}$$

This agrees with (4.4), since $H_0 = H = -\Delta$ by (7.5).

The general seminvariant of degree ν and weight w is thus

$$a_0^{\nu-w}\Delta^{w/2} = a_0^{\nu}(\xi_1 - \xi_2)^w, \tag{7.9}$$

for $\nu \geq w$ and w even (otherwise there is no such invariant).

7.6. **Further examples.** As examples of invariants of higher degree, we compute the basic invariants (covariants, seminvariants) for $n = 4$ (see Section 9 below) of f^2; these are clearly invariants (etc.) of f by Theorem 2.32:

$$A(f^2, f^2) = 4\Delta^2, \tag{7.10}$$

$$I(f^2) = \Delta^2, \tag{7.11}$$

$$J(f^2) = -2\Delta^3, \tag{7.12}$$

$$\Delta(f^2) = 0, \tag{7.13}$$

$$H_0(f^2) = -12\, a_0^2\, \Delta. \tag{7.14}$$

$$P(f^2) = -4\, a_0^2\, \Delta. \tag{7.15}$$

$$Q(f^2) = 0, \tag{7.16}$$

$$H(f^2) = -12\, \Delta\, f^2, \tag{7.17}$$

$$G_6(f^2) = 0. \tag{7.18}$$

8. Invariants of polynomials of degree 3

We consider invariants etc. of a polynomial $f(x) = a_0x^3 + a_1x^2 + a_2x + a_3$ of degree 3 (a *cubic* polynomial).

We give a table of covariants of low degree in Theorem 1, and the corresponding seminvariants in Theorem 2, using notation introduced below. (The tables give bases; further examples may be constructed by taking linear combinations of the covariants (seminvariants) in each entry.) It is easily checked that the dimensions agree with Theorem 2.37 (using for example [2, Table 14.3]). The invariants have $w = 3\nu/2$; these are all powers of Δ, and the only example in the tables is Δ.

8.1. **Invariants.** The discriminant is, see e.g. [11],

$$\Delta(f) = a_1^2a_2^2 - 4a_1^3a_3 - 4a_0a_2^3 + 18a_0a_1a_2a_3 - 27a_0^2a_3^2. \tag{8.1}$$

This is an invariant of degree 4 and weight 6.

Different normalizations are sometimes used. We have $\Delta = D_{[18]} = -27d_{[18]} = \frac{27}{2}R_{[8]}$.

As for $n = 2$, Δ is a basis for the invariants, i.e., every invariant is $c\Delta^\ell$ for some c and ℓ [18, Satz 2.8].

The apolar invariant $A(f, f)$ vanishes since $n = 3$ is odd.

	0	1	2	3	4	5	6	7	8	9
1	a_0									
2	a_0^2		P							
3	a_0^3		a_0P	Q						
4	a_0^4		a_0^2P	a_0Q	P^2		Δ			
5	a_0^5		a_0^3P	a_0^2Q	a_0P^2	PQ	$a_0\Delta$			
6	a_0^6		a_0^4P	a_0^3Q	$a_0^2P^2$	a_0PQ	$a_0^2\Delta, P^3; Q^2$		ΔP	
7	a_0^7		a_0^5P	a_0^4Q	$a_0^3P^2$	a_0^2PQ	$a_0^3\Delta, a_0P^3; a_0Q^2$	P^2Q	$a_0\Delta P$	ΔQ

Table 1. Invariants and seminvariants of low degree of cubic polynomials. Each entry gives either a basis for the linear space of seminvariants of given degree (row) and weight (column), or a basis separated by a semicolon from further examples of such seminvariants.

	0	1	2	3	4	5	6	7	8	9
1	f									
2	f^2		H							
3	f^3		fH	G						
4	f^4		f^2H	fG	H^2		Δ			
5	f^5		f^3H	f^2G	fH^2	HG	$f\Delta$			
6	f^6		f^4H	f^3G	f^2H^2	fHG	$f^2\Delta, H^3; G^2$		ΔH	
7	f^7		f^5H	f^4G	f^3H^2	f^2HG	$f^3\Delta, fH^3; fG^2$	H^2G	$f\Delta H$	ΔG

Table 2. Invariants and covariants of low degree of cubic polynomials. Each entry gives either a basis for the linear space of covariants of given degree (row) and weight (column), or a basis separated by a semicolon from further examples of such covariants.

8.2. **Reduced form.** The reduced form of f is

$$\widehat{f}(x) = a_0x^3 + px + q := f\Big(x - \frac{a_1}{3a_0}\Big), \tag{8.2}$$

which yields

$$p := \frac{3a_0a_2 - a_1^2}{3a_0}, \tag{8.3}$$

$$q := \frac{2a_1^3 + 27a_0^2a_3 - 9a_0a_1a_2}{27a_0^2}. \tag{8.4}$$

In terms of the coefficients of the reduced polynomial $\widehat{f}(x) = a_0x^3+px+q$, the discriminant is given by

$$\Delta(f) = \Delta(\widehat{f}) = -4a_0p^3 - 27a_0^2q^2 \tag{8.5}$$

8.3. **Seminvariants.** The coefficients p and q in (8.2) are rational seminvariants by Theorem 3.16. We conventionally denote the numerators in (8.3) and (8.4) by P and Q and have thus the seminvariants

$$P := 3a_0a_2 - a_1^2, \tag{8.6}$$
$$Q := 2a_1^3 + 27a_0^2a_3 - 9a_0a_1a_2. \tag{8.7}$$

P has degree 2 and weight 2; Q has degree 3 and weight 3. Conversely, we have

$$p = \frac{P}{3\,a_0}, \tag{8.8}$$
$$q = \frac{Q}{27\,a_0^2}. \tag{8.9}$$

(Other notations: $P = -P_{[3]}$, with opposite sign; $Q = U_{[3]}$.)

By (8.5) and (8.8)–(8.9), the discriminant is given by

$$\Delta = -\frac{4P^3}{27\,a_0^2} - \frac{Q^2}{27\,a_0^2}. \tag{8.10}$$

Hence, the relation (*syzygy*)

$$27\,a_0^2\,\Delta = -4P^3 - Q^2. \tag{8.11}$$

8.4. **Covariants.** The form f itself is a covariant of degree 1, weight 0 and order 3, see Example 2.12.

The Hessian covariant is the polynomial of degree $2(n-2) = 2$ given by (3.4), which yields

$$H(f;x) = \bigl(12\,a_0\,a_2 - 4\,a_1^2\bigr)\,x^2 + (36\,a_0\,a_3 - 4\,a_1\,a_2)\,x + 12\,a_1\,a_3 - 4\,a_2^2. \tag{8.12}$$

This is a covariant of degree 2, weight 2 and order 2.

The Hessian source H_0 (Example 2.27) is thus the seminvariant of degree 2 and weight 2

$$H_0 = 12\,a_0\,a_2 - 4\,a_1^2 = 12a_0\,p = 4P. \tag{8.13}$$

Conversely, P is the source of the covariant

$$\begin{aligned}\widetilde{H}(x) = \widetilde{H}(f;x) &:= \frac{1}{4}H(f;x)\\ &= \bigl(3\,a_0\,a_2 - a_1^2\bigr)\,x^2 + (9\,a_0\,a_3 - a_1\,a_2)\,x + 3\,a_1\,a_3 - a_2^2.\end{aligned} \tag{8.14}$$

(Other notations: $H(X)_{[3]} = -\widetilde{H}(X)$, so $H(X)_{[3]}$ has source $-P = P_{[3]}$. Further, $H = 18\Delta_{[8]} = 36h_{[18]}$; $\widetilde{H} = 6\Delta_{[8]} = 9h_{[18]}$.)

The only non-trivial Gundelfinger covariant, see Example 2.17, is $G_1(f) = H(f)$.

The Jacobian (see Example 2.14) of $f(x)$ and $H(f;x)$ is

$$\begin{aligned}J(f,H(f)) = \bigl(108\,a_0^2\,a_3 - 36\,a_0\,a_1\,a_2 + 8\,a_1^3\bigr)\,x^3 \\ + \bigl(108\,a_0\,a_1\,a_3 - 72\,a_0\,a_2^2 + 12\,a_1^2\,a_2\bigr)\,x^2 \\ + \bigl(-108\,a_0\,a_2\,a_3 + 72\,a_1^2\,a_3 - 12\,a_1\,a_2^2\bigr)\,x \\ + 36\,a_1\,a_2\,a_3 - 8\,a_2^3 - 108\,a_0\,a_3^2. \qquad (8.15)\end{aligned}$$

This is, by Theorem 2.32, a covariant, which has degree 3, order 3 and weight 3. Its source is

$$108\,a_0^2\,a_3 - 36\,a_0\,a_1\,a_2 + 8\,a_1^3 = 4Q. \qquad (8.16)$$

Conversely, the covariant corresponding to the seminvariant Q is

$$\begin{aligned}G(x) := \frac{1}{4}J(f,H(f)) = \bigl(27\,a_0^2\,a_3 - 9\,a_0\,a_1\,a_2 + 2\,a_1^3\bigr)\,x^3 \\ + \bigl(27\,a_0\,a_1\,a_3 - 18\,a_0\,a_2^2 + 3\,a_1^2\,a_2\bigr)\,x^2 \\ + \bigl(-27\,a_0\,a_2\,a_3 + 18\,a_1^2\,a_3 - 3\,a_1\,a_2^2\bigr)\,x \\ + 9\,a_1\,a_2\,a_3 - 2\,a_2^3 - 27\,a_0\,a_3^2 \qquad (8.17)\end{aligned}$$

(Other notations: $G(x) = G(x)_{[3]} = 27\,Q_{[8]} = 27\,j_{[18]}$; $T_{[14]} = J(H(f),f) = -4G$.)

The relation (8.11) corresponds to the similar relation (syzygy) between the corresponding covariants

$$27f(x)^2\Delta = -4(H(f;x)/4)^3 - G(x)^2 \qquad (8.18)$$

or

$$432\,\Delta\,f(x)^2 + H(f;x)^3 + 16\,G(x)^2 = 432\,\Delta\,f(x)^2 + H(f;x)^3 + J(f,H(f))^2 = 0 \qquad (8.19)$$

Theorem 8.1 ([18, Satz 2.24])**.** *The covariants $\{f,H,G,\Delta\}$ form a basis of all covariants for cubic polynomials. Equivalently, $\{a_0,P,Q,\Delta\}$ is a basis of all seminvariants.*

The basis is not algebraically independent since we have the syzygy (8.19), i.e. $432\Delta f^2 + H^3 + 16G^2 = 0$, or, equivalently, (8.11).

8.5. **Seminvariants of f'.** The discriminant of the quadratic polynomial f' is a seminvariant by Theorem 3.7; it is given by

$$\Delta_{\langle 2\rangle}(f') = \Delta_{\langle 2\rangle}(3a_0x^2 + 2a_1x + a_2) = 4\,a_1^2 - 12\,a_0\,a_2 = -4P. \qquad (8.20)$$

This has, cf. Remark 3.8, degree 2 and weight 2 as the discriminant for $n=2$, see Section 7; its order is 2.

Alternatively, by Example 3.9 and (8.13), we have

$$H_0(f') = H_0(f) = 4P. \qquad (8.21)$$

Since $\Delta = -H = -H_0$ for a quadratic polynomial, see (7.5), we obtain (8.20).

