ON THE DISCRIMINANT OF BINARY FORMS

MARKUS ROST

Contents

1. Introduction

Let

$$f(x,y) = a_0x^d + a_1x^{d-1}y + \cdots + a_dy^{d-1}$$

be a homogeneous form of degree d in 2 variables. The discriminant of f is a homogeneous polynomial of degree $2d-2$ in the coefficients a_i.

In this text we describe a presentation of the discriminant as the determinant of a $(2d-2)\times(2d-2)$-matrix which is linear in the coefficients of f. There are no denominators and the method works over any ring in a coordinate free way. The construction has a natural explanation in terms of jet spaces.

In brief, the discriminant of f is the determinant of the equation

$$af_x + bf_y - (a_x + b_y)f = 0$$

or, in a more compact form,

$$(a/f)_x + (b/f)_y = 0$$

where a, b are homogeneous forms of degree $d-2$.

As is well known, the determinant of

$$af_x + bf_y = 0$$

is the resultant of the derivatives f_x, f_y which is d^{d-2} times the discriminant. The extra term $-(a_x + b_y)f$ eliminates the factor d^{d-2}, so to speak.

Date: December 5, 2018.

2. Preliminaries and Examples

We use the notations disc(f) and

$$\Delta(f) = (-1)^{d(d-1)/2} \operatorname{disc}(f)$$

for the two common definitions of the discriminant (see [1, A IV.76], [2]). They are characterized by

$$\operatorname{disc}(f) = \prod_{i \neq j} (\alpha_i - \alpha_j)$$
$$\Delta(f) = \prod_{i<j} (\alpha_i - \alpha_j)^2$$

for a polynomial of the form $f(x,y) = \prod_{i=1}^{d} (x - \alpha_i y)$.

For $d = 2$, 3 one has

$$\Delta(f) = a_1^2 - 4a_0a_2$$
$$\Delta(f) = a_1^2 a_2^2 - 4a_0a_2^3 - 4a_3a_1^3 - 27a_0^2a_3^2 + 18a_0a_1a_2a_3$$

respectively, and $\operatorname{disc}(f) = -\Delta(f)$ in both cases.

A standard method to describe (or define) disc(f) is to consider the resultants of f and its derivatives. One has

$$\operatorname{disc}(f) = \frac{1}{a_0} \operatorname{Res}(f, f_x) = \frac{1}{d^{d-2}} \operatorname{Res}(f_x, f_y)$$

and the resultant may be computed with the Sylvester matrix [1, A IV.71].

For instance computing via $\operatorname{Res}(f, f_x)$ yields for $d = 2$

$$\operatorname{disc}(f) = \frac{1}{a_0} |f, xf_x, yf_x| = \frac{1}{a_0} \begin{vmatrix} a_0 & 2a_0 & 0 \\ a_1 & a_1 & 2a_0 \\ a_2 & 0 & a_1 \end{vmatrix} = 4a_0a_2 - a_1^2$$

and computing via $\operatorname{Res}(f_x, f_y)$ yields for $d = 3$

$$\operatorname{disc}(f) = \frac{1}{3} |xf_x, yf_x, xf_y, yf_y| = \frac{1}{3} \begin{vmatrix} 3a_0 & 0 & a_1 & 0 \\ 2a_1 & 3a_0 & 2a_2 & a_1 \\ a_2 & 2a_1 & 3a_3 & 2a_2 \\ 0 & a_2 & 0 & 3a_3 \end{vmatrix}$$

There are several methods to get rid of the denominators. A basic remark is Eulers relation

$$xf_x + yf_y = df$$

As for $\frac{1}{a_0} \operatorname{Res}(f, f_x)$, one first subtracts d-times the first column $x^{d-2}f$ from the d-th column $x^{d-1}f_x$. Then the first row has just a_0 in the first place and one drops the first line and column. This results in

$$\operatorname{disc}(f) = \left| x^{d-3}f, \dots, y^{d-3}f, -x^{d-2}f_y, x^{d-2}f_x, \dots, y^{d-2}f_x \right|$$

For $d = 3$ this gives

$$\operatorname{disc}(f) = |f, -xf_y, xf_x, yf_x|$$

To eliminate the factor d^{d-2} from $\mathrm{Res}(f_x, f_y)$ one may rearrange the matrix so that one can divide by d on a $(d-2)$-dimensional subspace.

