RESULTANT AND DISCRIMINANT OF POLYNOMIALS

SVANTE JANSON

ABSTRACT. This is a collection of classical results about resultants and discriminants for polynomials, compiled mainly for my own use. All results are well-known 19th century mathematics, but I have not investigated the history, and no references are given.

1. Resultant

Definition 1.1. Let $f(x) = a_n x^n + \dots + a_0$ and $g(x) = b_m x^m + \dots + b_0$ be two polynomials of degrees (at most) n and m, respectively, with coefficients in an arbitrary field F. Their *resultant* $R(f,g) = R_{n,m}(f,g)$ is the element of F given by the determinant of the $(m+n) \times (m+n)$ *Sylvester matrix* $\mathrm{Syl}(f,g) = \mathrm{Syl}_{n,m}(f,g)$ given by

$$\begin{pmatrix} a_n & a_{n-1} & a_{n-2} & \dots & 0 & 0 & 0 \\ 0 & a_n & a_{n-1} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a_1 & a_0 & 0 \\ 0 & 0 & 0 & \dots & a_2 & a_1 & a_0 \\ b_m & b_{m-1} & b_{m-2} & \dots & 0 & 0 & 0 \\ 0 & b_m & b_{m-1} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & b_1 & b_0 & 0 \\ 0 & 0 & 0 & \dots & b_2 & b_1 & b_0 \end{pmatrix} \tag{1.1}$$

where the m first rows contain the coefficients $a_n, a_{n-1}, \dots, a_0$ of f shifted $0, 1, \dots, m-1$ steps and padded with zeros, and the n last rows contain the coefficients $b_m, b_{m-1}, \dots, b_0$ of g shifted $0, 1, \dots, n-1$ steps and padded with zeros. In other words, the entry at (i,j) equals a_{n+i-j} if $1 \le i \le m$ and b_{i-j} if $m+1 \le i \le m+n$, with $a_i = 0$ if $i > n$ or $i < 0$ and $b_i = 0$ if $i > m$ or $i < 0$.

Date: September 22, 2007; revised August 16, 2010.

Example. If $n = 3$ and $m = 2$,

$$R(f,g) = \begin{vmatrix} a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & a_3 & a_2 & a_1 & a_0 \\ b_2 & b_1 & b_0 & 0 & 0 \\ 0 & b_2 & b_1 & b_0 & 0 \\ 0 & 0 & b_2 & b_1 & b_0 \end{vmatrix}.$$

In the exterior algebra over $F[x]$, we thus have

$$\begin{aligned}(x^{m-1}f(x)) \wedge (x^{m-2}f(x)) \wedge \cdots \wedge f(x) \\ \wedge (x^{n-1}g(x)) \wedge (x^{n-2}g(x)) \wedge \cdots \wedge g(x) \\ = R(f,g)x^{n+m-1} \wedge x^{n+m-2} \wedge \cdots \wedge 1, \quad (1.2)\end{aligned}$$

which can be used as an alternative form of Definition 1.1.

Remark 1.2. Typically, one assumes in Definition 1.1 that $n = \deg(f)$ and $m = \deg(g)$, i.e. that $a_n \neq 0$ and $b_m \neq 0$; this implies that $R(f,g)$ is completely determined by the polynomials f and g (and it excludes the case $f = 0$ or $g = 0$). It is, however, convenient to use the slightly more general version above which also allows n and m to be regarded as given and then $R(f,g)$ is defined for all polynomials f and g of degrees $\deg(f) \leq n$, $\deg(g) \leq m$. (See for example Remarks 1.9 and 3.4.) In this case, we may use the notation $R_{n,m}(f,g)$ to avoid ambiguity, but usually we write just $R(f,g)$.

Remark 1.3. It is sometimes convenient to regard a_i and b_j is indeterminates, thus regarding f and g as polynomials with coefficients in the field $F(a_n, \dots, a_0, b_m, \dots, b_0)$. Any formula or argument that requires $a_n \neq 0$ and $b_m \neq 0$ then can be used; if this results in, for example, a polynomial identity involving $R_{n,m}(f,g)$, then this formula holds also if we substitute any values in F for $a_n, \dots, b_0$.

The resultant is obviously a homogeneous polynomial of degree $n + m$ with integer coefficients in the coefficients a_i, b_j. More precisely, we have the following. We continue to use the notations a_i and b_j for the coefficients of f and g, respectively, as in Definition 1.1.

Theorem 1.4. *$R_{n,m}(f,g)$ is a homogeneous polynomial with integer coefficients in the coefficients a_i, b_j.*

(i) *$R_{n,m}(f,g)$ is homogeneous of degree m in $a_n, \dots, a_0$ and degree n in $b_m, \dots, b_0$.*
(ii) *If a_i and b_i are regarded as having degree i, then $R_{n,m}(f,g)$ is homogeneous of degree nm.*

Proofs of this and other results in this section are given in Section 2.

Remark 1.5. If we write $R_{n,m}(f,g)$ as a polynomial with integer coefficients for any field with characteristic 0, such as $\mathbb{Q}$ or $\mathbb{C}$, then the formula is valid (with the same coefficients) for every field F. (Because the coefficients are given by expanding the determinant of $\mathrm{Syl}_{n,m}(f,g)$ and thus have a combinatorial interpretation independent of F. Of course, for a field of characteristic $p \neq 0$, the coefficients may be reduced modulo p, so they are not unique in that case.)

The main importance of the resultant lies in the following formula, which often is taken as the definition.

Theorem 1.6. *Let $f(x) = a_n x^n + \dots + a_0$ and $g(x) = b_m x^m + \dots + b_0$ be two polynomials of degrees n and m, respectively, with coefficients in an arbitrary field F. Suppose that, in some extension of F, f has n roots $\xi_1, \dots, \xi_n$ and g has m roots $\eta_1, \dots, \eta_m$ (not necessarily distinct). Then*

$$R(f,g) = a_n^m b_m^n \prod_{i=1}^{n} \prod_{j=1}^{m} (\xi_i - \eta_j). \tag{1.3}$$

Here and below, the roots of a polynomial are listed with multiple roots repeated according to their multiplicities. Thus every polynomial of degree n has n roots in some extension field (for example in an algebraically closed extension). Combining Theorem 1.6 and Definition 1.1, we see that the product in (1.3) lies in coefficient field F, and that it does not depend on the choice of extension field.

Theorem 1.6 implies the perhaps most important result about resultants.

Corollary 1.7. *Let f and g be two non-zero polynomials with coefficients in a field F. Then f and g have a common root in some extension of F if and only if $R(f,g) = 0$.*

It is here implicit that $R = R_{n,m}$ with $n = \deg(f)$ and $m = \deg(g)$. Since f and g have a common root in some extension of F if and only if they have a common non-trivial (i.e., non-constant) factor in F, Corollary 1.7 can also be stated as follows.

Corollary 1.8. *Let f and g be two non-zero polynomials with coefficients in a field. Then f and g have a common non-trivial factor if and only if $R(f,g) = 0$. Equivalently, f and g are coprime if and only if $R(f,g) \neq 0$.*

Remark 1.9. If n and m are fixed, we can (in the style of projective geometry) say that a polynomial f with $\deg(f) \leq n$ has $n - \deg(f)$ roots at ∞, and similarly g has $m - \deg(g)$ roots at ∞. Thus f always has n roots and g has m, in $F_1 \cup \{\infty\}$ for some extension F_1. With this interpretation, Corollary 1.7 holds for all polynomials with $\deg(f) \leq n$ and $\deg(g) \leq m$. (Including $f = 0$ and $g = 0$ except in the trivial case $n = m = 0$.)