8.6. **The case $a_0 = 0$.** When $a_0 = 0$, i.e., considering the restriction to polynomials of degree 2, we have, cf. Example 4.8,

$$\Delta_{\langle 3\rangle}(a_{0\,\langle 2\rangle}x^2 + a_{1\,\langle 2\rangle}x + a_{2\,\langle 2\rangle}) = a_{0\,\langle 2\rangle}^2\Delta_{\langle 2\rangle}, \tag{8.22}$$

$$P_{\langle 3\rangle}(a_{0\,\langle 2\rangle}x^2 + a_{1\,\langle 2\rangle}x + a_{2\,\langle 2\rangle}) = -a_{0\,\langle 2\rangle}^2, \tag{8.23}$$

$$Q_{\langle 3\rangle}(a_{0\,\langle 2\rangle}x^2 + a_{1\,\langle 2\rangle}x + a_{2\,\langle 2\rangle}) = 2\,a_{0\,\langle 2\rangle}^3. \tag{8.24}$$

In particular, for $f \in \mathcal{P}_3$, using (8.20),

$$\Delta_{\langle 3\rangle}(f') = 9\,a_0^2\,\Delta_{\langle 2\rangle}(f') = -36\,a_0^2\,P, \tag{8.25}$$

$$P_{\langle 3\rangle}(f') = -9\,a_0^2, \tag{8.26}$$

$$Q_{\langle 3\rangle}(f') = 54\,a_0^3. \tag{8.27}$$

8.7. **Seminvariants and roots.** By Example 4.2,

$$\Delta = a_0^4(\xi_1-\xi_2)^2(\xi_1-\xi_3)^2(\xi_2-\xi_3)^2. \tag{8.28}$$

By (4.4),

$$\begin{aligned} H_0 &= -2\,a_0^2\big((\xi_1-\xi_2)^2 + (\xi_1-\xi_3)^2 + (\xi_2-\xi_3)^2\big) \\ &= -4\,a_0^2\left(\xi_1^2+\xi_2^2+\xi_3^2-\xi_1\,\xi_2-\xi_1\,\xi_3-\xi_2\,\xi_3\right), \end{aligned} \tag{8.29}$$

and thus, by (8.13),

$$\begin{aligned} P &= -\frac{a_0^2}{2}\big((\xi_1-\xi_2)^2 + (\xi_1-\xi_3)^2 + (\xi_2-\xi_3)^2\big) \\ &= -a_0^2\left(\xi_1^2+\xi_2^2+\xi_3^2-\xi_1\,\xi_2-\xi_1\,\xi_3-\xi_2\,\xi_3\right). \end{aligned} \tag{8.30}$$

Further, by a calculation or from (8.9) and (8.2), noting that $\widehat{f}$ has roots $\xi_i - (\xi_1+\xi_2+\xi_3)/3$,

$$\begin{aligned} Q &= -a_0^3\,(2\,\xi_1-\xi_2-\xi_3)\,(2\,\xi_2-\xi_1-\xi_3)\,(2\,\xi_3-\xi_1-\xi_2) \\ &= -a_0^3\big(2\,\xi_1^3+2\,\xi_2^3+2\,\xi_3^3-3\,\xi_1\xi_2^2-3\,\xi_1\xi_3^2-3\,\xi_1^2\xi_2-3\,\xi_1^2\xi_3 \\ &\qquad -3\,\xi_2\xi_3^2-3\,\xi_2^2\xi_3+12\xi_1\xi_2\xi_3\big). \end{aligned} \tag{8.31}$$

8.8. **Covariants and roots.** For the corresponding covariants we have first by Example 4.15, cf., (8.29),

$$H(f;x) = -2a_0^2\big((\xi_1-\xi_2)^2(x-\xi_3)^2 + (\xi_1-\xi_3)^2(x-\xi_2)^2 + (\xi_2-\xi_3)^2(x-\xi_1)^2\big). \tag{8.32}$$

For G we use (8.31) and Theorem 4.13. We have

$$\xi_1\xi_2\xi_3(2\xi_1^{-1}-\xi_2^{-1}-\xi_3^{-1}) = 2\xi_2\xi_3-\xi_1\xi_3-\xi_1\xi_2 = \xi_2(\xi_3-\xi_1)+\xi_3(\xi_2-\xi_1)$$

which after the substitution $\xi_i \mapsto x-\xi_i$ and permutation of the indices leads to

$$\begin{aligned} G(f;x) &= -a_0^3\Big((x-\xi_2)(\xi_1-\xi_3)+(x-\xi_3)(\xi_1-\xi_2)\Big) \\ &\cdot\Big((x-\xi_1)(\xi_2-\xi_3)+(x-\xi_3)(\xi_2-\xi_1)\Big)\Big((x-\xi_1)(\xi_3-\xi_2)+(x-\xi_2)(\xi_3-\xi_1)\Big). \end{aligned} \tag{8.33}$$

8.9. **Further examples.** The apolar invariant of the Hessian covariant is an invariant given by, see (8.12) and (7.3),

$$\begin{aligned} A(H(f),H(f)) &= 4\left(12\,a_0\,a_2-4\,a_1^2\right)\left(12\,a_1\,a_3-4\,a_2^2\right)-\left(36\,a_0\,a_3-4\,a_1\,a_2\right)^2 \\ &= -1296\,a_0^2\,a_3^2+864\,a_0\,a_1\,a_2\,a_3-192\,a_0\,a_2^3-192\,a_1^3\,a_3+48\,a_1^2\,a_2^2 \\ &= 48\,\Delta. \end{aligned} \tag{8.34}$$

This has degree 4 and weight 6.

The apolar invariant of the 6th degree polynomial f^2 is

$$\begin{aligned} A(f^2,f^2) &= 1296\,a_0^2\,a_3^2-864\,a_0\,a_1\,a_2\,a_3+192\,a_0\,a_2^3+192\,a_1^3\,a_3-48\,a_1^2\,a_2^2 \\ &= -48\,\Delta. \end{aligned} \tag{8.35}$$

Similarly, the apolar invariant of the 12th degree polynomial f^4 is

$$A(f^4,f^4) = 1244160\,\Delta^2 = 2^{10}\cdot 3^5\cdot 5\cdot \Delta^2. \tag{8.36}$$

Recall that every invariant is a constant times a power of Δ, so these formulas are no surprises.

The discriminant of the quadratic covariant $H(x)$ is

$$\begin{aligned} \Delta_{\langle 2\rangle}(H(x)) &= 1296\,a_0^2\,a_3^2-864\,a_0\,a_1\,a_2\,a_3+192\,a_0\,a_2^3+192\,a_1^3\,a_3-48\,a_1^2\,a_2^2 \\ &= -48\Delta, \end{aligned} \tag{8.37}$$

cf. (7.3) and (8.34). Thus the covariant $\widetilde{H}(x)$ in (8.14) corresponding to P has discriminant -3Δ.

The discriminant and covariants H and G of the cubic covariant $G(x)$ are

$$\Delta(G(x)) = 729\,\Delta^3, \tag{8.38}$$

$$H(G(x)) = 27\,\Delta\,H(x), \tag{8.39}$$

$$G(G(x)) = -729\,\Delta^2\,f(x). \tag{8.40}$$

We calculate also the resultants of f, $H(f)$ and $G(f)$:

$$R(f,H) = -64\Delta^2, \tag{8.41}$$

$$R(f,G) = 8\Delta^3, \tag{8.42}$$

$$R(H,G) = -1728\Delta^3, \tag{8.43}$$

where the first also follows by Example 4.11.

For the seminvariants in Examples 4.5–4.6, we have, recalling $\widehat{a}_2 = p = P/3a_0$, $\widehat{a}_3 = q = Q/27a_0^2$ and $\widehat{a}_4 = 0$, see (8.2) and (8.8)–(8.9),

$$a_0^2 S_2 = -\frac{2}{3}P, \tag{8.44}$$

$$a_0^3 S_3 = -\frac{1}{9}Q, \tag{8.45}$$

$$a_0^4 S_4 = \frac{2}{9}P^2, \tag{8.46}$$

and

$$a_0^2\chi_2 = -\frac{2}{9}P, \tag{8.47}$$
$$a_0^3\chi_3 = -\frac{1}{27}Q, \tag{8.48}$$
$$a_0^4\chi_4 = -\frac{2}{27}P^2. \tag{8.49}$$

8.10. **Vanishing invariants and covariants.**

Theorem 8.2. *Let f be a polynomial of degree* 3.

(i) $\Delta(f) = 0$ *if and only if f has a double (or triple) root; i.e., if and only if it has a square factor.*
(ii) $H(f) = 0$ *if and only if f has a triple root, i.e., if and only if $f(x) = c(x - x_0)^3$.*
(iii) $G(f) = 0$ *if and only if f has a triple root, i.e., if and only if $f(x) = c(x - x_0)^3$.*

Proof. Parts (i) and (ii) are Theorems 5.1 and 5.2. For (iii), suppose that $G = 0$. By (8.38), then $\Delta = 0$, so f has a double root ξ. By projective invariance, we may assume that $\xi = 0$, so $f(x) = a_0x^3 + a_1x^2$. Then, by (8.17), $G(x) = 2a_1^3x^3$, and thus $a_1 = 0$ too, and $\xi = 0$ is a triple root. (Alternatively, $G = 0$ and $\Delta = 0$ imply $H = 0$ by (8.19), and we may use (ii).)

The converse follows similarly from (8.17) and projective invariance, or from (8.19) and (i)+(ii). □

8.11. **Geometry of real cubics.** Let f be a real cubic, with $a_0 \neq 0$. Then f has an inflection point (x_0, y_0) given by $0 = f''(x_0) = 6a_0x_0 + 2a_1$, so $x_0 = -a_1/3a_0$ and, using (8.2),

$$y_0 = f(x_0) = \widehat{f}(0) = q. \tag{8.50}$$

Thus, by (8.9),

$$(x_0, y_0) = \Bigl(-\frac{a_1}{3a_0}, q\Bigr) = \Bigl(-\frac{a_1}{3a_0}, \frac{Q}{27\,a_0^2}\Bigr). \tag{8.51}$$

Note that f is symmetric about (x_0, y_0), cf. (8.2).

The extreme points $x_\pm$ are given by, using (8.2) again,

$$0 = f'(x) = \widehat{f}'(x - x_0) = 3a_0(x - x_0)^2 + p; \tag{8.52}$$

hence, using also (8.8),

$$x_\pm = x_0 \pm \sqrt{\frac{-p}{3a_0}} = x_0 \pm \frac{\sqrt{-P}}{3a_0} = \frac{-a_1 \pm \sqrt{-P}}{3a_0}. \tag{8.53}$$

Consequently, f has real (local) maximum and minimum points if $P < 0$, but not if $P \geq 0$; in the latter case, f is monotonously increasing (if $a_0 > 0$)

or decreasing (if $a_0 < 0$) on $(-\infty, \infty)$. (This includes the case $P = 0$, when $f'(x_0) = f''(x_0) = 0$.)

Moreover, the extreme values $y_\pm = f(x_\pm)$ are given by, using (8.2), (8.53) and (8.8)–(8.9),

$$\begin{aligned} y_\pm := f(x_\pm) &= \widehat{f}(x_\pm - x_0) = a_0(x_\pm - x_0)^3 + p(x_\pm - x_0) + q \\ &= a_0\left(\frac{\pm\sqrt{-P}}{3a_0}\right)^3 + p\,\frac{\pm\sqrt{-P}}{3a_0} + q = \frac{\pm 2P\sqrt{-P}}{27a_0^2} + \frac{Q}{27a_0^2} \\ &= \frac{Q \pm 2P\sqrt{-P}}{27a_0^2}. \end{aligned} \tag{8.54}$$

In particular, we see that f has three distinct real roots

$\iff$ $x_\pm$ are real and $y_- < 0 < y_+$ or $y_+ < 0 < y_-$

$\iff$ $P < 0$ and $|2P\sqrt{-P}| > |Q|$

$\iff$ $-4P^3 > Q^2$

$\iff$ $\Delta = -(4P^3 + Q^2)/27a_0^2 > 0$.

