AN ARITHMETIC APPLICATION OF INVARIANT THEORY

W. DUKE

Abstract. The classical invariant theory of binary forms and pairs of binary forms is applied to a problem about the representation of integers by certain binary quartic forms.

1. Introduction

A venerable problem in number theory is to describe the integers that are represented by an integral binary form. Much more is known here in the quadratic case than in higher degree cases, as is reflected in the well-known relationship between binary quadratic forms and (proper) ideals in orders of quadratic fields. A general tool that may be applied to this problem is the invariant theory of forms and pairs of forms. In this paper I will provide an illustration where invariant theory yields a result about binary quartic forms that is parallel to the following classical result about binary quadratic forms.

Theorem 1. Suppose that $F(x, y) = ax^2 + bxy + cy^2$ is a primitive binary quadratic form with integer coefficients and non-zero discriminant $D_F = b^2 - 4ac$. Then there exists a primitive binary quadratic form $G$ with $D_G = D_F$ that properly represents the square of each integer that is properly represented by $F$ and is prime to $D_F$.

Observe that the restriction to proper representations eliminates some obvious trivialities and that we are allowing forms that are reducible over $\mathbb{Q}$ with distinct roots. The simplest example is when

$$F(x, y) = x^2 + y^2$$

with $D_F = -4$, where $G = F$ and the result follows easily from the duplication formula

$$\left(x^2 - y^2\right)^2 + (2xy)^2 = (x^2 + y^2)^2.$$  

Note that $x^2 + y^2$ does not represent 4 properly although $2 = 1^2 + 1^2$; the theorem is not true without the assumption that the represented integer be prime to the discriminant.

Consider now a binary quartic form $F$ with integral coefficients that is even in that it has the form

$$F(x, y) = ax^4 + 2bx^3y + cx^2y^2 + 2dxy^3 + ey^4,$$

for integers $a, b, c, d, e$ and is primitive in the sense that $\gcd(a, b, c, d, e) = 1$. The set of primitive even $F$ is preserved under the usual action of $\Gamma = \text{SL}(2, \mathbb{Z})$ and there are two independent integral invariants, namely

$$A_F = 12ae - 12bd + c^2 \quad \text{and} \quad B_F = 36ace + 18bcd - 54b^2e - 54d^2a - c^3.$$  

The discriminant of $F$ is given by

$$D_F = \frac{1}{27}(A_F^3 - B_F^2) = (bcd)^2 + \cdots.$$ 

Given such a form with non-zero discriminant, it is not in general possible to find another binary quartic form $G$, even one with a different (non-zero) discriminant, that properly

Research supported by NSF grant DMS 1701638 and Simons Foundation Award Number 554649.
represents the square of every integer that is properly represented by $F$ and is prime to $D_F$. This can be seen already when
\[ F(x, y) = x^4 + y^4, \]
which has $D_F = 1$. It is known (see [5, Theorem 1]) that the number of integers $\leq T^2$ properly represented by $F$ is $\gg T^2$ for $T \geq 1$. Suppose that there were a $G$ that represented properly the square of every integer $\leq T^2$ properly represented by $F$. Then $G$ would properly represent $\gg T^2$ squares less than $T$, i.e. the equation
\[ z^2 = G(x, y) \]
would have $\gg T^2$ solutions with $\gcd(x, y) = 1$ and $|z| \leq T$. This leads to a contradiction since by Mordell’s theorem applied to the Jacobian of the genus one curve defined over $\mathbb{Q}$ by $y^2z^2 = G(x, y)$, that number is $\ll \epsilon T^2$, for any $\epsilon > 0$.

Our main result shows that a weaker statement does hold for certain $F$.

**Theorem 2.** Suppose that $F$ is a primitive even binary quartic form with $A_F = P^2$ for some $P \in \mathbb{Z}$, where $A_F$ was defined in (1.3), and with a non-zero discriminant $D_F$. Then there exists a primitive even binary quartic form $G$ that properly represents some multiple of the square of each integer that is properly represented by $F$ and is prime to $D_F$. In addition, $D_G = m^2D_F$ for some $m \in \mathbb{Z}$ with $4|m$.