We have also the following useful formulas related to (1.3).

Theorem 1.10. *Let $f(x) = a_n x^n + \dots + a_0$ and $g(x) = b_m x^m + \dots + b_0$ be two polynomials with coefficients in an arbitrary field F.*

(i) *Suppose that f has n roots $\xi_1, \dots, \xi_n$ in some extension of F. Then*

$$R(f, g) = a_n^m \prod_{i=1}^{n} g(\xi_i). \tag{1.4}$$

(ii) *Suppose that g has m roots $\eta_1, \dots, \eta_m$ in some extension of F. Then*

$$R(f, g) = (-1)^{nm} b_m^n \prod_{j=1}^{m} f(\eta_j). \tag{1.5}$$

In (i), f necessarily has degree n, while $\deg(g) \le m$ may be less than m. Similarly, in (ii), $\deg(g) = m$ and $\deg(f) \le n$.

The Sylvester matrix of g and f is obtained by permuting the rows of the Sylvester matrix of f and g. The number of inversions of the permutation is nm, and it follows immediately from Definition 1.1 that

$$R_{m,n}(g, f) = (-1)^{nm} R_{n,m}(f, g). \tag{1.6}$$

(If $\deg(f) = n$ and $\deg(g) = m$, (1.6) also follows from (1.3).) This kind of anti-symmetry explains why there is a factor $(-1)^{nm}$ in (1.5) but not in (1.4).

The factorization properties in Theorem 1.10 can also be expresed as follows.

Theorem 1.11. *If f_1, f_2 and g are polynomials with $\deg(f_1) \le n_1$, $\deg(f_2) \le n_2$ and $\deg(g) \le m$, then*

$$R_{n_1+n_2,m}(f_1 f_2, g) = R_{n_1,m}(f_1, g) R_{n_2,m}(f_2, g). \tag{1.7}$$

Similarly, if f, g_1 and g_2 are polynomials with $\deg(f) \le n$, $\deg(g_1) \le m_1$ and $\deg(g_2) \le m_2$, then

$$R_{n,m_1+m_2}(f, g_1 g_2) = R_{n,m_1}(f, g_1) R_{n,m_2}(f, g_2). \tag{1.8}$$

There is, besides (1.6), also another type of symmetry. With $f(x) = a_n x^n + \dots + a_0$ and $g(x) = b_m x^m + \dots + b_0$ as above, define the reversed polynomials by

$$f^*(x) = x^n f(1/x) = a_n + a_{n-1} x + \dots + a_0 x^n, \tag{1.9}$$

$$g^*(x) = x^m g(1/x) = b_m + b_{m-1} x + \dots + b_0 x^m. \tag{1.10}$$

Theorem 1.12. *With notations as above, for any two polynomials f and g with $\deg(f) \le n$ and $\deg(g) \le m$,*

$$R_{n,m}(f^*, g^*) = R_{m,n}(g, f) = (-1)^{nm} R_{n,m}(f, g).$$

As said in Remark 1.2, the standard case for the resultant is when $\deg(f) = n$ and $\deg(g) = m$. We can always reduce to that case by the following formulas.

Theorem 1.13. (i) *If $\deg(g) \le k \le m$, then*

$$R_{n,m}(f, g) = a_n^{m-k} R_{n,k}(f, g). \tag{1.11}$$

(ii) *If* $\deg(f) \le k \le n$, *then*

$$R_{n,m}(f,g) = (-1)^{(n-k)m} b_m^{n-k} R_{k,m}(f,g). \tag{1.12}$$

Note further that if both $\deg(f) < n$ and $\deg(g) < m$, then $R_{n,m}(f,g) = 0$. (Because the first column in (1.1) vanishes, or from (1.11) or (1.12).)

Theorem 1.14. *Let* f *and* g *be polynomials with* $\deg(f) \le n$ *and* $\deg(g) \le m$. *If* $n \ge m$ *and* h *is any polynomial with* $\deg(h) \le n-m$, *then*

$$R_{n,m}(f+hg,g) = R_{n,m}(f,g). \tag{1.13}$$

Similarly, if $n \le m$ *and* h *is any polynomial with* $\deg(h) \le m-n$, *then*

$$R_{n,m}(f,g+hf) = R_{n,m}(f,g). \tag{1.14}$$

Theorem 1.15. *If* f *and* g *are polynomials of degrees* n *and* m *as above, with roots* $\xi_1,\dots,\xi_n$ *and* $\eta_1,\dots,\eta_m$ *in some extension field, then the resultant* $R\bigl(f(x),g(y-x)\bigr)$ *(with* $g(y-x)$ *regarded as a polymial in* x*) is a polynomial in* y *of degree* nm *with roots* $\xi_i+\eta_j$, $1 \le i \le n$ *and* $1 \le j \le m$. *Further,* $R\bigl(f(x),g(y-x)\bigr)$ *has leading coefficient* $a_n^m b_m^n$. *In particular,* $R\bigl(f(x),g(y-x)\bigr)$ *is monic if both* f *and* g *are.*

If $\deg(f) < n$ or $\deg(g) < m$, but not both, then by Theorem 1.13 $R\bigl(f(x),g(y-x)\bigr)$ is still a polynomial whose roots are given by $\xi_i+\eta_j$, where ξ_i runs through the roots of f and η_j through the roots of g (with multiplicities). If $\deg(f) < n$ and $\deg(g) < m$, then $R\bigl(f(x),g(y-x)\bigr) = 0$.

Example 1.16. Let $n=1$ and $f(x) = ax+c$. Then, if $a \ne 0$, f has the single root $\xi = -c/a$ and (1.4) yields

$$R_{1,m}(f,g) = a^m g(-c/a) = \sum_{j=0}^{m} b_j (-c)^j a^{m-j}. \tag{1.15}$$

This formula (ignoring the middle expression) holds also if $a=0$ (and then simplifies to $R_{1,m}(c,g) = b_m(-c)^m$), for example by Remark 1.3.

Example 1.17. Let $n \ge 0$ and let $f(x)$ and $g(x)$ be two polynomials of degree $\le n$ with coefficients in a field F. Further, let $a,b,c,d \in F$.

Assume first that $d \ne 0$. Then, by Theorem 1.14 and Theorem 1.4,

$$\begin{aligned} R_{n,n}(af+bg,cf+dg) &= R_{n,n}\bigl(af+bg-(b/d)(cf+dg),\, cf+dg\bigr) \\ &= R_{n,n}\bigl((a-bc/d)f,\, cf+dg\bigr) \\ &= (a-bc/d)^n R_{n,n}(f,\, cf+dg) \\ &= (a-bc/d)^n R_{n,n}(f,dg) \\ &= (ad-bc)^n R_{n,n}(f,g). \end{aligned} \tag{1.16}$$

The final formula holds in the case $d=0$ too, for example by regarding d as an indeterminate. We may write the result as

$$R_{n,n}(af+bg,cf+dg) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}^n R_{n,n}(f,g). \tag{1.17}$$

1.1. **Trivial cases.** For completeness we allow $n = 0$ or $m = 0$ above. The case $m = n = 0$ is utterly trivial: $f(x)$ and $g(x)$ are constants, the Sylvester matrix (1.1) has 0 rows and columns (the empty matrix), and $R_{0,0}(f, g) = 1$ (by definition).

