Chapter 3

Resultants

In Chapter 2, we saw how Gröbner bases can be used in Elimination Theory. An alternate approach to the problem of elimination is given by *resultants*. The resultant of two polynomials is well known and is implemented in many computer algebra systems. In this chapter, we will review the properties of the resultant and explore its generalization to several polynomials in several variables. This *multipolynomial resultant* can be used to eliminate variables from three or more equations and, as we will see at the end of the chapter, it is a surprisingly powerful tool for finding solutions of equations.

§1 The Resultant of Two Polynomials

Given two polynomials $f, g \in k[x]$ of positive degree, say

$$
\begin{aligned}
f &= a_0 x^l + \cdots + a_l, \quad a_0 \neq 0, \quad l > 0 \\
g &= b_0 x^m + \cdots + b_m, \quad b_0 \neq 0, \quad m > 0.
\end{aligned}
\tag{1.1}
$$

Then the *resultant* of f and g, denoted $\operatorname{Res}(f,g)$, is the $(l+m)\times(l+m)$ determinant

$$
\operatorname{Res}(f,g) = \det \underbrace{\begin{pmatrix}
a_0 & & & & b_0 & & & \\
a_1 & a_0 & & & b_1 & b_0 & & \\
a_2 & a_1 & \ddots & & b_2 & b_1 & \ddots & \\
\vdots & a_2 & \ddots & a_0 & \vdots & b_2 & \ddots & b_0 \\
a_l & \vdots & \ddots & a_1 & b_m & \vdots & \ddots & b_1 \\
 & a_l & & a_2 & & b_m & & b_2 \\
 & & \ddots & \vdots & & & \ddots & \vdots \\
 & & & a_l & & & & b_m
\end{pmatrix}}_{m \text{ columns} \qquad\qquad l \text{ columns}}
\tag{1.2}
$$

where the blank spaces are filled with zeros. When we want to emphasize the dependence on x, we will write $\mathrm{Res}(f,g,x)$ instead of $\mathrm{Res}(f,g)$. As a simple example, we have

$$
(1.3)\quad \mathrm{Res}(x^3+x-1, 2x^2+3x+7) = \det\begin{pmatrix} 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 3 & 2 & 0 \\ 1 & 0 & 7 & 3 & 2 \\ -1 & 1 & 0 & 7 & 3 \\ 0 & -1 & 0 & 0 & 7 \end{pmatrix} = 159.
$$

Exercise 1. Show that $\mathrm{Res}(f,g) = (-1)^{lm}\mathrm{Res}(g,f)$. Hint: What happens when you interchange two columns of a determinant?

Three basic properties of the resultant are:

- (Integer Polynomial) $\mathrm{Res}(f,g)$ is an integer polynomial in the coefficients of f and g.
- (Common Factor) $\mathrm{Res}(f,g) = 0$ if and only if f and g have a nontrivial common factor in $k[x]$.
- (Elimination) There are polynomials $A, B \in k[x]$ such that $A\,f + B\,g = \mathrm{Res}(f,g)$. The coefficients of A and B are integer polynomials in the coefficients of f and g.

Proofs of these properties can be found in [CLO], Chapter 3, §5. The Integer Polynomial property says that there is a polynomial

$$
\mathrm{Res}_{l,m} \in \mathbb{Z}[u_0,\ldots,u_l,v_0,\ldots,v_m]
$$

such that if f, g are as in (1.1), then

$$
\mathrm{Res}(f,g) = \mathrm{Res}_{l,m}(a_0,\ldots,a_l,b_0,\ldots,b_m).
$$

Over the complex numbers, the Common Factor property tells us that $f, g \in \mathbb{C}[x]$ have a common root if and only if their resultant is zero. Thus (1.3) shows that x^3+x-1 and $2x^2+3x+7$ have no common roots in $\mathbb{C}$ since $159 \neq 0$, even though we don't know the roots themselves.

To understand the Elimination property, we need to explain how resultants can be used to eliminate variables from systems of equations. As an example, consider the equations

$$
\begin{aligned}
f &= xy - 1 = 0\\
g &= x^2 + y^2 - 4 = 0.
\end{aligned}
$$

Here, we have two variables to work with, but if we regard f and g as polynomials in x whose coefficients are polynomials in y, we can compute the resultant with respect to x to obtain

$$
\mathrm{Res}(f,g,x) = \det\begin{pmatrix} y & 0 & 1 \\ -1 & y & 0 \\ 0 & -1 & y^2-4 \end{pmatrix} = y^4 - 4y^2 + 1.
$$

By the Elimination property, there are polynomials $A, B \in k[x, y]$ with $A \cdot (xy - 1) + B \cdot (x^2 + y^2 - 4) = y^4 - 4y^2 + 1$. This means $\text{Res}(f, g, x)$ is in the elimination ideal $\langle f, g \rangle \cap k[y]$ as defined in §1 of Chapter 2, and it follows that $y^4 - 4y^2 + 1$ vanishes at any common solution of $f = g = 0$. Hence, by solving $y^4 - 4y^2 + 1 = 0$, we can find the y-coordinates of the solutions. Thus resultants relate nicely to what we did in Chapter 2.

Exercise 2. Use resultants to find all solutions of the above equations $f = g = 0$. Also find the solutions using $\text{Res}(f, g, y)$. In Maple, the command for resultant is `resultant`.

More generally, if f and g are *any* polynomials in $k[x, y]$ in which x appears to a positive power, then we can compute $\text{Res}(f, g, x)$ in the same way. Since the coefficients are polynomials in y, the Integer Polynomial property guarantees that $\text{Res}(f, g, x)$ is again a polynomial in y. Thus, we can use the resultant to eliminate x, and as above, $\text{Res}(f, g, x)$ is in the elimination ideal $\langle f, g \rangle \cap k[y]$ by the Elimination property. For a further discussion of the connection between resultants and elimination theory, the reader should consult Chapter 3 of [CLO] or Chapter XI of [vdW].

One interesting aspect of the resultant is that it can be expressed in many different ways. For example, given $f, g \in k[x]$ as in (1.1), suppose their roots are $\xi_1, \ldots, \xi_l$ and $\eta_1, \ldots, \eta_m$ respectively (note that these roots might lie in some bigger field). Then one can show that the resultant is given by

$$
\begin{aligned}
\text{Res}(f, g) &= a_0^m b_0^l \prod_{i=1}^{l} \prod_{j=1}^{m} (\xi_i - \eta_j) \\
&= a_0^m \prod_{i=1}^{l} g(\xi_i) \qquad (1.4) \\
&= (-1)^{lm} b_0^l \prod_{j=1}^{m} f(\eta_j).
\end{aligned}
$$

A proof of this is given in the exercises at the end of the section.

Exercise 3.

a. Show that the three products on the right hand side of (1.4) are all equal. Hint: $g = b_0(x - \eta_1) \cdots (x - \eta_m)$.

b. Use (1.4) to show that $\text{Res}(f_1 f_2, g) = \text{Res}(f_1, g)\text{Res}(f_2, g)$.

The formulas given in (1.4) may seem hard to use since they involve the roots of f or g. But in fact there is a relatively simple way to compute the above products. For example, to understand the formula $\text{Res}(f, g) = a_0^m \prod_{i=1}^{l} g(\xi_i)$, we will use the techniques of §2 of Chapter 2. Thus, consider the quotient ring $A_f = k[x]/\langle f\rangle$, and let the multiplication map m_g be defined by

$$m_g([h]) = [g]\cdot[h] = [gh] \in A_f,$$

where $[h] \in A_f$ is the coset of $h \in k[x]$. If we think in terms of remainders on division by f, then we can regard A_f as consisting of all polynomials h of degree $< l$, and under this interpretation, $m_g(h)$ is the remainder of gh on division by f. Then we can compute the resultant $\mathrm{Res}(f,g)$ in terms of m_g as follows.

(1.5) Proposition. $\mathrm{Res}(f,g) = a_0^m \det(m_g : A_f \to A_f)$.

PROOF. Note that A_f is a vector space over k of dimension l (this is clear from the remainder interpretation of A_f). Further, as explained in §2 of Chapter 2, $m_g : A_f \to A_f$ is a linear map. Recall from linear algebra that the determinant $\det(m_g)$ is defined to be the determinant of any matrix M representing the linear map m_g. Since M and m_g have the same eigenvalues, it follows that $\det(m_g)$ is the product of the eigenvalues of m_g, counted with multiplicity.

In the special case when $g(\xi_1),\ldots,g(\xi_l)$ are distinct, we can prove our result using the theory of Chapter 2. Namely, since $\{\xi_1,\ldots,\xi_l\} = \mathbf{V}(f)$, it follows from Theorem (4.5) of Chapter 2 that the numbers $g(\xi_1),\ldots,g(\xi_l)$ are the eigenvalues of m_g. Since these are distinct and A_f has dimension l, it follows that the eigenvalues have multiplicity one, so that $\det(m_g) = g(\xi_1)\cdots g(\xi_l)$, as desired. The general case will be covered in the exercises at the end of the section. $\square$

Exercise 4. For $f = x^3 + x - 1$ and $g = 2x^2 + 3x + 7$ as in (1.3), use the basis $\{1, x, x^2\}$ of A_f (thinking of A_f in terms of remainders) to show

$$\mathrm{Res}(f,g) = 1^2\det(m_g) = \det\begin{pmatrix} 7 & 2 & 3\\ 3 & 5 & -1\\ 2 & 3 & 5\end{pmatrix} = 159.$$

Note that the 3×3 determinant in this example is smaller than the 5×5 determinant required by the definition (1.2). In general, Proposition (1.5) tells us that $\mathrm{Res}(f,g)$ can be represented as an $l\times l$ determinant, while the definition of resultant uses an $(l+m)\times(l+m)$ matrix. The `getmatrix` procedure from Exercise 18 of Chapter 2, §4 can be used to construct the smaller matrix. Also, by interchanging f and g, we can represent the resultant using an $m\times m$ determinant.

For the final topic of this section, we will discuss a variation on $\mathrm{Res}(f,g)$ which will be important for §2. Namely, instead of using polynomials in the single variable x, we could instead work with *homogeneous* polynomials in variables x, y. Recall that a polynomial is homogeneous if every term has the same total degree. Thus, if $F, G \in k[x,y]$ are homogeneous polynomials of total degrees l, m respectively, then we can write

$$(1.6)\qquad \begin{aligned} F &= a_0x^l + a_1x^{l-1}y + \cdots + a_ly^l \\ G &= b_0x^m + b_1x^{m-1}y + \cdots + b_my^m. \end{aligned}$$

Note that a_0 or b_0 (or both) might be zero. Then we define $\mathrm{Res}(F, G) \in k$ using the same determinant as in (1.2).

Exercise 5. Show that $\mathrm{Res}(x^l, y^m) = 1$.

If we homogenize the polynomials f and g of (1.1) using appropriate powers of y, then we get F and G as in (1.6). In this case, it is obvious that $\mathrm{Res}(f, g) = \mathrm{Res}(F, G)$. However, going the other way is a bit more subtle, for if F and G are given by (1.6), then we can dehomogenize by setting $y = 1$, but we might fail to get polynomials of the proper degrees since a_0 or b_0 might be zero. Nevertheless, the resultant $\mathrm{Res}(F, G)$ still satisfies the following basic properties.

(1.7) Proposition. *Fix positive integers l and m.*

a. *There is a polynomial* $\mathrm{Res}_{l,m} \in \mathbb{Z}[a_0, \ldots, a_l, b_0, \ldots, b_m]$ *such that*

$$\mathrm{Res}(F, G) = \mathrm{Res}_{l,m}(a_0, \ldots, a_l, b_0, \ldots, b_m)$$

for all F, G as in (1.6).

b. *Over the field of complex numbers,* $\mathrm{Res}(F, G) = 0$ *if and only if the equations $F = G = 0$ have a solution $(x, y) \neq (0, 0)$ in $\mathbb{C}^2$ (this is called a* **nontrivial** *solution).*

Proof. The first statement is an obvious consequence of the determinant formula for the resultant. As for the second, first observe that if $(u, v) \in \mathbb{C}^2$ is a nontrivial solution, then so is $(\lambda u, \lambda v)$ for any nonzero complex number λ. We now break up the proof into three cases.

First, if $a_0 = b_0 = 0$, then note that the resultant vanishes and that we have the nontrivial solution $(x, y) = (1, 0)$. Next, suppose that $a_0 \neq 0$ and $b_0 \neq 0$. If $\mathrm{Res}(F, G) = 0$, then, when we dehomogenize by setting $y = 1$, we get polynomials $f, g \in \mathbb{C}[x]$ with $\mathrm{Res}(f, g) = 0$. Since we're working over the complex numbers, the Common Factor property implies f and g must have a common root $x = u$, and then $(x, y) = (u, 1)$ is the desired nontrivial solution. Going the other way, if we have a nontrival solution (u, v), then our assumption $a_0b_0 \neq 0$ implies that $v \neq 0$. Then $(u/v, 1)$ is also a solution, which means that u/v is a common root of the dehomogenized polynomials. From here, it follows easily that $\mathrm{Res}(F, G) = 0$.

The final case is when exactly one of a_0, b_0 is zero. The argument is a bit more complicated and will be covered in the exercises at the end of the section. □

We should also mention that many other properties of the resultant, along with proofs, are contained in Chapter 12 of [GKZ].

Additional Exercises for §1

Exercise 6. As an example of how resultants can be used to eliminate variables from equations, consider the parametric equations

$$\begin{aligned} x &= 1 + s + t + st \\ y &= 2 + s + st + t^2 \\ z &= s + t + s^2. \end{aligned}$$

Our goal is to eliminate s, t from these equations to find an equation involving only x, y, z.

a. Use Gröbner basis methods to find the desired equation in x, y, z.
b. Use resultants to find the desired equations. Hint: Let $f = 1 + s + t + st - x$, $g = 2 + s + st + t^2 - y$ and $h = s + t + s^2 - z$. Then eliminate t by computing $\mathrm{Res}(f, g, t)$ and $\mathrm{Res}(f, h, t)$. Now what resultant do you use to get rid of s?
c. How are the answers to parts a and b related?

Exercise 7. Let f, g be as in (1.1). If we divide g by f, we get $g = q\,f + r$, where $\deg(r) < \deg(g) = m$. Then, assuming that r is nonconstant, show that

$$\mathrm{Res}(f, g) = a_0^{m-\deg(r)}\mathrm{Res}(f, r).$$

Hint: Let $g_1 = g - (b_0/a_0)x^{m-l}f$ and use column operations to subtract b_0/a_0 times the first l columns in the f part of the matrix from the columns in the g part. Expanding repeatedly along the first row gives $\mathrm{Res}(f, g) = a_0^{m-\deg(g_1)}\mathrm{Res}(f, g_1)$. Continue this process to obtain the desired formula.

Exercise 8. Our definition of $\mathrm{Res}(f, g)$ requires that f, g have positive degrees. Here is what to do when f or g is constant.

a. If $\deg(f) > 0$ but g is a nonzero constant b_0, show that the determinant (1.2) still makes sense and gives $\mathrm{Res}(f, b_0) = b_0^l$.
b. If $\deg(g) > 0$ and $a_0 \neq 0$, what is $\mathrm{Res}(a_0, g)$? Also, what is $\mathrm{Res}(a_0, b_0)$? What about $\mathrm{Res}(f, 0)$ or $\mathrm{Res}(0, g)$?
c. Exercise 7 assumes that the remainder r has positive degree. Show that the formula of Exercise 7 remains true even if r is constant.

Exercise 9. By Exercises 1, 7 and 8, resultants have the following three properties: $\mathrm{Res}(f, g) = (-1)^{lm}\mathrm{Res}(g, f)$; $\mathrm{Res}(f, b_0) = b_0^l$; and $\mathrm{Res}(f, g) = a_0^{m-\deg(r)}\mathrm{Res}(f, r)$ when $g = q\,f + r$. Use these properties to describe an algorithm for computing resultants. Hint: Your answer should be similar to the Euclidean algorithm.

Exercise 10. This exercise will give a proof of (1.4).

a. Given f, g as usual, define $\operatorname{res}(f,g) = a_0^m \prod_{i=1}^l g(\xi_i)$, where $\xi_1, \ldots, \xi_l$ are the roots of f. Then show that $\operatorname{res}(f,g)$ has the three properties of resultants mentioned in Exercise 9.
b. Show that the algorithm for computing $\operatorname{res}(f,g)$ is the same as the algorithm for computing $\operatorname{Res}(f,g)$, and conclude that the two are equal for all f, g.

