

Polynomial Resultants

Henry Woody

*Department of Mathematics and Computer Science
University of Puget Sound*

Copyright ©2016 Henry Woody. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license can be found at <http://www.gnu.org/copyleft/fdl.html>

Abstract

This paper covers various aspects of the resultant of polynomials. Starting with a definition, we move to a practical method of calculating the resultant, specifically through the use of the Sylvester matrix, whose entries are the coefficients of the two polynomials, and whose determinant gives the resultant of two polynomials. We focus on whether or not two univariate polynomials have a common factor, and finding solutions to homogeneous systems of multivariate polynomials. We cover, as well, the discriminant of polynomials, which is the resultant of a polynomial and its derivative.

1 Introduction

Polynomials are covered in elementary algebra, so many people are familiar with their general properties and principles, but, as is the case with almost all topics covered in elementary algebra, they have much more going on than is shown in high school mathematics. The study of polynomials often centers around the search for roots, or zeros, of those polynomials. An integral part of the study of abstract algebra, polynomials possess interesting qualities and give rise to important topics, such as field extensions. Throughout this paper we let F be a field. We start with a simple definition.

Definition 1.1 (Polynomial). A polynomial $f(x) \in F[x]$ is defined as

$$f(x) = \prod_{i=0}^n (x - \alpha_i) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0, \quad a_n \neq 0$$

where each $a_i \in F$ is a coefficient, each α_i is a root of $f(x)$, and the degree of $f(x)$, $\deg(f(x))$, is n .

For more information about polynomials, their properties, and polynomial rings, we refer the reader to [2].

2 Resultants

We are often concerned with knowing whether or not two polynomials share a root or common factor. One way to determine this is to factor both polynomials completely and compare the sets of roots. While perhaps the most obvious method, it is not the most efficient since factoring polynomials entirely, especially high degree or multivariate polynomials, can be difficult if not impossible through algebraic methods. A better way to accomplish this is to calculate the greatest common divisor of the two polynomials using the Euclidean algorithm, but this requires that the polynomials be in a Euclidean domain. This is the case for polynomials over fields, however not all polynomials are defined over fields and not all polynomial rings are Euclidean domains. Thus we would like a method of determining if two polynomials share a common factor that will work efficiently for any polynomial. Resultants satisfy these criteria.

Definition 2.1 (Resultant). Given two polynomials $f(x) = a_n x^n + \dots + a_1 x + a_0$, $g(x) = b_m x^m + \dots + b_1 x + b_0 \in F[x]$, their resultant relative to the variable x is a polynomial over the field of coefficients of $f(x)$ and $g(x)$, and is defined as

$$\text{Res}(f, g, x) = a_n^m b_m^n \prod_{i,j} (\alpha_i - \beta_j),$$

where $f(\alpha_i) = 0$ for $1 \leq i \leq n$, and $g(\beta_j) = 0$ for $1 \leq j \leq m$.

There are a few things to take away from this definition. First, $\text{Res}(f, g, x)$ is an element of F . Second, $\text{Res}(f, g, x) = a_n^m \prod_i g(\alpha_i) = (-1)^{nm} b_m^n \prod_j f(\beta_j)$, which is to say the resultant is the product of either polynomial evaluated at each root, including multiplicity, of the other polynomial. Finally, the reader should note what happens in the case that the two polynomials share a root, as outlined in the following lemma.

Lemma 2.2. *The resultant of $f(x)$ and $g(x)$ is equal to zero if and only if the two polynomials have a root in common.*

Proof. To see this, suppose γ is a root shared by both polynomials, then one term of the product is $(\gamma - \gamma) = 0$, hence the whole product vanishes. Conversely, suppose $\text{Res}(f, g, x) = 0$, then, since F is a field and therefore has no zero divisors, at least one term of the product zero. Suppose that term is $(\alpha_k - \beta_l)$, this implies $\alpha_k = \beta_l$, so this is a root shared by $f(x)$ and $g(x)$. \square

It is clear that the resultant of two polynomials will tell us whether or not two polynomials share a root. However, by this definition, calculating the resultant requires us to factor each polynomial completely in order to find roots, and doing so would render the resultant redundant. We will now work toward a more efficient way of determining the resultant.