In the case $m = 0$, $g(x) = b_0$ is constant. The Sylvester matrix is the diagonal matrix $b_0 I_n$, where I_n is the $n \times n$ identity matrix, and thus $R_{n,0}(f, g) = b_0^n$.

Similarly, or by (1.6), if $n = 0$, then $f(x) = a_0$ and $R_{0,m}(f, g) = a_0^m$. (These formulas are special cases of (1.5) and (1.4).)

Note that the formulas (1.3), (1.4), (1.5) in Theorems 1.6 and 1.10 hold also for $n = 0$ and $m = 0$, with empty products defined to be 1.

1.2. **Another determinant formula.**

Theorem 1.18. *Let f and g be polynomials with* $\deg(f) = n$ *and* $\deg(g) \leq m$. *Let, for $k \geq 0$, $r_k(x) = r_{k,n-1}x^{n-1} + \dots + r_{k,0}$ be the remainder of $x^k g(x)$ modulo $f(x)$, i.e., $x^k g(x) = q_k(x) f(x) + r_k(x)$ for some polynomial q_k and* $\deg(r_k) \leq n - 1$. *Then (where as above a_n is the leading coefficient of f),*

$$R_{n,m}(f, g) = a_n^m \begin{vmatrix} r_{n-1,n-1} & \dots & r_{n-1,0} \\ \vdots & & \vdots \\ r_{0,n-1} & \dots & r_{0,0} \end{vmatrix}. \tag{1.18}$$

1.3. **More on the Sylvester matrix.** The following theorem extends Corollary 1.8, since it in particular says that the Sylvester matrix of f and g is singular if and only if their greatest common divisor has degree ≥ 1.

Theorem 1.19. *Let f and g be two polynomials with* $\deg(f) = n$ *and* $\deg(g) = m$, *and let* $h := \mathrm{GCD}(f, g)$ *be their greatest common divisor (i.e., a common divisor of highest degree). Then* $\deg(h)$ *is the corank of the Sylvester matrix* $\mathrm{Syl}(f, g)$. *In other words, the Sylvester matrix* $\mathrm{Syl}(f, g)$ *has rank* $n + m - \deg(h)$.

There is also an explicit description of the left null space.

Theorem 1.20. *Let f and g be two polynomials with* $\deg(f) \leq n$ *and* $\deg(g) \leq m$. *Let $v := (\alpha_{m-1}, \dots, \alpha_0, \beta_{n-1}, \dots, \beta_0)$ be a row vector of dimension $m + n$. Then $v\,\mathrm{Syl}(f, g) = 0$ if and only if $pf + qg = 0$, where $p(x) = \alpha_{m-1}x^{m-1} + \dots + \alpha_0 x^0$ and $q(x) = \beta_{n-1}x^{n-1} + \dots + \beta_0 x^0$.*

1.4. **Further examples.** If $n = m = 2$,

$$R(f, g) = \begin{vmatrix} a_2 & a_1 & a_0 & 0 \\ 0 & a_2 & a_1 & a_0 \\ b_2 & b_1 & b_0 & 0 \\ 0 & b_2 & b_1 & b_0 \end{vmatrix} = (a_2 b_0 - b_2 a_0)^2 - (a_2 b_1 - b_2 a_1)(a_1 b_0 - b_1 a_0).$$

If $n = m = 3$,

$$R(f,g) = \begin{vmatrix} a_3 & a_2 & a_1 & a_0 & 0 & 0 \\ 0 & a_3 & a_2 & a_1 & a_0 & 0 \\ 0 & 0 & a_3 & a_2 & a_1 & a_0 \\ b_3 & b_2 & b_1 & b_0 & 0 & 0 \\ 0 & b_3 & b_2 & b_1 & b_0 & 0 \\ 0 & 0 & b_3 & b_2 & b_1 & b_0 \end{vmatrix}.$$

More generally, if $m = n$, then

$$R(f,g) = \begin{vmatrix} a_n & a_{n-1} & \dots & a_0 & 0 & 0 & \dots & 0 \\ 0 & a_n & \dots & a_1 & a_0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{n-1} & a_{n-2} & 0 & \dots & a_0 \\ b_n & b_{n-1} & \dots & b_0 & 0 & 0 & \dots & 0 \\ 0 & b_n & \dots & b_1 & b_0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & b_{n-1} & b_{n-2} & 0 & \dots & b_0 \end{vmatrix}.$$

2. Proofs

We begin by noting that, as said above, (1.6) holds by a permutation of the rows in (1.1).

Proof of Theorem 1.4. It is obvious that $R_{n,m}(f,g)$ is a homogeneous polynomial with integer coefficients in the coefficients a_i, b_j, of total degree $m+n$. Moreover, to replace a_i by ta_i and b_j by ub_j in the Sylvester matrix means that we multiply each of the first m rows by t and each of the last n by u, and thus the determinant $R_{n,m}(f,g)$ by $t^m u^n$, which shows (i). (It is here best to treat a_i, b_j, t and u as different indeterminates, and do the calculations in $F(a_0, \dots, a_n, b_0, \dots, b_m, t, u)$.)

Similarly, (ii) follows because to replace each a_i by $t^i a_i$ and each b_j by $t^j b_j$ in $\mathrm{Syl}_{n,m}(f,g)$ yields the same result as multiplying the i:th row by t^{n+i} for $i = 1, \dots, m$ and by t^i for $i = m+1, \dots, m+n$, and the $j : th$ column by t^{-j}; this multiplies the determinant $R_{n,m}(f,g)$ by $t^{mn+\sum_1^{n+m} i - \sum_1^{n+m} j} = t^{nm}$. ☐

Proof of Theorem 1.14. First, assume $n \geq m$ and $\deg(h) \leq n - m$. The Sylvester matrix $\mathrm{Syl}_{n,m}(f + hg, g)$ is obtained from $\mathrm{Syl}_{n,m}(f,g)$ by row operations that do not change its determinant $R_{n,m}$. (If $h(x) = c_l x^l + \dots + c_0$, add c_k times row $n + i - k$ to row i, for $i = 1, \dots, m$ and $k = 0, \dots, l$.)

The second part follows similarly, or by the first part and (1.6). ☐

Proof of Theorem 1.13. (i). Suppose that $\deg(g) < m$, so $b_m = 0$. Then the first column of the Sylvester matrix (1.1) is 0 except for its first element a_n, and the submatrix of $\mathrm{Syl}_{n,m}(f,g)$ obtained by deleting the first row and column equals $\mathrm{Syl}_{n,m-1}(f,g)$. Hence, by expanding the determinant $R_{n,m}(f,g)$ along the first column,

$$R_{n,m}(f,g) = a_n R_{n,m-1}(f,g).$$

The formula (1.11) now follows for $k = m, m-1, \dots, 0$ by backwards induction.