We will give proofs of Theorems 1 and 2 that are completely analogous. The proof of Theorem 1 uses what is essentially Gaussian duplication while that of Theorem 2 uses a kind of higher composition law where an even binary quartic is composed with itself and with another even binary quartic, a covariant. These composition laws are given in Propositions 1 and 2.

A simple example to illustrate Theorem 2 is given by
\[ F(x, y) = x^4 + 2x^3y + 2xy^3 + y^4, \]
which has $A_F = 0$ and $D_F = -108$. The proof of Theorem 2 shows that we may take
\[ G(x, y) = x^4 + 6x^2y^2 - 3y^4, \]
for which $A_G = 0$ and $D_G = -8^2 \cdot 108$. As particular instances, we have
\[ F(1, 2) = 37 \quad \text{and} \quad G(18, 7) = 141 \cdot 37^2, \]
\[ F(2, 3) = 253 \quad \text{and} \quad G(90, 19) = 1293 \cdot 253^2. \]

We cannot apply Theorem 2 directly to study the even binary quartic
\[ (ax^2 + bxy + cy^2)^2 = a^2x^4 + 2abx^3y + (b^2 + 2ac)x^2y^2 + 2bcxy^3 + c^2y^4, \]
for which $A_F = (b^2 - 4ac)^2$, since this quartic has zero discriminant. However, a modification of the proof of Theorem 2 applied to this quartic shows that in Theorem 1 we can replace the word “square” by “cube”, provided we assume that the discriminant of the quadratic form is square-free. This is possible since the covariant that occurs in the composition law for quartics becomes proportional to the quartic, which results in the cube after taking a square root. In this way we get a proof of this well-known result that does not use the class group, at least directly. This is shown below after the proof of Theorem 2.

Of course, much more can be said about such questions in the quadratic case by using the above-mentioned correspondence between classes of forms and ideal classes and applying algebraic number theory. My purpose here is to illustrate the use of tools that are also applicable to basic representation problems by higher degree binary forms.
2. Preliminaries

First we set notation, make some definitions and record some facts from the invariant theory of binary forms and pairs of binary forms. Let

\[ F(x, y) = \sum_{n=0}^{m} a_n x^{m-n} y^n \]

where \( a_0, \ldots, a_m \in \mathbb{Z} \) be an integral binary form of degree \( m \). We will use the abbreviation

\[ F(x, y) = [a_0, \ldots, a_m]. \]

Such an \( F \) is acted on by \( g = (a \ b; c \ d) \in \Gamma = \text{SL}(2, \mathbb{Z}) \) through

\[ f \mapsto F|g = F(ax + by, cx + dy). \]

For each pair of non-negative integers \( m, n \) fix \( c_{m,n} \in \mathbb{Z}^+ \) such that \( c_{m,0} = 1 \) and \( c_{m,n} = c_{m,m-n} \). Define \( F \) to be the set of all \( F \) that can be written in the form

\[ F(x, y) = \sum_{n=0}^{m} c_{m,n} a_n x^{m-n} y^n \quad a_0, \ldots, a_m \in \mathbb{Z}. \]

Say that \( F \) is admissible if it is preserved under (2.2) and if, for \( F \in \mathcal{F} \) as given in (2.3), \( \gcd(a_0, \ldots, a_m) \) is left invariant under (2.2). In addition to the set \( \mathcal{F}_1 \) of ordinary integral forms where \( c_{m,n} = 1 \) for all \( m, n \), another admissible set is the set of Gaussian forms \( \mathcal{F}_3 \) with \( c_{m,n} = \binom{m}{n} \). We abbreviate a Gaussian form by using open parentheses, e.g.

\[ ax^2 + 2xy + cy^2 = (a, b, c). \]

The set \( \mathcal{F}_2 \) of even forms, with \( c_{m,1} = c_{m,-1} = 2 \) for \( m \) even and \( c_{m,n} = 1 \) otherwise, is also admissible. Note that admissibility is easily checked by expanding \( F(x + y, y) \) and \( F(-y, x) \). Clearly \( \mathcal{F}_3 \subset \mathcal{F}_2 \subset \mathcal{F}_1 \). For an admissible \( \mathcal{F} \) the form \( F \in \mathcal{F} \) given by (2.3) is called primitive (for \( \mathcal{F} \)) if \( \gcd(a_0, \ldots, a_m) = 1 \) and properly primitive if

\[ \gcd(c_{m,0} a_0, \ldots, c_{m,m} a_m) = 1. \]