Exercise 11. Let $f = a_0x^l + a_1x^{l-1} + \cdots + a_l \in k[x]$ be a polynomial with $a_0 \neq 0$, and let $A_f = k[x]/\langle f \rangle$. Given $g \in k[x]$, let $m_g : A_f \to A_f$ be multiplication by g.

a. Use the basis $\{1, x, \ldots, x^{l-1}\}$ of A_f (so we are thinking of A_f as consisting of remainders) to show that the matrix of m_x is

$$C_f = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_l/a_0 \\ 1 & 0 & \cdots & 0 & -a_{l-1}/a_0 \\ 0 & 1 & \cdots & 0 & -a_{l-2}/a_0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1/a_0 \end{pmatrix}.$$

   This matrix (or more commonly, its transpose) is called the *companion matrix* of f.
b. If $g = b_0x^m + \cdots + b_m$, then explain why the matrix of m_g is given by

$$g(C_f) = b_0C_f^m + b_1C_f^{m-1} + \cdots + b_mI,$$

   where I is the $l \times l$ identity matrix. Hint: By Proposition (4.2) of Chapter 2, the map sending $g \in k[x]$ to $m_g \in M_{l\times l}(k)$ is a ring homomorphism.
c. Conclude that $\operatorname{Res}(f,g) = a_0^m \det(g(C_f))$.

Exercise 12. In Proposition (1.5), we interpreted $\operatorname{Res}(f,g)$ as the determinant of a linear map. It turns out that the original definition (1.2) of resultant has a similar interpretation. Let P_n denote the vector space of polynomials of degree $\leq n$. Since such a polynomial can be written $a_0x^n + \cdots + a_n$, it follows that $\{x^n, \ldots, 1\}$ is a basis of P_n.

a. Given f, g as in (1.1), show that if $(A, B) \in P_{m-1} \oplus P_{l-1}$, then $A\,f + B\,g$ is in P_{l+m-1}. Conclude that we get a linear map $\Phi_{f,g} : P_{m-1} \oplus P_{l-1} \to P_{l+m-1}$.
b. If we use the bases $\{x^{m-1}, \ldots, 1\}$ of P_{m-1}, $\{x^{l-1}, \ldots, 1\}$ of P_{l-1} and $\{x^{l+m-1}, \ldots, 1\}$ of P_{l+m-1}, show that the matrix of the linear map $\Phi_{f,g}$ from part a is exactly the matrix used in (1.2). Thus, $\operatorname{Res}(f,g) = \det(\Phi_{f,g})$, provided we use the above bases.
c. If $\operatorname{Res}(f,g) \neq 0$, conclude that every polynomial of degree $\leq l + m - 1$ can be written uniquely as $A\,f + B\,g$ where $\deg(A) < m$ and $\deg(B) < l$.

Exercise 13. In the text, we only proved Proposition (1.5) in the special case when $g(\xi_1), \dots, g(\xi_l)$ are distinct. For the general case, suppose $f = a_0(x-\xi_1)^{a_1}\cdots(x-\xi_r)^{a_r}$, where $\xi_1, \dots, \xi_r$ are distinct. Then we want to prove that $\det(m_g) = \prod_{i=1}^r g(\xi_i)^{a_i}$.

a. First, suppose that $f = (x-\xi)^a$. In this case, we can use the basis of A_f given by $\{(x-\xi)^{a-1}, \dots, x-\xi, 1\}$ (as usual, we think of A_f as consisting of remainders). Then show that the matrix of m_g with respect to the above basis is upper triangular with diagonal entries all equal to $g(\xi)$. Conclude that $\det(m_g) = g(\xi)^a$. Hint: Write $g = b_0x^m + \cdots + b_m$ in the form $g = c_0(x-\xi)^m + \cdots + c_{m-1}(x-\xi) + c_m$ by replacing x with $(x-\xi)+\xi$ and using the binomial theorem. Then let $x = \xi$ to get $c_m = g(\xi)$.
b. In general, when $f = a_0(x-\xi_1)^{a_1}\cdots(x-\xi_r)^{a_r}$, show that there is a well-defined map

$$A_f \longrightarrow (k[x]/\langle(x-\xi_1)^{a_1}\rangle) \oplus \cdots \oplus (k[x]/\langle(x-\xi_r)^{a_r}\rangle)$$

which preserves sums and products. Hint: This is where working with cosets is a help. It is easy to show that the map sending $[h] \in A_f$ to $[h] \in k[x]/\langle(x-\xi_i)^{a_i}\rangle$ is well-defined since $(x-\xi_i)^{a_i}$ divides f.
c. Show that the map of part b is a ring isomorphism. Hint: First show that the map is one-to-one, and then use linear algebra and a dimension count to show it is onto.
d. By considering multiplication by g on

$$(k[x]/\langle(x-\xi_1)^{a_1}\rangle) \oplus \cdots \oplus (k[x]/\langle(x-\xi_r)^{a_r}\rangle)$$

and using part a, conclude that $\det(m_g) = \prod_{i=1}^r g(\xi_i)^{a_i}$ as desired.

Exercise 14. This exercise will complete the proof of Proposition (1.7). Suppose that F, G are given by (1.6) and assume $a_0 \neq 0$ and $b_0 = \cdots = b_{r-1} = 0$ but $b_r \neq 0$. If we dehomogenize by setting $y = 1$, we get polynomials f, g of degree $l, m-r$ respectively.

a. Show that $\mathrm{Res}(F,G) = a_0^r\mathrm{Res}(f,g)$.
b. Show that $\mathrm{Res}(F,G) = 0$ if and only $F = G = 0$ has a nontrivial solution. Hint: Modify the argument given in the text for the case when a_0 and b_0 were both nonzero.

§2 Multipolynomial Resultants

In §1, we studied the resultant of two homogeneous polynomials F, G in variables x, y. Generalizing this, suppose we are given $n+1$ homogeneous polynomials $F_0, \dots, F_n$ in variables $x_0, \dots, x_n$, and assume that each F_i has positive total degree. Then we get $n+1$ equations in $n+1$ unknowns:

$$F_0(x_0, \dots, x_n) = \cdots = F_n(x_0, \dots, x_n) = 0. \tag{2.1}$$

Because the F_i are homogeneous of positive total degree, these equations always have the solution $x_0 = \cdots = x_n = 0$, which we call the *trivial* solution. Hence, the crucial question is whether there is a *nontrivial* solution. For the rest of this chapter, we will work over the complex numbers, so that a nontrivial solution will be a point in $\mathbb{C}^{n+1} \setminus \{(0, \ldots, 0)\}$.

In general, the existence of a nontrivial solution depends on the coefficients of the polynomials $F_0, \ldots, F_n$: for most values of the coefficients, there are no nontrivial solutions, while for certain special values, they exist.

One example where this is easy to see is when the polynomials F_i are all linear, i.e., have total degree 1. Since they are homogeneous, the equations (2.1) can be written in the form:

$$
\begin{aligned}
F_0 &= c_{00}x_0 + \cdots + c_{0n}x_n = 0\\
&\vdots\\
F_n &= c_{n0}x_0 + \cdots + c_{nn}x_n = 0.
\end{aligned}
\tag{2.2}
$$

This is an $(n+1) \times (n+1)$ system of linear equations, so that by linear algebra, there is a nontrivial solution if and only if the determinant of the coefficient matrix vanishes. Thus we get the *single* condition $\det(c_{ij}) = 0$ for the existence of a nontrivial solution. Note that this determinant is a polynomial in the coefficients c_{ij}.

Exercise 1. There was a single condition for a nontrivial solution of (2.2) because the number of equations $(n+1)$ equaled the number of unknowns (also $n+1$). When these numbers are different, here is what can happen.

a. If we have $r < n+1$ linear equations in $n+1$ unknowns, explain why there is *always* a nontrivial solution, no matter what the coefficients are.
b. When we have $r > n+1$ linear equations in $n+1$ unknowns, things are more complicated. For example, show that the equations

$$
\begin{aligned}
F_0 &= c_{00}x + c_{01}y = 0\\
F_1 &= c_{10}x + c_{11}y = 0\\
F_2 &= c_{20}x + c_{21}y = 0
\end{aligned}
$$

have a nontrivial solution if and only if the *three* conditions

$$
\det\begin{pmatrix} c_{00} & c_{01}\\ c_{10} & c_{11}\end{pmatrix} = \det\begin{pmatrix} c_{00} & c_{01}\\ c_{20} & c_{21}\end{pmatrix} = \det\begin{pmatrix} c_{10} & c_{11}\\ c_{20} & c_{21}\end{pmatrix} = 0
$$

are satisfied.

In general, when we have $n+1$ homogeneous polynomials $F_0, \ldots, F_n \in \mathbb{C}[x_0, \ldots, x_n]$, we get the following Basic Question: *What conditions must the coefficients of $F_0, \ldots, F_n$ satisfy in order that $F_0 = \cdots = F_n = 0$ has a nontrivial solution?* To state the answer precisely, we need to introduce some notation. Suppose that d_i is the total degree of F_i, so that F_i can be written

$$F_i = \sum_{|\alpha|=d_i} c_{i,\alpha} x^\alpha.$$

For each possible pair of indices i, α, we introduce a variable $u_{i,\alpha}$. Then, given a polynomial $P \in \mathbb{C}[u_{i,\alpha}]$, we let $P(F_0, \ldots, F_n)$ denote the number obtained by replacing each variable $u_{i,\alpha}$ in P with the corresponding coefficient $c_{i,\alpha}$. This is what we mean by a *polynomial in the coefficients of the* F_i. We can now answer our Basic Question.

(2.3) Theorem. *If we fix positive degrees $d_0, \ldots, d_n$, then there is a unique polynomial* Res $\in \mathbb{Z}[u_{i,\alpha}]$ *which has the following properties:*

a. *If $F_0, \ldots, F_n \in \mathbb{C}[x_0, \ldots, x_n]$ are homogeneous of degrees $d_0, \ldots, d_n$, then the equations (2.1) have a nontrivial solution over $\mathbb{C}$ if and only if* $\mathrm{Res}(F_0, \ldots, F_n) = 0$.
b. $\mathrm{Res}(x_0^{d_0}, \ldots, x_n^{d_n}) = 1$.
c. Res *is irreducible, even when regarded as a polynomial in* $\mathbb{C}[u_{i,\alpha}]$.

Proof. A complete proof of the existence of the resultant is beyond the scope of this book. See Chapter 13 of [GKZ] or §78 of [vdW] for proofs. At the end of this section, we will indicate some of the intuition behind the proof when we discuss the geometry of the resultant. The question of uniqueness will be considered in Exercise 5. □

We call $\mathrm{Res}(F_0, \ldots, F_n)$ the *resultant* of $F_0, \ldots, F_n$. Sometimes we write $\mathrm{Res}_{d_0,\ldots,d_n}$ instead of Res if we want to make the dependence on the degrees more explicit. In this notation, if each $F_i = \sum_{j=0}^n c_{ij} x_j$ is linear, then the discussion following (2.2) shows that

$$\mathrm{Res}_{1,\ldots,1}(F_0, \ldots, F_n) = \det(c_{ij}).$$

Another example is the resultant of two polynomials, which was discussed in §1. In this case, we know that $\mathrm{Res}(F_0, F_1)$ is given by the determinant (1.2). Theorem (2.3) tells us that this determinant is an irreducible polynomial in the coefficients of F_0, F_1.

Before giving further examples of multipolynomial resultants, we want to indicate their usefulness in applications. Let's consider the *implicitization problem*, which asks for the equation of a parametric curve or surface. For concreteness, suppose a surface is given parametrically by the equations

$$\begin{aligned} x &= f(s,t) \\ y &= g(s,t) \\ z &= h(s,t), \end{aligned} \tag{2.4}$$

where $f(s,t), g(s,t), h(s,t)$ are polynomials (not necessarily homogeneous) of total degrees d_0, d_1, d_2. There are several methods to find the equation $p(x,y,z) = 0$ of the surface described by (2.4). For example, Chapter 3 of

[CLO] uses Gröbner bases for this purpose. We claim that in many cases, multipolynomial resultants can be used to find the equation of the surface.

To use our methods, we need homogeneous polynomials, and hence we will homogenize the above equations with respect to a third variable u. For example, if we write $f(s,t)$ in the form

$$f(s,t) = f_{d_0}(s,t) + f_{d_0-1}(s,t) + \cdots + f_0(s,t),$$

where f_j is homogeneous of total degree j in s,t, then we get

$$F(s,t,u) = f_{d_0}(s,t) + f_{d_0-1}(s,t)u + \cdots + f_0(s,t)u^{d_0},$$

which is now homogeneous in s,t,u of total degree d_0. Similarly, $g(s,t)$ and $h(s,t)$ homogenize to $G(s,t,u)$ and $H(s,t,u)$, and the equations (2.4) become

$$(2.5) \qquad F(s,t,u) - xu^{d_0} = G(s,t,u) - yu^{d_1} = H(s,t,u) - zu^{d_2} = 0.$$

Note that x,y,z are regarded as coefficients in these equations.

We can now solve the implicitization problem for (2.4) as follows.

(2.6) Proposition. *With the above notation, assume that the system of homogeneous equations*

$$f_{d_0}(s,t) = g_{d_1}(s,t) = h_{d_2}(s,t) = 0$$

has only the trivial solution. Then, for a given triple $(x,y,z) \in \mathbb{C}^3$, the equations (2.4) have a solution $(s,t) \in \mathbb{C}^2$ if and only if

$$\mathrm{Res}_{d_0,d_1,d_2}(F - xu^{d_0}, G - yu^{d_1}, H - zu^{d_2}) = 0.$$

Proof. By Theorem (2.3), the resultant vanishes if and only if (2.5) has a nontrivial solution (s,t,u). If $u \neq 0$, then $(s/u, t/u)$ is a solution to (2.4). However, if $u = 0$, then (s,t) is a nontrivial solution of $f_{d_0}(s,t) = g_{d_1}(s,t) = h_{d_2}(s,t) = 0$, which contradicts our hypothesis. Hence, $u = 0$ can't occur. Going the other way, note that a solution (s,t) of (2.4) gives the nontrivial solution $(s,t,1)$ of (2.5). $\square$

Since the resultant is a polynomial in the coefficients, it follows that

$$(2.7) \qquad p(x,y,z) = \mathrm{Res}_{d_0,d_1,d_2}(F - xu^{d_0}, G - yu^{d_1}, H - zu^{d_2})$$

is a polynomial in x,y,z which, by Proposition (2.6), vanishes *precisely* on the image of the parametrization. In particular, this means that the parametrization covers *all* of the surface $p(x,y,z) = 0$, which is not true for all polynomial parametrizations—the hypothesis that $f_{d_0}(s,t) = g_{d_1}(s,t) = h_{d_2}(s,t) = 0$ has only the trivial solution is important here.

Exercise 2.

a. If $f_{d_0}(s,t) = g_{d_1}(s,t) = h_{d_2}(s,t) = 0$ has a nontrivial solution, show that the resultant (2.7) vanishes identically. Hint: Show that (2.5) always has a nontrivial solution, no matter what x,y,z are.

b. Show that the parametric equations $(x, y, z) = (st, s^2t, st^2)$ define the surface $x^3 = yz$. By part a, we know that the resultant (2.7) can't be used to find this equation. Show that in this case, it is also true that the parametrization is not onto—there are points on the surface which don't come from any s, t.

We should point out that for some systems of equations, such as

$$\begin{aligned} x &= 1 + s + t + st \\ y &= 2 + s + 3t + st \\ z &= s - t + st, \end{aligned}$$

the resultant (2.7) vanishes identically by Exercise 2, yet a resultant can still be defined—this is one of the *sparse resultants* which we will consider in Chapter 7.

One difficulty with multipolynomial resultants is that they tend to be *very* large expressions. For example, consider the system of equations given by 3 quadratic forms in 3 variables:

$$\begin{aligned} F_0 &= c_{01}x^2 + c_{02}y^2 + c_{03}z^2 + c_{04}xy + c_{05}xz + c_{06}yz = 0 \\ F_1 &= c_{11}x^2 + c_{12}y^2 + c_{13}z^2 + c_{14}xy + c_{15}xz + c_{16}yz = 0 \\ F_2 &= c_{21}x^2 + c_{22}y^2 + c_{23}z^2 + c_{24}xy + c_{25}xz + c_{26}yz = 0. \end{aligned}$$

Classically, this is a system of "three ternary quadrics". By Theorem (2.3), the resultant $\mathrm{Res}_{2,2,2}(F_0, F_1, F_2)$ vanishes exactly when this system has a nontrivial solution in x, y, z.

The polynomial $\mathrm{Res}_{2,2,2}$ is very large: it has 18 variables (one for each coefficient c_{ij}), and the theory of §3 will tell us that it has total degree 12. Written out in its full glory, $\mathrm{Res}_{2,2,2}$ has 21,894 terms (we are grateful to Bernd Sturmfels for this computation). Hence, to work effectively with this resultant, we need to learn some more compact ways of representing it. We will study this topic in more detail in §3 and §4, but to whet the reader's appetite, we will now give one of the many interesting formulas for $\mathrm{Res}_{2,2,2}$.