(ii). By (1.6) and part (i),

$$\begin{aligned} R_{n,m}(f,g) &= (-1)^{nm} R_{m,n}(g,f) = (-1)^{nm} b_m^{n-k} R_{m,k}(g,f) \\ &= (-1)^{nm-km} b_m^{n-k} R_{k,m}(f,g). \end{aligned}$$ □

Proof of Theorems 1.6 and 1.10. We prove these theorems together by induction on $n+m$.

Assume that $\deg(f) = n$ and $\deg(g) = m$. Then, at least in some extension field, $f(x) = a_n \prod_{i=1}^n (x - \xi_i)$ and $g(x) = b_m \prod_{j=1}^m (x - \eta_j)$, and (1.3) is equivalent to both (1.4) and (1.5). Assume by induction that these formulas hold for all smaller values of $n+m$ (and all polynomials of these degrees).

Case 1. First, suppose $0 < n = \deg(f) \le m = \deg(g)$. Divide g by f to obtain polynomials q and r with $g = qf + r$ and $\deg(r) < \deg(f) = n$. Note that

$$\deg(q) = \deg(qf) - \deg(f) = \deg(g - r) - n = m - n.$$

By Theorem 1.14,

$$R_{n,m}(f,g) = R_{n,m}(f, g - qf) = R_{n,m}(f,r). \tag{2.1}$$

Case 1a. Suppose that $r \ne 0$ and let $k := \deg(r) \ge 0$. By Theorem 1.13 and the inductive hypothesis in the form (1.4),

$$R_{n,m}(f,r) = a_n^{m-k} R_{n,k}(f,r) = a_n^{m-k} a_n^k \prod_{i=1}^n r(\xi_i) = a_n^m \prod_{i=1}^n g(\xi_i),$$

since $g(\xi_i) = q(\xi_i) f(\xi_i) + r(\xi_i) = r(\xi_i)$, which verifies (1.4) and thus (1.3).

Case 1b. Suppose now that $r = 0$, so $g = qf$, but $n > 0$. Then $\mathrm{Syl}_{n,m}(f,r) = \mathrm{Syl}_{n,m}(f,0)$ has the last n rows identically 0, so $R_{n,m}(f,r) = 0$, and $R_{n,m}(f,g) = 0$ by (2.1). Further, $g(\xi_1) = q(\xi_1) f(\xi_1) = 0$ so ξ_1 is a root of g too, and the right hand side of (1.3) vanishes too. Hence, (1.3) holds.

Case 2. Suppose that $n = 0$. As remarked in Subsection 1.1, $R_{0,m}(f,g) = a_0^m$, which agrees with (1.3). (This includes the case $n = m = 0$ that starts the induction.)

Case 3. Suppose that $m = \deg(g) < n = \deg(f)$. This is reduced to Case 1 or 2 by (1.6).

This completes the induction, and the proof of Theorem 1.6. It remains to verify Theorem 1.10(i),(ii) also in the cases $\deg(g) < m$ and $\deg(f) < n$, respectively. This follows by Theorem 1.13, as in the proof of Case 1a above, or by Remark 1.3. □

Proof of Corollaries 1.7 and 1.8. Immediate from (1.3), using the fact on common factors stated before Corollary 1.8. □

Proof of Theorem 1.11. By the argument in Remark 1.3, we may assume that $\deg(g) = m$, so g has m roots in some extension of F, and then (1.7) follows from (1.5). Similarly, (1.8) follows from (1.4). □

Proof of Theorem 1.12. The Sylvester matrix $\mathrm{Syl}_{n,m}(f^*, g^*)$ is obtained from $\mathrm{Syl}_{m,n}(g, f)$ by reversing the order of both rows and columns, and thus they have the same determinant. □

Proof of Theorem 1.15. We have $g(x) = b_m \prod_{j=1}^m (x - \eta_j)$ and thus

$$g(y-x) = b_m \prod_{j=1}^{m} (y - x - \eta_j) = (-1)^m b_m \prod_{j=1}^{m} (x - y + \eta_j).$$

Thus, by (1.4) (or Theorem 1.6),

$$\begin{aligned} R\big(f(x), g(y-x)\big) &= a_n^m (-1)^{nm} b_m^n \prod_{i=1}^{n} \prod_{j=1}^{m} (\xi_i - y + \eta_j) \\ &= a_n^m b_m^n \prod_{i=1}^{n} \prod_{j=1}^{m} (y - \xi_i - \eta_j). \end{aligned}$$ □

Proof of Theorem 1.18. We work in the exterior algebra over $F(x)$, using (1.2). Let D be the determinant in (1.18); thus

$$r_{n-1}(x) \wedge \cdots \wedge r_0(x) = D x^{n-1} \wedge \cdots \wedge x^0. \tag{2.2}$$

For $k \le n-1$, $x^k g(x) - r_k(x) = q_k(x) f(x)$ has degree $\le n+m-1$ and is thus a linear combination of $f(x)$, $xf(x)$, ..., $x^{m-1} f(x)$; hence, using (2.2) and (1.2) (with $g(x)$ replaced by 1),

$$\begin{aligned} &(x^{m-1} f(x)) \wedge \cdots \wedge f(x) \wedge (x^{n-1} g(x)) \wedge \cdots \wedge g(x) \\ &\quad = (x^{m-1} f(x)) \wedge \cdots \wedge f(x) \wedge r_{n-1}(x) \wedge \cdots \wedge r_0(x) \\ &\quad = (x^{m-1} f(x)) \wedge \cdots \wedge f(x) \wedge D x^{n-1} \wedge \cdots \wedge x^0 \\ &\quad = D R_{n,m}(f, 1) x^{n+m-1} \wedge \cdots \wedge 1. \end{aligned}$$

Consequently, using (1.2) again,

$$R_{n,m}(f, g) = D R_{n,m}(f, 1).$$

Finally, by Theorem 1.13, or by (1.4),

$$R_{n,m}(f, 1) = a_n^m R_{n,0}(f, 1) = a_n^m.$$ □

Proof of Theorem 1.19. Note first that $\mathrm{Syl}_{n,m}(f, g)$ and $\mathrm{Syl}_{m,n}(g, f)$ have the same rank and corank, so we may interchange f and g. We may thus assume $n \ge m$.

In this case, we may as in the proof of Theorem 1.14 for any polynomial q with $\deg(q) \le n - m$ obtain $\mathrm{Syl}_{n,m}(f - qg, g)$ from $\mathrm{Syl}_{n,m}(f, g)$ by row operations that do not change the rank and corank. In particular, we may replace f by the remainder r obtained when dividing f by g. Then $\deg(r) < m \le n$ and the first column of $\mathrm{Syl}_{n,m}(r,g)$ has a single non-zero element, b_m in row $m+1$. We may thus delete the first column and the $m+1$:th row without changing the corank, and this yields $\mathrm{Syl}_{n-1,m}(r,g)$. (Cf. the proof of Theorem 1.13.) Repeating, we see that if $r \neq 0$ and $k = \deg(r)$, then $\mathrm{Syl}_{n,m}(f,g)$ has the same corank as $\mathrm{Syl}_{k,m}(r,g)$ and $\mathrm{Syl}_{m,k}(g,r)$. We repeat from the start, by dividing g by r and so on; this yields the Euclidean algorithm for finding the GCD h, and we finally end up with the Sylvester matrix $\mathrm{Syl}_{k,l}(0,h)$, for some $k \ge 0$ and $l = \deg(h)$, which evidently has corank l since the first l rows are 0 and the last k are independent, as is witnessed by the lower left $k \times k$ minor which is triangular.