An (arithmetic) covariant for \( F \in \mathcal{F}_1 \) is a binary form \( P_F \) whose coefficients are integral polynomials of the coefficients of \( F \) that satisfies

\[ P_F|g = P_{F|g} \]

for all \( g \in \Gamma \). If \( P_F|g = P_F \) for all \( g \in \Gamma \) then \( P_F \) gives an invariant of \( F \). The discriminant \( D_F = \text{disc}_{x,y}(F) \) is an invariant of \( F \). The Hessian of \( F \)

\[ H_F(x, y) = \det \begin{pmatrix} F_{xx} & F_{xy} \\ F_{yx} & F_{yy} \end{pmatrix} \]

is a covariant of \( F \). The Jacobian of two forms \( F, G \) is given by

\[ J_{F,G}(x, y) := \det \begin{pmatrix} F_{x} & F_{y} \\ G_{x} & G_{y} \end{pmatrix}. \]

We have that \( J_{F,H_F}(x, y) \) is a covariant of \( F \). The discriminant form of two forms \( F, G \) of the same degree is given by

\[ D_{F,G}(x, y) = \text{disc}_{u,v}(xF(u, v) + yG(u, v)). \]

If \( F \) and \( H_F \) have the same degree, \( D_{F,H_F}(x, y) \) is an invariant of \( F \).

A pair \((f_1, f_2)\) of binary forms of the same degree, each in some \( \mathcal{F} \), is acted on in two ways by \( g \in \Gamma \). One action is through

\[ (f_1, f_2) \mapsto (f_1, f_2)|g = (f_1|g, f_2|g) \]
and the second through
\[(f_1, f_2) \mapsto (f_1, f_2)g = (af_1 + cf_2, bf_1 + df_2).\]

Invariants and covariants of pairs of forms with a given action are defined similarly to the case of a single form.

The Jacobian \(J_{f_1, f_2}(x, y)\) is an invariant of \((f_1, f_2)\) with respect to (2.7) and a covariant of \((f_1, f_2)\) with respect to (2.6). The discriminant form \(D_{f_1, f_2}(x, y)\) is an invariant of \((f_1, f_2)\) with respect to (2.6) and a covariant of \((f_1, f_2)\) with respect to (2.7).

### 3. Pairs of binary quadratic forms

The proof of Theorem 1 makes use of the invariant theory of pairs of binary (Gaussian) quadratic forms, whose theory we will now review. For \(f_1 = (a_1, b_1, c_1)\) and \(f_2 = (a_2, b_2, c_2)\) we have
\[(3.1) \quad D_{f_1, f_2} = [b_1^2 - a_1c_1, 2b_1b_2 - a_2c_2 - a_2c_1, b_2^2 - a_2c_2].\]

Set
\[(3.2) \quad F(x, y) = \frac{1}{4}J_{f_1, f_2}(x, y) = [(a_1, b_2), (a_1, c_2), (b_1, c_2)],\]
where we are using the convenient notation from [9]:
\[(3.3) \quad (a_i, b_j) = a_ib_j - a_jb_i.\]

A direct calculation verifies the following identity (“syzygy”):

**Proposition 1.** For any pair \((f_1, f_2)\) with discriminant form \(D_{f_1, f_2}\) we have the identity
\[(3.4) \quad D_{f_1, f_2}(-f_2(x, y), f_1(x, y)) = F^2(x, y).\]

**Proof of Theorem 1.** Now we may give a proof of Theorem 1, one that serves as a model for our proof of Theorem 2.

Suppose that \(F(x, y) = ax^2 + bxy + cy^2\) is a primitive binary quadratic form with integer coefficients and non-zero discriminant \(D_F\). First we show that \(F\) may be represented as the (normalized) Jacobian of two binary quadratic forms of Gaussian type. This is an immediate consequence of [6, art. 279]. An easy way to see this directly is to complete \((a, b, c)\) to a matrix \(A \in \text{SL}(3, \mathbb{Z})\), which is possible since \(\gcd(a, b, c) = 1\), and extract the coefficients of \(f_1 = (a_1, b_1, c_1)\) and \(f_2 = (a_2, b_2, c_2)\) from the entries of \(A^{-1}\).