First, let J denote the Jacobian determinant of F_0, F_1, F_2:

$$J = \det \begin{pmatrix} \dfrac{\partial F_0}{\partial x} & \dfrac{\partial F_0}{\partial y} & \dfrac{\partial F_0}{\partial z} \\ \dfrac{\partial F_1}{\partial x} & \dfrac{\partial F_1}{\partial y} & \dfrac{\partial F_1}{\partial z} \\ \dfrac{\partial F_2}{\partial x} & \dfrac{\partial F_2}{\partial y} & \dfrac{\partial F_2}{\partial z} \end{pmatrix},$$

which is a cubic homogeneous polynomial in x, y, z. This means that the partial derivatives of J are quadratic and hence can be written in the following form:

$$\begin{aligned}\frac{\partial J}{\partial x} &= b_{01}x^2 + b_{02}y^2 + b_{03}z^2 + b_{04}xy + b_{05}xz + b_{06}yz\\ \frac{\partial J}{\partial y} &= b_{11}x^2 + b_{12}y^2 + b_{13}z^2 + b_{14}xy + b_{15}xz + b_{16}yz\\ \frac{\partial J}{\partial z} &= b_{21}x^2 + b_{22}y^2 + b_{23}z^2 + b_{24}xy + b_{25}xz + b_{26}yz.\end{aligned}$$

Note that each b_{ij} is a cubic polynomial in the c_{ij}. Then, by a classical formula of Salmon (see [Sal], Art. 90), the resultant of three ternary quadrics is given by the 6×6 determinant

$$(2.8)\quad \mathrm{Res}_{2,2,2}(F_0, F_1, F_2) = \frac{-1}{512}\det\begin{pmatrix} c_{01} & c_{02} & c_{03} & c_{04} & c_{05} & c_{06}\\ c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16}\\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26}\\ b_{01} & b_{02} & b_{03} & b_{04} & b_{05} & b_{06}\\ b_{11} & b_{12} & b_{13} & b_{14} & b_{15} & b_{16}\\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} & b_{26}\end{pmatrix}.$$

Exercise 3.

a. Use (2.8) to explain why $\mathrm{Res}_{2,2,2}$ has total degree 12 in the variables $c_{01}, \ldots, c_{26}$.

b. Why is the fraction $-1/512$ needed in (2.8)? Hint: Compute the resultant $\mathrm{Res}_{2,2,2}(x^2, y^2, z^2)$.

c. Use (2.7) and (2.8) to find the equation of the surface defined by the equations

$$\begin{aligned} x &= 1 + s + t + st\\ y &= 2 + s + st + t^2\\ z &= s + t + s^2.\end{aligned}$$

Note that $st = st + t^2 = s^2 = 0$ has only the trivial solution, so that Proposition (2.6) applies. You should compare your answer to Exercise 6 of §1.

In §4 we will study the general question of how to find a formula for a given resultant. Here is an example which illustrates one of the methods we will use. Consider the following system of three homogeneous equations in three variables:

$$(2.9)\qquad \begin{aligned} F_0 &= a_1x + a_2y + a_3z = 0\\ F_1 &= b_1x + b_2y + b_3z = 0\\ F_2 &= c_1x^2 + c_2y^2 + c_3z^2 + c_4xy + c_5xz + c_6yz = 0.\end{aligned}$$

Since F_0 and F_1 are linear and F_2 is quadratic, the resultant involved is $\mathrm{Res}_{1,1,2}(F_0, F_1, F_2)$. We get the following formula for this resultant.

(2.10) Proposition. $\mathrm{Res}_{1,1,2}(F_0, F_1, F_2)$ *is given by the polynomial*

$$
\begin{aligned}
&a_1^2b_2^2c_3 - a_1^2b_2b_3c_6 + a_1^2b_3^2c_2 - 2a_1a_2b_1b_2c_3 + a_1a_2b_1b_3c_6\\
&+ a_1a_2b_2b_3c_5 - a_1a_2b_3^2c_4 + a_1a_3b_1b_2c_6 - 2a_1a_3b_1b_3c_2 - a_1a_3b_2^2c_5\\
&+ a_1a_3b_2b_3c_4 + a_2^2b_1^2c_3 - a_2^2b_1b_3c_5 + a_2^2b_3^2c_1 - a_2a_3b_1^2c_6\\
&+ a_2a_3b_1b_2c_5 + a_2a_3b_1b_3c_4 - 2a_2a_3b_2b_3c_1 + a_3^2b_1^2c_2 - a_3^2b_1b_2c_4 + a_3^2b_2^2c_1.
\end{aligned}
$$

PROOF. Let R denote the above polynomial, and suppose we have a nontrivial solution (x, y, z) of (2.9). We will first show that this forces a slight variant of R to vanish. Namely, consider the six equations

$$x \cdot F_0 = y \cdot F_0 = z \cdot F_0 = y \cdot F_1 = z \cdot F_1 = 1 \cdot F_2 = 0, \tag{2.11}$$

which we can write as

$$
\begin{array}{ccccccccccccc}
a_1x^2 & + & 0 & + & 0 & + & a_2xy & + & a_3xz & + & 0 & = & 0\\
0 & + & a_2y^2 & + & 0 & + & a_1xy & + & 0 & + & a_3yz & = & 0\\
0 & + & 0 & + & a_3z^2 & + & 0 & + & a_1xz & + & a_2yz & = & 0\\
0 & + & b_2y^2 & + & 0 & + & b_1xy & + & 0 & + & b_3yz & = & 0\\
0 & + & 0 & + & b_3z^3 & + & 0 & + & b_1xz & + & b_2yz & = & 0\\
c_1x^2 & + & c_2y^2 & + & c_3z^2 & + & c_4xy & + & c_5xz & + & c_6yz & = & 0.
\end{array}
$$

If we regard $x^2, y^2, z^2, xy, xz, yz$ as "unknowns", then this system of six linear equations has a nontrivial solution, which implies that the determinant D of its coefficient matrix is zero. Using a computer, one easily checks that the determinant is $D = -a_1R$.

Thinking geometrically, we have proved that in the 12 dimensional space $\mathbb{C}^{12}$ with $a_1, \ldots, c_6$ as coordinates, the polynomial D vanishes on the set

$$\{(a_1, \ldots, c_6) : (2.9) \text{ has a nontrivial solution}\} \subset \mathbb{C}^{12}. \tag{2.12}$$

However, by Theorem (2.3), having a nontrivial solution is equivalent to the vanishing of the resultant, so that D vanishes on the set

$$\mathbf{V}(\mathrm{Res}_{1,1,2}) \subset \mathbb{C}^{12}.$$

This means that $D \in \mathbf{I}(\mathbf{V}(\mathrm{Res}_{1,1,2})) = \sqrt{\langle \mathrm{Res}_{1,1,2} \rangle}$, where the last equality is by the Nullstellensatz (see §4 of Chapter 1). But $\mathrm{Res}_{1,1,2}$ is irreducible, which easily implies that $\sqrt{\langle \mathrm{Res}_{1,1,2} \rangle} = \langle \mathrm{Res}_{1,1,2} \rangle$. This proves that $D \in \langle \mathrm{Res}_{1,1,2} \rangle$, so that $D = -a_1R$ is a multiple of $\mathrm{Res}_{1,1,2}$. Irreducibility then implies that $\mathrm{Res}_{1,1,2}$ divides either a_1 or R. The results of §3 will tell us that $\mathrm{Res}_{1,1,2}$ has total degree 5. It follows that $\mathrm{Res}_{1,1,2}$ divides R, and since R also has total degree 5, it must be a constant multiple of $\mathrm{Res}_{1,1,2}$. By computing the value of each when $(F_0, F_1, F_2) = (x, y, z^2)$, we see that the constant must be 1, which proves that $R = \mathrm{Res}_{1,1,2}$, as desired. □

Exercise 4. Verify that $R = 1$ when $(F_0, F_1, F_2) = (x, y, z^2)$.

The equations (2.11) may seem somewhat unmotivated. In §4 we will see that there is a systematic reason for choosing these equations.

The final topic of this section is the geometric interpretation of the resultant. We will use the same framework as in Theorem (2.3). This means that we consider homogeneous polynomials of degree $d_0, \ldots, d_n$, and for each monomial x^α of degree d_i, we introduce a variable $u_{i,\alpha}$. Let M be the total number of these variables, so that $\mathbb{C}^M$ is an affine space with coordinates $u_{i,\alpha}$ for all $0 \leq i \leq n$ and $|\alpha| = d_i$. A point of $\mathbb{C}^M$ will be written $(c_{i,\alpha})$. Then consider the "universal" polynomials

$$\mathbf{F}_i = \sum_{|\alpha|=d_i} u_{i,\alpha} x^\alpha, \quad i = 0, \ldots, n.$$

Note that the coefficients of the x^α are the variables $u_{i,\alpha}$. If we evaluate $\mathbf{F}_0, \ldots, \mathbf{F}_n$ at $(c_{i,\alpha}) \in \mathbb{C}^M$, we get the polynomials $F_0, \ldots, F_n$, where $F_i = \sum_{|\alpha|=d_i} c_{i,\alpha} x^\alpha$. Thus, we can think of points of $\mathbb{C}^M$ as parametrizing all possible $(n+1)$-tuples of homogeneous polynomials of degrees $d_0, \ldots, d_n$.

To keep track of nontrivial solutions of these polynomials, we will use projective space $\mathbb{P}^n(\mathbb{C})$, which we write as $\mathbb{P}^n$ for short. Recall the following:

- A point in $\mathbb{P}^n$ has homogeneous coordinates $(a_0, \ldots, a_n)$, where $a_i \in \mathbb{C}$ are not all zero, and another set of coordinates $(b_0, \ldots, b_n)$ gives the same point in $\mathbb{P}^n$ if and only if there is a complex number $\lambda \neq 0$ such that $(b_0, \ldots, b_n) = \lambda(a_0, \ldots, a_n)$.
- If $F(x_0, \ldots, x_n)$ is homogeneous of degree d and $(b_0, \ldots, b_n) = \lambda(a_0, \ldots, a_n)$ are two sets of homogeneous coordinates for some point $p \in \mathbb{P}^n$, then

$$F(b_0, \ldots, b_n) = \lambda^d F(a_0, \ldots, a_n).$$

  Thus, we can't define the value of F at p, but the equation $F(p) = 0$ makes perfect sense. Hence we get the *projective variety* $\mathbf{V}(F) \subset \mathbb{P}^n$, which is the set of points of $\mathbb{P}^n$ where F vanishes.

For a homogeneous polynomial F, notice that $\mathbf{V}(F) \subset \mathbb{P}^n$ is determined by the *nontrivial* solutions of $F = 0$. For more on projective space, see Chapter 8 of [CLO].

Now consider the product $\mathbb{C}^M \times \mathbb{P}^n$. A point $(c_{i,\alpha}, a_0, \ldots, a_n) \in \mathbb{C}^M \times \mathbb{P}^n$ can be regarded as $n+1$ homogeneous polynomials and a point of $\mathbb{P}^n$. The "universal" polynomials $\mathbf{F}_i$ are actually polynomials on $\mathbb{C}^M \times \mathbb{P}^n$, which gives the subset $W = \mathbf{V}(\mathbf{F}_0, \ldots, \mathbf{F}_n)$. Concretely, this set is given by

$$\begin{aligned}
W = \{(c_{i,\alpha}, a_0, \ldots, a_n) \in \mathbb{C}^M \times \mathbb{P}^n :\ & (a_0, \ldots, a_n) \text{ is a} \\
& \text{nontrivial solution of } F_0 = \cdots = F_n = 0, \text{ where} \\
& F_0, \ldots, F_n \text{ are determined by } (c_{i,\alpha})\} \\
= \{\text{all possible pairs consisting of a set of equations}\ & \\
& F_0 = \cdots = F_n = 0 \text{ of degrees } d_0, \ldots, d_n \text{ and} \\
& \text{a nontrivial solution of the equations}\}.
\end{aligned} \tag{2.13}$$

Now comes the interesting part: there is a natural projection map

$$\pi : \mathbb{C}^M \times \mathbb{P}^n \longrightarrow \mathbb{C}^M$$

defined by $\pi(c_{i,\alpha}, a_0, \ldots, a_n) = (c_{i,\alpha})$, and under this projection, the variety $W \subset \mathbb{C}^M \times \mathbb{P}^n$ maps to

$$\begin{aligned}\pi(W) = \{(c_{i,\alpha}) \in \mathbb{C}^M : &\text{ there is } (a_0, \ldots, a_n) \in \mathbb{P}^n \\ &\text{such that } (c_{i,\alpha}, a_0, \ldots, a_n) \in W\} \\ = \{\text{all possible}&\text{ sets of equations } F_0 = \cdots = F_n = 0 \text{ of} \\ &\text{degrees } d_1, \ldots, d_n \text{ which have a nontrivial solution}\}.\end{aligned}$$

Note that when the degrees are $(d_0, d_1, d_2) = (1, 1, 2)$, $\pi(W)$ is as in (2.12).

The essential content of Theorem (2.3) is that the set $\pi(W)$ is defined by the *single irreducible* equation $\mathrm{Res}_{d_0,\ldots,d_n} = 0$. To prove this, first note that $\pi(W)$ is a variety in $\mathbb{C}^M$ by the following result of elimination theory.

- (Projective Extension Theorem) Given a variety $W \subset \mathbb{C}^M \times \mathbb{P}^n$ and the projection map $\pi : \mathbb{C}^M \times \mathbb{P}^n \to \mathbb{C}^M$, the image $\pi(W)$ is a variety in $\mathbb{C}^M$.

(See, for example, §5 of Chapter 8 of [CLO].) This is one of the key reasons we work with projective space (the corresponding assertion for affine space is false in general). Hence $\pi(W)$ is defined by the vanishing of certain polynomials on $\mathbb{C}^M$. In other words, the existence of a nontrivial solution of $F_0 = \cdots = F_n = 0$ is determined by polynomial conditions on the coefficients of $F_0, \ldots, F_n$.

The second step in the proof is to show that we need only one polynomial and that this polynomial is irreducible. Here, a rigorous proof requires knowing certain facts about the dimension and irreducible components of a variety (see, for example, [Sha], §6 of Chapter I). If we accept an intuitive idea of dimension, then the basic idea is to show that the variety $\pi(W) \subset \mathbb{C}^M$ is irreducible (can't be decomposed into smaller pieces which are still varieties) of dimension $M-1$. In this case, the theory will tell us that $\pi(W)$ must be defined by exactly one irreducible equation, which is the resultant $\mathrm{Res}_{d_0,\ldots,d_n} = 0$.

To prove this, first note that $\mathbb{C}^M \times \mathbb{P}^n$ has dimension $M + n$. Then observe that $W \subset \mathbb{C}^M \times \mathbb{P}^n$ is defined by the $n+1$ equations $\mathbf{F}_0 = \cdots = \mathbf{F}_n = 0$. Intuitively, each equation drops the dimension by one, though strictly speaking, this requires that the equations be "independent" in an appropriate sense. In our particular case, this is true because each equation involves a disjoint set of coefficient variables $u_{i,\alpha}$. Thus the dimension of W is $(M+n) - (n+1) = M - 1$. One can also show that W is irreducible (see Exercise 9 below). From here, standard arguments imply that $\pi(W)$ is irreducible. The final part of the argument is to show that the map $W \to \pi(W)$ is one-to-one "most of the time". Here, the idea is that if $F_0 = \cdots = F_n = 0$ do happen to have a nontrivial solution, then this solution is usually unique (up to a scalar multiple). For the special case when all of the F_i are linear, we will prove this in Exercise 10 below. For the general case, see Proposition 3.1 of Chapter 3 of [GKZ]. Since $W \to \pi(W)$ is onto and one-to-one most of the time, $\pi(W)$ also has dimension $M - 1$.

Additional Exercises for §2

Exercise 5. To prove the uniqueness of the resultant, suppose there are two polynomials Res and Res$'$ satisfying the conditions of Theorem (2.3).
a. Adapt the argument used in the proof of Proposition (2.10) to show that Res divides Res$'$ and Res$'$ divides Res. Note that this uses conditions a and c of the theorem.
b. Now use condition b of Theorem (2.3) to conclude that Res $=$ Res$'$.

Exercise 6. A homogeneous polynomial in $\mathbb{C}[x]$ is written in the form ax^d. Show that $\mathrm{Res}_d(ax^d) = a$. Hint: Use Exercise 5.

Exercise 7. When the hypotheses of Proposition (2.6) are satisfied, the resultant (2.7) gives a polynomial $p(x, y, z)$ which vanishes precisely on the parametrized surface. However, p need not have the smallest possible total degree: it can happen that $p = q^d$ for some polynomial q of smaller total degree. For example, consider the (fairly silly) parametrization given by $(x, y, z) = (s, s, t^2)$. Use the formula of Proposition (2.10) to show that in this case, p is the square of another polynomial.