(Alternatively, Theorem 1.19 follows easily from Theorem 1.20.) □

Proof of Theorem 1.20. Let $v\,\mathrm{Syl}_{n,m}(f,g) = (\gamma_1,\dots,\gamma_{n+m})$. Then, for $j = 1,\dots,m+n$, with a_k and b_k defined for all integers k as at the end of Definition 1.1,

$$\gamma_j = \sum_{i=1}^{m} \alpha_{m-i} a_{n+i-j} + \sum_{i=m+1}^{m+n} \beta_{m+n-i} b_{i-j},$$

which equals the coefficient of x^{m+n-j} in $pf + qg$. □

3. Discriminant

Several different normalizations of the discriminant of a polynomial are used by different authors, differing in sign and in factors that are powers of the leading coefficient of the polynomial. One natural choice is the following.

Definition 3.1. Let f be a polynomial of degree $n \ge 1$ with coefficients in an arbitrary field F. Let F_1 be an extension of F where f splits, and let $\xi_1,\dots,\xi_n$ be the roots of f in F_1 (taken with multiplicities). Then the (normalized) *discriminant* of f is

$$\Delta_0(f) := \prod_{1 \le i < j \le n} (\xi_i - \xi_j)^2. \tag{3.1}$$

Note that such a field F_1 always exists, for example an algebraic closure of F will do, and that, e.g. by Theorem 3.3 below, $\Delta_0(f) \in F$ and does not depend on the choice of F_1. (This also follows from the fact that $\Delta_0(f)$ is a symmetric polynomial in $\xi_1,\dots,\xi_n$, and thus by a well-known fact a polynomial in the elementary symmetric polynomials $\sigma_k(\xi_1,\dots,\xi_n) = (-1)^k a_{n-k}/a_n$, $k = 1,\dots,n$.)

Note further that $\Delta_0(cf) = \Delta_0(f)$ for any constant $c \neq 0$.

However, while the definition of Δ_0 is simple and natural, Δ_0 is particularly useful for monic polynomials. In general, it is often more convenient to use the following version, which by Theorem 3.5 below is a polynomial in the coefficients of f. (This is the most common version of the discriminant. The names 'normalized' and 'standard' are my own.)

Definition 3.2. Let $f = a_n x^n + \cdots + a_0$ be a polynomial of degree $n \geq 1$ with coefficients in an arbitrary field F. Then the (standard) *discriminant* of f is

$$\Delta(f) := a_n^{2n-2}\Delta_0(f) = a_n^{2n-2} \prod_{1\leq i<j\leq n} (\xi_i - \xi_j)^2, \tag{3.2}$$

where as above $\xi_1, \dots, \xi_n$ are the roots of f in some extension F_1 of F.

The discriminant can also be defined as the resultant of f and its derivative f', with a suitable normalizing factor, as is stated more precisely in the following theorem.

Theorem 3.3. *Let $f = a_n x^n + \cdots + a_0$ be a polynomial of degree $n \geq 1$ with coefficients in an arbitrary field F. Then the discriminant of f is given by*

$$\Delta(f) = (-1)^{n(n-1)/2} a_n^{-1} R(f, f') \tag{3.3}$$

and thus

$$\Delta_0(f) = (-1)^{n(n-1)/2} a_n^{-(2n-1)} R(f, f'). \tag{3.4}$$

Remark 3.4. To be precise, we should write $R_{n,n-1}(f, f')$ in this theorem. Typically, $\deg(f') = \deg(f) - 1 = n - 1$ and we may then write $R(f, f')$ without any ambiguity. (For example, always when F has characteristic 0, such as $\mathbb{R}$ and $\mathbb{C}$.) However, if F has characteristic $p > 0$ and $p|n$, then $\deg(f') < n - 1$. In this case, if $\deg(f') = k$, then by Theorem 1.13, $R_{n,n-1}(f, f') = a_n^{n-1-k} R_{n,k}(f, f')$ and thus

$$\Delta(f) = (-1)^{n(n-1)/2} a_n^{n-k-2} R_{n,k}(f, f'), \tag{3.5}$$

$$\Delta_0(f) = (-1)^{n(n-1)/2} a_n^{-n-k} R_{n,k}(f, f'). \tag{3.6}$$

Proof of Theorem 3.3. Let f have roots $\xi_1, \dots, \xi_n$ (in some extension field). Then $f(x) = a_n \prod_{i=1}^n (x - \xi)$, and thus $f'(\xi_i) = a_n \prod_{j\neq i}(\xi_i - \xi_j)$. Consequently, by (1.4),

$$\begin{aligned} R_{n,n-1}(f, f') &= a_n^{n-1+n} \prod_{i=1}^{n} \prod_{j\neq i} (\xi_i - \xi_j) = a_n^{2n-1} \prod_{1\leq i<j\leq n} (\xi_i - \xi_j)(\xi_j - \xi_i) \\ &= (-1)^{n(n-1)/2} a_n^{2n-1} \Delta_0(f) = (-1)^{n(n-1)/2} a_n \Delta(f). \qquad \square \end{aligned}$$

Theorem 3.5. *$\Delta(f)$ is a homogeneous polynomial with integer coefficients in the coefficients $a_0, \dots, a_n$ of f. Further, with $n = \deg(f)$,*

(i) *$\Delta(f)$ is homogeneous of degree $2n - 2$ in $a_0, \dots, a_n$.*

(ii) *If a_i is regarded as having degree i, then $\Delta(f)$ is homogeneous of degree $n(n-1)$.*

Proof. The derivative $f'(x) = b_{n-1}x^{n-1} + \cdots + b_0$ with $b_j = (j+1)a_{j+1}$. Hence all entries of the Sylvester matrix $\mathrm{Syl}_{n,n-1}(f, f')$ are integer multiples of $a_0, \dots, a_n$, and thus $R(f, f')$ is a homogeneous polynomial with integer coefficients in $a_0, \dots, a_n$. Moreover, the only (possibly) non-zero entries in the first column of $\mathrm{Syl}_{n,n-1}(f, f')$ are a_n and $b_{n-1} = na_n$; hence $R(f, f')$ is a multiple of a_n, and $\Delta(f, f')$ is also such a polynomial by (3.3).

Since $R(f, f')$ has total degree $n + n - 1$, this also shows that $\Delta(f)$ has degree $2n-2$. Alternatively, this follows from Definition 3.2, since replacing a_i by ta_i for all i does not change the roots $\xi_1, \dots, \xi_n$ of f.

For (ii), note that if $f_t(x) = \sum_{i=0}^n a_i t^i x^i$, for an indeterminate t, then f_t has roots $t^{-1}\xi_1, \dots, t^{-1}\xi_n$, and Definition 3.2 yields

$$\Delta(f_t) = (t^n a_n)^{2n-2} t^{-n(n-1)} \prod_{1 \le i < j \le n} (\xi_i - \xi_j)^2 = t^{n(n-1)} \Delta(f).$$

(Alternatively, (ii) is easily derived from (3.3) and Theorem 1.4(i),(ii).) □

Remark 3.6. As for the resultant, see Remark 1.5, the integer coefficients of $\Delta(f)$ do not depend on the field F (except for the obvious non-uniqueness when $\mathrm{char}(F) \neq 0$).