Similarly, there is a real double root if $P < 0$ and $\Delta = 0$, and a triple root if $P = 0 = Q$. We thus have found the following classical result, which also follows directly from (4.3), see [11, 12]:

Theorem 8.3. *Let f be a real cubic.*

(i) *If $\Delta > 0$, then f has 3 distinct real roots.*
(ii) *If $\Delta = 0$, then f has either one double and one simple root, both real ($P < 0$), or a real triple root ($P = Q = 0$).*
(iii) *If $\Delta < 0$, then f has one real root and a pair of two (non-real) conjugate complex roots.*

Remark 8.4. More generally, it follows from (4.3) that if f is a real polynomial of degree n with only simple roots, having $n - 2m$ real roots and m pairs of conjugate complex (non-real) roots, then $\operatorname{sign}(\Delta(f)) = (-1)^m$.

We also have, by (8.54) and (8.11), the quantitative relation

$$y_+ y_- = \frac{Q^2 + 4P^3}{729\, a_0^4} = -\frac{\Delta}{27\, a_0^2}. \tag{8.55}$$

In fact, since $\Delta = -a_0^{-1} R(f, f')$, where R is the resultant, this follows immediately from a standard property of the resultant; more generally, for a polynomial of arbitrary degree n, with stationary points (roots of f') $\eta_1, \dots, \eta_{n-1}$,

$$\Delta(f) = (-1)^{n(n-1)/2} n^n a_0^{n-1} \prod_{j=1}^{n-1} f(\eta_j), \tag{8.56}$$

see [11].

Note also the corresponding formula, by (8.53) and (8.6) or directly from $f'(x) = 3a_0x^2 + 2a_1x + a_2$,

$$x_+ x_- = \frac{a_2}{3a_0}. \tag{8.57}$$

We can further study the location of the roots. We have for example the following criteria for positive roots.

Theorem 8.5. *Let f be a real cubic with $a_0 > 0$.*

(i) *f has three distinct positive roots in $(0, \infty)$*
$\iff$ *$\Delta > 0$ (which implies $P < 0$), $a_3 < 0$ and $-a_1 > \sqrt{-P}$*
$\iff$ *$\Delta > 0$, $a_1 < 0$, $a_2 > 0$, $a_3 < 0$.*

(ii) *f has three roots (not necessarily distinct) in $[0, \infty)$*
$\iff$ *$\Delta \geq 0$ (which implies $P \leq 0$), $a_3 \leq 0$ and $-a_1 \geq \sqrt{-P}$*
$\iff$ *$\Delta \geq 0$, $a_1 \leq 0$, $a_2 \geq 0$, $a_3 \leq 0$.*

Proof. Consider for example (i). We may suppose that f has three real roots, so $\Delta > 0$, and then $P < 0$ by (8.11). A geometric consideration shows that the roots are all positive $\iff$ $x_\pm > 0$ and $a_3 = f(0) < 0$, and the result follows by (8.53) and (8.6). Case (ii) is similar, considering also cases with a double or triple root. □

Note that (4.1) immediately implies that if $\xi_1, \xi_2, \xi_3 \geq 0$, and $a_0 > 0$, then $a_1 \leq 0$, $a_2 \geq 0$, $a_3 \leq 0$, but the converse is less obvious.

9. Invariants of polynomials of degree 4

We consider invariants etc. of a polynomial $f(x) = a_0x^4 + a_1x^3 + a_2x^2 + a_3x + a_4$ of degree 4 (a *quartic* polynomial).

We give a table of covariants of low degree in Theorem 3, and corresponding seminvariants in Theorem 4, using notation introduced below. Again, it is easily checked that the dimensions agree with Theorem 2.37 (using for example [2, Table 14.3]). (The examples given in the table is a rather arbitrary selection when the dimension is > 1. For example, note that when $\nu = w$, there is by Theorem 3.19 always a basis for the seminvariants consisting of monomials in P, Q, R; for example, for $\nu = w = 5$, $\{P^3, Q^2, PR\}$.) The invariants have $w = 2\nu$; for each such ν and w, there is a basis consisting of monomials in I and J (but for $\nu = 6$, $w = 12$, $\{\Delta, I^3\}$ is another example).

9.1. Invariants.

The discriminant is, see [11],

$$\begin{aligned}\Delta(f) = {} & 256\, a_0^3\, a_4^3 - 192\, a_0^2\, a_1\, a_3\, a_4^2 - 128\, a_0^2\, a_2^2\, a_4^2 \\ & + 144\, a_0^2\, a_2\, a_3^2\, a_4 - 27\, a_0^2\, a_3^4 + 144\, a_0\, a_1^2\, a_2\, a_4^2 \\ & - 6\, a_0\, a_1^2\, a_3^2\, a_4 - 80\, a_0\, a_1\, a_2^2\, a_3\, a_4 + 18\, a_0\, a_1\, a_2\, a_3^3 \\ & + 16\, a_0\, a_2^4\, a_4 - 4\, a_0\, a_2^3\, a_3^2 - 27\, a_1^4\, a_4^2 \\ & + 18\, a_1^3\, a_2\, a_3\, a_4 - 4\, a_1^3\, a_3^3 - 4\, a_1^2\, a_2^3\, a_4 + a_1^2\, a_2^2\, a_3^2. \end{aligned} \tag{9.1}$$

(See also (9.18) below.) This is an invariant of degree 6 and weight 12. (Other notations: $\Delta_{[3]} = \Delta_{[4]} = 27\Delta$; $\Delta_{0[3]} = \Delta$; $D_{[18]} = \Delta$.)

There are simpler invariants, however. The apolar invariant, see Example 2.8, is

$$A(f, f) = 48\, a_0\, a_4 - 12\, a_1\, a_3 + 4\, a_2^2 = 4I, \tag{9.2}$$

	0	1	2	3	4	5	6	7	8	9	10	11	12
1	f												
2	f^2		H		I								
3	f^3		fH	G_6	fI		J						
4	f^4		f^2H	fG_6	f^2I, H^2		fJ, IH		I^2				
5	f^5		f^3H	f^2G_6	f^3I, fH^2	HG_6	f^2J, fIH	IG_6	fI^2, JH		IJ		
6	f^6		f^4H	f^3G_6	f^4I, f^2H^2	fHG_6	$f^3J, f^2IH, H^3; G_6^2$	fIG_6	f^2I^2, fJH, IH^2	JG_6	fIJ, I^2H		$I^3, J^2; \Delta$

TABLE 3. Invariants and covariants of low degree of quartic polynomials. Each entry gives either a basis for the linear space of covariants of given degree (row) and weight (column), or a basis separated by a semicolon from further examples of such covariants.

	0	1	2	3	4	5	6	7	8	9	10	11	12
1	a_0												
2	a_0^2		P		I								
3	a_0^3		a_0P	Q	a_0I		J						
4	a_0^4		a_0^2P	a_0Q	$a_0^2I, P^2; R$		a_0J, IP		I^2				
5	a_0^5		a_0^3P	a_0^2Q	a_0^3I, a_0P^2	PQ	a_0^2J, a_0IP	IQ	a_0I^2, JP		IJ		
6	a_0^6		a_0^4P	a_0^3Q	$a_0^4I, a_0^2P^2$	a_0PQ	$a_0^3J, a_0^2IP, P^3; Q^2$	a_0IQ	$a_0^2I^2, a_0JP, IP^2$	JQ	a_0IJ, I^2P		$I^3, J^2; \Delta$

TABLE 4. Invariants and seminvariants of low degree of quartic polynomials. Each entry gives either a basis for the linear space of seminvariants of given degree (row) and weight (column), or a basis separated by a semicolon from further examples of such seminvariants.

where I is the conveniently normalized invariant

$$I = 12\,a_0\,a_4 - 3\,a_1\,a_3 + a_2^2. \tag{9.3}$$

The apolar invariant and I are invariants of degree 2 and weight 4. (Other notations: $A = 4!\,A_{[18]} = 24\,A_{[18]}$; $I = 6\,i_{[8]} = 12\,P_{[18]}$.)

The Hankel determinant (catalecticant), see Example 2.9, is an invariant of degree 3 and weight 6. It is, by a calculation, in our normalization,

$$\begin{aligned}\operatorname{Han}(f) &= \begin{vmatrix}\breve a_0 & \breve a_1 & \breve a_2\\ \breve a_1 & \breve a_2 & \breve a_3\\ \breve a_2 & \breve a_3 & \breve a_4\end{vmatrix} = \begin{vmatrix} a_0 & \frac14 a_1 & \frac16 a_2\\ \frac14 a_1 & \frac16 a_2 & \frac14 a_3\\ \frac16 a_2 & \frac14 a_3 & a_4\end{vmatrix}\\ &= \frac{72\,a_0\,a_2\,a_4 - 27\,a_0\,a_3^2 - 27\,a_1^2\,a_4 + 9\,a_1\,a_2\,a_3 - 2\,a_2^3}{432}\\ &= \frac{J}{432},\end{aligned} \tag{9.4}$$

where we thus define

$$J := 72\,a_0\,a_2\,a_4 - 27\,a_0\,a_3^2 - 27\,a_1^2\,a_4 + 9\,a_1\,a_2\,a_3 - 2\,a_2^3. \tag{9.5}$$

J is thus an invariant of degree 3 and weight 6. (Other notation: $Q_{[18]} = \operatorname{Han}(f) = J_{[7]} = J/432$; $J_{[8]} = J/72$.)

By (2.26), the second Gundelfinger covariant in Example 2.17 is the invariant

$$G_2(f) = 24^3\operatorname{Han}(f) = 32\,J. \tag{9.6}$$

Another way to construct J is by taking the joint apolar invariant $A(H(f), f)$; this invariant of degree 3 and weight 6 equals $24J$, see (9.81).

Theorem 9.1 ([18, Satz 2.9])**.** *I and J form a basis for the invariants of quartic polynomials. Furthermore, I and J are algebraically independent.*

Thus, informally speaking, I and J are the only invariants. More precisely, every invariant is an isobaric polynomial in I and J. For example, the discriminant is such a polynomial; a calculation reveals that

$$\Delta = \frac{4}{27}I^3 - \frac{1}{27}J^2. \tag{9.7}$$

See Subsection 9.10 for further examples.

Example 9.2. Since I^3 and J^2 both are invariants of degree 6 and weight 12, the quotient I^3/J^2 is an absolute invariant. Similarly, J^2/I^3, I^3/Δ, J^2/Δ, etc. are absolute invariants; these are all simple rational functions of each other. In fact, since I and J form a basis for the invariants, it is easy to see that every absolute invariant is a rational function of I^3/J^2, or of any other of the absolute invariants just given.

9.2. **Covariants.** The form f itself is a covariant of degree 1, weight 0 and order 4, see Example 2.12.