Let us first consider the example $d = 3$. To get rid of the factor 3 one uses again Eulers relation

$$xf_x + yf_y = 3f$$

It shows that one may simply replace the first or the last column by f and drop the denominator 3. This results for instance in

$$\mathrm{disc}(f) = |xf_x, yf_x, xf_y, f|$$

However, to get more symmetry it is better to replace the columns xf_x, yf_y by

$$u = xf_x - f, \quad v = yf_y - f$$

Indeed, since $u + v = f$, one has

$$\begin{aligned} xf_x &= 2u + v \\ yf_y &= v + 2u \end{aligned}$$

and the factor $\left|\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right| = 3$ gets eliminated. Thus

$$\mathrm{disc}(f) = |xf_x - f, yf_x, xf_y, yf_y - f| = \begin{vmatrix} 2a_0 & 0 & a_1 & -a_0 \\ a_1 & 3a_0 & 2a_2 & 0 \\ 0 & 2a_1 & 3a_3 & a_2 \\ -a_3 & a_2 & 0 & 2a_3 \end{vmatrix}$$

3. Description of the discriminant as determinant

The last example generalizes to any degree d as follows. Let S^k denote the space of homogeneous polynomials of degree k in x, y. One has $\mathrm{rank}\, S^k = k + 1$. Given $f \in S^d$, consider the map

$$\begin{aligned} \Phi_f \colon S^{d-2} \oplus S^{d-2} &\to S^{2d-3} \\ \Phi_f(a, b) &= af_x + bf_y - (a_x + b_y)f \end{aligned}$$

Since both spaces have the same dimension, one can take the determinant. If one chooses the respective basis over the ground ring **Z**, that determinant is well-defined up to sign. With the basis as in [1, A IV.71] one gets

$$\mathrm{disc}(f) = \det(\Phi_f) \tag{1}$$

Without the extra term $-(a_x + b_y)f$, the map Φ_f would give the usual Sylvester matrix computing the resultant $\mathrm{Res}(f_x, f_y)$ (as above for $d = 3$).

We will prove (1) in two ways. A first step is to define Φ_f in a coordinate free way.

Let V be a locally free module of rank 2 and let

$$\begin{aligned} D \colon S^k V &\to V \otimes S^{k-1} V \\ D(v_1 \cdots v_k) &= \sum_i v_i \otimes v_1 \cdots v_{i-1} v_{i+1} \cdots v_k \end{aligned}$$

be the derivative. For $f \in S^d V$ the map Φ_f reads as

$$\begin{aligned} \Phi_f \colon V^* \otimes S^{d-2} V &\to S^{2d-3} V \\ \Phi_f(\varphi \otimes g) &= (Df)(\varphi) \cdot g - (Dg)(\varphi) \cdot f \end{aligned}$$

It follows in particular that $f \to \Phi_f$ is SL(2)-invariant. In this setting, (1) reads as

$$\det(\Phi_{v_1\cdots v_d}) = \prod_{i\neq j}(v_i \wedge v_j) \in (\Lambda^2 V)^{d(d-1)}$$

Note further that

$$\Phi_f(a,b) = (af_x - a_x f) + (bf_y - b_y f)$$

Hence a compact way to write Φ_f is

$$-\Phi_f(a,b)/f^2 = (a/f)_x + (b/f)_y$$

If f has a double root, say $f = x^2 g$, then Φ_f has the zero $(a,b) = (0,g)$.

To prove (1) recall the formula

$$\Delta(xg) = g(0,1)^2\Delta(g)$$

for $g \in S^{d-1}$ which is another way of characterizing Δ. The corresponding statement for $\det(\Phi)$ follows from the commutative diagram with exact rows

$$\text{(2)}\qquad \begin{array}{ccccc} S^{d-3}\oplus S^{d-3} & \xrightarrow{\cdot(x,x)} & S^{d-2}\oplus S^{d-2} & \longrightarrow & \langle y^{d-2}\rangle \oplus \langle y^{d-2}\rangle \\ {\scriptstyle\Phi_g}\downarrow & & {\scriptstyle\Phi_{xg}}\downarrow & & \downarrow {\scriptstyle\begin{pmatrix} * & g(0,1) \\ g(0,1) & 0\end{pmatrix}} \\ S^{2d-5} & \xrightarrow{\cdot x^2} & S^{2d-3} & \longrightarrow & \langle xy^{2d-4}, y^{2d-3}\rangle \end{array}$$

By induction one concludes (1).