Theorem 3.7. *Let $f = a_n x^n + \dots + a_0$ be a polynomial of degree $n \ge 1$ with coefficients in an arbitrary field F, and let the roots of $f'(x) = 0$ be $\eta_1, \dots, \eta_{n-1}$ (in some extension of F). Then*

$$\Delta(f) = (-1)^{n(n-1)/2} n^n a_n^{n-1} \prod_{j=1}^{n-1} f(\eta_j), \tag{3.7}$$

and thus

$$\Delta_0(f) = (-1)^{n(n-1)/2} n^n a_n^{-(n-1)} \prod_{j=1}^{n-1} f(\eta_j). \tag{3.8}$$

Proof. By Theorem 3.3 and (1.5), since $f'(x) = na_n x^{n-1} + \dots$. □

The roots η_j of f' are the *stationary points* of f, and the function values $f(\eta_j)$ there the *stationary values*. Thus, assuming for simplicity $a_n = 1$, Theorem 3.7 says that the discriminant is a constant times the product of the stationary values.

The perhaps most important use of the discriminant is the following immediate consequence of Definitions 3.1 and 3.2.

Theorem 3.8. *Let f be a polynomial of degree $n \ge 1$ with coefficients in a field F. Then*

$$\Delta_0(f) = 0 \iff \Delta(f) = 0 \iff f \text{ has a double root in some extension of } F.$$

Equivalently, f has n distinct roots in some extension field if and only if the discriminant is $\neq 0$.

By Theorem 3.5, $\Delta(f)$ for f of a given degree $n \ge 1$ is a polynomial in $a_0, \dots, a_n$; we can apply this polynomial also when $a_n = 0$, i.e., to polynomials f of degree $< n$. To avoid confusion, we denote this polynomial in the coefficients $a_0, \dots, a_n$ by $\Delta^{(n)}(f)$, defined for all polynomials $f = a_n x^n + \dots + a_0$

of degree $\leq n$. Thus $\Delta^{(n)}(f) = \Delta(f)$ when $a_n \neq 0$. This polynomial has a simple symmetry.

Theorem 3.9. *If $f = a_n x^n + \cdots + a_0$ is a polynomial of degree $\leq n$ and f^* is defined by (1.9), then*

$$\Delta^{(n)}(f^*) = \Delta^{(n)}(f). \tag{3.9}$$

In particular, if f has degree n and $a_0 \neq 0$, then

$$\Delta(f^*) = \Delta(f). \tag{3.10}$$

Proof. Suppose first that $a_n \neq 0$ and $a_0 \neq 0$. Let f have roots $\xi_1, \dots, \xi_n$ in some extension field; these roots are non-zero and f^* has the roots $\xi_1^{-1}, \dots, \xi_n^{-1}$ and leading coefficient $a_0 = a_n \xi_1 \cdots \xi_n$. Hence, by Definition 3.2,

$$\Delta(f^*) = a_0^{2n-2} \prod_{1 \leq i < j \leq n} \bigl(\xi_i^{-1} - \xi_j^{-1}\bigr)^2 = a_n^{2n-2} \prod_{1 \leq i < j \leq n} \bigl(\xi_j - \xi_i\bigr)^2 = \Delta(f),$$

which proves (3.10). In particular, this holds if we regard $a_0, \dots, a_n$ as indeterminates, and thus (3.9) follows in general because both sides are polynomials in $a_0, \dots, a_n$. □

We give another simple consequence of the definition.

Theorem 3.10. *If f and g are polynomials of degrees n and $m \geq 1$, then*

$$\Delta(fg) = \Delta(f)\Delta(g)R(f,g)^2. \tag{3.11}$$

Proof. This follows from Definition 3.2 and Theorem 1.6.

Alternatively, by Theorems 1.11, 1.14 and 1.11 again,

$$\begin{aligned} R(fg,(fg)') &= R(fg, f'g + fg') = R(f, f'g + fg')R(g, f'g + fg') \\ &= R(f, f'g)R(g, fg') = R(f, f')R(f, g)R(g, f)R(g, g'), \end{aligned}$$

and the result follows by (3.3) and (1.6). □

As said above, $\Delta^{(n)}(f) = \Delta(f)$ when $a_n \neq 0$. In the opposite case $a_n = 0$, we have the following simple formula, which can be regarded as a relation between discriminants for polynomials of different degrees. (See the examples in (4.1) and (4.3), or, more complicated, in Examples 4.7 and 4.3.)

Theorem 3.11. *If $a_n = 0$, then*

$$\Delta^{(n)}(f) = a_{n-1}^2 \Delta^{(n-1)}(f). \tag{3.12}$$

In particular, if $a_n = a_{n-1} = 0$, then $\Delta^{(n)}(f) = 0$.

Proof. Assume first $a_{n-1} \neq 0$ and $a_0 \neq 0$. Let $g(x) = a_{n-1}x^{n-1} + \cdots + a_0$ (this is the same as $f(x)$, but we regard it as a polynomial of degree $n-1$), and define f^* by (1.9) and g^* by (1.10), with m replaced by $n-1$. Then $f^*(x) = xg^*(x)$, where f^* has degree n and $g^*(x)$ degree $n-1$. Trivially,

$\Delta(x) = 1$, and Example 1.16 shows $R_{1,n-1}(x, g^*) = g^*(0) = a_{n-1}$. Hence, Theorems 3.9 and 3.10 yield

$$\Delta^{(n)}(f) = \Delta(f^*) = \Delta(xg^*) = \Delta(g^*)a_{n-1}^2 = a_{n-1}^2 \Delta^{(n-1)}(g),$$

which shows (3.12) in the case $a_{n-1}, a_0 \neq 0$. The general case follows, because both sides of (3.12) are polynomials in $a_0, \dots, a_{n-1}$. (An alternative proof without using inversion and Theorem 3.9 is givven in Appendix A.) $\square$

Remark 3.12. If we fix $n \geq 1$ and as in Remark 1.9 say that a polynomial f with $\deg(f) \leq n$ has $n - \deg(f)$ roots at ∞, then Theorems 3.8 and 3.11 show that $\Delta^{(n)}(f) = 0$ if and only if f has a double root (or more precisely, a multiple root) in $F_1 \cup \{\infty\}$ for some extension F_1 (and in any extension where f splits).

Remark 3.13. If $f = a_n x^n + \dots + a_0$ is a polynomial of degree n with non-zero $a_0, \dots, a_{n-1}$, define

$$\Delta^*(f) := \prod_{i=0}^{n-1} a_i^{-2} \cdot \Delta(f). \tag{3.13}$$

Theorem 3.11 shows that if we, more generally, for f of degree $\leq n$ define

$$\Delta^{(n)*}(f) := \prod_{i=0}^{n-1} a_i^{-2} \cdot \Delta^{(n)}(f), \tag{3.14}$$

then, whenever $\deg(f) < n$,

$$\Delta^{(n)*}(f) = \Delta^{(n-1)*}(f). \tag{3.15}$$

It is here best to regard the coefficients a_i as indeterminates; then $\Delta^{(n)*}(f)$ is a Laurent polynomial in $a_0, \dots, a_n$, and (3.15) shows that there is a single Laurent series $\boldsymbol{\Delta}^*$ in the infinitely many indeterminates $a_0, a_1, \dots$ such that if f is a polynomial of any degree $n \geq 1$, then $\Delta^*(f)$ is obtained from this series $\boldsymbol{\Delta}^*$ by substituting $a_i = 0$ for $i > n$. (This has to be done with some care since also negative powers appear, but each term containing a negative power $a_i^{-\alpha_i}$ with $i > n$ contains also a positive power $a_j^{\alpha_j}$ with $j > i > n$, so there is no real problem; we simply delete all terms containing some non-zero power of some a_i with $i > n$.) We may regard $\boldsymbol{\Delta}^*$ as a universal discriminant. (Or as a mere curiosity.)