The Hessian covariant is the polynomial of degree $2(n-2) = 4$ given by (3.4), which yields

$$\begin{aligned}H(f;x) = {}&\bigl(24\,a_0\,a_2 - 9\,a_1^2\bigr)\,x^4 + (72\,a_0\,a_3 - 12\,a_1\,a_2)\,x^3\\ &+ \bigl(144\,a_0\,a_4 + 18\,a_1\,a_3 - 12\,a_2^2\bigr)\,x^2 + (72\,a_1\,a_4 - 12\,a_2\,a_3)\,x\\ &+ (24\,a_2\,a_4 - 9\,a_3^2).\end{aligned} \tag{9.8}$$

This is a covariant of degree 2, weight 2 and order 4. We also define

$$\widetilde{H}(f;x) := \frac{1}{3}H(f;x) = \left(8\,a_0\,a_2 - 3\,a_1^2\right)x^4 + (24\,a_0\,a_3 - 4\,a_1\,a_2)\,x^3 \\ + \left(48\,a_0\,a_4 + 6\,a_1\,a_3 - 4\,a_2^2\right)x^2 + (24\,a_1\,a_4 - 4\,a_2\,a_3)\,x + 8\,a_2\,a_4 - 3\,a_3^2. \tag{9.9}$$

$\widetilde{H}$ too has degree 2, weight 2 and order 4. (Other notations: $g_{4[3]} = g_{4[4]} = -\widetilde{H}$; $h_{[18]} = \widetilde{H}/48 = H/144$; $H_{[18]} = H$.)

The Jacobian determinant, see Example 2.14, of f and $H(f)$ is a covariant of order $4+4-2=6$ given by

$$\begin{aligned} G_6(f) = {} & \left(288\,a_0^2\,a_3 - 144\,a_0\,a_1\,a_2 + 36\,a_1^3\right)x^6 \\ & + \left(1152\,a_0^2\,a_4 + 144\,a_0\,a_1\,a_3 - 288\,a_0\,a_2^2 + 72\,a_1^2\,a_2\right)x^5 \\ & + \left(1440\,a_0\,a_1\,a_4 - 720\,a_0\,a_2\,a_3 + 180\,a_1^2\,a_3\right)x^4 \\ & + \left(-720\,a_0\,a_3^2 + 720\,a_1^2\,a_4\right)x^3 \\ & + \left(-1440\,a_0\,a_3\,a_4 + 720\,a_1\,a_2\,a_4 - 180\,a_1\,a_3^2\right)x^2 \\ & + \left(-1152\,a_0\,a_4^2 - 144\,a_1\,a_3\,a_4 + 288\,a_2^2\,a_4 - 72\,a_2\,a_3^2\right)x \\ & + 144\,a_2\,a_3\,a_4 - 36\,a_3^3 - 288\,a_1\,a_4^2. \end{aligned} \tag{9.10}$$

We normalize this to $\widetilde{G_6}(f;x) := G_6(f;x)/36$, where thus

$$\begin{aligned} \widetilde{G_6}(f) = {} & \left(8\,a_0^2\,a_3 - 4\,a_0\,a_1\,a_2 + a_1^3\right)x^6 \\ & + \left(32\,a_0^2\,a_4 + 4\,a_0\,a_1\,a_3 - 8\,a_0\,a_2^2 + 2\,a_1^2\,a_2\right)x^5 \\ & + \left(40\,a_0\,a_1\,a_4 - 20\,a_0\,a_2\,a_3 + 5\,a_1^2\,a_3\right)x^4 \\ & + \left(-20\,a_0\,a_3^2 + 20\,a_1^2\,a_4\right)x^3 \\ & + \left(-40\,a_0\,a_3\,a_4 + 20\,a_1\,a_2\,a_4 - 5\,a_1\,a_3^2\right)x^2 \\ & + \left(-32\,a_0\,a_4^2 - 4\,a_1\,a_3\,a_4 + 8\,a_2^2\,a_4 - 2\,a_2\,a_3^2\right)x \\ & + 4\,a_2\,a_3\,a_4 - a_3^3 - 8\,a_1\,a_4^2 \end{aligned} \tag{9.11}$$

(Other notations: $g_{6[3]} = g_{6[4]} = \widetilde{G_6}$; $j_{[18]} = \widetilde{G_6}/32 = G_6/1152$.) G_6 and $\widetilde{G_6}$ have degree 3, weight 3 and order 6.

Theorem 9.3 ([18, Satz 2.25])**.** *The invariants I and J and the covariants f, H and G_6 form a basis for the covariants of quartic polynomials.*

The basic covariants satisfy the relation (syzygy)

$$\widetilde{H}^3 - 48\,If^2\widetilde{H} + 64\,Jf^3 + 27\,\widetilde{G_6}^2 = 0. \tag{9.12}$$

or

$$2^4H^3 - 2^8 3^3 If^2\,H + 2^{10}3^3 Jf^3 + 3^2 G_6^2 = 0. \tag{9.13}$$

The only non-trivial Gundelfinger covariants are $G_1(f) = H(f;x)$ and $G_2(f) = 24^3\,\mathrm{Han}(f) = 32\,J$, see (9.6).

9.3. **Reduced form.** The reduced form of f is

$$\widehat{f}(x) = a_0x^4 + px^2 + qx + r := f\Bigl(x - \frac{a_1}{4a_0}\Bigr); \tag{9.14}$$

thus $p := \widehat{a}_2$, $q := \widehat{a}_3$, $r := \widehat{a}_4$. These rational seminvariants are given by

$$p = \frac{8\,a_0\,a_2 - 3\,a_1^2}{8\,a_0} \tag{9.15}$$

$$q = \frac{8\,a_0^2\,a_3 - 4\,a_0\,a_1\,a_2 + a_1^3}{8\,a_0^2} \tag{9.16}$$

$$r = \frac{256\,a_0^3\,a_4 - 64\,a_0^2\,a_1\,a_3 + 16\,a_0\,a_1^2\,a_2 - 3\,a_1^4}{256a_0^3} \tag{9.17}$$

In terms of the coefficients of the reduced polynomial $\widehat{f}$, the discriminant is given by

$$\Delta(f) = \Delta(\widehat{f}) = -4\,a_0p^3q^2 + 16\,a_0p^4r - 27\,a_0^2q^4 + 144a_0^2\,pq^2r - 128\,a_0^2p^2r^2 + 256\,a_0^3r^3. \tag{9.18}$$

9.4. **Seminvariants.** We denote the numerators of (9.15)–(9.17) by P, Q, R, respectively, and have thus

$$p = \frac{P}{8\,a_0}, \tag{9.19}$$

$$q = \frac{Q}{8\,a_0^2}, \tag{9.20}$$

$$r = \frac{R}{256a_0^3}, \tag{9.21}$$

with

$$P := 8\,a_0\,a_2 - 3\,a_1^2, \tag{9.22}$$

$$Q := 8\,a_0^2\,a_3 - 4\,a_0\,a_1\,a_2 + a_1^3, \tag{9.23}$$

$$R := 256\,a_0^3\,a_4 - 64\,a_0^2\,a_1\,a_3 + 16\,a_0\,a_1^2\,a_2 - 3\,a_1^4. \tag{9.24}$$

(R should not be confused with the resultant in Section 4. Other notations: $P = H_{[3]} = -p_{[4]}$; $Q = R_{[3]} = r_{[4]}$.) These are seminvariants of degree and weight $(2,2)$, $(3,3)$ and $(4,4)$. We have

$$R = \frac{1}{3}\bigl(64\,a_0^2\,I - P^2\bigr). \tag{9.25}$$

The Hessian source H_0 is by (9.8) the seminvariant of degree 2 and weight 4

$$H_0 = 24\,a_0\,a_2 - 9\,a_1^2 = 3\,P. \tag{9.26}$$

Thus the source of the covariant $\widetilde{H}$ is P.

By (9.10)–(9.11) and (9.23), Q is the source of $\widetilde{G_6}$, while the source of G_6 is $36\,Q$.

Theorem 9.3 and the syzygy (9.12) translate to the following.

Theorem 9.4. *The invariants I and J and the seminvariants a_0, P and Q form a basis for the seminvariants of quartic polynomials. These satisfy the syzygy*

$$P^3 - 48\,I\,a_0^2\,P + 64\,J\,a_0^3 + 27\,Q^2 = 0. \tag{9.27}$$

9.5. **Cubic resolvent.** Let $\tilde{p} := p/a_0$, $\tilde{q} := q/a_0$, $\tilde{r} := r/a_0$, the coefficients of the reduced monic polynomial $\widehat{f}/a_0$. The *cubic resolvent* of f is the cubic polynomial

$$\begin{aligned}\operatorname{Res}(f;x) &:= x^3 + 2\tilde{p}x^2 + (\tilde{p}^2 - 4\tilde{r})x - \tilde{q}^2 \\ &= x^3 + \frac{P}{4\,a_0^2}x^2 + \frac{P^2 - R}{64\,a_0^4}x - \frac{Q^2}{64\,a_0^6},\end{aligned} \tag{9.28}$$

see e.g. [12]. The numerator $P^2 - R$ is a seminvariant of degree and weight 4, and we have by (9.25)

$$\frac{P^2 - R}{4} = \frac{P^2 - 16\,a_0^2\,I}{3} = -64\,a_0^3\,a_4 + 16\,a_0^2\,a_1\,a_3 + 16\,a_0^2\,a_2^2 - 16\,a_0\,a_1^2\,a_2 + 3\,a_1^4. \tag{9.29}$$

This seminvariant is used in [3, 4] with the notations

$$Q_{[3]} = q_{[4]} := \frac{1}{3}\big(P^2 - 16\,a_0^2\,I\big) = \frac{P^2 - R}{4}. \tag{9.30}$$

The reduced form of the cubic resolvent is, after some calculations,

$$\widehat{\operatorname{Res}}(f;x) := \operatorname{Res}\Big(f; x - \frac{P}{12\,a_0^2}\Big) = x^3 - \frac{I}{3\,a_0^2}\,x + \frac{J}{27\,a_0^3}. \tag{9.31}$$

Changing the variable to clear the denominators, we find

$$(3\,a_0)^3\,\widehat{\operatorname{Res}}(f; x/3a_0) = 27\,a_0^3\,\operatorname{Res}\Big(f; \frac{4a_0\,x - P}{12\,a_0^2}\Big) = x^3 - 3I\,x + J. \tag{9.32}$$

Thus the cubic polynomial $\widetilde{\operatorname{Res}}(f;x) := x^3 - 3I\,x + J$ is also a form of the resolvent.