Now we claim that we may take
\[G(x, y) = D_{f_1, f_2}(x, y) = (a_1c_1 - b_1^2)x^2 + (a_1c_2 + c_1a_2 - 2b_1b_2)xy + (a_2c_2 - b_2^2)y^2.\]

The discriminant of \(G(x, y)\) is given by
\[D_G = (a_1c_2 + c_1a_2 - 2b_1b_2)^2 - 4(a_1c_1 - b_1^2)(a_2c_2 - b_2^2)\]
\[= (a_1c_2)^2 - 4(a_1b_2)(b_1c_2) = D_F\]
by (3.2). It follows easily that \(G\) is primitive.

The resultant of \(f_1\) and \(f_2\) is given by
\[\text{result}(f_1, f_2) = (a_1c_2)^2 - 4(a_1b_2)(b_1c_2) = D_F.\]

Thus for a prime \(p\), if \(p|f_1(x, y)\) and \(p|f_2(x, y)\) for \(\gcd(x, y) = 1\) then \(p|D_F\). Combining this with (3.4) proves Theorem 1. \(\square\)
4. Pairs of binary cubic forms

The proof of Theorem 2 makes use of the invariant theory of pairs of binary (Gaussian) cubic forms. We review this next. A good reference is [9, p.204–218].

A pair \((f_1, f_2)\) of Gaussian binary cubic forms has the discriminant form

\[ D_{f_1,f_2}(x,y) = [\delta_0, \delta_1, \delta_2, \delta_3, \delta_4], \]

an invariant binary quartic under \((2.6)\). For \(f_1 = (a_1, b_1, c_1, d_1)\) and \(f_2 = (a_2, b_2, c_2, d_2)\) we have

\[ \delta_0 = D_{f_1} = -a_1^2d_1^2 + 6a_1b_1c_1d_1 + 3b_1^2c_1^2 - 4a_1c_1^3 - 4d_1b_1^3, \quad \delta_4 = D_{f_2} \]

and

\[ \delta_1 = 6b_1b_2c_1^2 - 4a_2c_1^2 + 6b_1^2c_1c_2 - 12a_1c_1^2c_2 - 12b_1^2b_2d_1 + 6a_2b_1c_1d_1 + 6a_1b_2c_1d_1 + 6a_1b_1c_2d_1 - 2a_1^2d_1^2 - 6a_1b_1c_1d_2 - 2a_1^2d_1^2 \]

\[ \delta_2 = 3b_1^2c_1^2 + 12b_1b_2c_1c_2 - 12a_2c_1^2c_2 + 6b_1^2c_1c_2 - 12a_1c_1^2c_2 - 12b_1^2b_2d_1 + 6a_2b_1c_1d_1 + 6a_1b_2c_1d_1 + 6a_1b_1c_2d_1 - 2a_1^2d_1^2 - 12b_1^2d_2d_1 + 6a_2b_1c_1d_2 + 6a_1b_2c_1d_2 - 2a_1a_2d_1d_2 - 2a_1^2d_1^2 \]

\[ \delta_3 = 6b_2^2c_1c_2 + 6b_1b_2c_2^2 - 12a_2c_1c_2^2 - 4a_1c_1^2c_2 - 4b_2^2d_1 + 6a_2b_2c_2d_1 - 12b_1^2d_2d_1 + 6a_2b_2c_1d_2 + 6a_1b_2c_2d_2 - 2a_2^2d_2d_2 - 2a_1a_2d_2^2. \]

A full set of (scalar) invariants of the pair \((f_1, f_2)\) is provided by \(\delta_0, \delta_1, \delta_2, \delta_3, \delta_4\) together with \(P\) and \(Q\), which are given by (recall the convention (3.3))

\[ P = P_{f_1,f_2} = (a_1d_2) - 3(b_1c_2) \]

\[ Q = Q_{f_1,f_2} = -(b_1c_2)^3 - (c_1a_2)^2(c_1d_2) - (a_1b_2)(b_1d_2)^2 + (b_1c_2)^2(a_1d_2) + 3(a_1b_2)(b_1c_2)(c_1d_2) + (a_1d_2)(a_1b_2)(c_1d_2). \]