Exercise 8. The method used in the proof of Proposition (2.10) can be used to explain how the determinant (1.2) arises from nontrivial solutions $F = G = 0$, where F, G are as in (1.6). Namely, if (x, y) is a nontrivial solution of (1.6), then consider the $l + m$ equations

$$
\begin{aligned}
x^{m-1} \cdot F &= 0\\
x^{m-2}y \cdot F &= 0\\
&\;\vdots\\
y^{m-1} \cdot F &= 0\\
x^{l-1} \cdot G &= 0\\
x^{l-2}y \cdot G &= 0\\
&\;\vdots\\
y^{l-1} \cdot G &= 0.
\end{aligned}
$$

Regarding this as a system of linear equations in unknowns x^{l+m-1}, $x^{l+m-2}y, \ldots, y^{l+m-1}$, show that the coefficient matrix is exactly the transpose of (1.2), and conclude that the determinant of this matrix must vanish whenever (1.6) has a nontrivial solution.

Exercise 9. In this exercise, we will give a rigorous proof that the set W from (2.13) is irreducible of dimension $M-1$. For convenience, we will write a point of $\mathbb{C}^M$ as $(F_0, \ldots, F_n)$.

a. If $p = (a_0, \ldots, a_n)$ are fixed homogeneous coordinates for a point $p \in \mathbb{P}^n$, show that the map $\mathbb{C}^M \to \mathbb{C}^{n+1}$ defined by $(F_0, \ldots, F_n) \mapsto (F_0(p), \ldots, F_n(p))$ is linear and onto. Conclude that the kernel of this map has dimension $M-n-1$. Denote this kernel by $K(p)$.
b. Besides the projection $\pi : \mathbb{C}^M \times \mathbb{P}^n \to \mathbb{C}^M$ used in the text, we also have a projection map $\mathbb{C}^M \times \mathbb{P}^n \to \mathbb{P}^n$, which is projection on the second factor. If we restrict this map to W, we get a map $\tilde{\pi} : W \to \mathbb{P}^n$ defined by $\tilde{\pi}(F_0, \ldots, F_n, p) = p$. Then show that

$$\tilde{\pi}^{-1}(p) = K(p) \times \{p\},$$

where as usual $\tilde{\pi}^{-1}(p)$ is the inverse image of $p \in \mathbb{P}^n$ under $\tilde{\pi}$, i.e., the set of all points of W which map to p under $\tilde{\pi}$. In particular, this shows that $\tilde{\pi} : W \to \mathbb{P}^n$ is onto and that all inverse images of points are irreducible (being linear subspaces) of the same dimension.
c. Use Theorem 8 of [Sha], §6 of Chapter 1, to conclude that W is irreducible.
d. Use Theorem 7 of [Sha], §6 of Chapter 1, to conclude that W has dimension $M-1 = n$ (dimension of $\mathbb{P}^n$) $+ M - n - 1$ (dimension of the inverse images).

Exercise 10. In this exercise, we will show that the map $W \to \pi(W)$ is usually one-to-one in the special case when $F_0, \ldots, F_n$ have degree 1. Here, we know that if $F_i = \sum_{j=0}^n c_{ij} x_j$, then $\mathrm{Res}(F_0, \ldots, F_n) = \det(A)$, where $A = (c_{ij})$. Note that A is an $(n+1) \times (n+1)$ matrix.

a. Show that $F_0 = \cdots = F_n = 0$ has a nontrivial solution if and only if A has rank $< n+1$.
b. If A has rank n, prove that there is a unique nontrivial solution (up to a scalar multiple).
c. Given $0 \le i, j \le n$, let $A^{i,j}$ be the $n \times n$ matrix obtained from A by deleting row i and column j. Prove that A has rank $< n$ if and only if $\det(A^{i,j}) = 0$ for all i, j. Hint: To have rank $\ge n$, it must be possible to find n columns which are linearly independent. Then, looking at the submatrix formed by these columns, it must be possible to find n rows which are linearly independent. This leads to one of the matrices $A^{i,j}$.
d. Let $Y = \mathbf{V}(\det(A^{i,j}) : 0 \le i, j \le n)$. Show that $Y \subset \pi(W)$ and that $Y \ne \pi(W)$. Since $\pi(W)$ is irreducible, standard arguments show that Y has dimension strictly smaller than $\pi(W)$ (see, for example, Corollary 2 to Theorem 4 of [Sha], §6 of Chapter I).
e. Show that if $a, b \in W$ and $\pi(a) = \pi(b) \in \pi(W) \setminus Y$, then $a = b$. Since Y has strictly smaller dimension than $\pi(W)$, this is a precise version of what we mean by saying the map $W \to \pi(W)$ is "usually one-to-one". Hint: Use parts b and c.

§3 Properties of Resultants

In Theorem (2.3), we saw that the resultant $\mathrm{Res}(F_0, \ldots, F_n)$ vanishes if and only if $F_0 = \cdots = F_n = 0$ has a nontrivial solution, and is irreducible over $\mathbb{C}$ when regarded as a polynomial in the coefficients of the F_i. These conditions characterize the resultant up to a constant, but they in no way exhaust the many properties of this remarkable polynomial. This section will contain a summary of the other main properties of the resultant. No proofs will be given, but complete references will be provided.

Throughout this section, we will fix total degrees $d_0, \ldots, d_n > 0$ and let $\mathrm{Res} = \mathrm{Res}_{d_0,\ldots,d_n} \in \mathbb{Z}[u_{i,\alpha}]$ be the resultant polynomial from §2.

We begin by studying the degree of the resultant.

(3.1) Theorem. *For a fixed j between 0 and n,* Res *is homogeneous in the variables $u_{j,\alpha}$, $|\alpha| = d_j$, of degree $d_0 \cdots d_{j-1}d_{j+1} \cdots d_n$. This means that*

$$\mathrm{Res}(F_0, \ldots, \lambda F_j, \ldots, F_n) = \lambda^{d_0 \cdots d_{j-1}d_{j+1} \cdots d_n}\mathrm{Res}(F_0, \ldots, F_n).$$

Furthermore, the total degree of Res *is* $\sum_{j=0}^{n} d_0 \cdots d_{j-1}d_{j+1} \cdots d_n$.

PROOF. A proof can be found in §2 of [Jou1] or Chapter 13 of [GKZ]. □

Exercise 1. Show that the final assertion of Theorem (3.1) is an immediate consequence of the formula for $\mathrm{Res}(F_0, \ldots, \lambda F_j, \ldots, F_n)$. Hint: What is $\mathrm{Res}(\lambda F_0, \ldots, \lambda F_n)$?

Exercise 2. Show that formulas (1.2) and (2.8) for $\mathrm{Res}_{l,m}$ and $\mathrm{Res}_{2,2,2}$ satisfy Theorem (3.1).

We next study the symmetry and multiplicativity of the resultant.

(3.2) Theorem.

a. *If $i < j$, then*

$$\mathrm{Res}(F_0, \ldots, F_i, \ldots, F_j, \ldots, F_n) = (-1)^{d_0 \cdots d_n}\mathrm{Res}(F_0, \ldots, F_j, \ldots, F_i, \ldots, F_n),$$

where the bottom resultant is for degrees $d_0, \ldots, d_j, \ldots, d_i, \ldots, d_n$.

b. *If $F_j = F_j' F_j''$ is a product of homogeneous polynomials of degrees d_j' and d_j'', then*

$$\mathrm{Res}(F_0, \ldots, F_j, \ldots, F_n) = \mathrm{Res}(F_0, \ldots, F_j', \ldots, F_n) \cdot \mathrm{Res}(F_0, \ldots, F_j'', \ldots, F_n),$$

where the resultants on the bottom are for degrees $d_0, \ldots, d_j', \ldots, d_n$ and $d_0, \ldots, d_j'', \ldots, d_n$.

Proof. A proof of the first assertion of the theorem can be found in §5 of [Jou1]. As for the second, we can assume $j = n$ by part a. This case will be covered in Exercise 9 at the end of the section. □

Exercise 3. Prove that formulas (1.2) and (2.8) for $\mathrm{Res}_{l,m}$ and $\mathrm{Res}_{2,2,2}$ satisfy part a of Theorem (3.2).

Our next task is to show that the analog of Proposition (1.5) holds for general resultants. We begin with some notation. Given homogeneous polynomials $F_0, \ldots, F_n \in \mathbb{C}[x_0, \ldots, x_n]$ of degrees $d_0, \ldots, d_n$, let

$$\begin{aligned} f_i(x_0, \ldots, x_{n-1}) &= F_i(x_0, \ldots, x_{n-1}, 1) \\ \overline{F}_i(x_0, \ldots, x_{n-1}) &= F_i(x_0, \ldots, x_{n-1}, 0). \end{aligned} \tag{3.3}$$

Note that $\overline{F}_0, \ldots, \overline{F}_{n-1}$ are homogeneous in $\mathbb{C}[x_0, \ldots, x_{n-1}]$ of degrees $d_0, \ldots, d_{n-1}$.

(3.4) Theorem. *If* $\mathrm{Res}(\overline{F}_0, \ldots, \overline{F}_{n-1}) \neq 0$, *then the quotient ring* $A = \mathbb{C}[x_0, \ldots, x_{n-1}]/\langle f_0, \ldots, f_{n-1}\rangle$ *has dimension* $d_0 \cdots d_{n-1}$ *as a vector space over* $\mathbb{C}$, *and*

$$\mathrm{Res}(F_0, \ldots, F_n) = \mathrm{Res}(\overline{F}_0, \ldots, \overline{F}_{n-1})^{d_n} \det(m_{f_n} : A \to A),$$

where $m_{f_n} : A \to A$ *is the linear map given by multiplication by* f_n.

Proof. Although we will not prove this result (see [Jou1], §§2, 3 and 4 for a complete proof), we will explain (non-rigorously) why the above formula is reasonable. The first step is to show that the ring A is a finite-dimensional vector space over $\mathbb{C}$ when $\mathrm{Res}(\overline{F}_0, \ldots, \overline{F}_{n-1}) \neq 0$. The crucial idea is to think in terms of the projective space $\mathbb{P}^n$. We can decompose $\mathbb{P}^n$ into two pieces using x_n: the affine space $\mathbb{C}^n \subset \mathbb{P}^n$ defined by $x_n = 1$, and the "hyperplane at infinity" $\mathbb{P}^{n-1} \subset \mathbb{P}^n$ defined by $x_n = 0$. Note that the other variables $x_0, \ldots, x_{n-1}$ play two roles: they are ordinary coordinates for $\mathbb{C}^n \subset \mathbb{P}^n$, and they are homogeneous coordinates for the hyperplane at infinity.

The equations $F_0 = \cdots = F_{n-1} = 0$ determine a projective variety $V \subset \mathbb{P}^n$. By (3.3), $f_0 = \cdots = f_{n-1} = 0$ defines the "affine part" $\mathbb{C}^n \cap V \subset V$, while $\overline{F}_0 = \cdots = \overline{F}_{n-1} = 0$ defines the "part at infinity" $\mathbb{P}^{n-1} \cap V \subset V$. Hence, the hypothesis $\mathrm{Res}(\overline{F}_0, \ldots, \overline{F}_{n-1}) \neq 0$ implies that there are no solutions at infinity. In other words, the projective variety V is contained in $\mathbb{C}^n \subset \mathbb{P}^n$. Now we can apply the following result from algebraic geometry:

- (Projective Varieties in Affine Space) If a projective variety in $\mathbb{P}^n$ is contained in an affine space $\mathbb{C}^n \subset \mathbb{P}^n$, then the projective variety must consist of a finite set of points.

(See, for example, [Sha], §5 of Chapter I.) Applied to V, this tells us that V must be a finite set of points. Since $\mathbb{C}$ is algebraically closed and $V \subset \mathbb{C}^n$

is defined by $f_0 = \cdots = f_{n-1} = 0$, the Finiteness Theorem from §2 of Chapter 2 implies that $A = \mathbb{C}[x_0, \ldots, x_{n-1}]/\langle f_0, \ldots, f_{n-1}\rangle$ is finite dimensional over $\mathbb{C}$. Hence $\det(m_{f_n} : A \to A)$ is defined, so that the formula of the theorem makes sense.

We also need to know the dimension of the ring A. The answer is provided by Bézout's Theorem:

- (Bézout's Theorem) If the equations $F_0 = \cdots = F_{n-1} = 0$ have degrees $d_0, \ldots, d_{n-1}$ and finitely many solutions in $\mathbb{P}^n$, then the number of solutions (counted with multiplicity) is $d_0 \cdots d_{n-1}$.

(See [Sha], §2 of Chapter II.) This tells us that V has $d_0 \cdots d_{n-1}$ points, counted with multiplicity. Because $V \subset \mathbb{C}^n$ is defined by $f_0 = \cdots = f_{n-1} = 0$, Theorem (2.2) from Chapter 4 implies that the number of points in V, counted with multiplicity, is the dimension of $A = \mathbb{C}[x_0, \ldots, x_{n-1}]/\langle f_0, \ldots, f_{n-1}\rangle$. Thus, Bézout's Theorem shows that $\dim A = d_0 \cdots d_{n-1}$.

We can now explain why $\mathrm{Res}(\overline{F}_0, \ldots, \overline{F}_{n-1})^{d_n} \det(m_{f_n})$ behaves like a resultant. The first step is to prove that $\det(m_{f_n})$ vanishes if and only if $F_0 = \cdots = F_n = 0$ has a solution in $\mathbb{P}^n$. If we have a solution p, then $p \in V$ since $F_0(p) = \cdots = F_{n-1}(p) = 0$. But $V \subset \mathbb{C}^n$, so we can write $p = (a_0, \ldots, a_{n-1}, 1)$, and $f_n(a_0, \ldots, a_{n-1}) = 0$ since $F_n(p) = 0$. Then Theorem (2.6) of Chapter 2 tells us that $f_n(a_0, \ldots, a_{n-1}) = 0$ is an eigenvalue of m_{f_n}, which proves that $\det(m_{f_n}) = 0$. Conversely, if $\det(m_{f_n}) = 0$, then one of its eigenvalues must be zero. Since the eigenvalues are $f_n(p)$ for $p \in V$ (Theorem (2.6) of Chapter 2 again), we have $f_n(p) = 0$ for some p. Writing p in the form $(a_0, \ldots, a_{n-1}, 1)$, we get a nontrivial solution of $F_0 = \cdots = F_n = 0$, as desired.

Finally, we will show that $\mathrm{Res}(\overline{F}_0, \ldots, \overline{F}_{n-1})^{d_n} \det(m_{f_n})$ has the homogeneity properties predicted by Theorem (3.1). If we replace F_j by λF_j for some $j < n$ and $\lambda \in \mathbb{C} \setminus \{0\}$, then $\overline{\lambda F_j} = \lambda \overline{F}_j$, and neither A nor m_{f_n} are affected. Since

$$\begin{aligned}\mathrm{Res}(\overline{F}_0, \ldots, \lambda\overline{F}_j, \ldots, \overline{F}_{n-1}) &= \\ &\lambda^{d_0 \cdots d_{j-1} d_{j+1} \cdots d_{n-1}} \mathrm{Res}(\overline{F}_0, \ldots, \overline{F}_j, \ldots, \overline{F}_{n-1}),\end{aligned}$$

we get the desired power of λ because of the exponent d_n in the formula of the theorem. On the other hand, if we replace F_n with λF_n, then $\mathrm{Res}(\overline{F}_0, \ldots, \overline{F}_{n-1})$ and A are unchanged, but m_{f_n} becomes $m_{\lambda f_n} = \lambda m_{f_n}$. Since

$$\det(\lambda m_{f_n}) = \lambda^{\dim A} \det(m_{f_n})$$

it follows that we get the correct power of λ because, as we showed above, A has dimension $d_0 \cdots d_{n-1}$.

This discussion shows that the formula $\mathrm{Res}(\overline{F}_0, \ldots, \overline{F}_{n-1})^{d_n} \det(m_{f_n})$ has many of the properties of the resultant, although some important points were left out (for example, we didn't prove that it is a polynomial in the coefficients of the F_i). We also know what this formula means geometrically: it asserts that the resultant is a product of two terms, one coming from the behavior of $F_0, \ldots, F_{n-1}$ at infinity and the other coming from the behavior of $f_n = F_n(x_0, \ldots, x_{n-1}, 1)$ on the affine variety determined by vanishing of $f_0, \ldots, f_{n-1}$. □

Exercise 4. When $n = 2$, show that Proposition (1.5) is a special case of Theorem (3.4). Hint: Start with f, g as in (1.1) and homogenize to get (1.6). Use Exercise 6 of §2 to compute $\mathrm{Res}(\overline{F})$.

Exercise 5. Use Theorem (3.4) and `getmatrix` to compute the resultant of the polynomials $x^2 + y^2 + z^2, xy + xz + yz, xyz$.

The formula given in Theorem (3.4) is sometimes called the *Poisson Formula*. Some further applications of this formula will be given in the exercises at the end of the section.