Lemma 2.3. *Let $f(x), g(x) \in F[x]$ have degrees n and m , both greater than zero, respectively. Then $f(x)$ and $g(x)$ have a non-constant common factor if and only if there exist nonzero polynomials $A(x), B(x) \in F[x]$ such that $\deg(A(x)) \leq m - 1$, $\deg(B(x)) \leq n - 1$ and $A(x)f(x) + B(x)g(x) = 0$.*

Proof. Assume $f(x)$ and $g(x)$ have a common, non-constant, factor $h(x) \in F[x]$. Then $f(x) = h(x)f_1(x)$ and $g(x) = h(x)g_1(x)$, for some $f_1(x), g_1(x) \in F[x]$.

Consider,

$$g_1(x)f(x) + (-f_1(x))g(x) = g_1(x)(h(x)f_1(x)) - f_1(x)(h(x)g_1(x)) = 0.$$

Notice $\deg(g_1(x)) \leq m - 1$ and $\deg(f_1(x)) \leq n - 1$ since $h(x)$ has at least degree 1, therefore these are the polynomials we seek. Note $A(x)$ is equal to the product of factors of $g(x)$ that are not shared by $f(x)$ and $B(x)$ is the product of factors of $f(x)$ that $g(x)$ does not share (multiplied by -1).

We will prove the left implication by contradiction. First we assume the polynomials $A(x)$ and $B(x)$ that satisfy the criteria above exist. Now suppose, to the contrary of the lemma, that $f(x)$ and $g(x)$ have no non-constant common factors, then $\gcd(f(x), g(x)) = 1$. Hence there exist polynomials $r(x)$ and $s(x)$ in $F[x]$ such that $r(x)f(x) + s(x)g(x) = 1$. We also know that $A(x)f(x) + B(x)g(x) = 0 \implies A(x)f(x) = -B(x)g(x)$. Then

$$\begin{aligned} A(x) &= 1A(x) \\ &= (r(x)f(x) + s(x)g(x))A(x) \\ &= r(x)f(x)A(x) + s(x)g(x)A(x) \\ &= r(x)(-B(x)g(x)) + s(x)g(x)A(x) \\ &= (s(x)A(x) - r(x)B(x))g(x) \end{aligned}$$

Since $A(x) \neq 0$, we know $(s(x)A(x) - r(x)B(x))g(x) \neq 0$. We also know $\deg(g(x)) = m$, which means the degree of $A(x)$ is at least m , but this is a contradiction since $A(x)$ was defined to have degree strictly less than m . \square

Lemma 2.3 can be translated from polynomials to the integers. If we consider the irreducible polynomial factors of polynomials as prime factors of integers, then if we have two integers a and b that share at least one prime factor, we can find two other integers c and d such that $ac + bd = 0$. Then c is the product of prime factors of a that are not shared with b , and d is the product of prime factors that b has but a lacks, and one will be negative if necessary.

We now look more in depth at the equation $A(x)f(x) + B(x)g(x) = 0$ from Lemma 2.3, and let

$$\begin{aligned} f(x) &= a_n x^n + \dots + a_1 x + a_0, \quad a_n \neq 0 & g(x) &= b_m x^m + \dots + b_1 x + b_0, \quad b_m \neq 0 \\ A(x) &= c_{m-1} x^{m-1} + \dots + c_1 x + c_0, & B(x) &= d_{n-1} x^{n-1} + \dots + d_1 x + d_0. \end{aligned}$$

By the definition of polynomial equality, the equation $A(x)f(x) + B(x)g(x) = 0$ is a homogeneous system of $n + m$ equations in $n + m$ variables, which are the c_i and d_j . This system is shown below.