The commutativity of (2) follows from a direct computation. One has

$$\Phi_{xg} = a(g + xg_x) + bxg_y - (a_x + b_y)xg$$

Taking this mod xS^{2d-4} yields

$$\Phi_{xg}(a,b) \equiv a(0,1)g(0,1)y^{2d-3} \qquad (\text{mod } xS^{2d-4})$$

This explains the second row in the matrix on the right of (2).

Assume then $a(0,1) = 0$ so that $a = x\tilde{a}$. One has

$$\Phi_{xg}(x\tilde{a},b) = x\tilde{a}(g + xg_x) + bxg_y - (\tilde{a} + x\tilde{a}_x + b_y)xg$$

which yields

$$\text{(3)}\qquad \Phi_{xg}(x\tilde{a},b) = (bg_y - b_y g)x + (\tilde{a}g_x - \tilde{a}_x g)x^2$$

Calculating mod x^2S^{2d-3} yields

$$\begin{aligned} \Phi_{xg}(x\tilde{a},b) &\equiv (bg_y - b_y g)x \\ &\equiv b(0,1)g(0,1)\big(y^{d-2}(y^{d-1})_y - (y^{d-2})_y y^{d-1}\big)x \\ &\equiv b(0,1)g(0,1)y^{2d-4}x \end{aligned}$$

This explains the first row in the matrix on the right of (2).

Another consequence of (3) is

$$\Phi_{xg}(x\tilde{a},x\tilde{b}) = x^2\Phi_g(\tilde{a},\tilde{b})$$

which means the commutativity of the first square of (2).

4. Demonstration via the resultant

Another proof of (1) stems from a direct comparison with the Sylvester matrix for the resultant which is given by

$$\Psi_f \colon S^{d-2} \oplus S^{d-2} \to S^{2d-3}$$
$$\Psi_f(a,b) = af_x + bf_y$$

or, without coordinates, by

$$\Psi_f \colon V^* \otimes S^{d-2}V \to S^{2d-3}V$$
$$\Psi_f(\varphi \otimes g) = (Df)(\varphi) \cdot g$$

It is convenient to tensor with the line bundle $\Lambda^2 V$ and to rewrite Ψ_f and Φ_f as

$$\Psi_f, \Phi_f \colon V \otimes S^{d-2}V \to \Lambda^2 V \otimes S^{2d-3}V$$
$$\Psi_f(v \otimes g) = v \wedge Df \cdot g$$
$$\Phi_f(v \otimes g) = v \wedge Df \cdot g - v \wedge Dg \cdot f$$

Thus

$$\Phi_f = \Psi_f - \lambda \cdot f$$

with

$$\lambda \colon V \otimes S^{d-2}V \to \Lambda^2 V \otimes S^{d-3}V$$
$$\lambda(v \otimes g) = v \wedge Dg$$

There is the standard exact sequence

$$0 \to \Lambda^2 V \otimes S^{d-3}V \xrightarrow{\kappa} V \otimes S^{d-2}V \xrightarrow{\mu} S^{d-1}V \to 0$$

with

$$\kappa\big((v \wedge w) \otimes h\big) = v \otimes wh - w \otimes vh$$

and with μ the multiplication in the symmetric algebra. One has $\lambda \circ D = 0$ (in derivative notation this is $h_{xy} = h_{yx}$) and the following variants of Eulers relation

$$\lambda \circ \kappa = d - 1$$
$$\kappa \circ \lambda + D \circ \mu = d - 1$$
$$\mu \circ D = d - 1$$

Now a key observation is that

$$\Psi_f \circ \kappa \colon \Lambda^2 V \otimes S^{d-3}V \to \Lambda^2 V \otimes S^{2d-3}V$$

is multiplication by df:

$$\Psi_f \circ \kappa = df$$

Thus if we put (assuming for a moment that d is a non zero divisor)

$$M = \frac{1}{d}\kappa(\Lambda^2 V \otimes S^{d-3}V) + V \otimes S^{d-2}V \subset \frac{1}{d}(V \otimes S^{d-2}V)$$