Remark 9.5. The roots of the cubic resolvent $\operatorname{Res}(f)$ are γ_1^2, γ_2^2, γ_3^2, where

$$\gamma_1 := \tfrac{1}{2}(\xi_1 + \xi_2 - \xi_3 - \xi_4), \tag{9.33}$$
$$\gamma_2 := \tfrac{1}{2}(\xi_1 - \xi_2 + \xi_3 - \xi_4), \tag{9.34}$$
$$\gamma_3 := \tfrac{1}{2}(\xi_1 - \xi_2 - \xi_3 + \xi_4), \tag{9.35}$$

The quartic equation $f(x) = 0$ can thus be solved by finding the roots of $\operatorname{Res}(f)$, taking the square roots to find $\gamma_1, \gamma_2, \gamma_3$, with the signs satisfying $\gamma_1\gamma_2\gamma_3 = -\tilde{q}$, and finally inverting (9.33)–(9.35) together with $\xi_1 + \xi_2 + \xi_3 + \xi_4 = -\tilde{p}$, see [12]. Alternatively, one can first find the roots of $\widehat{\operatorname{Res}}(f)$

or $\widetilde{\operatorname{Res}}(f)$; for example, if the roots of $\widetilde{\operatorname{Res}}(f)$ are z_1, z_2, z_3, we take $\gamma_i = \pm\frac12 a_0^{-1}\sqrt{(4a_0z_i - P)/3}$. Equivalently, the roots of $\widetilde{\operatorname{Res}}(f)$ are

$$z_i := 3a_0\gamma_i^2 + \frac{P}{4a_0} = 3a_0\gamma_i^2 + 2p, \qquad i = 1,2,3, \tag{9.36}$$

while the roots of $\widehat{\operatorname{Res}}(f)$ are

$$\frac{z_i}{3a_0} = \gamma_i^2 + \frac{P}{12\,a_0^2} = \gamma_i^2 + \frac{2p}{3a_0}, \qquad i = 1,2,3. \tag{9.37}$$

Remark 9.6. Another common version of the cubic resolvent is (see [12])

$$\begin{aligned} \operatorname{Res}^*(f;x) &:= \operatorname{Res}\Bigl(f; x - \tilde p - \frac{a_1^2}{8a_0^2}\Bigr) = \operatorname{Res}\Bigl(f; x - \frac{4a_0a_2 - a_1^2}{4a_0^2}\Bigr) \\ &= x^3 - \frac{a_2}{a_0}x^2 + \frac{a_1a_3 - 4a_0a_2}{a_0^2}x + \frac{4a_0a_2a_4 - a_0a_3^2 - a_1^2a_4}{a_0^3}. \end{aligned} \tag{9.38}$$

This has the roots $\xi_1\xi_2 + \xi_3\xi_4$, $\xi_1\xi_3 + \xi_2\xi_4$ and $\xi_1\xi_4 + \xi_2\xi_3$. However, these roots are *not* translation invariant, so the coefficients of Res^* are *not* seminvariants.

We have, by (9.31) and (8.5), (9.7) and (4.3),

$$\Delta_{\langle 3\rangle}(\operatorname{Res}(f)) = \Delta_{\langle 3\rangle}(\widehat{\operatorname{Res}}(f)) = \frac{4I^3}{27a_0^6} - \frac{J^2}{27a_0^6} = a_0^{-6}\Delta = \Delta_0. \tag{9.39}$$

Further, by (9.31) and (8.6) or (8.8),

$$P_{\langle 3\rangle}(\operatorname{Res}(f)) = P_{\langle 3\rangle}(\widehat{\operatorname{Res}}(f)) = -a_0^{-2}I, \tag{9.40}$$

$$Q_{\langle 3\rangle}(\operatorname{Res}(f)) = Q_{\langle 3\rangle}(\widehat{\operatorname{Res}}(f)) = a_0^{-3}J. \tag{9.41}$$

For the version $\widetilde{\operatorname{Res}}(f) = x^3 - 3I\,x + J$ we have, directly from (8.5)–(8.9), the corresponding

$$\Delta_{\langle 3\rangle}(\widetilde{\operatorname{Res}}(f)) = 4\cdot 27\,I^3 - 27\,J^2 = 3^6\Delta, \tag{9.42}$$

$$P_{\langle 3\rangle}(\widetilde{\operatorname{Res}}(f)) = -9\,I, \tag{9.43}$$

$$Q_{\langle 3\rangle}(\widetilde{\operatorname{Res}}(f)) = 27\,J. \tag{9.44}$$

9.6. **Seminvariants of f'.** We calculate the basic seminvariants of the cubic polynomial f':

$$\begin{aligned} \Delta_{\langle 3\rangle}(f') &= -432\,a_0^2\,a_3^2 + 432\,a_0\,a_1\,a_2\,a_3 - 128\,a_0\,a_2^3 - 108\,a_1^3\,a_3 + 36\,a_1^2\,a_2^2 \\ &= 16\,a_0\,J - 12\,IP, \end{aligned} \tag{9.45}$$

$$P_{\langle 3\rangle}(f') = 24\,a_0\,a_2 - 9\,a_1^2 = 3P, \tag{9.46}$$

$$Q_{\langle 3\rangle}(f') = 432\,a_0^2\,a_3 - 216\,a_0\,a_1\,a_2 + 54\,a_1^3 = 54\,Q. \tag{9.47}$$

9.7. **The case $a_0 = 0$.** When $a_0 = 0$, i.e., considering the restriction to polynomials of degree 3, we have, cf. Example 4.8, for any polynomial $f \in \mathcal{P}_3$,

$$\Delta_{\langle 4\rangle} = a_{0\,\langle 3\rangle}^2 \Delta_{\langle 3\rangle}, \tag{9.48}$$

$$I_{\langle 4\rangle} = -P_{\langle 3\rangle}, \tag{9.49}$$

$$J_{\langle 4\rangle} = -Q_{\langle 3\rangle}, \tag{9.50}$$

$$P_{\langle 4\rangle} = -3\,a_{0\,\langle 3\rangle}^2, \tag{9.51}$$

$$Q_{\langle 4\rangle} = a_{0\,\langle 3\rangle}^3, \tag{9.52}$$

In particular, for $f \in \mathcal{P}_4$, using (9.45)–(9.47),

$$\Delta_{\langle 4\rangle}(f') = 16\,a_0^2\,\Delta_{\langle 3\rangle}(f') = 256\,a_0^3\,J - 192\,a_0^2\,I\,P, \tag{9.53}$$

$$I_{\langle 4\rangle}(f') = -3\,P, \tag{9.54}$$

$$J_{\langle 4\rangle}(f') = -54\,Q, \tag{9.55}$$

$$P_{\langle 4\rangle}(f') = -48\,a_0^2, \tag{9.56}$$

$$Q_{\langle 4\rangle}(f') = 64\,a_0^3. \tag{9.57}$$

9.8. **Seminvariants and roots.** By Example 4.2,

$$\Delta = a_0^6(\xi_1-\xi_2)^2(\xi_1-\xi_3)^2(\xi_1-\xi_4)^2(\xi_2-\xi_3)^2(\xi_2-\xi_4)^2(\xi_3-\xi_4)^2. \tag{9.58}$$

For I and J we obtain by calculations, using $\sum^*$ to denote a sum over different indices, where moreover identical terms are counted only once (thus, for example, $\sum_{i,j}^* \xi_i\xi_j = \sum_{i<j}\xi_i\xi_j$),

$$I = a_0^2\Bigl(\sum_{i,j}^{*}\xi_i^2\xi_j^2 - \sum_{i,j,k}^{*}\xi_i^2\xi_j\xi_k + 6\xi_1\xi_2\xi_3\xi_4\Bigr) \tag{9.59}$$

where the first sum has 6 terms and the second 12, and

$$J = a_0^3\Bigl(-2\sum_{i,j}^{*}\xi_i^3\xi_j^3 + 3\sum_{i,j,k}^{*}\xi_i^3\xi_j^2\xi_k - 12\sum_{i,j,k,l}^{*}\xi_i^3\xi_j\xi_k\xi_l - 12\sum_{i,j,k}^{*}\xi_i^2\xi_j^2\xi_k^2 + 6\sum_{i,j,k,l}^{*}\xi_i^2\xi_j^2\xi_k\xi_l\Bigr), \tag{9.60}$$

where the sums have 6, 24, 4, 4 and 6 terms.

For the seminvariants we have first, by (4.4),

$$H_0 = -3\,a_0^2\sum_{1\le i<j\le 4}(\xi_i-\xi_j)^2 = -3\,a_0^2\Bigl(3\sum_{i=1}^{4}\xi_i^2 - 2\sum_{1\le i<j\le 4}\xi_i\xi_j\Bigr) \tag{9.61}$$

and thus, by (9.26),

$$P = -\,a_0^2\sum_{1\le i<j\le 4}(\xi_i-\xi_j)^2 = -a_0^2\Bigl(3\sum_{i}\xi_i^2 - 2\sum_{i,j}^{*}\xi_i\xi_j\Bigr) \tag{9.62}$$

where the sums have 4 and 6 terms. Further, by calculation,

$$Q = -a_0^3\Bigl(\sum_i \xi_i^3 - \sum_{i,j}^{*} \xi_i^2\xi_j + 2\sum_{i,j,k}^{*} \xi_i\xi_j\xi_k\Bigr) \tag{9.63}$$

where the sums have 4, 12 and 4 terms.

The formulas (9.59) and (9.62) for I and P cannot be factorized further, but for J and Q we have

$$\begin{aligned} J = -a_0^3\bigl((\xi_1-\xi_3)(\xi_2-\xi_4)+(\xi_1-\xi_4)(\xi_2-\xi_3)\bigr) \\ \cdot\bigl((\xi_1-\xi_2)(\xi_3-\xi_4)+(\xi_1-\xi_4)(\xi_3-\xi_2)\bigr) \\ \cdot\bigl((\xi_1-\xi_2)(\xi_4-\xi_3)+(\xi_1-\xi_3)(\xi_4-\xi_2)\bigr) \end{aligned} \tag{9.64}$$

and

$$Q = -a_0^3(\xi_1+\xi_2-\xi_3-\xi_4)(\xi_1-\xi_2+\xi_3-\xi_4)(\xi_1-\xi_2-\xi_3+\xi_4). \tag{9.65}$$

Explicit formulas for the covariants H and G_6 in terms of the roots can be obtained from Example 4.15 and Theorem 4.13 together with (9.63) or (9.65). We leave these to the reader.

9.9. **Cross ratio.** Let F be a field and $F^* := F \cup \{\infty\}$. The *cross ratio* $[x_1, x_2; x_3, x_4]$ is defined for $x_1, x_2, x_3, x_4 \in F^*$ by

$$[x_1, x_2; x_3, x_4] := \frac{(x_1-x_3)(x_2-x_4)}{(x_1-x_4)(x_2-x_3)} \in F^*. \tag{9.66}$$

More precisely, the cross ratio is well-defined by (9.66) if $x_1, x_2, x_3, x_4 \in F$ are distinct, and more generally if $x_1, x_2, x_3, x_4 \in F^*$ are distinct with the natural interpretations

$$\begin{aligned} [\infty, x_2; x_3, x_4] &= \frac{x_2-x_4}{x_2-x_3}, & [x_1, \infty; x_3, x_4] &= \frac{x_1-x_3}{x_1-x_4}, \\ [x_1, x_2; \infty, x_4] &= \frac{x_2-x_4}{x_1-x_4}, & [x_1, x_2; x_3, \infty] &= \frac{x_1-x_3}{x_2-x_3}. \end{aligned} \tag{9.67}$$

Furthermore, the cross ratio is also defined when two of $x_1, \dots, x_4 \in F^*$ coincide, and even when two different pairs of them coincide. (In these cases, the cross ratio is always 0, 1 or ∞; it is 0 if $x_1 = x_3$ or $x_2 = x_4$, 1 if $x_1 = x_2$ or $x_3 = x_4$, and ∞ if $x_1 = x_4$ or $x_2 = x_3$.) In the remaining cases, when three or four of $x_1, \dots, x_4$ coincide, the cross ratio is undefined.

If $x_2, x_3, x_4 \in F^*$ are distinct, then $x \mapsto [x, x_2; x_3, x_4]$ is the unique projective (= fractional linear) map $F^* \to F^*$ that maps $x_2 \mapsto 1$, $x_3 \mapsto 0$, $x_4 \mapsto \infty$.