In the special case when $F_0, \ldots, F_n$ all have the same total degree $d > 0$, the resultant $\mathrm{Res}_{d,\ldots,d}$ has degree d^n in the coefficients of each F_i, and its total degree is $(n + 1)d^n$. Besides all of the properties listed so far, the resultant has some other interesting properties in this case:

(3.5) Theorem. $\mathrm{Res} = \mathrm{Res}_{d,\ldots,d}$ *has the following properties:*

a. *If F_j are homogeneous of total degree d and $G_i = \sum_{j=0}^n a_{ij}F_j$, where (a_{ij}) is an invertible matrix with entries in $\mathbb{C}$, then*
$$\mathrm{Res}(G_0, \ldots, G_n) = \det(a_{ij})^{d^n}\,\mathrm{Res}(F_0, \ldots, F_n).$$

b. *If we list all monomials of total degree d as $x^{\alpha(1)}, \ldots, x^{\alpha(N)}$ and pick $n+1$ distinct indices $1 \le i_0 < \cdots < i_n \le N$, the* **bracket** *$[i_0 \ldots i_n]$ is defined to be the determinant*
$$[i_0 \ldots i_n] = \det(u_{i,\alpha(i_j)}) \in \mathbb{Z}[u_{i,\alpha(j)}].$$
Then Res *is a polynomial in the brackets $[i_0 \ldots i_n]$.*

Proof. See Proposition 5.11.2 of [Jou1] for a proof of part a. For part b, note that if (a_{ij}) has determinant 1, then part a implies $\mathrm{Res}(G_0, \ldots, G_n) = \mathrm{Res}(F_0, \ldots, F_n)$, so Res is invariant under the action of $\mathrm{SL}(n + 1, \mathbb{C}) = \{A \in M_{(n+1)\times(n+1)}(\mathbb{C}) : \det(A) = 1\}$ on $(n + 1)$-tuples of homogeneous polynomials of degree d. If we regard the coefficients of the universal polynomials $\mathbf{F}_i$ as an $(n + 1) \times N$ matrix $(u_{i,\alpha(j)})$, then this action is matrix multiplication by elements of $\mathrm{SL}(n+1, \mathbb{C})$. Since Res is invariant under this action, the *First Fundamental Theorem of Invariant Theory* (see [Stu1], Section 3.2) asserts that Res is a polynomial in the $(n + 1) \times (n + 1)$ minors of $(u_{i,\alpha(j)})$, which are exactly the brackets $[i_0 \ldots i_n]$. □

Exercise 6. Show that each bracket $[i_0 \dots i_n] = \det(u_{i,\alpha(i_j)})$ is invariant under the action of $\mathrm{SL}(n+1, \mathbb{C})$.

We should mention that the expression of Res in terms of the brackets $[i_0 \dots i_n]$ is not unique. The different ways of doing this are determined by the algebraic relations among the brackets, which are described by the *Second Fundamental Theorem of Invariant Theory* (see Section 3.2 of [Stu1]).

As an example of Theorem (3.5), consider the resultant of three ternary quadrics

$$
\begin{aligned}
F_0 &= c_{01}x^2 + c_{02}y^2 + c_{03}z^2 + c_{04}xy + c_{05}xz + c_{06}yz = 0\\
F_1 &= c_{11}x^2 + c_{12}y^2 + c_{13}z^2 + c_{14}xy + c_{15}xz + c_{16}yz = 0\\
F_2 &= c_{21}x^2 + c_{22}y^2 + c_{23}z^2 + c_{24}xy + c_{25}xz + c_{26}yz = 0.
\end{aligned}
$$

In §2, we gave a formula for $\mathrm{Res}_{2,2,2}(F_0, F_1, F_2)$ as a certain 6×6 determinant. Using Theorem (3.5), we get quite a different formula. If we list the six monomials of total degree 2 as $x^2, y^2, z^2, xy, xz, yz$, then the bracket $[i_0 i_1 i_2]$ is given by

$$
[i_0 i_1 i_2] = \det \begin{pmatrix} c_{0i_0} & c_{0i_1} & c_{0i_2} \\ c_{1i_0} & c_{1i_1} & c_{1i_2} \\ c_{2i_0} & c_{2i_1} & c_{2i_2} \end{pmatrix}.
$$

By [KSZ], the resultant $\mathrm{Res}_{2,2,2}(F_0, F_1, F_2)$ is the following polynomial in the brackets $[i_0 i_1 i_2]$:

$$
\begin{aligned}
&[145][246][356][456] - [146][156][246][356] - [145][245][256][356]\\
&- [145][246][346][345] + [125][126][356][456] - 2[124][156][256][356]\\
&- [134][136][246][456] - 2[135][146][346][246] + [235][234][145][456]\\
&- 2[236][345][245][145] - [126]^2[156][356] - [125]^2[256][356]\\
&- [134]^2[246][346] - [136]^2[146][246] - [145][245][235]^2\\
&- [145][345][234]^2 + 2[123][124][356][456] - [123][125][346][456]\\
&- [123][134][256][456] + 2[123][135][246][456] - 2[123][145][246][356]\\
&- [124]^2[356]^2 + 2[124][125][346][356] - 2[124][134][256][356]\\
&- 3[124][135][236][456] - 4[124][135][246][356] - [125]^2[346]^2\\
&+ 2[125][135][246][346] - [134]^2[256]^2 + 2[134][135][246][256]\\
&- 2[135]^2[246]^2 - [123][126][136][456] + 2[123][126][146][356]\\
&- 2[124][136]^2[256] - 2[125][126][136][346] + [123][125][235][456]\\
&- 2[123][125][245][356] - 2[124][235]^2[156] - 2[126][125][235][345]\\
&- [123][234][134][456] + 2[123][234][346][145] - 2[236][134]^2[245]\\
&- 2[235][234][134][146] + 3[136][125][235][126] - 3[126][135][236][125]
\end{aligned}
$$

$$
\begin{aligned}
&- [136][125]^2[236] - [126]^2[135][235] - 3[134][136][126][234] \\
&+ 3[124][134][136][236] + [134]^2[126][236] + [124][136]^2[234] \\
&- 3[124][135][234][235] + 3[134][234][235][125] - [135][234]^2[125] \\
&- [124][235]^2[134] - [136]^2[126]^2 - [125]^2[235]^2 \\
&- [134]^2[234]^2 + 3[123][124][135][236] + [123][134][235][126] \\
&+ [123][135][126][234] + [123][134][236][125] + [123][136][125][234] \\
&+ [123][124][235][136] - 2[123]^2[126][136] + 2[123]^2[125][235] \\
&- 2[123]^2[134][234] - [123]^4.
\end{aligned}
$$

This expression for $\mathrm{Res}_{2,2,2}$ has total degree 4 in the brackets since the resultant has total degree 12 and each bracket has total degree 3 in the c_{ij}. Although this formula is rather complicated, its 68 terms are a lot simpler than the 21,894 terms we get when we express $\mathrm{Res}_{2,2,2}$ as a polynomial in the c_{ij}!

Exercise 7. When $F_0 = a_0x^2 + a_1xy + a_2y^2$ and $F_1 = b_0x^2 + b_1xy + b_2y^2$, the only brackets to consider are $[01] = a_0b_1 - a_1b_0$, $[02] = a_0b_2 - a_2b_0$ and $[12] = a_1b_2 - a_2b_1$ (why?). Express $\mathrm{Res}_{2,2}$ as a polynomial in these three brackets. Hint: In the determinant (1.2), expand along the first row and then expand along the column containing the zero.

Theorem (3.5) also shows that the resultant of two homogeneous polynomials $F_0(x, y), F_1(x, y)$ of degree d can be written in terms of the brackets $[ij]$. The resulting formula is closely related to the *Bézout Formula* described in Chapter 12 of [GKZ].

For further properties of resultants, the reader should consult Chapter 13 of [GKZ] or Section 5 of [Jou1].

Additional Exercises for §3

Exercise 8. The product formula (1.4) can be generalized to arbitrary resultants. With the same hypotheses as Theorem (3.4), let $V = \mathbf{V}(f_0, \ldots, f_{n-1})$ be as in the proof of the theorem. Then

$$\mathrm{Res}(F_0, \ldots, F_n) = \mathrm{Res}(\overline{F}_0, \ldots, \overline{F}_{n-1})^{d_n} \prod_{p \in V} f_n(p)^{m(p)},$$

where $m(p)$ is the multiplicity of p in V. This concept is defined in [Sha], §2 of Chapter II, and §2 of Chapter 4. For this exercise, assume that V consists of $d_0 \cdots d_{n-1}$ distinct points (which means that all of the multiplicities $m(p)$ are equal to 1) and that f_n takes distinct values on these points. Then use Theorem (2.6) of Chapter 2, together with Theorem (3.4), to show that the above formula for the resultant holds in this case.

Exercise 9. In Theorem (3.4), we assumed that the field was $\mathbb{C}$. It turns out that the result is true over any field k. In this exercise, we will use this version of the theorem to prove part b of Theorem (3.2) when $F_n = F_n' F_n''$. The trick is to choose k appropriately: we will let k be the field of rational functions in the coefficients of $F_0, \ldots, F_{n-1}, F_n', F_n''$. This means we regard each coefficient as a separate variable and then k is the field of rational functions in these variables with coefficients in $\mathbb{Q}$.

a. Explain why $\overline{F}_0, \ldots, \overline{F}_{n-1}$ are the "universal" polynomials of degrees $d_0, \ldots, d_{n-1}$ in $x_0, \ldots, x_{n-1}$, and conclude that $\mathrm{Res}(\overline{F}_0, \ldots, \overline{F}_{n-1})$ is nonzero.
b. Use Theorem (3.4) (over the field k) to show that

$$\mathrm{Res}(F_0, \ldots, F_n) = \mathrm{Res}(F_0, \ldots, F_n') \cdot \mathrm{Res}(F_0, \ldots, F_n'').$$

Notice that you need to use the theorem three times. Hint: $m_{f_n} = m_{f_n'} \circ m_{f_n''}$.

Exercise 10. The goal of this exercise is to generalize Proposition (2.10) by giving a formula for $\mathrm{Res}_{1,1,d}$ for any $d > 0$. The idea is to apply Theorem (3.4) when the field k consists of rational functions in the coefficients of F_0, F_1, F_2 (so we are using the version of the theorem from Exercise 9). For concreteness, suppose that

$$\begin{aligned} F_0 &= a_1 x + a_2 y + a_3 z = 0 \\ F_1 &= b_1 x + b_2 y + b_3 z = 0. \end{aligned}$$

a. Show that $\mathrm{Res}(\overline{F}_0, \overline{F}_1) = a_1 b_2 - a_2 b_1$ and that the only solution of $f_0 = f_1 = 0$ is

$$x_0 = \frac{a_2 b_3 - a_3 b_2}{a_1 b_2 - a_2 b_1} \qquad y_0 = -\frac{a_1 b_3 - a_3 b_1}{a_1 b_2 - a_2 b_1}.$$

b. By Theorem (3.4), $k[x,y]/\langle f_0, f_1 \rangle$ has dimension one over $\mathbb{C}$. Use Theorem (2.6) of Chapter 2 to show that

$$\det(m_{f_2}) = f_2(x_0, y_0).$$

c. Since $f_2(x,y) = F_2(x,y,1)$, use Theorem (3.4) to conclude that

$$\mathrm{Res}_{1,1,d}(F_0, F_1, F_2) = F_2(a_2 b_3 - a_3 b_2, -(a_1 b_3 - a_3 b_1), a_1 b_2 - a_2 b_1).$$

Note that $a_2 b_3 - a_3 b_2, a_1 b_3 - a_3 b_1, a_1 b_2 - a_2 b_1$ are the 2×2 minors of the matrix

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}.$$

d. Use part c to verify the formula for $\mathrm{Res}_{1,1,2}$ given in Proposition (2.10).
e. Formulate and prove a formula similar to part c for the resultant $\mathrm{Res}_{1,\ldots,1,d}$. Hint: Use Cramer's Rule. The formula (with proof) can be found in Proposition 5.4.4 of [Jou1].

Exercise 11. Consider the elementary symmetric functions $\sigma_1, \ldots, \sigma_n \in \mathbb{C}[x_1, \ldots, x_n]$. These are defined by

$$\begin{aligned} \sigma_1 &= x_1 + \cdots + x_n \\ &\;\;\vdots \\ \sigma_r &= \sum_{i_1 < i_2 < \cdots < i_r} x_{i_1} x_{i_2} \cdots x_{i_r} \\ &\;\;\vdots \\ \sigma_n &= x_1 x_2 \cdots x_n. \end{aligned}$$

Since σ_i is homogeneous of total degree i, the resultant $\mathrm{Res}(\sigma_1, \ldots, \sigma_n)$ is defined. The goal of this exercise is to prove that this resultant equals -1 for all $n > 1$. Note that this exercise deals with n polynomials and n variables rather than $n+1$.

a. Show that $\mathrm{Res}(x + y, xy) = -1$.
b. To prove the result for $n > 2$, we will use induction and Theorem (3.4). Thus, let

$$\begin{aligned} \overline{\sigma}_i &= \sigma_i(x_1, \ldots, x_{n-1}, 0) \\ \tilde{\sigma}_i &= \sigma_i(x_1, \ldots, x_{n-1}, 1) \end{aligned}$$

as in (3.3). Prove that $\overline{\sigma}_i$ is the ith elementary symmetric function in $x_1, \ldots, x_{n-1}$ and that $\tilde{\sigma}_i = \overline{\sigma}_i + \overline{\sigma}_{i-1}$ (where $\overline{\sigma}_0 = 1$).
c. If $A = \mathbb{C}[x_1, \ldots, x_{n-1}]/\langle \tilde{\sigma}_1, \ldots, \tilde{\sigma}_{n-1} \rangle$, then use part b to prove that the multiplication map $m_{\tilde{\sigma}_n} : A \to A$ is multiplication by $(-1)^n$. Hint: Observe that $\tilde{\sigma}_n = \overline{\sigma}_{n-1}$.
d. Use induction and Theorem (3.5) to show that $\mathrm{Res}(\sigma_1, \ldots, \sigma_n) = -1$ for all $n > 1$.

Exercise 12. Using the notation of Theorem (3.4), show that

$$\mathrm{Res}(F_0, \ldots, F_{n-1}, x_n^d) = \mathrm{Res}(\overline{F}_0, \ldots, \overline{F}_{n-1})^d.$$

§4 Computing Resultants

Our next task is to discuss methods for computing resultants. While Theorem (3.4) allows one to compute resultants inductively (see Exercise 5 of §3 for an example), it is useful to have other tools for working with resultants. In this section, we will give some further formulas for the resultant and then discuss the practical aspects of computing $\mathrm{Res}_{d_0,\ldots,d_n}$. We will begin by generalizing the method used in Proposition (2.10) to find a formula for $\mathrm{Res}_{1,1,2}$. Recall that the essence of what we did in (2.11) was to multiply each equation by appropriate monomials so that we got a square matrix whose determinant we could take.

To do this in general, suppose we have $F_0, \ldots, F_n \in \mathbb{C}[x_0, \ldots, x_n]$ of total degrees $d_0, \ldots, d_n$. Then set

$$d = \sum_{i=0}^{n} (d_i - 1) + 1 = \sum_{i=0}^{n} d_i - n.$$

For instance, when $(d_0, d_1, d_2) = (1, 1, 2)$ as in the example in Section 2, one computes that $d = 2$, which is precisely the degree of the monomials on the left hand side of the equations following (2.11).

Exercise 1. Monomials of total degree d have the following special property which will be very important below: each such monomial is divisible by $x_i^{d_i}$ for at least one i between 0 and n. Prove this. Hint: Argue by contradiction.

Now take the monomials $x^\alpha = x_0^{a_0} \cdots x_n^{a_n}$ of total degree d and divide them into $n + 1$ sets as follows:

$$\begin{aligned}
S_0 &= \{x^\alpha : |\alpha| = d,\ x_0^{d_0} \text{ divides } x^\alpha\}\\
S_1 &= \{x^\alpha : |\alpha| = d,\ x_0^{d_0} \text{ doesn't divide } x^\alpha \text{ but } x_1^{d_1} \text{ does}\}\\
&\ \ \vdots\\
S_n &= \{x^\alpha : |\alpha| = d,\ x_0^{d_0}, \ldots, x_{n-1}^{d_{n-1}} \text{ don't divide } x^\alpha \text{ but } x_n^{d_n} \text{ does}\}.
\end{aligned}$$

By Exercise 1, every monomial of total degree d lies in one of $S_0, \ldots, S_n$. Note also that these sets are mutually disjoint. One observation we will need is the following:

$$\text{if } x^\alpha \in S_i, \text{ then we can write } x^\alpha = x_i^{d_i} \cdot x^\alpha / x_i^{d_i}.$$

Notice that $x^\alpha / x_i^{d_i}$ is a monomial of total degree $d - d_i$ since $x^\alpha \in S_i$.

Exercise 2. When $(d_0, d_1, d_2) = (1, 1, 2)$, show that $S_0 = \{x^2, xy, xz\}$, $S_1 = \{y^2, yz\}$, and $S_2 = \{z^2\}$, where we are using x, y, z as variables. Write down *all* of the $x^\alpha / x_i^{d_i}$ in this case and see if you can find these monomials in the equations (2.11).