It follows from Theorem 3.5 that the monomials that appear in $\boldsymbol{\Delta}^*$ have integer coefficients and all have the form $\prod_{i=0}^k a_i^{\alpha_i}$ with $\sum_i \alpha_i = -2$ and $\sum_i i\alpha_i = 0$; except for the term a_0^{-2}, they further have $\alpha_k > 0$ if k is chosen minimal. See Example 4.11.

4. Examples of discriminants

Example 4.1 ($n = 1$)**.** If $f = ax + b$, then trivially $\Delta(f) = \Delta(f_0) = 1$.

Example 4.2 $(n = 2)$**.** If $f(x) = ax^2 + bx + c$, then Theorem 3.3 yields

$$\Delta(f) = -a^{-1}R(f, f') = -a^{-1} \begin{vmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{vmatrix} = b^2 - 4ac \tag{4.1}$$

and

$$\Delta_0(f) = a^{-2}\Delta(f) = \frac{b^2 - 4ac}{a^2} = \left(\frac{b}{a}\right)^2 - 4\frac{c}{a}. \tag{4.2}$$

Note that the standard formula for finding the roots of $ax^2 + bx + c = 0$ can be written

$$x_\pm = \frac{b}{2a} \pm \frac{1}{2}\sqrt{\Delta_0(f)} = \frac{b \pm \sqrt{\Delta(f)}}{2a}.$$

Example 4.3 $(n = 3)$**.** If $f(x) = ax^3 + bx^2 + cx + d$, then Theorem 3.3 yields

$$\begin{aligned} \Delta(f) = -a^{-1}R(f, f') &= -a^{-1} \begin{vmatrix} a & b & c & d & 0 \\ 0 & a & b & c & d \\ 3a & 2b & c & 0 & 0 \\ 0 & 3a & 2b & c & 0 \\ 0 & 0 & 3a & 2b & c \end{vmatrix} \\ &= b^2c^2 - 4ac^3 - 4b^3d + 18abcd - 27a^2d^2. \end{aligned} \tag{4.3}$$

Example 4.4 $(n = 3)$**.** If $f(x) = x^3 + bx^2 + cx + d$ is monic, (4.3) simplifies to

$$\Delta_0(f) = \Delta(f) = -R(f, f') = b^2c^2 - 4c^3 - 4b^3d + 18bcd - 27d^2.$$

Example 4.5 $(n = 3)$**.** For $f(x) = x^3 + px + q$, without second degree term, (4.3) simplifies further to

$$\Delta_0(f) = \Delta(f) = -4p^3 - 27q^2.$$

Example 4.6 $(n = 3)$**.** The polynomial $f(x) = 4x^3 - g_2x - g_3$ is important in the theory of the Weierstrass elliptic functions. Its discriminant is, by (4.3),

$$\Delta(4x^3 - g_2x - g_3) = 16g_2^3 - 432g_3^2.$$

Equivalently, $\Delta_0(4x^3 - g_2x - g_3) = \frac{1}{16}g_2^3 - \frac{27}{16}g_3^2$. In this context, it is customary to change the normalization and define the discriminant as

$$16\Delta_0(f) = \frac{1}{16}\Delta(f) = g_2^3 - 27g_3^2.$$

Example 4.7 ($n = 4$)**.** If $f(x) = ax^4 + bx^3 + cx^2 + dx + e$, then Theorem 3.3 yields

$$\Delta(f) = a^{-1}R(f, f') = a^{-1} \begin{vmatrix} a & b & c & d & e & 0 & 0 \\ 0 & a & b & c & d & e & 0 \\ 0 & 0 & a & b & c & d & e \\ 4a & 3b & 2c & d & 0 & 0 & 0 \\ 0 & 4a & 3b & 2c & d & 0 & 0 \\ 0 & 0 & 4a & 3b & 2c & d & 0 \\ 0 & 0 & 0 & 4a & 3b & 2c & d \end{vmatrix}$$

$$\begin{aligned} &= b^2c^2d^2 - 4\,b^2c^3e - 4\,b^3d^3 + 18\,b^3cde - 27\,b^4e^2 - 4\,ac^3d^2 + 16\,ac^4e \\ &\quad + 18\,abcd^3 - 80\,abc^2de - 6\,ab^2d^2e + 144\,ab^2ce^2 - 27\,a^2d^4 \\ &\quad + 144\,a^2cd^2e - 128\,a^2c^2e^2 - 192\,a^2bde^2 + 256\,a^3e^3. \end{aligned} \tag{4.4}$$

Example 4.8 ($n = 4$)**.** If $f(x) = x^4 + bx^3 + cx^2 + dx + e$ is monic, then (4.4) simplifies slightly to

$$\begin{aligned} \Delta_0(f) = \Delta(f) &= R(f, f') \\ &= b^2c^2d^2 - 4\,b^2c^3e - 4\,b^3d^3 + 18\,b^3cde - 27\,b^4e^2 - 4\,c^3d^2 \\ &\quad + 16\,c^4e + 18\,bcd^3 - 80\,bc^2de - 6\,b^2d^2e + 144\,b^2ce^2 \\ &\quad - 27\,d^4 + 144\,cd^2e - 128\,c^2e^2 - 192\,bde^2 + 256\,e^3. \end{aligned}$$

Example 4.9 ($n = 4$)**.** If $f(x) = x^4 + px^2 + qx + r$ is monic and without third degree term, then (4.4) simplifies further to

$$\Delta_0(f) = \Delta(f) = -4\,p^3q^2 - 27\,q^4 + 16\,p^4r + 144\,pq^2r - 128\,p^2r^2 + 256\,r^3.$$

Example 4.10. Let $f(x) = x^n + px + q$ for some $n \geq 2$. Then $f'(x) = nx^{n-1} + p$ and, using Theorem 1.14 with $h(x) = -x/n$, Theorem 1.13 and Example 1.16, at least if F has characteristic 0,

$$\begin{aligned} (-1)^{n(n-1)/2}\Delta(f) &= R_{n,n-1}(f, f') = R_{n,n-1}(x^n + px + q, nx^{n-1} + p) \\ &= R_{n,n-1}\big(p(1 - 1/n)x + q, nx^{n-1} + p\big) \\ &= (-1)^{(n-1)^2} n^{n-1} R_{1,n-1}\big(p(1 - 1/n)x + q, nx^{n-1} + p\big) \\ &= (-1)^{(n-1)} n^{n-1} \Big(n(-q)^{n-1} + p\big(p(1 - 1/n)\big)^{n-1}\Big) \\ &= n^n q^{n-1} + (-1)^{n-1}(n-1)^{n-1}p^n. \end{aligned}$$