The cross ratio depends on the order of $x_1, \dots, x_4$, and the 24 different permutations give, in general, 6 different values. These values determine each other; if $[x_1, x_2; x_3, x_4] = \lambda$, then,

$$[x_1, x_2; x_3, x_4] = [x_2, x_1; x_4, x_3] = [x_3, x_4; x_1, x_2] = [x_4, x_3; x_2, x_1] = \lambda,$$

$$[x_1, x_2; x_4, x_3] = [x_2, x_1; x_3, x_4] = [x_3, x_4; x_2, x_1] = [x_4, x_3; x_1, x_2] = \frac{1}{\lambda},$$

$$[x_1, x_3; x_2, x_4] = [x_2, x_4; x_1, x_3] = [x_3, x_1; x_4, x_2] = [x_4, x_2; x_3, x_1] = 1 - \lambda,$$
$$[x_1, x_3; x_4, x_2] = [x_2, x_4; x_3, x_1] = [x_3, x_1; x_2, x_4] = [x_4, x_2; x_1, x_3] = \frac{1}{1-\lambda},$$
$$[x_1, x_4; x_2, x_3] = [x_2, x_3; x_1, x_4] = [x_3, x_2; x_4, x_1] = [x_4, x_1; x_3, x_2] = \frac{\lambda - 1}{\lambda},$$
$$[x_1, x_4; x_3, x_2] = [x_2, x_3; x_4, x_1] = [x_3, x_2; x_1, x_4] = [x_4, x_1; x_2, x_3] = \frac{\lambda}{\lambda - 1}.$$

The symmetric group S_4 thus acts on the space F^*. The functions of λ above are all projective maps, and thus we have a homomorphism of S_4 into the group $PGL(1, F)$ of projective maps; the kernel is the four-group and the image is a subgroup of $PGL(1, F)$ of order 6, isomorphic to S_3 (for example, by their permutations of $\{0, 1, \infty\}$). The orbits have in general 6 elements, but orbits including a fixpoint of one of the non-trivial maps above are smaller; there are two or three such exceptional orbits, viz. $\{0, 1, \infty\}$, $\{-1, \frac{1}{2}, 2\}$, and, provided $\sqrt{-3} \in F$, $\{\frac{1}{2} \pm \frac{\sqrt{-3}}{2}\}$.

We have $[x_1, x_2; x_3, x_4] \in \{0, 1, \infty\}$ if and only if two of $x_1, \dots, x_4$ coincide.

Quadruples $x_1, \dots, x_4$ with $[x_1, x_2; x_3, x_4] \in \{-1, \frac{1}{2}, 2\}$ are called *harmonic quadruples*. (For example, one point at infinity and three points in an aritmetic sequence, such as $-1, 0, 1, \infty$. Another example is four points equally spaced on a circle, such as $1, i, -1, -i$.)

Quadruples $x_1, \dots, x_4$ with $[x_1, x_2; x_3, x_4] \in \{\frac{1}{2} \pm \frac{\sqrt{-3}}{2}\}$ are called *self-apolar* or *equianharmonic*. (For example, three points evenly spaced on a circle, together with either the centre or infinity, such as $0, 1, e^{2\pi i/3}, e^{4\pi i/3}$.)

If $f \in \mathcal{P}_4$, let $\xi_1, \dots, \xi_4$ be its roots, and $\lambda := [\xi_1, \xi_2; \xi_3, \xi_4]$. Then λ depends on the ordering of the roots, as explained above, but the polynomial

$$\Lambda(z) := (z - \lambda)\Big(z - \frac{1}{\lambda}\Big)\Big(z - (1-\lambda)\Big)\Big(z - \frac{1}{1-\lambda}\Big)\Big(z - \frac{\lambda}{\lambda-1}\Big)\Big(z - \frac{\lambda-1}{\lambda}\Big) \tag{9.68}$$

does not depend on the order, so it depends on f only. The coefficients of $\Lambda(z)$ are symmetric rational functions of $\xi_1, \dots, \xi_4$, and are thus rational functions of the coefficients $a_0, \dots, a_4$ of f. Moreover, $\Lambda(z)$ is invariant under projective transformations, and is thus an absolute invariant of f. A calculation yields, using (9.7),

$$\begin{aligned}\Lambda(z) &= z^6 - 3\,z^5 - \frac{I^3 + 2\,J^2}{9\,\Delta} z^4 + \frac{26\,I^3 + 7\,J^2}{27\,\Delta} z^3 - \frac{I^3 + 2\,J^2}{9\,\Delta} z^2 - 3\,z + 1 \\ &= z^6 - 3\,z^5 - \frac{3\,I^3 + 6\,J^2}{4\,I^3 - J^2} z^4 + \frac{26\,I^3 + 7\,J^2}{4\,I^3 - J^2} z^3 - \frac{3\,I^3 + 6\,J^2}{4\,I^3 - J^2} z^2 - 3\,z + 1,\end{aligned} \tag{9.69}$$

where we recognize (slightly disguised) the absolute invariant I^3/J^2, see Example 9.2. We have $\Lambda(\lambda) = 0$, i.e.,

$$\left(4I^3 - J^2\right)\lambda^6 + \left(-12I^3 + 3J^2\right)\lambda^5 + \left(-3I^3 - 6J^2\right)\lambda^4$$

$$
\begin{aligned}
&+ \big(26I^3 + 7J^2\big)\lambda^3 + \big(-3I^3 - 6J^2\big)\lambda^2 + \big(-12I^3 + 3J^2\big)\lambda + 4I^3 - J^2 \\
&= \big(4\lambda^6 - 12\lambda^5 - 3\lambda^4 + 26\lambda^3 - 3\lambda^2 - 12\lambda + 4\big)I^3 \\
&\qquad\qquad - \big(\lambda^6 - 3\lambda^5 + 6\lambda^4 - 7\lambda^3 + 6\lambda^2 - 3\lambda + 1\big)J^2 \\
&= 0, \qquad (9.70)
\end{aligned}
$$

which after a rearrangement yields

$$
\begin{aligned}
\frac{J^2}{I^3} &= \frac{4\,\lambda^6 - 12\,\lambda^5 - 3\,\lambda^4 + 26\,\lambda^3 - 3\,\lambda^2 - 12\,\lambda + 4}{\lambda^6 - 3\,\lambda^5 + 6\,\lambda^4 - 7\,\lambda^3 + 6\,\lambda^2 - 3\,\lambda + 1} \\
&= \frac{(\lambda-2)^2\,(2\,\lambda-1)^2\,(\lambda+1)^2}{(\lambda^2-\lambda+1)^3} \qquad (9.71)
\end{aligned}
$$

and equivalently, using (9.7) again,

$$
\frac{I^3}{\Delta} = \frac{\lambda^6 - 3\,\lambda^5 + 6\,\lambda^4 - 7\,\lambda^3 + 6\,\lambda^2 - 3\,\lambda + 1}{\lambda^4 - 2\,\lambda^3 + \lambda^2} = \frac{\big(\lambda^2-\lambda+1\big)^3}{\lambda^2\,(\lambda-1)^2} \qquad (9.72)
$$

and

$$
\frac{J^2}{\Delta} = \frac{4\,\lambda^6 - 12\,\lambda^5 - 3\,\lambda^4 + 26\,\lambda^3 - 3\,\lambda^2 - 12\,\lambda + 4}{\lambda^4 - 2\,\lambda^3 + \lambda^2} = \frac{(\lambda-2)^2\,(2\,\lambda-1)^2\,(\lambda+1)^2}{\lambda^2\,(\lambda-1)^2}. \qquad (9.73)
$$

We have really proved these formulas for the case of four distinct roots in F, but it is easy to see that they hold also in the case of one or two double roots (in this case $\Delta = 0$ and $\lambda \in \{0, 1, \infty\}$), and (by projective invariance) also if there is a single or double root at ∞. Note that if two of I, J and Δ vanish, then so do all three because of (9.7); this happens if and only there is a triple (or quadruple) root (see Theorem 5.4), and then cross ratio λ is undefined. In this case thus both sides of (9.71)–(9.73) are undefined. Otherwise, if there is no triple root, at most one of I, J and Δ vanishes, and both sides of (9.71)–(9.73) are defined as elements of F^* (they may be ∞, viz. when the denominator vanishes or, for (9.72)–(9.73), when $\lambda = \infty$), and they are equal.

9.10. **Further examples.** A simple example of higher invariants is $A(f^\nu, f^\nu)$ for $\nu \geq 1$. This has degree 2ν and weight 4ν. We have $A(f, f) = 4I$ by (9.2) and, for example,

$$
\begin{aligned}
A_{\langle 8\rangle}(f^2, f^2) &= 82944\,a_0^2\,a_4^2 - 41472\,a_0\,a_1\,a_3\,a_4 + 13824\,a_0\,a_2^2\,a_4 \\
&\qquad + 5184\,a_1^2\,a_3^2 - 3456\,a_1\,a_2^2\,a_3 + 576\,a_2^4 \\
&= 576\,I^2, \qquad (9.74) \\
A_{\langle 12\rangle}(f^3, f^3) &= 564480\,I^3 - 11520\,J^2. \qquad (9.75)
\end{aligned}
$$

The apolar invariant of the Hessian covariant is an invariant given by, see (9.8) and (2.10),

$$\begin{aligned} A_{\langle 4\rangle}(H(f),H(f)) &= 82944\, a_0^2\, a_4^2 - 41472\, a_0\, a_1\, a_3\, a_4 + 13824\, a_0\, a_2^2\, a_4 \\ &\qquad + 5184\, a_1^2\, a_3^2 - 3456\, a_1\, a_2^2\, a_3 + 576\, a_2^4 \\ &= 576\, I^2. \end{aligned} \tag{9.76}$$

Equivalently, $A(\widetilde{H}(f),\widetilde{H}(f)) = 64\, I^2$.

Similarly, omitting the details,

$$A(\widetilde{H}(f)^2,\widetilde{H}(f)^2) = 147456\, I^4, \tag{9.77}$$

$$\begin{aligned} A(\widetilde{H}(f)^3,\widetilde{H}(f)^3) &= 2123366400\, I^6 + 188743680\, I^3 J^2 - 47185920\, J^4 \\ &= 47185920(5\, I^3 + J^2)(9\, I^3 - J^2), \end{aligned} \tag{9.78}$$

with the coefficients $147456 = 2^{14}3^2$ and $47185920 = 2^{20}3^2 5$.

Equivalently,

$$A(H(f)^2,H(f)^2) = 11943936\, I^4, \tag{9.79}$$

$$A(H(f)^3,H(f)^3) = 34398535680\,(5\, I^3 + J^2)(9\, I^3 - J^2), \tag{9.80}$$

where $11943936 = 2^{14}3^6$ and $34398535680 = 2^{20}3^8 5$.