Exercise 3. Prove that the number of monomials in S_n is exactly $d_0 \cdots d_{n-1}$. This fact will play an extremely important role in what follows. Hint: Given integers $a_0, \ldots, a_{n-1}$ with $0 \le a_i \le d_i - 1$, prove that there is a unique a_n such that $x_0^{a_0} \cdots x_n^{a_n} \in S_n$. Exercise 1 will also be useful.

Now we can write down a system of equations that generalizes (2.11). Namely, consider the equations

$$
\begin{aligned}
x^\alpha/x_0^{d_0} \cdot F_0 &= 0 \quad \text{for all } x^\alpha \in S_0 \\
&\;\vdots \\
x^\alpha/x_n^{d_n} \cdot F_n &= 0 \quad \text{for all } x^\alpha \in S_n.
\end{aligned}
\tag{4.1}
$$

Exercise 4. When $(d_0, d_1, d_2) = (1, 1, 2)$, check that the system of equations given by (4.1) is *exactly* what we wrote down in (2.11).

Since F_i has total degree d_i, it follows that $x^\alpha/x_i^{d_i} \cdot F_i$ has total degree d. Thus each polynomial on the left side of (4.1) can be written as a linear combination of monomials of total degree d. Suppose that there are N such monomials. (In the exercises at the end of the section, you will show that N equals the binomial coefficient $\binom{d+n}{n}$.) Then observe that the total number of equations is the number of elements in $S_0 \cup \cdots \cup S_n$, which is also N. Thus, regarding the monomials of total degree d as unknowns, we get a system of N linear equations in N unknowns.

(4.2) Definition. The determinant of the coefficient matrix of the $N \times N$ system of equations given by (4.1) is denoted D_n.

For example, if we have

$$
\begin{aligned}
F_0 &= a_1x + a_2y + a_3z = 0 \\
F_1 &= b_1x + b_2y + b_3z = 0 \\
F_2 &= c_1x^2 + c_2y^2 + c_3z^2 + c_4xy + c_5xz + c_6yz = 0,
\end{aligned}
\tag{4.3}
$$

then the equations following (2.11) imply that

$$
D_2 = \det \begin{pmatrix}
a_1 & 0 & 0 & a_2 & a_3 & 0 \\
0 & a_2 & 0 & a_1 & 0 & a_3 \\
0 & 0 & a_3 & 0 & a_1 & a_2 \\
0 & b_2 & 0 & b_1 & 0 & b_3 \\
0 & 0 & b_3 & 0 & b_1 & b_2 \\
c_1 & c_2 & c_3 & c_4 & c_5 & c_6
\end{pmatrix}.
\tag{4.4}
$$

Exercise 5. When we have polynomials $F_0, F_1 \in \mathbb{C}[x, y]$ as in (1.6), show that the coefficient matrix of (4.1) is exactly the transpose of the matrix (1.2). Thus, $D_1 = \mathrm{Res}(F_0, F_1)$ in this case.

Here are some general properties of D_n:

Exercise 6. Since D_n is the determinant of the coefficient matrix of (4.1), it is clearly a polynomial in the coefficients of the F_i.

a. For a fixed i between 0 and n, show that D_n is homogeneous in the coefficients of F_i of degree equal to the number μ_i of elements in S_i. Hint: Show that replacing F_i by λF_i has the effect of multiplying a certain number (how many?) equations of (4.1) by λ. How does this affect the determinant of the coefficient matrix?
b. Use Exercise 3 to show that D_n has degree $d_0 \cdots d_{n-1}$ as a polynomial in the coefficients of F_n. Hint: If you multiply each coefficient of F_n by $\lambda \in \mathbb{C}$, show that D_n gets multiplied by $\lambda^{d_0 \cdots d_{n-1}}$.
c. What is the total degree of D_n? Hint: Exercise 19 will be useful.

Exercise 7. In this exercise, you will prove that D_n is divisible by the resultant.

a. Prove that D_n vanishes whenever $F_0 = \cdots = F_n = 0$ has a nontrivial solution. Hint: If the F_i all vanish at $(c_0, \ldots, c_n) \neq (0, \ldots, 0)$, then show that the monomials of total degree d in $c_0, \ldots, c_n$ give a nontrivial solution of (4.1).
b. Using the notation from the end of §2, we have $\mathbf{V}(\mathrm{Res}) \subset \mathbb{C}^N$, where $\mathbb{C}^N$ is the affine space whose variables are the coefficients $u_{i,\alpha}$ of $F_0, \ldots, F_n$. Explain why part a implies that D_n vanishes on $\mathbf{V}(\mathrm{Res})$.
c. Adapt the argument of Proposition (2.10) to prove that $D_n \in \langle \mathrm{Res} \rangle$, so that Res divides D_n.

Exercise 7 shows that we are getting close to the resultant, for it enables us to write

$$D_n = \mathrm{Res} \cdot \text{extraneous factor}. \tag{4.5}$$

We next show that the extraneous factor doesn't involve the coefficients of F_n and in fact uses only some of the coefficients of $F_0, \ldots, F_{n-1}$.

(4.6) Proposition. *The extraneous factor in (4.5) is an integer polynomial in the coefficients of* $\overline{F}_0, \ldots, \overline{F}_{n-1}$, *where* $\overline{F}_i = F_i(x_0, \ldots, x_{n-1}, 0)$.

PROOF. Since D_n is a determinant, it is a polynomial in $\mathbb{Z}[u_{i,\alpha}]$, and we also know that $\mathrm{Res} \in \mathbb{Z}[u_{i,\alpha}]$. Exercise 7 took place in $\mathbb{C}[u_{i,\alpha}]$ (because of the Nullstellensatz), but in fact, the extraneous factor (let's call it E_n) must lie in $\mathbb{Q}[u_{i,\alpha}]$ since dividing D_n by Res produces at worst rational coefficients. Since Res is irreducible in $\mathbb{Z}[u_{i,\alpha}]$, standard results about polynomial rings over $\mathbb{Z}$ imply that $E_n \in \mathbb{Z}[u_{i,\alpha}]$ (see Exercise 20 for details).

Since $D_n = \mathrm{Res} \cdot E_n$ is homogeneous in the coefficients of F_n, Exercise 20 at the end of the section implies that Res and E_n are also homogeneous in these coefficients. But by Theorem (3.1) and Exercise 6, both Res and D_n have degree $d_0 \cdots d_{n-1}$ in the coefficients of F_n. It follows immediately that E_n has degree zero in the coefficients of F_n, so that it depends only on the coefficients of $F_0, \ldots, F_{n-1}$.

To complete the proof, we must show that E_n depends only on the coefficients of the $\overline{F}_i$. This means that coefficients of $F_0, \ldots, F_{n-1}$ with x_n to a positive power don't appear in E_n. To prove this, we use the following clever argument of Macaulay (see [Mac1]). As above, we think of Res, D_n and E_n as polynomials in the $u_{i,\alpha}$, and we define the *weight* of $u_{i,\alpha}$ to be the exponent a_n of x_n (where $\alpha = (a_0, \ldots, a_n)$). Then, the weight of a monomial in the $u_{i,\alpha}$, say $u_{i_1,\alpha_1}^{m_1} \cdots u_{i_l,\alpha_l}^{m_l}$, is defined to be the sum of the weights of each u_{i_j,α_j} multiplied by the corresponding exponents. Finally, a polynomial in the $u_{i,\alpha}$ is said to be *isobaric* if every term in the polynomial has the same weight.

In Exercise 23 at the end of the section, you will prove that every term in D_n has weight $d_0 \cdots d_n$, so that D_n is isobaric. The same exercise will show that $D_n = \mathrm{Res} \cdot E_n$ implies that Res and E_n are isobaric and that the weight of D_n is the sum of the weights of Res and E_n. Hence, it suffices to prove that E_n has weight zero (be sure you understand this). To simplify notation, let u_i be the variable representing the coefficient of $x_i^{d_i}$ in F_i. Note that $u_0, \ldots, u_{n-1}$ have weight zero while u_n has weight d_n. Then Theorems (2.3) and (3.1) imply that one of the terms of Res is

$$\pm u_0^{d_1 \cdots d_n} u_1^{d_0 d_2 \cdots d_n} \cdots u_n^{d_0 \cdots d_{n-1}}$$

(see Exercise 23). This term has weight $d_0 \cdots d_n$, which shows that the weight of Res is $d_0 \cdots d_n$. We saw above that D_n has the same weight, and it follows that E_n has weight zero, as desired. $\square$

Although the extraneous factor in (4.5) involves fewer coefficients than the resultant, it can have a very large degree, as shown by the following example.

Exercise 8. When $d_i = 2$ for $0 \le i \le 4$, show that the resultant has total degree 80 while D_4 has total degree 420. What happens when $d_i = 3$ for $0 \le i \le 4$? Hint: Use Exercises 6 and 19.

Notice that Proposition (4.6) also gives a method for computing the resultant: just factor D_n into irreducibles, and the only irreducible factor in which all variables appear is the resultant! Unfortunately, this method is wildly impractical owing to the slowness of multivariable factorization (especially for polynomials as large as D_n).

In the above discussion, the sets $S_0, \ldots, S_n$ and the determinant D_n depended on how the variables $x_0, \ldots, x_n$ were ordered. In fact, the notation D_n was chosen to emphasize that the variable x_n came last. If we fix i between 0 and $n-1$ and order the variables so that x_i comes last, then we get slightly different sets $S_0, \ldots, S_n$ and a slightly different system of equations (4.1). We will let D_i denote the determinant of this system of equations. (Note that there are many different orderings of the variables for which x_i is the last. We pick just one when computing D_i.)

Exercise 9. Show that D_i is homogeneous in the coefficients of each F_j and in particular, is homogeneous of degree $d_0 \cdots d_{i-1} d_{i+1} \cdots d_n$ in the coefficients of F_i.

We can now prove the following classical formula for Res.

(4.7) Proposition. *When $\mathbf{F}_0, \ldots, \mathbf{F}_n$ are universal polynomials as at the end of §2, the resultant is the greatest common divisor of the polynomials $D_0, \ldots, D_n$ in the ring $\mathbb{Z}[u_{i,\alpha}]$, i.e.,*

$$\mathrm{Res} = \pm \mathrm{GCD}(D_0, \ldots, D_n).$$

Proof. For each i, there are many choices for D_i (corresponding to the $(n-1)!$ ways of ordering the variables with x_i last). We need to prove that no matter which of the various D_i we pick for each i, the greatest common divisor of $D_0, \ldots, D_n$ is the resultant (up to a sign).

By Exercise 7, we know that Res divides D_n, and the same is clearly true for $D_0, \ldots, D_{n-1}$. Furthermore, the argument used in the proof of Proposition (4.6) shows that $D_i = \mathrm{Res} \cdot E_i$, where $E_i \in \mathbb{Z}[u_{i,\alpha}]$ doesn't involve the coefficients of F_i. It follows that

$$\mathrm{GCD}(D_0, \ldots, D_n) = \mathrm{Res} \cdot \mathrm{GCD}(E_0, \ldots, E_n).$$

Since each E_i doesn't involve the variables $u_{i,\alpha}$, the GCD on the right must be constant, i.e., an integer. However, since the coefficients of D_n are relatively prime (see Exercise 10 below), this integer must be ± 1, and we are done. Note that GCD's are only determined up to invertible elements, and in $\mathbb{Z}[u_{i,\alpha}]$, the only invertible elements are ± 1. □

Exercise 10. Show that $D_n(x_0^{d_0}, \ldots, x_n^{d_n}) = \pm 1$, and conclude that as a polynomial in $\mathbb{Z}[u_{i,\alpha}]$, the coefficients of D_n are relatively prime. Hint: If you order the monomials of total degree d appropriately, the matrix of (4.1) will be the identity matrix when $F_i = x_i^{d_i}$.

While the formula of Proposition (4.7) is very pretty, it is not particularly useful in practice. This brings us to our final resultant formula, which will tell us exactly how to find the extraneous factor in (4.5). The key idea, due to Macaulay, is that the extraneous factor is in fact a minor (i.e., the determinant of a submatrix) of the $N \times N$ matrix from (4.1). To describe this minor, we need to know which rows and columns of the matrix to delete. Recall also that we can label the rows and columns of the matrix of (4.1) using all monomials of total degree $d = \sum_{i=0}^n d_i - n$. Given such a monomial x^α, Exercise 1 implies that $x_i^{d_i}$ divides x^α for at least one i.

(4.8) Definition. Let $d_0, \ldots, d_n$ and d be as usual.

a. A monomial x^α of total degree d is *reduced* if $x_i^{d_i}$ divides x^α for *exactly* one i.

b. D'_n is the determinant of the submatrix of the coefficient matrix of (4.1) obtained by deleting all rows and columns corresponding to reduced monomials x^α.

Exercise 11. When $(d_0, d_1, d_2) = (1, 1, 2)$, we have $d = 2$. Show that all monomials of degree 2 are reduced except for xy. Then show that the $D'_3 = a_1$ corresponding to the submatrix (4.4) obtained by deleting everything but row 2 and column 4.

Exercise 12. Here are some properties of reduced monomials and D'_n.

a. Show that the number of reduced monomials is equal to

$$\sum_{j=0}^{n} d_0 \cdots d_{j-1} d_{j+1} \cdots d_n.$$

Hint: Adapt the argument used in Exercise 3.

b. Show that D'_n has the same total degree as the extraneous factor in (4.5) and that it doesn't depend on the coefficients of F_n. Hint: Use part a and note that all monomials in S_n are reduced.

Macaulay's observation is that the extraneous factor in (4.5) is exactly D'_n up to a sign. This gives the following formula for the resultant as a quotient of two determinants.

(4.9) Theorem. *When $F_0, \ldots, F_n$ are universal polynomials, the resultant is given by*

$$\mathrm{Res} = \pm \frac{D_n}{D'_n}.$$

Further, if k is any field and $F_0, \ldots, F_n \in k[x_0, \ldots, x_n]$, then the above formula for Res *holds whenever $D'_n \neq 0$.*

Proof. This is proved in Macaulay's paper [Mac2]. For a modern proof, see [Jou2]. □

Exercise 13. Using x_0, x_1, x_2 as variables with x_0 regarded as last, write $\mathrm{Res}_{1,2,2}$ as a quotient D_0/D'_0 of two determinants and write down the matrices involved (of sizes 10×10 and 2×2 respectively). The reason for using D_0/D'_0 instead of D_2/D'_2 will become clear in Exercise 2 of §5. A similar example is worked out in detail in [BGW].

While Theorem (4.9) applies to all resultants, it has some disadvantages. In the universal case, it requires dividing two very large polynomials, which can be very time consuming, and in the numerical case, we have the awkward situation where both D'_n and D_n vanish, as shown by the following exercise.

Exercise 14. Give an example of polynomials of degrees $1, 1, 2$ for which the resultant is nonzero yet the determinants D_2 and D_2' both vanish. Hint: See Exercise 10.

Because of this phenomenon, it would be nice if the resultant could be expressed as a single determinant, as happens with $\mathrm{Res}_{l,m}$. It is not known if this is possible in general, though many special cases have been found. We saw one example in the formula (2.8) for $\mathrm{Res}_{2,2,2}$. This can be generalized (in several ways) to give formulas for $\mathrm{Res}_{l,l,l}$ and $\mathrm{Res}_{l,l,l,l}$ when $l \geq 2$ (see [GKZ], Chapter 3, §4 and Chapter 13, §1, and [Sal], Arts. 90 and 91). As an example of these formulas, the following exercise will show how to express $\mathrm{Res}_{l,l,l}$ as a single determinant of size $2l^2 - l$ when $l \geq 2$.

Exercise 15. Suppose that $F_0, F_1, F_2 \in \mathbb{C}[x, y, z]$ have total degree $l \geq 2$. Before we can state our formula, we need to create some auxiliary equations. Given nonnegative integers a, b, c with $a + b + c = l - 1$, show that every monomial of total degree l in x, y, z is divisible by either x^{a+1}, y^{b+1}, or z^{c+1}, and conclude that we can write F_0, F_1, F_2 in the form

$$
\begin{aligned}
F_0 &= x^{a+1}P_0 + y^{b+1}Q_0 + z^{c+1}R_0\\
F_1 &= x^{a+1}P_1 + y^{b+1}Q_1 + z^{c+1}R_1\\
F_2 &= x^{a+1}P_2 + y^{b+1}Q_2 + z^{c+1}R_2.
\end{aligned}
\tag{4.10}
$$

There may be many ways of doing this. We will regard F_0, F_1, F_2 as universal polynomials and pick one particular choice for (4.10). Then set

$$
F_{a,b,c} = \det\begin{pmatrix} P_0 & Q_0 & R_0\\ P_1 & Q_1 & R_1\\ P_2 & Q_2 & R_2 \end{pmatrix}.
$$

You should check that $F_{a,b,c}$ has total degree $2l - 2$.