Note that \(D_{f_1,f_2}(x,y)\) is even. It follows by direct calculation\(^1\) that for \(A_D, B_D\) from (1.3) we have

\[ A_D = P(P^3 - 24Q) \quad \text{and} \quad B_D = P^6 - 36P^3Q + 216Q^2. \]

The normalized Jacobian of \((f_1, f_2)\) is the even form given by

\[ \frac{1}{9} J_{f_1,f_2}(x,y) = [(a_1 b_2), (2a_1 c_2), (a_1 d_2) + 3(b_1 c_2), 2(b_1 d_2), (c_1 d_2)]. \]

Write

\[ F = \frac{1}{9} J_{f_1,f_2} = [a, 2b, c, 2d, e] \]

and let

\[ \frac{1}{12} H_F(x,y) = [2ac - 3b^2, 12ad - 2bc, 12ae + 6bd - c^2, 12be - 2cd, 2ce - 3d^2] \]

be its Hessian. By another direct calculation we have

\[ A_F = P^2 \quad \text{and} \quad B_F = 54Q - P^3. \]

Define the covariant quartic

\[ K = K_{f_1,f_2} = -2PF - H_F = (3a_1^2c_2^2 + \cdots)x^4 + \cdots. \]

Once more, a calculation justifies the next result.

\(^1\)Use a computer or, as is done in [9, p. 211], proceed by hand after reducing to canonical forms.
**Proposition 2.** Notation as above, for any pair \((f_1, f_2)\) of binary cubic forms with discriminant form \(D_{f_1, f_2}\), we have the syzygy

\[
D_{f_1, f_2}(-f_2(x, y), f_1(x, y)) = K(x, y)F^2(x, y),
\]

where \(F\) is defined below (4.5) and \(K\) is defined in (4.8).

**Proof of Theorem 2.** For

\[
F(x, y) = ax^4 + 2bx^3y + cx^2y^2 + 2dxy^3 + ey^4,
\]

where \(A_F = P^2\), we will next show that there exist \(f_1 = (a_1, b_1, c_1, d_1)\) and \(f_2 = (a_2, b_2, c_2, d_2)\) such that

\[
F(x, y) = \frac{1}{5} J_{f_1, f_2}(x, y).
\]

Since by (1.3) we have

\[
P^2 = 12ae - 12bd + c^2,
\]

we may choose the sign of \(P\) so that \(P \equiv c \pmod{6}\). We must show that \((a_1, b_1, c_1, d_1)\) and \((a_2, b_2, c_2, d_2)\) exist so that

\[
a = (a_1 b_2), \quad b = (a_1 c_2), \quad \frac{1}{2}(c + P) = (a_1 d_2), \quad \frac{1}{2}(c - P) = (b_1 c_2), \quad d = (b_1 d_2), \quad e = (c_1 d_2).
\]

We have the necessary condition that \(ae - bd + \frac{1}{12}(c^2 - P^2) = 0\) from (4.10). As before, the needed existence is due to Gauss, who proved a similar result in his treatment of the composition of two binary quadratic forms. The exact statement we need was given by Arndt [1] and a simple proof is given in [3, p. 65]. Again a matrix-theoretic proof may be given.

As in the case of two binary quadratic forms, we take

\[
G(x, y) = D_{f_1, f_2}(x, y).
\]

Now the resultant of \(f_1\) and \(f_2\) is by [9, p. 205] or a direct calculation

\[
\text{result}(f_1, f_2) = P^3 - 27Q.
\]

On the other hand, by (1.4) and (4.7)

\[
D_F = Q(P^3 - 27Q).
\]

Finally, from (4.4) we have

\[
D_G = (4Q)^2 D_F.
\]

Thus Theorem 2 follows from Proposition 2, after possibly redefining \(G\) to make it primitive. \(\square\)