The (joint) apolar invariant $A(H(f),f)$ is an invariant given by, see (9.8) and (2.10),

$$\begin{aligned} A(H(f),f) &= 1728\, a_0\, a_2\, a_4 - 648\, a_0\, a_3^2 - 648\, a_1^2\, a_4 + 216\, a_1\, a_2\, a_3 - 48\, a_2^3 \\ &= 24J, \end{aligned} \tag{9.81}$$

see (9.5). This has degree 3 and weight 6. Equivalently, $A(\widetilde{H}(f),f) = 8J$. We can also form, for example,

$$A_{\langle 8\rangle}\bigl(f\widetilde{H}(f), f\widetilde{H}(f)\bigr) = 192\,(20\, I^3 + 7J^2). \tag{9.82}$$

Further invariants (etc.) of the Hessian covariant are

$$I(\widetilde{H}(f)) = 16\, I^2, \tag{9.83}$$

$$J(\widetilde{H}(f)) = 64\, J^2 - 128\, I^3 = 64\bigl(J^2 - 2\, I^3\bigr), \tag{9.84}$$

$$\Delta(\widetilde{H}(f)) = 2^{12} J^2\, \Delta, \tag{9.85}$$

$$P(\widetilde{H}(f)) = 64\, a_0\, J - 16\, I\, P, \tag{9.86}$$

$$Q(\widetilde{H}(f)) = -64\, J\, Q, \tag{9.87}$$

$$\widetilde{H}(\widetilde{H}(f)) = 64\, J\, f - 16\, I\, \widetilde{H}, \tag{9.88}$$

$$\widetilde{G_6}(\widetilde{H}(f)) = -64\, J\, \widetilde{G_6}. \tag{9.89}$$

The apolar invariant $A(\widetilde{G_6},\widetilde{G_6})$ of the sextic polynomial $\widetilde{G_6}(f)$ is an invariant of degree 6 and weight 12 given by

$$A_{\langle 6\rangle}(\widetilde{G_6}(f),\widetilde{G_6}(f)) = 960\, \Delta. \tag{9.90}$$

The discriminant $\Delta(\widetilde{G_6})$ is an invariant of degree 30 and weight 60 given by

$$\Delta_{\langle 6\rangle}(\widetilde{G_6}(f)) = -2^{18}\,\Delta^5. \tag{9.91}$$

We calculate also the resultants of f, $\widetilde{H}(f)$ and $\widetilde{G_6}(f)$:

$$R(f,\widetilde{H}) = 81\Delta^2, \tag{9.92}$$

$$R(f,\widetilde{G_6}) = \Delta^3, \tag{9.93}$$

$$R(\widetilde{H},\widetilde{G_6}) = 2^{12}\Delta^3 J^2, \tag{9.94}$$

where the first also follows by Example 4.11.

For the seminvariants in Examples 4.5–4.6, we have, recalling $\widehat{a}_2 = p = P/8a_0$, $\widehat{a}_3 = q = Q/8a_0^2$ and $\widehat{a}_4 = r = R/256a_0^3$, see (9.14) and (9.19)–(9.21), and using also (9.25),

$$a_0^2 S_2 = -\frac{1}{4}P, \tag{9.95}$$

$$a_0^3 S_3 = -\frac{3}{8}Q, \tag{9.96}$$

$$a_0^4 S_4 = -\frac{1}{64}R + \frac{1}{32}P^2 = -\frac{1}{3}a_0^2 I + \frac{7}{192}P^2, \tag{9.97}$$

and

$$a_0^2 \chi_2 = -\frac{1}{16}P, \tag{9.98}$$

$$a_0^3 \chi_3 = -\frac{3}{32}Q, \tag{9.99}$$

$$a_0^4 \chi_4 = -\frac{1}{256}R - \frac{1}{256}P^2 = -\frac{1}{384}\big(32a_0^2 I + P^2\big). \tag{9.100}$$

9.11. **Vanishing invariants and covariants.** Let f be a quartic polynomial, with roots $\xi_1,\dots,\xi_4$. (The results extend to the case $a_0 = 0$ when one or several roots are ∞ with no or trivial modifications.)

Since I and J form a basis for the invariants (Theorem 9.1), Theorem 5.4 shows that $I = J = 0$ if and only if f has a triple root. Moreover, since $\Delta = \frac{4}{27}I^3 - \frac{1}{27}J^2$ by (9.7), we have the following:

Theorem 9.7. *Let f be a quartic polynomial. If f has a triple (or quadruple) root then $\Delta = I = J = 0$.*

Conversely, if there is no triple root, then at most one of Δ, I and J vanishes.

If there is no triple root, the cross ratio $[\xi_1,\xi_2;\xi_3,\xi_4]$ of the roots is well-defined by Subsection 9.9, and the vanishing of the basis invariants I and J, as well as Δ, can be characterised by this cross ratio.

Theorem 9.8. *Let f be a quartic polynomial with roots ξ_1,ξ_2,ξ_3,ξ_4, and assume that there is no triple (or quadruple) root.*

(i) $I = 0$ *if and only if the cross-ratio* $[\xi_1, \xi_2; \xi_3, \xi_4] = \frac{1}{2} \pm \frac{\sqrt{3}}{2}\mathrm{i}$, *i.e., if and only if the roots form a equianharmonic (self-apolar) quadruple.*
(ii) $J = 0$ *if and only if the cross-ratio* $[\xi_1, \xi_2; \xi_3, \xi_4] \in \{-1, \frac{1}{2}, 2\}$, *i.e., if and only if the roots form a harmonic quadruple.*
(iii) $\Delta = 0$ *if and only if the cross-ratio* $[\xi_1, \xi_2; \xi_3, \xi_4] \in \{0, 1, \infty\}$, *i.e., if and only there is a double root.*

Note that the three conditions use the three exceptional orbits of cross ratios, see Subsection 9.9.

Proof. When there is no triple root, the cross ratio $\lambda = [\xi_1, \xi_2; \xi_3, \xi_4]$ is well-defined by Subsection 9.9 and at most one of I, J and Δ vanishes by Theorem 9.7; the results now follow from (9.71)–(9.73). (The result for Δ is of course an immediate consequence of (5.1).) □

Theorem 9.8(ii) also follows from (9.64), which shows that $J = 0$ if and only if one of the three factors in the brackets there vanishes, or equivalently that one of the three cross-ratios

$$\frac{(\xi_1 - \xi_3)(\xi_2 - \xi_4)}{(\xi_1 - \xi_4)(\xi_2 - \xi_3)}, \qquad \frac{(\xi_1 - \xi_2)(\xi_3 - \xi_4)}{(\xi_1 - \xi_4)(\xi_3 - \xi_2)}, \qquad \frac{(\xi_1 - \xi_2)(\xi_4 - \xi_3)}{(\xi_1 - \xi_3)(\xi_4 - \xi_2)}$$

equals -1. (These are $[\xi_1, \xi_2; \xi_3, \xi_4]$, $[\xi_1, \xi_3; \xi_2, \xi_4]$ and $[\xi_1, \xi_4; \xi_2, \xi_3]$.)

Corollary 5.9 gives another interpretation of $J = 0$, since J is a multiple of G_2 (the catalecticant when $n = 4$):

Theorem 9.9. *The following are equivalent for a quartic polynomial* f*:*
(i) $J = 0$.
(ii) f *belongs to the closure* $\overline{\mathcal{P}}_{4,2}$ *of the set* $\mathcal{P}_{4,2} := \{c_1(x - x_1)^2 + c_2(x - x_2)^4\}$.
(iii) f *has one of the forms* $c_1(x - x_1)^4 + c_2(x - x_2)^4$, $c_1(x - x_1)^4 + c_2$, $c_1(x - x_1)^4 + c_2(x - x_1)^3$, $c_1 + c_2 x$. *(The last two comprise the cases when* f *has a triple root, finite or infinite).*

For the covariants H and G_6 we have the following. The first is just an instance of the general Theorem 5.2.

Theorem 9.10. *The following are equivalent for a quartic polynomial* f*.*
(i) $H(f) = 0$.
(ii) f *has a single, quadruple root, i.e.,* $\xi_1 = \xi_2 = \xi_3 = \xi_4$.
(iii) $f(x) = c(x - x_0)^4$ *for some* c *and* x_0.

Theorem 9.11. *The following are equivalent for a quartic polynomial* f*.*
(i) $G_6(f) = 0$.
(ii) *Every root is (at least) a double root.*
(iii) *The roots coincide in two pairs* $\xi_1 = \xi_2$ *and* $\xi_3 = \xi_4$ *(up to labelling); this includes the case when all four roots coincide.*
(iv) $f = cg^2$ *for some quadratic polynomial* g.

Proof. It is easy to see that (ii), (iii) and (iv) are equivalent.

Suppose now (i), i.e., $G_6(f) = 0$. By (9.90), then $\Delta = 0$, so f has a double root ξ. By projective invariance, we may assume that $\xi = 0$, so $f(x) = a_0x^4 + a_1x^3 + a_2x^2$. For f of this form, with $a_3 = a_4 = 0$, (9.10) reduces to

$$\begin{aligned}\widetilde{G_6}(f) &= \bigl(-4\,a_0\,a_1\,a_2 + a_1^3\bigr)\,x^6 + \bigl(-8\,a_0\,a_2^2 + 2\,a_1^2\,a_2\bigr)\,x^5 \\ &= \bigl(a_1^2 - 4\,a_0\,a_2\bigr)\bigl(a_1x^6 + 2\,a_2x^5\bigr). \end{aligned} \tag{9.101}$$

Hence either $a_1^2 - 4a_0a_2 = 0$ or $a_1 = a_2 = 0$; in both cases $\Delta_{\langle 2\rangle}(a_0x^2 + a_1x + a_2) = a_1^2 - 4a_0a_2 = 0$. Hence, $a_0x^2 + a_1x + a_2$ has a double root ξ, and f has the roots $0, 0, \xi, \xi$.

Conversely, if f has only double roots, we may again by projective invariance assume that 0 is a root, and then $f(x) = a_0x^4 + a_1x^3 + a_2x^2$, where we now know that also $a_0x^2 + a_1x + a_2$ has a double root, and thus its discriminant $a_1^2 - 4a_0a_2 = 0$. Hence, $G_6(f) = 0$ by (9.101). □

9.12. **Roots and resolvent of a real quartic.** Consider a real quartic f, with $a_0 \neq 0$. Then f has either 0, 2 or 4 real roots (counted with multiplicities). The discriminant partly discriminates between these cases, by the following simple and classic result, which is a simple consequence of (4.3), see Remark 8.4. (In this subsection, "complex" means non-real.)

Theorem 9.12. *Let f be a real quartic polynomial.*

(i) $\Delta(f) > 0 \iff$ *f has either 4 distinct real roots,* or *4 complex roots in two conjugate pairs.*
(ii) $\Delta(f) < 0 \iff$ *f has 2 real roots and 2 conjugate complex roots.*
(iii) $\Delta(f) = 0 \iff$ *f has a double (or triple or quadruple) root. In this case, f has 1 quadruple real root,* or *2 real roots, one triple and one single,* or *2 double real roots,* or *3 real roots, one double and two single,* or *1 double real root and 2 conjugate complex roots,* or *2 conjugate complex double roots.*

To completely distinguish between the different cases we employ further seminvariants and covariants. (In the following theorem the roots are assumed to be distinct except as explicitly stated.) Note that $I \geq 0$ when $\Delta \geq 0$ by (9.7), so $\sqrt{I} \geq 0$ in this case.