Then consider the equations

$$
\begin{aligned}
x^\alpha \cdot F_0 &= 0, &\quad x^\alpha &\text{ of total degree } l-2\\
x^\alpha \cdot F_1 &= 0, &\quad x^\alpha &\text{ of total degree } l-2\\
x^\alpha \cdot F_2 &= 0, &\quad x^\alpha &\text{ of total degree } l-2\\
F_{a,b,c} &= 0, &\quad x^a y^b z^c &\text{ of total degree } l-1.
\end{aligned}
\tag{4.11}
$$

Each polynomial on the left hand side has total degree $2l - 2$, and you should prove that there are $2l^2 - l$ monomials of this total degree. Thus we can regard the equations in (4.11) as having $2l^2 - l$ unknowns. You should also prove that the number of equations is $2l^2 - l$. Thus the coefficient matrix of (4.11), which we will denote C_l, is a $(2l^2 - l) \times (2l^2 - l)$ matrix.

In the following steps, you will prove that the resultant is given by

$$
\mathrm{Res}_{l,l,l}(F_0, F_1, F_2) = \pm \det(C_l).
$$

a. If $(u, v, w) \neq (0, 0, 0)$ is a solution of $F_0 = F_1 = F_2 = 0$, show that $F_{a,b,c}$ vanishes at (u, v, w). Hint: Regard (4.10) as a system of equations in unknowns $x^{a+1}, y^{b+1}, z^{c+1}$.
b. Use standard arguments to show that $\mathrm{Res}_{l,l,l}$ divides $\det(C_l)$.
c. Show that $\det(C_l)$ has degree l^2 in the coefficients of F_0. Show that the same is true for F_1 and F_2.
d. Conclude that $\mathrm{Res}_{l,l,l}$ is a multiple of $\det(C_l)$.
e. When $(F_0, F_1, F_2) = (x^l, y^l, z^l)$, show that $\det(C_l) = \pm 1$. Hint: Show that $F_{a,b,c} = x^{l-1-a}y^{l-1-b}z^{l-1-c}$ and that all monomials of total degree $2l-2$ not divisible by x^l, y^l, z^l can be written uniquely in this form. Then show that C_l is the identity matrix when the equations and monomials in (4.11) are ordered appropriately.
f. Conclude that $\mathrm{Res}_{l,l,l}(F_0, F_1, F_2) = \pm \det(C_l)$.

Exercise 16. Use Exercise 15 to compute the following resultants.
a. $\mathrm{Res}(x^2 + y^2 + z^2, xy + xz + yz, x^2 + 2xz + 3y^2)$.
b. $\mathrm{Res}(st + su + tu + u^2(1 - x), st + su + t^2 + u^2(2 - y), s^2 + su + tu - u^2 z)$, where the variables are s, t, u, and x, y, z are part of the coefficients. Note that your answer should agree with what you found in Exercise 3 of §2.

Other determinantal formulas for resultants can be found in [DD], [SZ], and [WZ]. We should also mention that besides the quotient formula given in Theorem (4.9), there are other ways to represent resultants as quotients. These go under a variety of names, including *Morley forms* [Jou1], *Bezoutians* [ElM1], and *Dixon matrices* [KSY]. See [EmM] for a survey. Computer implementations of resultants are available in [Lew] (for the Dixon formulation of [KSY]) and [WEM] (for the Macaulay formulation of Theorem (4.9)). Also, the Maple package `MR` implementing Theorem (4.9) can be found at `http://minimair.org/MR.mpl`.

We will end this section with a brief discussion of some of the practical aspects of computing resultants. All of the methods we've seen involve computing determinants or ratios of determinants. Since the usual formula for an $N \times N$ determinant involves $N!$ terms, we will need some clever methods for computing large determinants.

As Exercise 16 illustrates, the determinants can be either *numerical*, with purely numerical coefficients (as in part a of the exercise), or *symbolic*, with coefficients involving other variables (as in part b). Let's begin with numerical determinants. In most cases, this means determinants whose entries are rational numbers, which can be reduced to integer entries by clearing denominators. The key idea here is to reduce modulo a prime p and do arithmetic over the finite field $\mathbb{F}_p$ of the integers mod p. Computing the determinant here is easier since we are working over a field, which allows us to use standard algorithms from linear algebra (using row and column operations) to find the determinant. Another benefit is that we don't have to worry how big the numbers are getting (since we always reduce mod p). Hence we can compute the determinant mod p fairly easily. Then we do this for several primes $p_1, \ldots, p_r$ and use the Chinese Remainder Theorem to recover the original determinant. Strategies for how to choose the size and number of primes p_i are discussed in [CM] and [Man2], and the sparseness properties of the matrices in Theorem (4.9) are exploited in [CKL].

This method works fine provided that the resultant is given as a single determinant or a quotient where the denominator is nonzero. But when we have a situation like Exercise 14, where the denominator of the quotient is zero, something else is needed. One way to avoid this problem, due to Canny [Can1], is to prevent determinants from vanishing by making some coefficients symbolic. Suppose we have $F_0, \ldots, F_n \in \mathbb{Z}[x_0, \ldots, x_n]$. The determinants D_n and D_n' from Theorem (4.9) come from matrices we will denote M_n and M_n'. Thus the formula of the theorem becomes

$$\mathrm{Res}(F_0, \ldots, F_n) = \pm \frac{\det(M_n)}{\det(M_n')}$$

provided $\det(M_n') \neq 0$. When $\det(M_n') = 0$, Canny's method is to introduce a new variable u and consider the resultant

$$\mathrm{Res}(F_0 - u\,x_0^{d_0}, \ldots, F_n - u\,x_n^{d_n}). \tag{4.12}$$

Exercise 17. Fix an ordering of the monomials of total degree d. Since each equation in (4.1) corresponds to such a monomial, we can order the equations in the same way. The ordering of the monomials and equations determines the matrices M_n and M_n'. Then consider the new system of equations we get by replacing F_i by $F_i - u\,x_i^{d_i}$ in (4.1) for $0 \leq i \leq n$.

a. Show that the matrix of the new system of equations is $M_n - u\,I$, where I is the identity matrix of the same size as M_n.
b. Show that the matrix we get by deleting all rows and columns corresponding to reduced monomials, show that the matrix we get is $M_n' - u\,I$ where I is the appropriate identity matrix.

This exercise shows that the resultant (4.12) is given by

$$\mathrm{Res}(F_0 - u\,x_0^{d_0}, \ldots, F_n - u\,x_n^{d_n}) = \pm \frac{\det(M_n - u\,I)}{\det(M_n' - u\,I)}$$

since $\det(M_n' - u\,I) \neq 0$ (it is the characteristic polynomial of M_n'). It follows that the resultant $\mathrm{Res}(F_0, \ldots, F_n)$ is the constant term of the polynomial obtained by dividing $\det(M_n - u\,I)$ by $\det(M_n' - u\,I)$. In fact, as the following exercise shows, we can find the constant term directly from these polynomials:

Exercise 18. Let F and G be polynomials in u such that F is a multiple of G. Let $G = b_r u^r +$ higher order terms, where $b_r \neq 0$. Then $F = a_r u^r +$ higher order terms. Prove that the constant term of F/G is a_r/b_r.

It follows that the problem of finding the resultant is reduced to computing the determinants $\det(M_n - u\,I)$ and $\det(M_n' - u\,I)$. These are called *generalized characteristic polynomials* in [Can1].

This brings us to the second part of our discussion, the computation of symbolic determinants. The methods described above for the numerical case don't apply here, so something new is needed. One of the most interesting methods involves interpolation, as described in [CM]. The basic idea is that one can reconstruct a polynomial from its values at a sufficiently large number of points. More precisely, suppose we have a symbolic determinant, say involving variables $u_0, \ldots, u_n$. The determinant is then a polynomial $D(u_0, \ldots, u_n)$. Substituting $u_i = a_i$, where $a_i \in \mathbb{Z}$ for $0 \le i \le n$, we get a numerical determinant, which we can evaluate using the above method. Then, once we determine $D(a_0, \ldots, a_n)$ for sufficiently many points $(a_0, \ldots, a_n)$, we can reconstruct $D(u_0, \ldots, u_n)$. Roughly speaking, the number of points chosen depends on the degree of D in the variables $u_0, \ldots, u_n$. There are several methods for choosing points $(a_0, \ldots, a_n)$, leading to various interpolation schemes (Vandermonde, dense, sparse, probabilistic) which are discussed in [CM]. We should also mention that in the case of a single variable, there is a method of Manocha [Man2] for finding the determinant without interpolation.

Now that we know how to compute resultants, it's time to put them to work. In the next section, we will explain how resultants can be used to solve systems of polynomial equations. We should also mention that a more general notion of resultant, called the *sparse resultant*, will be discussed in Chapter 7.

Additional Exercises for §4

Exercise 19. Show that the number of monomials of total degree d in $n+1$ variables is the binomial coefficient $\binom{d+n}{n}$.

Exercise 20. This exercise is concerned with the proof of Proposition (4.6).

a. Suppose that $E \in \mathbb{Z}[u_{i,\alpha}]$ is irreducible and nonconstant. If $F \in \mathbb{Q}[u_{i,\alpha}]$ is such that $D = EF \in \mathbb{Z}[u_{i,\alpha}]$, then prove that $F \in \mathbb{Z}[u_{i,\alpha}]$. Hint: We can find a positive integer m such that $mF \in \mathbb{Z}[u_{i,\alpha}]$. Then apply unique factorization to $m \cdot D = E \cdot mF$.
b. Let $D = EF$ in $\mathbb{Z}[u_{i,\alpha}]$, and assume that for some j, D is homogeneous in the $u_{j,\alpha}$, $|\alpha| = d_j$. Then prove that E and F are also homogeneous in the $u_{j,\alpha}$, $|\alpha| = d_j$.

Exercise 21. In this exercise and the next we will prove the formula for $\mathrm{Res}_{2,2,2}$ given in equation (2.8). Here we prove two facts we will need.

a. Prove Euler's formula, which states that if $F \in k[x_0, \ldots, x_n]$ is homogeneous of total degree d, then
$$d\,F = \sum_{i=0}^{n} x_i \,\frac{\partial F}{\partial x_i}\,.$$
Hint: First prove it for a monomial of total degree d and then use linearity.

b. Suppose that
$$M = \det\begin{pmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{pmatrix},$$
where $A_1, \ldots, C_3$ are in $k[x_0, \ldots, x_n]$. Then prove that
$$\frac{\partial M}{\partial x_i} = \det\begin{pmatrix} \partial A_1/\partial x_i & A_2 & A_3 \\ \partial B_1/\partial x_i & B_2 & B_3 \\ \partial C_1/\partial x_i & C_2 & C_3 \end{pmatrix} + \det\begin{pmatrix} A_1 & \partial A_2/\partial x_i & A_3 \\ B_1 & \partial B_2/\partial x_i & B_3 \\ C_1 & \partial C_2/\partial x_i & C_3 \end{pmatrix}$$
$$+ \det\begin{pmatrix} A_1 & A_2 & \partial A_3/\partial x_i \\ B_1 & B_2 & \partial B_3/\partial x_i \\ C_1 & C_2 & \partial C_3/\partial x_i \end{pmatrix}.$$

Exercise 22. We can now prove formula (2.8) for $\mathrm{Res}_{2,2,2}$. Fix $F_0, F_1, F_2 \in \mathbb{C}[x, y, z]$ of total degree 2. As in §2, let J be the Jacobian determinant
$$J = \det\begin{pmatrix} \partial F_0/\partial x & \partial F_0/\partial y & \partial F_0/\partial z \\ \partial F_1/\partial x & \partial F_1/\partial y & \partial F_1/\partial z \\ \partial F_2/\partial x & \partial F_2/\partial y & \partial F_2/\partial z \end{pmatrix}.$$

a. Prove that J vanishes at every nontrivial solution of $F_0 = F_1 = F_2 = 0$. Hint: Apply Euler's formula (part a of Exercise 21) to F_0, F_1, F_2.

b. Show that
$$x \cdot J = 2\det\begin{pmatrix} F_0 & \partial F_0/\partial y & \partial F_0/\partial z \\ F_1 & \partial F_1/\partial y & \partial F_1/\partial z \\ F_2 & \partial F_2/\partial y & \partial F_2/\partial z \end{pmatrix},$$
and derive similar formulas for $y \cdot J$ and $z \cdot J$. Hint: Use column operations and Euler's formula.

c. By differentiating the formulas from part b for $x \cdot J$, $y \cdot J$ and $z \cdot J$ with respect to x, y, z, show that the partial derivatives of J vanish at all nontrivial solutions of $F_0 = F_1 = F_2 = 0$. Hint: Part b of Exercise 21 and part a of this exercise will be useful.

d. Use part c to show that the determinant in (2.8) vanishes at all nontrivial solutions of $F_0 = F_1 = F_2 = 0$.

e. Now prove (2.8). Hint: The proof is similar to what we did in parts b–f of Exercise 15.

Exercise 23. This exercise will give more details needed in the proof of Proposition (4.6). We will use the same terminology as in the proof. Let the weight of the variable $u_{i,\alpha}$ be $w(u_{i,\alpha})$.

a. Prove that a polynomial $P(u_{i,\alpha})$ is isobaric of weight m if and only if $P(\lambda^{w(u_{i,\alpha})}u_{i,\alpha}) = \lambda^m P(u_{i,\alpha})$ for all nonzero $\lambda \in \mathbb{C}$.
b. Prove that if $P = QR$ is isobaric, then so are Q and R. Also show that the weight of P is the sum of the weights of Q and R. Hint: Use part a.
c. Prove that D_n is isobaric of weight $d_0 \cdots d_n$. Hint: Assign the variables $x_0, \ldots, x_{n-1}, x_n$ respective weights $0, \ldots, 0, 1$. Let x^γ be a monomial with $|\gamma| = d$ (which indexes a column of D_n), and let $\alpha \in S_i$ (which indexes a row in D_n). If the corresponding entry in D_n is $c_{\gamma,\alpha,i}$, then show that
$$\begin{aligned} w(c_{\gamma,\alpha,i}) &= w(x^\gamma) - w(x^\alpha/x_i^{d_i}) \\ &= w(x^\gamma) - w(x^\alpha) + \begin{cases} 0 & i < n \\ d_n & i = n. \end{cases} \end{aligned}$$
Note that x^γ and x^α range over *all* monomials of total degree d.
d. Use Theorems (2.3) and (3.1) to prove that if u_i represents the coefficient of $x_i^{d_i}$ in F_i, then $\pm u_0^{d_1\cdots d_n} \cdots u_n^{d_0\cdots d_{n-1}}$ is in Res.

§5 Solving Equations via Resultants

In this section, we will show how resultants can be used to solve polynomial systems. To start, suppose we have n homogeneous polynomials $F_1, \ldots, F_n$ of degree $d_1, \ldots, d_n$ in variables $x_0, \ldots, x_n$. We want to find the nontrivial solutions of the system of equations

$$F_1 = \cdots = F_n = 0. \tag{5.1}$$

But before we begin our discussion of finding solutions, we first need to review Bézout's Theorem and introduce the important idea of *genericity*.

As we saw in §3, Bézout's Theorem tells us that when (5.1) has finitely many solutions in $\mathbb{P}^n$, the number of solutions is $d_1 \cdots d_n$, counting multiplicities. In practice, it is often convenient to find solutions in affine space. In §3, we dehomogenized by setting $x_n = 1$, but in order to be compatible with Chapter 7, we now dehomogenize using $x_0 = 1$. Hence, we define:

$$\begin{aligned} f_i(x_1, \ldots, x_n) &= F_i(1, x_1, \ldots, x_n) \\ \overline{F}_i(x_1, \ldots, x_n) &= F_i(0, x_1, \ldots, x_n). \end{aligned} \tag{5.2}$$

Note that f_i has total degree at most d_i. Inside $\mathbb{P}^n$, we have the affine space $\mathbb{C}^n \subset \mathbb{P}^n$ defined by $x_0 = 1$, and the solutions of the affine equations

$$f_1 = \cdots = f_n = 0 \tag{5.3}$$

are precisely the solutions of (5.1) which lie in $\mathbb{C}^n \subset \mathbb{P}^n$. Similarly, the nontrivial solutions of the homogeneous equations

$$\overline{F}_1 = \cdots = \overline{F}_n = 0$$

may be regarded as the solutions which lie "at ∞". We say that (5.3) has *no solutions at* ∞ if $\overline{F}_1 = \cdots = \overline{F}_n = 0$ has no nontrivial solutions. By Theorem (2.3), this is equivalent to the condition

$$\mathrm{Res}_{d_1,\ldots,d_n}(\overline{F}_1, \ldots, \overline{F}_n) \neq 0. \tag{5.4}$$

The proof of Theorem (3.4) implies the following version of Bézout's Theorem.

(5.5) Theorem (Bézout's Theorem). *Assume that $f_1, \ldots, f_n$ are defined as in (5.2) and that the affine equations (5.3) have no solutions at ∞. Then these equations have $d_1 \cdots d_n$ solutions (counted with multiplicity), and the ring*

$$A = \mathbb{C}[x_1, \ldots, x_n]/\langle f_1, \ldots, f_n \rangle$$

has dimension $d_1 \cdots d_n$ as a vector space over $\mathbb{C}$.