$$\begin{aligned} a_n c_{m-1} + b_m d_{n-1} &= 0 && \text{(coefficients of } x^{n+m-1}) \\ a_n c_{m-2} + a_{n-1} c_{m-1} + b_m d_{n-2} + b_{m-1} d_{n-1} &= 0 && \text{(coefficients of } x^{n+m-2}) \\ a_n c_{m-3} + a_{n-1} c_{m-2} + a_{n-2} c_{m-1} \\ &+ b_m d_{n-3} + b_{m-1} d_{n-2} + b_{m-2} d_{n-1} &= 0 && \text{(coefficients of } x^{n+m-3}) \\ &&& \vdots \\ a_0 c_0 + b_0 d_0 &= 0 && \text{(coefficients of } x^0) \end{aligned}$$

Theorem 2.4 gives us a way to determine whether or not two polynomials have a common factor and all we need to know is the coefficients of those polynomials. This is a powerful result because of how simple it is to set up. So now we can easily tell if two polynomials share a factor, but the information conveyed by the resultant is binary, meaning it will tell us only if a factor is shared or not, but no more information than this. The following theorem gives us a way to find more information regarding the common factor of two polynomials, still without using the Euclidean algorithm.

Theorem 2.7. *If $f(x)$ is the characteristic polynomial of a square matrix M , and $g(x)$ is any polynomial, then the degree of the common factor of $f(x)$ and $g(x)$ is the nullity of the matrix $g(M)$.*

Proof. Proof of this theorem can be found in [4].

This theorem may seem as though it is only applicable in a narrow range of cases, but it is more general than it appears. For any monic polynomial $f(x)$, we can construct a matrix M so that $f(x)$ is the characteristic polynomial. Note also that if we start with two polynomials that are both not monic, we can simply factor out the leading coefficient from one of the polynomials, which will not affect the degree of the common factor. Of course this requires division in the ring of the coefficients, so this theorem will only hold for polynomials over a division ring, which is still less restrictive than a Euclidean domain, or polynomials over any ring if at least one of those polynomials is monic.

Let $f(x) = x^n + a_{n-1}x^{n-1} \dots + a_1x + a_0$, then the square matrix M is given below.

$$M = \begin{bmatrix} -a_{n-1} & -a_{n-2} & \cdots & -a_1 & -a_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

This method is easily implemented and will tell us the degree of the polynomial factor shared by two polynomials. It does not, however, actually give us the common factor, but if we are working in a Euclidean domain, the Euclidean algorithm can be used to find the greatest common factor.

3 Applications

We have already seen the most straightforward application of the resultant, namely that it indicates whether or not two polynomials share any non-constant factors, but we will now introduce two more. Not surprisingly, both of these focus on roots of polynomials.

3.1 Discriminant

The first application involves the discriminant of a polynomial and the resultant's connection to it. The discriminant, D , of a polynomial gives some insight into the nature of the polynomial's roots. For example the discriminant of a quadratic of the

form $f(x) = ax^2 + bx + c \in \mathbb{R}[x]$ is $D = b^2 - 4ac$. In this case, if D is positive, then $f(x)$ has two distinct real roots. If $D = 0$, then $f(x)$ has one real root with multiplicity 2. If D is negative, then $f(x)$ has no real roots, but has two complex roots that are conjugate. For higher degree polynomials, we can tell if a polynomial has a multiple root, meaning a root with multiplicity greater than 1, if the discriminant vanishes.

A polynomial of degree n is separable if it has n distinct roots in its splitting field. It is also the case that a polynomial is separable if the polynomial and its derivative are relatively prime. We can related these ideas to the discriminant of a polynomial.

Lemma 3.1. *A polynomial $f(x) \in F[x]$ is separable if and only if its discriminant is nonzero.*

Since the discriminant is zero only if a polynomial has a multiple root, and, equivalently if the polynomial and its derivative share a common factor, there is evidence to suggest the discriminant and the resultant are in some way related.

Lemma 3.2. *A polynomial $f(x) \in F[x]$ has a zero discriminant if and only if $\text{Res}(f, f', x) = 0$, where $f'(x)$ is the formal derivative of $f(x)$.*

The previous lemma indicates that the resultant and the discriminant are closely connected, in fact the discriminant is a multiple of a specific kind of resultant, specifically that of a polynomial and its derivative, as shown in the following definition.