Theorem 9.13. *Let f be a real quartic polynomial.*

(i) *f has 4 real roots* $\iff \Delta > 0$, *$P \leq 0$ and $P^2 - 16a_0^2I \geq 0$* $\iff \Delta > 0$ *and $P \leq -4a_0\sqrt{I}$.*
(ii) *f has 2 pairs of conjugate complex roots* $\iff \Delta > 0$ *and either $P > 0$ or $P^2 - 16a_0^2I < 0$* $\iff \Delta > 0$ *and $P > -4a_0\sqrt{I}$.*
(iii) *f has 2 real roots and 2 conjugate complex roots* $\iff \Delta < 0$.
(iv) *f has 1 quadruple real root* $\iff \Delta = I = J = P = 0 \iff H(x) \equiv 0$. *In this case also $Q = 0$ and $G_6(x) \equiv 0$.*

(v) *f has 1 triple and 1 single real root $\iff \Delta = I = J = 0$ but $P \neq 0$. In this case $P < 0$, $Q \neq 0$, $H(x) \not\equiv 0$, $G_6(x) \not\equiv 0$.*

(vi) *f has 2 double real roots $\iff \Delta = P^2 - 16a_0^2 I = 0$ and $P < 0$ $\iff G_6(x) \equiv 0$ and $P < 0$. In this case also $Q = 0$.*

(vii) *f has 2 conjugate complex double roots $\iff \Delta = P^2 - 16a_0^2 I = Q = 0$ and $P > 0 \iff G_6(x) \equiv 0$ and $P > 0$.*

(viii) *f has 3 real roots, one double and two single $\iff \Delta = 0$, $I > 0$, $P < 0$ and $P^2 - 16a_0^2 I > 0$.*

(ix) *f has 1 double real root and 2 conjugate complex roots $\iff \Delta = 0$ and either $P^2 - 16a_0^2 I < 0$ or $P > 0$ but not $P^2 - 16a_0^2 I = Q = 0$.*

Proof. (i),(ii): By Theorem 9.12, these cases are characterized by $\Delta > 0$. To distinguish the two cases, we note that by Remark 9.5, f has 4 real roots $\iff \gamma_1, \gamma_2, \gamma_3 \in \mathbb{R} \iff \gamma_1^2, \gamma_2^2, \gamma_3^2 \in [0, \infty)$. Since $\gamma_1^2, \gamma_2^2, \gamma_3^2$ are the roots of the cubic resolvent $\operatorname{Res}(f)$, it follows from Theorem 8.5 and (9.28) that f has 4 real roots $\iff \Delta(\operatorname{Res}(f)) \geq 0$, $P \leq 0$, $P^2 - R \geq 0$, and $-Q^2 \leq 0$. Since $\Delta(\operatorname{Res}(f)) = a_0^{-6}\Delta$ by (9.39) and $P^2 - R = \frac{4}{3}(P^2 - 16a_0^2 I)$ by (9.29), the results follow.

(iii): By Theorem 9.12.

In the remaining cases f has a multiple root and $\Delta = 0$. Note that then $4I^3 = J^2$ by (9.7); in particular, $I \geq 0$. We calculate the seminvariants and covariants by (9.3), (9.5), (9.22), (9.23), (9.8), (9.10) in the different cases to verify the direct parts of the assertions:

(iv): We may by invariance assume $f = a_0 x^4$, and then $I = J = P = Q = H(x) = G_6(x) = 0$.

(v): We may by invariance assume $f = a_0 x^3 (x - u)$ where $u \in \mathbb{R}$ with $u \neq 0$, and then $I = J = 0$, $P = -3a_0^2 u^2 < 0$, $Q = -a_0^3 u^3 \neq 0$, $H(x) = -9a_0^2 u^2 x^2$, $G_6(x) = -36 a_0^3 u^3 x^6$.

(vi): We may by invariance assume $f = a_0 x^2 (x - u)^2$ where $u \in \mathbb{R}$ with $u \neq 0$, and then $I = a_0 u^2$, $P = -4a_0^2 u^2 < 0$, $P^2 - 16a_0^2 I = 0$. Further, $G_6(x) = 0$ and thus $Q = 0$ by Theorem 9.11.

(vii): We may by invariance assume $f = a_0 (x - u - \mathrm{i}v)^2 (x - u + \mathrm{i}v)^2$ for some real u and $v \neq 0$, and then $I = 16a_0^2 v^4$, $P = 16a_0^2 v^2 > 0$, $P^2 - 16a_0^2 I = 0$, Further, $G_6(x) = 0$ and thus $Q = 0$ by Theorem 9.11.

(viii): We may by invariance assume $f = a_0 x^2 (x - u)(x - v)$ for some real $u, v \neq 0$, and then $I = a_0^2 u^2 v^2 > 0$, $P = -a_0^2 (3u^2 + 3v^2 - 2uv) = -a_0^2 (2u^2 + 2v^2 + (u - v)^2) < 0$, $P^2 - 16a_0^2 I = 3a_0^4 (u - v)^2 (3u^2 + 3v^2 + 2uv) > 0$.

(ix): We may by invariance assume $f = a_0 x^2 (x - u - \mathrm{i}v)^2 (x - u + \mathrm{i}v)^2$ for some real u and $v \neq 0$, and then $I = a_0^2 (u^2 + v^2)^2$, $P = 4a_0^2 (2v^2 - u^2)$, $P = 8a_0^3 v^2 u$, $P^2 - 16a_0^2 I = 48 a_0^4 v^2 (v^2 - 2u^2)$. If $v^2 \geq u^2$, then $P > 0$, and if $v^2 < u^2$, then $P^2 - 16a_0^2 I < 0$. Further, $P^2 - 16a_0^2 I = 0 \iff v^2 = 2u^2$, and $Q = 0 \iff u = 0$, which cannot hold simultaneously.

The converse implications in (iv),(v) now follow by Theorem 9.7 and Theorem 9.10.

If $G_6(x) \equiv 0$, then we have (iv), (vi) or (vii) by Theorem 9.11, and they are by the calculations above distinguished by the sign of P, which shows the converse implications assuming $G_6(x) \equiv 0$.

It is easily verified that the other conclusions in (vi)–(ix) are mutually exclusive, and also exclusive of (iv)–(v). Hence the converse implications follow. □

The proof used some properties of the cubic resolvent. Let us study its geometry further. The cubic resolvent $\mathrm{Res}(f)$ of the real quartic f has by (9.28), (9.31) and (8.50)–(8.51) an inflection point at

$$(x_0, y_0) = \Bigl(\frac{-P}{12a_0^2}, \frac{J}{27a_0^3}\Bigr) \tag{9.102}$$

and, by (8.53)–(8.54) and (9.40)–(9.41), extreme points at

$$(x_\pm, y_\pm) = \Bigl(\frac{-P \pm 4a_0\sqrt{I}}{12a_0^2}, \frac{J \mp 2I^{3/2}}{27a_0^3}\Bigr). \tag{9.103}$$

For the version $\widetilde{\mathrm{Res}}(f)$ in (9.32) we have simpler formulas: an inflection point at

$$(\widetilde{x}_0, \widetilde{y}_0) = \bigl(0, J\bigr) \tag{9.104}$$

and extreme points at

$$(\widetilde{x}_\pm, \widetilde{y}_\pm) = \bigl(\pm\sqrt{I}, J \mp 2I^{3/2}\bigr). \tag{9.105}$$

By (9.103) or (9.105), the resolvent has two distinct real extreme points if and only if $I > 0$, while the resolvent is strictly increasing if $I \le 0$, cf. Subsection 8.11 and (9.40), (9.43). We further see again that the resolvent has three distinct real roots if and only if $I > 0$ and $J - 2I^{3/2} < 0 < J + 2I^{3/2}$, or, equivalently, if and only if $4I^3 > J^2$, i.e., if and only if $\Delta = \frac{1}{27}(4I^3 - J^2) > 0$, cf. Theorem 8.3 and (9.39), (9.42).

Further, using Remark 9.5 and (9.103), f has 4 distinct real roots

$\iff$ $\mathrm{Res}(f)$ has 3 roots in $[0, \infty)$

$\iff$ $x_+ > x_- \ge 0$ and $y_- > 0 > y_+$

$\iff$ $I > 0$, $-P \ge 4a_0\sqrt{I}$ and $J < 2I^{3/2}$,

which by (9.7) yields another proof of Theorem 9.13(i).

Finally we note that, by (9.103) and (9.29),

$$x_+x_- = \frac{P^2 - 16a_0^2 I}{144a_0^4} = \frac{P^2 - R}{192a_0^4} \tag{9.106}$$

and, using also (9.7) again,

$$y_+y_- = \frac{J^2 - 4I^3}{729a_0^6} = -\frac{\Delta}{27a_0^3}; \tag{9.107}$$

hence, as observed by Nickalls [15], the seminvariants $P^2 - 16a_0^2 I$ and $J^2 - 4I^3 = -\Delta/27$ (ignoring normalizations) play a symmetric role in the geometry of the cubic resolvent. Recall from Theorem 9.13 that these (together with P) are the most important seminvariants when determining the number of real roots, at least when the roots are simple.

References

[1] A. Abdesselam & J. Chipalkatti, The higher transvectants are redundant. *Ann. Inst. Fourier (Grenoble)* **59** (2009), no. 5, 1671–1713.

[2] G. E. Andrews, *The Theory of Partitions.* Addison-Wesley, 1976. (Reprinted, Cambridge Univ. Press, Cambridge, 1984, 1998.)

[3] J. E. Cremona, Reduction of binary cubic and quartic forms. *LMS J. Comput. Math.* **2** (1999), 64–94.

[4] J. E. Cremona, Classical invariants and 2-descent on elliptic curves. *J. Symbolic Comput.* **31** (2001), no. 1–2, 71–87.

[5] L. E. Dickson, General theory of modular invariants. *Trans. Amer. Math. Soc.* **10** (1909), no. 2, 123–158.

[6] L. E. Dickson, *On Invariants and the Theory of Numbers.* The Madison Colloquium (1913), Vol. IV, pp. 1–110, Amer. Math. Soc., New York, 1914. Reprinted by Dover, New York, 1966.

[7] E. B. Elliott, *An Introduction to the Algebra of Quantics.* 2nd ed, Oxford University Press, Oxford, 1913.

[8] Oliver E. Glenn, *A Treatise on the Theory of Invariants.* Ginn, Boston, 1915.

[9] S. Gundelfinger, Zur theorie der binären Formen. *J. Reine Angew. Math.* **100** (1886), 413–424.

[10] D. Hilbert, *Theory of Algebraic Invariants.* Cambridge University Press, Cambridge, 1993.

[11] S. Janson, Resultant and discriminant of polynomials. Note N5, 2007. `http://www.math.uu.se/~svante/papers/#NOTES`

[12] S. Janson, Roots of polynomials of degrees 3 and 4. Resultant and discriminant of polynomials. Note N7, 2009. `arXiv:1009.2373v1.`

[13] J. P. S. Kung, Gundelfinger's theorem on binary forms. *Stud. Appl. Math.* **75** (1986), no. 2, 163–169.

[14] J. P. S. Kung & G.-C. Rota, The invariant theory of binary forms. *Bull. Amer. Math. Soc. (N.S.)* **10** (1984), no. 1, 27–85.

[15] R. W. D. Nickalls, The quartic equation: invariants and Euler's solution revealed. *Math. Gazette* **93** (2009), no. 526, 66–75.

[16] P. J. Olver, *Classical Invariant Theory.* Cambridge University Press, Cambridge, 1999.

[17] A. Salden, Euclidean invariants of linear scale-spaces. *Computer Vision ACCV'98.* pp. 65–72, Eds. Roland Chin and Ting-Chuen Po, Lecture Notes in Computer Science 1352, 1997.

[18] I. Schur, *Vorlesungen über Invariantentheorie.* Edited by Helmut Grunsky. Springer-Verlag, Berlin, 1968.

[19] W. L. G. Williams, Fundamental systems of formal modular seminvariants of the binary cubic. *Trans. Amer. Math. Soc.* **22** (1921), no. 1, 56–79.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

E-mail address: svante.janson@math.uu.se

URL: http://www.math.uu.se/∼svante/