there is the induced morphism

$$\widetilde{\Psi}_f \colon M \to S^{2d-3}V$$

For the determinants one has

$$\det(\Psi_f) = d^{\operatorname{rank} S^{d-3}V} \det(\widetilde{\Psi}_f)$$

so that

$$\det(\widetilde{\Psi}_f) = \operatorname{disc}(f)$$

The module M can be defined without assuming that d is a non zero divisor by means of extensions:

$$\begin{array}{ccccc} \Lambda^2 V \otimes S^{d-3}V & \xrightarrow{\kappa} & V \otimes S^{d-2}V & \xrightarrow{\mu} & S^{d-1}V \\ {\scriptstyle d}\big\downarrow & & {\scriptstyle p}\big\downarrow & & \big\| \\ \Lambda^2 V \otimes S^{d-3}V & \xrightarrow{\kappa_M} & M & \longrightarrow & S^{d-1}V \end{array}$$

In other words

$$M = \frac{\Lambda^2 V \otimes S^{d-3}V \oplus V \otimes S^{d-2}V}{(d, -\kappa)(\Lambda^2 V \otimes S^{d-3}V)}$$

and $\widetilde{\Psi}_f$ is given by

$$\widetilde{\Psi}_f([\beta, \alpha]) = \Psi_f(\alpha) + \beta \cdot f$$

A further observation is that there is the isomorphism

$$r \colon \Lambda^2 V \otimes S^{d-2}V \to M$$
$$r(\alpha) = p - \kappa_M \circ \lambda$$

(note that $r \circ \kappa = d\kappa_M - (d-1)\kappa_M = \kappa_M$). Finally one finds

$$\Phi_f = \widetilde{\Psi}_f \circ r$$

which results in (1).

5. Interpretation with jet bundles

Formula (1) can be interpreted geometrically by means of jet bundles on $\mathbf{P}^1$ (I am indebted to P. Deligne for explanations). In brief, the map Φ_f is given by

$$\varphi_f \colon H^0\big(\mathbf{P}^1, J_d(-1)\big) \xrightarrow{\wedge j(f)} H^0\big(\mathbf{P}^1, (\Lambda^2 J_d)(-1)\big)$$

where J_d is the jet bundle for $\mathcal{O}(d)$ and $j(f)$ is the jet of f. It is easy to see that $\Lambda^2 J_d \simeq \mathcal{O}(2d-2)$. The computation of the global sections of $J_d(-1)$ is more delicate.

If one assumes that d is invertible, the jet bundle splits as

$$J_d \simeq \mathcal{O}(d-1) \oplus \mathcal{O}(d-1)$$

This way φ_f becomes a map

$$\varphi_f \colon H^0\big(\mathbf{P}^1, \mathcal{O}(d-2) \oplus \mathcal{O}(d-2)\big) \to H^0\big(\mathbf{P}^1, \mathcal{O}(2d-3)\big)$$

or

$$\varphi_f \colon S^{d-2} \oplus S^{d-2} \to S^{2d-3}$$

Computing φ_f this way yields Ψ_f (the Sylvester matrix of the derivatives).

In the general case (over $\mathbf{Z}$) one finds

$$H^0\big(\mathbf{P}^1, J_d(-1)\big) \simeq M$$

and together with the isomorphism r above one gets again

$$H^0\big(\mathbf{P}^1, J_d(-1)\big) \simeq S^{d-2} \oplus S^{d-2}$$

However, computing φ_f this way yields $\widetilde{\Psi}_f$ resp. Φ_f.

The principle result on discriminants and jet bundles involved here is [3, Theorem 2.5, p. 56]. The book [3] works generally over $\mathbf{C}$ and gives as application the Sylvester formula with the denominator d^{d-2} [3, formula (2.9), p. 60].

References

[1] N. Bourbaki, *Éléments de mathématique*, Masson, Paris, 1981, Algèbre. Chapitres 4 à 7. [Algebra. Chapters 4–7].

[2] M. Demazure, *Résultant, discriminant*, Enseign. Math. (2) **58** (2012), no. 3-4, 333–373.

[3] I. Gelfand, M. Kapranov, and A. Zelevinsky, *Discriminants, resultants and multidimensional determinants*, Modern Birkhäuser Classics, Birkhäuser Boston Inc., Boston, MA, 2008, Reprint of the 1994 edition.

Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany

E-mail address: `rost` *at* `math.uni-bielefeld.de`

URL: `www.math.uni-bielefeld.de/~rost`