Definition 3.3. For a polynomial $f(x) \in F[x]$, where $f(x) = a_n x^n + \dots + a_1 x + a_0$, the discriminant is given by

$$D = \frac{(-1)^{n(n-1)/2}}{a_n} \text{Res}(f, f', x),$$

where $f'(x)$ is the derivative of $f(x)$.

At this point, it should not come as much of a surprise that the determinant is defined in terms of the resultant. The determinant is zero if $f(x)$ and $f'(x)$ share a root, which is only possible if $f(x)$ has a multiple root. In fact, the term determinant was coined by James Joseph Sylvester, for whom the Sylvester matrix is named.

Example 3.4. Let $f(x) = ax^2 + bx + c$, then $f'(x) = 2ax + b$, thus we have the equation for the determinant as follows

$$\begin{aligned} D &= \frac{(-1)^{2(2-1)/2}}{a} \begin{vmatrix} a & 2a & 0 \\ b & b & 2a \\ c & 0 & b \end{vmatrix} = \frac{-1}{a} (a(b^2) - b(2ab) + c(4a^2)) \\ &= \frac{-1}{a} (ab^2 - 2ab^2 + 4a^2c) \\ &= \frac{-1}{a} (-ab^2 + 4a^2c) \\ &= b^2 - 4ac, \end{aligned}$$

which is the explicit formula for the determinant of a quadratic.

Finally we conclude this subsection with a corollary that connects determinants, resultants, and the study of fields.

Corollary 3.5. *A polynomial $f(x) \in F[x]$ is separable if and only if $\text{Res}(f, f', x) \neq 0$. Equivalently, $f(x)$ is separable if and only if $\text{Syl}(f, f', x)$ is nonsingular.*

3.2 Elimination

So far we have been working exclusively with polynomials in one variable, in other words those in $F[x]$, but multivariate polynomials deserve consideration as well, and they are often more difficult to analyze than their univariate cousins. One important application of the resultant is in elimination theory, which focuses on eliminating variables from systems of multivariate polynomials. As we have seen in previous examples, and from the definition, the resultant takes two polynomials in the variable x and gives one polynomial with no variables. In multiple variables this outcome is generalized to remove one variable. First we look at an example.

Example 3.6. Let $f(x, y) = x^2y + x^2 + 3x - 1$, $g(x, y) = xy^2 + y - 5 \in F[x, y]$. We will examine the resultant of these two polynomials by considering them as polynomials in x with coefficients that are polynomials in y . We compute the resultant relative to x through the Sylvester matrix.

$$\begin{aligned} \text{Res}(f, g, x) &= \begin{vmatrix} y+1 & y^2 & 0 \\ 3 & y-5 & y^2 \\ -1 & 0 & y-5 \end{vmatrix} \\ &= (y+1)(y-5)^2 - 3y^2(y-5) + (-1)y^4 \\ &= -y^4 - 2y^3 + 6y^2 + 15y + 25 \end{aligned}$$

The resultant of these two bivariate polynomials is a single univariate polynomial, so the variable x has been eliminated. This resulting polynomial shares properties with its parents, $f(x, y)$ and $g(x, y)$, but is easier to analyze than its parents.

The example above illustrates an interesting property of the resultant of two multivariate polynomials. Whereas in the single variable case the resultant is a polynomial in the coefficients of the two starting polynomials, in the two variable case, the resultant relative to one variable is a polynomial in the other variable. This follows from the fact that $F[x, y] = F[y][x]$, which is to say the polynomial ring $F[x, y]$ of polynomials with coefficients from F and with variables x and y can be thought of as the polynomial ring $F[y][x]$, which has polynomials with coefficients from $F[y]$, meaning polynomials in y , in the variable x . Hence the placement of the x in $\text{Res}(f, g, x)$ to indicate the variable to be eliminated.