**Cubes represented by a binary quadratic form.** Now we justify the claim made after the statement of Theorem 2. For

\[
F(x, y) = (ax^2 + bxy + cy^2)^2 = [a^2, 2ab, b^2 + 2ac, 2bc, c^2],
\]

we have from (1.3) that

\[
A_F = (b^2 - 4ac)^2, \quad B_F = -(b^2 - 4ac)^3.
\]

Now (4.7), upon choosing the sign of \(P\) so that \(P \equiv b^2 + 2ac \pmod{6}\) via (4.13), yields

\[
P = b^2 - 4ac, \quad Q = 0.
\]
Thus from (4.11) we have
\[ \text{result}(f_1, f_2) = (b^2 - 4ac)^3. \]
A calculation using (4.6) shows that
\[ \frac{1}{12} H_F(x, y) = -(b^2 - 4ac)(ax^2 + bxy + cy^2)^2. \]
Now from (4.8) and (4.15) we get
\[ K(x, y) = -(b^2 - 4ac)(ax^2 + bxy + cy^2)^2. \]
Proposition 2 gives the formula
\[ D_{f_1, f_2}(-f_2(x, y), f_1(x, y)) = -(b^2 - 4ac)(ax^2 + bxy + cy^2)^6. \]
Since \( D_F = 0 \) this implies that we must have a quadratic form \( Q(x, y) = a'x^2 + b'xy + c'y^2 \) with integers \( a', b', c' \) such that \( \gcd(a', b', c') = 1 \) and an integer \( \kappa \) with
\[ (4.16) \quad \kappa Q(-f_2(x, y), f_1(x, y)) = (b^2 - 4ac)(ax^2 + bxy + cy^2)^3. \]
Now
\[ -(b^2 - 4ac)D_{f_1, f_2} = \kappa^2 Q^2 \]
so from (4.4) and (4.14)
\[ (4.17) \quad (b^2 - 4ac)^2 A_D = (b^2 - 4ac)^6 = \kappa^4 (b'^2 - 4a'c')^2. \]
If \( b^2 - 4ac \) is square-free we must have \( \kappa = b^2 - 4ac \) from (4.16) and so by (4.17)
\[ b^2 - 4ac = \pm (b'^2 - 4a'c'). \]
But \( b^2 - 4ac \) square-free implies that the plus sign must hold.

\[ \square \]

**Remark.** It is interesting to apply the identity of Proposition 2 to the pair of binary cubic forms \((f_1, f_2)\), where the first form \( f_1 = (a, b, c, d) \) has non-zero discriminant \( D = D_{f_1} \) (as computed in (4.1)), but otherwise is arbitrary and the second \( f_2 \) is the cubic covariant of \( f_1 \). Explicitly,
\[ f_2 = (-2b^3 + 3abc - a^2d, -b^2c + 2ac^2 - abd, bc^2 - 2b^2d + acd, 2c^3 - 3bcd + ad^2). \]
The Hessian of \( f_1 \) is
\[ H_{f_1} = [ac - b^2, ad - bc, bd - c^2]. \]
The discriminant of \( H_{f_1} \) is \(-D\). A calculation shows that
\[ (4.18) \quad D_{f_1, f_2}(x, y) = DQ^2(x, y), \]
where \( Q(x, y) = x^2 + Dy^2 \). The Jacobian of \((f_1, f_2)\) is
\[ \frac{1}{b} J_{f_1, f_2} = -2H^2_{f_1}. \]
Now Proposition 2 and (4.18) give, after computing the Hessian of \( H_{f_1}^2 \), the identity
\[ Q(-f_2(x, y), f_1(x, y))^2 = 16H_{f_1}(x, y)^6. \]
By combining this with a single evaluation, we see that the identity of Proposition 2 reduces to the classical identity of Eisenstein [4, §5, eq. 1] and Cayley [2]:
\[ f_2^2 + Df_2^2 = -4H^3_{f_1}. \]
This identity was used by Mordell [7] [8] to characterize all of the proper representations of cubes by the principal form when \( 4|D \).
REFERENCES


UCLA Mathematics Department, Box 951555, Los Angeles, CA 90095-1555

Email address: wdduke@ucla.edu