Since the right hand side is a polynomial in p and q with integer coefficient, the final formula holds for all fields and all $n \geq 2$. Consequently,

$$\Delta_0(f) = \Delta(f) = (-1)^{(n-1)(n-2)/2}(n-1)^{n-1}p^n + (-1)^{n(n-1)/2}n^n q^{n-1}.$$

Note the special cases in Examples 4.2, 4.5 and 4.9 (with $p = 0$). The next case is the quintic in Bring's form:

$$\Delta_0(x^5 + px + q) = \Delta(x^5 + px + q) = 4^4p^5 + 5^5q^4.$$

Example 4.11. It follows from Remark 3.13 and Example 4.7 that

$$\begin{aligned}\mathbf{\Delta}^* = {} & a_0^{-2} - 4\,a_2a_1^{-2}a_0^{-1} - 4\,a_3a_2^{-2}a_1a_0^{-2} + 18\,a_3a_2^{-1}a_1^{-1}a_0^{-1} - 27\,a_3^2a_2^{-2}a_1^{-2} \\ & - 4\,a_4a_3^{-2}a_2a_0^{-2} + 16\,a_4a_3^{-2}a_2^2a_1^{-2}a_0^{-1} + 18\,a_4a_3^{-1}a_2^{-1}a_1a_0^{-2} \\ & - 80\,a_4a_3^{-1}a_1^{-1}a_0^{-1} - 6\,a_4a_2^{-2}a_0^{-1} + 144\,a_4a_2^{-1}a_1^{-2} \\ & - 27\,a_4^2a_3^{-2}a_2^{-2}a_1^2a_0^{-2} + 144\,a_4^2a_3^{-2}a_2^{-1}a_0^{-1} - 128\,a_4^2a_3^{-2}a_1^{-2} \\ & - 192\,a_4^2a_3^{-1}a_2^{-2}a_1^{-1} + 256\,a_4^3a_3^{-2}a_2^{-2}a_1^{-2}a_0 + \dots, \end{aligned} \tag{4.5}$$

where the omitted terms have at least one factor a_k with $k \geq 5$. In particular, if $f(x) = a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$ has degree at most 4, this formula without "$\dots$" is an exact formula for $\Delta^*(f)$. Setting $a_4 = 0$ we find that if $\deg(f) \leq 3$, then

$$\Delta^*(f) = a_0^{-2} - 4\,a_2a_1^{-2}a_0^{-1} - 4\,a_3a_2^{-2}a_1a_0^{-2} + 18\,a_3a_2^{-1}a_1^{-1}a_0^{-1} - 27\,a_3^2a_2^{-2}a_1^{-2},$$

which is equivalent to (4.3). Similarly, setting also $a_3 = 0$, if $\deg(f) \leq 2$, then

$$\Delta^*(f) = a_0^{-2} - 4\,a_2a_1^{-2}a_0^{-1},$$

which is equivalent to (4.1).

5. Discriminants for real polynomials

If f is real and of degree n, then its n roots in $\mathbb{C}$ consist of $n-2\nu$ real roots and ν pairs $\xi_i, \overline{\xi_i}$ of complex (non-real) roots, for some ν with $0 \leq \nu \leq n/2$.

Theorem 5.1. *If f is a real polynomial of degree $n \geq 1$ with $n - 2\nu$ real roots and ν pairs of complex (non-real) roots, and all roots are distinct and thus $\Delta(f) \neq 0$, then*

$$\operatorname{sign}(\Delta(f)) = \operatorname{sign}(\Delta_0(f)) = (-1)^\nu. \tag{5.1}$$

Proof. This is easily seen directly from Definitions 3.1 and 3.2, by suitably pairing terms.

Alternatively, we may factor f into its irreducible real factors $f_1, \dots, f_{n-2\nu}$, $g_1, \dots, g_\nu$, where $\deg(f_i) = 1$ and $\deg(g_j) = 2$, and note that Theorem 3.10 and induction shows

$$\operatorname{sign}(\Delta(f)) = \prod_{i=1}^{n-2\nu} \operatorname{sign}(\Delta(f_i)) \prod_{j=1}^{\nu} \operatorname{sign}(\Delta(g_j)).$$

Further, each $\Delta(f_i) = 1$, while $\Delta(g_j) < 0$ by Definition 3.2, since g_j has two roots ξ and $\overline{\xi}$ with $(\xi - \overline{\xi})^2 < 0$. □

For example, this leads to the following classifications for low degrees. (In these examples, "complex" means "non-real".)

Example 5.2 ($n = 2$)**.** For a real quadratic polynomial f,

- $\Delta(f) > 0 \iff \Delta_0(f) > 0 \iff f$ has two distinct real roots;
- $\Delta(f) < 0 \iff \Delta_0(f) < 0 \iff f$ has no real root and two conjugate complex roots.

- $\Delta(f) = 0 \iff \Delta_0(f) = 0 \iff f$ has a double real root;

Example 5.3 ($n = 3$)**.** For a real cubic polynomial f,

- $\Delta(f) > 0 \iff \Delta_0(f) > 0 \iff f$ has 3 distinct real roots;
- $\Delta(f) < 0 \iff \Delta_0(f) < 0 \iff f$ has 1 real root and 2 conjugate complex roots.
- $\Delta(f) = 0 \iff \Delta_0(f) = 0 \iff f$ has either a triple real root, *or* one double real root and one single real root;

Example 5.4 ($n = 4$)**.** For a real quartic polynomial f,

- $\Delta(f) > 0 \iff \Delta_0(f) > 0 \iff f$ has either 4 distinct real roots, *or* 4 complex roots (in two conjugate pairs);
- $\Delta(f) < 0 \iff \Delta_0(f) < 0 \iff f$ has 2 real roots and 2 conjugate complex roots.
- $\Delta(f) = 0 \iff \Delta_0(f) = 0 \iff f$ 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.

Appendix A

Alternative proof of Theorem 3.11. Assume first $a_{n-1} \neq 0$; thus $f(x) = a_{n-1}x^{n-1} + \cdots + a_0$ has degree $n - 1$.

Let $f_\varepsilon(x) = (-\varepsilon x+1)f(x)$ for an indeterminate ε. Trivially, $\Delta(-\varepsilon x+1) = 1$, and Example 1.16 shows

$$R(-\varepsilon x + 1, f) = (-\varepsilon)^{n-1} f(1/\varepsilon) = (-1)^{n-1} f^*(\varepsilon)$$

with f^* defined by (1.9) with n replaced by $n - 1$. Hence, Theorem 3.10 yields

$$\Delta^{(n)}(f_\varepsilon) = \Delta(f_\varepsilon) = \Delta(f)(f^*(\varepsilon))^2 = \Delta^{(n-1)}(f)f^*(\varepsilon)^2.$$

Both sides are polynomials in $a_0, \dots, a_{n-1}$ and ε, so we may here put $\varepsilon = 0$ and obtain $\Delta^{(n)}(f) = \Delta^{(n-1)}(f)f^*(0)^2$. Since $f^*(0) = a_{n-1}$, this proves the result when $a_{n-1} \neq 0$. The general case follows, because both sides of (3.12) are polynomials in $a_0, \dots, a_{n-1}$. ☐

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/