Since two polynomials share a root if and only if their resultant is zero, we take the resultant polynomial and set it equal to zero. By finding roots of the resultant relative to x , we are finding values of y that will make the resultant relative to x zero, therefore making the two polynomials share a non-constant factor in the variable x .

After taking the resultant of two polynomials $f(x, y)$ and $g(x, y)$ relative to x in $F[x, y]$, and solving for roots of the resulting polynomial, we can take the resultant of these polynomials again, but this time relative to y and solve for this resultant's roots. Now we have two sets of partial solutions, and we can take combinations of solutions to $\text{Res}(f, g, x)(y) = 0$ and $\text{Res}(f, g, y)(x) = 0$ and test them in $f(x, y)$ and $g(x, y)$. Alternatively, we can evaluate $f(x, y)$ and $g(x, y)$ at each of the partial solutions given by the first resultant. This will eliminate the variable y , and we can then attempt to find roots of each of the simplified polynomials in $F[x]$. Neither of these is an elegant solution, but they are solutions nonetheless.

Example 3.7. Suppose $f(x, y) = x^2y^2 - 25x^2 + 9$ and $g(x, y) = 4x + y$ are two polynomials in $F[x, y]$. We now employ the resultant in order to find a common root of $f(x, y)$ and $g(x, y)$.

$$\begin{aligned} \text{Res}(f, g, x) &= \begin{vmatrix} y^2 - 25 & 4 & 0 \\ 0 & y & 4 \\ 9 & 0 & y \end{vmatrix} = y^4 - 25y^2 + 144 \\ \text{Res}(f, g, y) &= \begin{vmatrix} x^2 & 1 & 0 \\ 0 & 4x & 1 \\ -25x^2 + 9 & 0 & 4x \end{vmatrix} = 16x^4 - 25x^2 + 9 \end{aligned}$$

The four roots of $\text{Res}(f, g, x)$ are $y = \pm 3, \pm 4$, and those of $\text{Res}(f, g, y)$ are $x = \pm \frac{3}{4}, \pm 1$. By testing each of the 16 possible combinations of partial solutions, we find that the solutions to the homogeneous system

$$\begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0 \end{aligned}$$

are $(x, y) = (1, -4), (-1, 4), (\frac{3}{4}, -3), (-\frac{3}{4}, 3)$.

So we can use resultants to simplify polynomial systems in multiple variables, but how many polynomials are required to actually solve a system in n variables? In order to answer this question, we must first consider the manner in which homogeneous polynomial systems can be solved using the resultant for polynomials in more than two variables.

Variable elimination can be generalized further from polynomials in two variables to polynomials in n variables. If we consider the resultant as a function explicitly, then $\text{Res}_i : F[x_1, \dots, x_n] \times F[x_1, \dots, x_n] \rightarrow F[x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n]$, where Res_i is the resultant relative to the variable x_i . In other words, the resultant takes two polynomials in n variables and returns one polynomial in $n - 1$ variables. If we have a sufficient number of polynomials to start with, we can continue to eliminate variables until there is only one remaining, where it is significantly easier to solve for roots.

In order to solve a homogeneous polynomial system in n variables, n polynomials are required. Any fewer than n polynomials will cause the resultants to become zero before the univariate stage. Solving a homogeneous system of n polynomials in n variables in a manner similar to Example 3.7 is possible, however it can become extremely computationally expensive, especially in more than three variables. Since many determinants will be taken, the powers on the other variables can get quite large, thus we can end up with a univariate polynomial of high degree (5 or greater) which may not be solvable by radicals. In this case, an approximate computational method must be employed. For example, for a polynomial in four variables, where each variable has a power of at most two, the single variable polynomial obtained through multiple resultants can easily have degree greater than 100.

For polynomials in more than three variables, it is more practical to eliminate variables using the multivariate resultant, which takes n polynomials in n variables as an argument, rather than just two univariate polynomials. Note that by univariate here, we mean the single variable resultant only considers one variable, so a polynomial can be in multiple variables, but will be considered as a polynomial in one variable with

coefficients that are themselves polynomials. In combination with Gröbner bases, the multivariate resultant is one of the main tools utilized in elimination theory. For more on the multivariate resultant we refer the reader to [3].

Returning to the single variable resultant, assuming we have n polynomials in n variables, we now have a new question. Can we even be sure that we will find a common root of these polynomials? The next theorem answers this question in the affirmative.

Theorem 3.8. *If $(\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_n)$ is a solution to a homogeneous system of polynomials in $F[x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n]$ obtained by taking resultants of polynomials in $F[x_1, \dots, x_n]$ with respect to x_i , then there exists $\alpha_i \in E$, where E is the field in which all polynomials in the system split, such that $(\alpha_1, \dots, \alpha_i, \dots, \alpha_n)$ is a solution to the system in $F[x_1, \dots, x_n]$.*

Proof. Proof can be found in [1].

We can take Theorem 3.8 and essentially work from the bottom up. So if we have the appropriate number of polynomials in $F[x_1, \dots, x_n]$ then we can eliminate all variables except x_1 and find roots of the polynomial $f_1(x_1) \in F[x_1]$, which was obtained through repeated resultants. Then let α_1 be a root of this polynomial and consider the partially eliminated system with only x_1 and x_2 , i.e. polynomials in $F[x_1, x_2]$, where we know there exists an α_2 such that (α_1, α_2) is a root shared by all polynomials in the system in $F[x_1, x_2]$. We then continue this process until we have a full solution. Example 3.7 illustrates this result as there is one full solution for each of the partial solutions given by the each of the resultant polynomials.

4 Conclusion

What started as a simple question - do two polynomials share a root? - has led to some interesting results. The resultant has several different forms, each of which enhance our understanding of its properties and how it can be used. Considering the resultant as the determinant of the Sylvester matrix gives us a method for computing it. This method, along with Theorem 2.7, gives us some information about shared factors of two polynomials, without having to compute the greatest common factor using the Euclidean algorithm. This is useful if we are only interested in some properties of the shared factors of two polynomials, but do not actually need to know the greatest common factor explicitly. The form of the resultant given in Definition 2.1, or the equivalent definitions given below it, lends itself to a more insightful conceptual understanding of what the resultant is, and why it is equal to zero if the two polynomials have a common factor.

The applications of the resultant are both interesting and far-reaching. The resultant of a polynomial and its derivative is a multiple of the determinant of that polynomial. The determinant gives information about the nature of one polynomial's roots rather than about the roots of two polynomials together. The second application covered in this paper, variable elimination, shows the power of the resultant. The resultant can be employed to reshape complicated polynomial systems in multiple variables into univariate polynomials, which are significantly easier to analyze and solve for roots. In this application, the resultant represents a common theme in mathematics - taking complicated objects and ideas and translating them into versions that are simpler and better understood.

References

- [1] Cox, David, John Little, and Donal O'Shea. *Ideals, Varieties, and Algorithms*. 3rd ed. New York: Springer, 2007. Print. ISBN-10: 0-387-35650-9
- [2] Judson, Thomas. *Abstract Algebra: Theory and Applications*. 2015 ed. Ann Arbor, MI: Orthogonal Publishing, 2015. Print. ISBN 978-0-9898975-9-4
- [3] Kalorkoti, K. "On Macaulay's Form of the Resultant," *School of Informatics, University of Edinburgh*. Web. 25 April 2016.
- [4] Parker, W. V. "The Degree of the Highest Common Factor of Two Polynomials," *The American Mathematical Monthly* 42(3) (1935): 164-166. Web. 12 April 2016.
- [5] Sylvester, J.J. "On a remarkable discovery in the Theory of Canonical Forms and of Hyperdeterminants," *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science* 4 (1851): 391-410. Print.
- [6] Van Der Waerden, B. L. *Modern Algebra*. Vol. I. Trans. Fred Blum. New York: Frederick Ungar Publishing Co., 1953. Print.
- [7] Van Der Waerden, B. L. *Modern Algebra*. Vol. II. Trans. Theodore J. Benac. New York: Frederick Ungar Publishing Co., 1950. Print.