Note that this result does not hold for all systems of equations (5.3). In general, we need a language which allows us to talk about properties which are true for most but not necessarily all polynomials $f_1, \ldots, f_n$. This brings us to the idea of genericity.

(5.6) Definition. A property is said to *hold generically* for polynomials $f_1, \ldots, f_n$ of degree at most $d_1, \ldots, d_n$ if there is a nonzero polynomial in the coefficients of the f_i such that the property holds for all $f_1, \ldots, f_n$ for which the polynomial is nonvanishing.

Intuitively, a property of polynomials is *generic* if it holds for "most" polynomials $f_1, \ldots, f_n$. Our definition makes this precise by defining "most" to mean that some polynomial in the coefficients of the f_i is nonvanishing. As a simple example, consider a single polynomial $ax^2 + bx + c$. We claim that the property "$ax^2 + bx + c = 0$ has two solutions, counting multiplicity" holds generically. To prove this, we must find a polynomial in the coefficients a, b, c whose nonvanishing implies the desired property. Here, the condition is easily seen to be $a \neq 0$ since we are working over the complex numbers.

Exercise 1. Show that the property "$ax^2 + bx + c = 0$ has two distinct solutions" is generic. Hint: By the quadratic formula, $a(b^2 - 4ac) \neq 0$ implies the desired property.

A more relevant example is given by Theorem (5.5). Having no solutions at ∞ is equivalent to the nonvanishing of the resultant (5.4), and since $\mathrm{Res}_{d_1,\ldots,d_n}(\overline{F}_1,\ldots,\overline{F}_n)$ is a nonzero polynomial in the coefficients of the f_i, it follows that this version of Bézout's Theorem holds generically. Thus, for most choices of the coefficients, the equations $f_1 = \cdots = f_n = 0$ have $d_1 \cdots d_n$ solutions, counting multiplicity. In particular, if we choose polynomials $f_1, \ldots, f_n$ with random coefficients (say given by some random number generator), then, with a very high probability, Bézout's Theorem will hold for the corresponding system of equations.

In general, genericity comes in different "flavors". For instance, consider solutions of the equation $ax^2 + bx + c = 0$:

- Generically, $ax^2 + bx + c = 0$ has two solutions, counting multiplicity. This happens when $a \neq 0$.
- Generically, $ax^2 + bx + c = 0$ has two distinct solutions. By Exercise 1, this happens when $a(b^2 - 4ac) \neq 0$.

Similarly, there are different versions of Bézout's Theorem. In particular, one can strengthen Theorem (5.5) to prove that generically, the equations $f_1 = \cdots = f_n = 0$ have $d_1 \cdots d_n$ distinct solutions. This means that generically, (5.3) has no solutions at ∞ and all solutions have multiplicity one. A proof of this result will be sketched in Exercise 6 at the end of the section.

With this genericity assumption on $f_1, \ldots, f_n$, we know the number of distinct solutions of (5.3), and our next task is to find them. We could use the methods of Chapter 2, but it is also possible to find the solutions using resultants. This section will describe two closely related methods, u-resultants and hidden variables, for solving equations. The next section will discuss further methods which use eigenvalues and eigenvectors.

The u-Resultant

The basic idea of van der Waerden's *u-resultant* (see [vdW]) is to start with the homogeneous equations $F_1 = \cdots = F_n = 0$ of (5.1) and add another equation $F_0 = 0$ to (5.1), so that we have $n+1$ homogeneous equations in $n+1$ variables. We will use

$$F_0 = u_0 x_0 + \cdots + u_n x_n,$$

where $u_0, \ldots, u_n$ are independent variables. Because the number of equations equals the number of variables, we can form the resultant

$$\mathrm{Res}_{1,d_1,\ldots,d_n}(F_0, F_1, \ldots, F_n),$$

which is called the *u-resultant*. Note that the u-resultant is a polynomial in $u_0, \ldots, u_n$.

As already mentioned, we will sometimes work in the affine situation, where we dehomogenize $F_0, \ldots, F_n$ to obtain $f_0, \ldots, f_n$. This is the notation of (5.2), and in particular, observe that

$$f_0 = u_0 + u_1 x_1 + \cdots + u_n x_n. \tag{5.7}$$

Because $f_0, \ldots, f_n$ and $F_0, \ldots, F_n$ have the same coefficients, we write the u-resultant as $\mathrm{Res}(f_0, \ldots, f_n)$ instead of $\mathrm{Res}(F_0, \ldots, F_n)$ in this case.

Before we work out the general theory of the u-resultant, let's do an example. The following exercise will seem like a lot of work at first, but its surprising result will be worth the effort.

Exercise 2. Let

$$\begin{aligned} F_1 &= x_1^2 + x_2^2 - 10x_0^2 = 0 \\ F_2 &= x_1^2 + x_1 x_2 + 2x_2^2 - 16x_0^2 = 0 \end{aligned}$$

be the intersection of a circle and an ellipse in $\mathbb{P}^2$. By Bézout's Theorem, there are four solutions. To find the solutions, we add the equation

$$F_0 = u_0 x_0 + u_1 x_1 + u_2 x_2 = 0.$$

a. The theory of §4 computes the resultant using 10×10 determinants D_0, D_1 and D_2. Using D_0, Theorem (4.9) implies

$$\mathrm{Res}_{1,2,2}(F_0, F_1, F_2) = \pm \frac{D_0}{D_0'}.$$

If the variables are ordered x_2, x_1, x_0, show that $D_0 = \det(M_0)$, where M_0 is the matrix

$$M_0 = \begin{pmatrix} u_0 & u_1 & u_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & u_0 & 0 & u_2 & u_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & u_0 & u_1 & 0 & u_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & u_0 & 0 & 0 & 0 & u_1 & u_2 & 0 \\ -10 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -10 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & -10 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ -16 & 0 & 0 & 1 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & -16 & 0 & 0 & 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & -16 & 0 & 0 & 0 & 0 & 1 & 1 & 2 \end{pmatrix}.$$

Also show that $D_0' = \det(M_0')$, where M_0' is given by

$$M_0' = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Hint: Using the order x_2, x_1, x_0 gives $S_0 = \{x_0^3, x_0^2 x_1, x_0^2 x_2, x_0 x_1 x_2\}$, $S_1 = \{x_0 x_1^2, x_1^3, x_1^2 x_2\}$ and $S_2 = \{x_0 x_2^2, x_1 x_2^2, x_2^3\}$. The columns in M_0 correspond to the monomials x_0^3, $x_0^2 x_1$, $x_0^2 x_2$, $x_0 x_1 x_2$, $x_0 x_1^2$, $x_0 x_2^2$, x_1^3, $x_1^2 x_2$, $x_1 x_2^2$, x_2^3. Exercise 13 of §4 will be useful.

b. Conclude that

$$\begin{aligned}\mathrm{Res}_{1,2,2}(F_0, F_1, F_2) = \pm\, &(2u_0^4 + 16u_1^4 + 36u_2^4 - 80u_1^3u_2 + 120u_1u_2^3\\ &- 18u_0^2u_1^2 - 22u_0^2u_2^2 + 52u_1^2u_2^2 - 4u_0^2u_1u_2).\end{aligned}$$

c. Using a computer to factor this, show that $\mathrm{Res}_{1,2,2}(F_0, F_1, F_2)$ equals

$$(u_0 + u_1 - 3u_2)(u_0 - u_1 + 3u_2)(u_0^2 - 8u_1^2 - 2u_2^2 - 8u_1u_2)$$

up to a constant. By writing the quadratic factor as $u_0^2 - 2(2u_1 + u_2)^2$, conclude that $\mathrm{Res}_{1,2,2}(F_0, F_1, F_2)$ equals

$$(u_0 + u_1 - 3u_2)(u_0 - u_1 + 3u_2)(u_0 + 2\sqrt{2}u_1 + \sqrt{2}u_2)(u_0 - 2\sqrt{2}u_1 - \sqrt{2}u_2)$$

times a nonzero constant. Hint: If you are using Maple, let the resultant be `res` and use the command `factor(res)`. Also, the command `factor(res,RootOf(x^2-2))` will do the complete factorization.

d. The coefficients of the linear factors of $\mathrm{Res}_{1,2,2}(F_0, F_1, F_2)$ give four points

$$(1, 1, -3),\ (1, -1, 3),\ (1, 2\sqrt{2}, \sqrt{2}),\ (1, -2\sqrt{2}, -\sqrt{2})$$

in $\mathbb{P}^2$. Show that these points are the four solutions of the equations $F_1 = F_2 = 0$. Thus the solutions in $\mathbb{P}^2$ are *precisely* the coefficients of the linear factors of $\mathrm{Res}_{1,2,2}(F_0, F_1, F_2)$!

In this exercise, all of the solutions lay in the affine space $\mathbb{C}^2 \subset \mathbb{P}^2$ defined by $x_0 = 1$. In general, we will study the u-resultant from the affine point of view. The key fact is that when all of the multiplicities are one, the solutions of (5.3) can be found using $\mathrm{Res}_{1,d_1,\ldots,d_n}(f_0, \ldots, f_n)$.

(5.8) Proposition. *Assume that $f_1 = \cdots = f_n = 0$ have total degrees bounded by $d_1, \ldots, d_n$, no solutions at ∞, and all solutions of multiplicity one. If $f_0 = u_0 + u_1x_1 + \cdots + u_nx_n$, where $u_0, \ldots, u_n$ are independent variables, then there is a nonzero constant C such that*

$$\mathrm{Res}_{1,d_1,\ldots,d_n}(f_0, \ldots, f_n) = C \prod_{p \in \mathbf{V}(f_1,\ldots,f_n)} f_0(p).$$

Proof. Let $C = \mathrm{Res}_{d_1,\ldots,d_n}(\overline{F}_1, \ldots, \overline{F}_n)$, which is nonzero by hypothesis. Since the coefficients of f_0 are the variables $u_0, \ldots, u_n$, we need to work over the field $K = \mathbb{C}(u_0, \ldots, u_n)$ of rational functions in $u_0, \ldots, u_n$. Hence, in this proof, we will work over K rather than over $\mathbb{C}$. Fortunately, the results we need are true over K, even though we proved them only over $\mathbb{C}$.

Adapting Theorem (3.4) to the situation of (5.2) (see Exercise 8) yields

$$\mathrm{Res}_{1,d_1,\ldots,d_n}(f_0, \ldots, f_n) = C\ \det(m_{f_0}),$$

where $m_{f_0} : A \to A$ is the linear map given by multiplication by f_0 on the quotient ring

$$A = K[x_1, \dots, x_n]/\langle f_1, \dots, f_n\rangle.$$

By Theorem (5.5), A is a vector space over K of dimension $d_1 \cdots d_n$, and Theorem (4.5) of Chapter 2 implies that the eigenvalues of m_{f_0} are the values $f_0(p)$ for $p \in \mathbf{V}(f_1, \dots, f_n)$. Since all multiplicities are one, there are $d_1 \cdots d_n$ such points p, and the corresponding values $f_0(p)$ are distinct since $f_0 = u_0 + u_1x_1 + \cdots + u_nx_n$ and $u_0, \dots, u_n$ are independent variables. Thus m_{f_0} has $d_1 \cdots d_n$ distinct eigenvalues $f_0(p)$, so that

$$\det(m_{f_0}) = \prod_{p\in\mathbf{V}(f_1,\dots,f_n)} f_0(p).$$

This proves the proposition. □

To see more clearly what the proposition says, let the points of $\mathbf{V}(f_1, \dots, f_n)$ be p_i for $1 \le i \le d_1 \cdots d_n$. If we write each point as $p_i = (a_{i1}, \dots, a_{in}) \in \mathbb{C}^n$, then (5.7) implies

$$f_0(p_i) = u_0 + a_{i1}u_1 + \cdots + a_{in}u_n,$$

so that by Proposition (5.8), the u-resultant is given by

$$\text{(5.9)} \qquad \mathrm{Res}_{1,d_1,\dots,d_n}(f_0, \dots, f_n) = C \prod_{i=1}^{d_1\cdots d_n} \big(u_0 + a_{i1}u_1 + \cdots + a_{in}u_n\big).$$

We see clearly that the u-resultant is a polynomial in $u_0, \dots, u_n$. Furthermore, we get the following method for finding solutions of (5.3): compute $\mathrm{Res}_{1,d_1,\dots,d_n}(f_0, \dots, f_n)$, factor it into linear factors, and then read off the solutions! Hence, once we have the u-resultant, solving (5.3) is reduced to a problem in multivariable factorization.

To compute the u-resultant, we use Theorem (4.9). Because of our emphasis on f_0, we represent the resultant as the quotient

$$\text{(5.10)} \qquad \mathrm{Res}_{1,d_1,\dots,d_n}(f_0, \dots, f_n) = \pm\frac{D_0}{D_0'}.$$

This is the formula we used in Exercise 2. In §4, we got the determinant D_0 by working with the homogenizations F_i of the f_i, regarding x_0 as the last variable, and decomposing monomials of degree $d = 1 + d_1 + \cdots + d_n - n$ into disjoint subsets $S_0, \dots, S_n$. Taking x_0 last means that S_0 consists of the $d_1 \cdots d_n$ monomials

$$\text{(5.11)} \quad S_0 = \{x_0^{a_0}x_1^{a_1}\cdots x_n^{a_n} : 0 \le a_i \le d_i - 1 \text{ for } i > 0,\ \textstyle\sum_{i=0}^n a_i = d\}.$$

Then D_0 is the determinant of the matrix M_0 representing the system of equations (4.1). We saw an example of this in Exercise 2.

The following exercise simplifies the task of computing u-resultants.

Exercise 3. Assuming that $D_0' \neq 0$ in (5.10), prove that D_0' does not involve $u_0, \dots, u_n$ and conclude that $\mathrm{Res}_{1,d_1,\dots,d_n}(f_0, \dots, f_n)$ and D_0 differ by a constant factor when regarded as polynomials in $\mathbb{C}[u_0, \dots, u_n]$.

We will write D_0 as $D_0(u_0,\ldots,u_n)$ to emphasize the dependence on $u_0,\ldots,u_n$. We can use $D_0(u_0,\ldots,u_n)$ only when $D_0' \neq 0$, but since D_0' is a polynomial in the coefficients of the f_i, Exercise 3 means that generically, the linear factors of the determinant $D_0(u_0,\ldots,u_n)$ give the solutions of our equations (5.3). In this situation, we will apply the term *u-resultant* to both $\mathrm{Res}_{1,d_1,\ldots,d_n}(f_0,\ldots,f_n)$ and $D_0(u_0,\ldots,u_n)$.

Unfortunately, the u-resultant has some serious limitations. First, it is not easy to compute symbolic determinants of large size (see the discussion at the end of §4). And even if we can find the determinant, multivariable factorization as in (5.9) is very hard, especially since in most cases, floating point numbers will be involved.

There are several methods for dealing with this situation. We will describe one, as presented in [CM]. The basic idea is to specialize some of the coefficients in $f_0 = u_0 + u_1x_1 + \cdots + u_nx_n$. For example, the argument of Proposition (5.8) shows that when the x_n-coordinates of the solution points are distinct, the specialization $u_1 = \cdots = u_{n-1} = 0, u_n = -1$ transforms (5.9) into the formula

$$(5.12) \qquad \mathrm{Res}_{1,d_1,\ldots,d_n}(u_0 - x_n, f_1,\ldots,f_n) = C\prod_{i=1}^{d_1\cdots d_n}\bigl(u_0 - a_{in}\bigr),$$

where a_{in} is the x_n-coordinate of $p_i = (a_{i1},\ldots,a_{in}) \in \mathbf{V}(f_1,\ldots,f_n)$. This resultant is a univariate polynomial in u_0 whose roots are precisely the x_n-coordinates of solutions of (5.3). There are similar formulas for the other coordinates of the solutions.

If we use the numerator $D_0(u_0,\ldots,u_n)$ of (5.10) as the u-resultant, then setting $u_1 = \cdots = u_n = 0, u_n = -1$ gives $D_0(u_0,0,\ldots,0,-1)$, which is a polynomial in u_0. The argument of Exercise 3 shows that generically, $D_0(u_0,0,\ldots,0,-1)$ is a constant multiple $\mathrm{Res}(u_0 - x_n, f_1,\ldots,f_n)$, so that its roots are also the x_n-coordinates. Since $D_0(u_0,0,\ldots,0,-1)$ is given by a symbolic determinant depending on the single variable u_0, it is *much* easier to compute than in the multivariate case. Using standard techniques (discussed in Chapter 2) for finding the roots of univariate polynomials such as $D_0(u_0,0,\ldots,0,-1)$, we get a computationally efficient method for finding the x_n-coordinates of our solutions. Similarly, we can find the other coordinates of the solutions by this method.

Exercise 4. Let $D_0(u_0,u_1,u_2)$ be the determinant in Exercise 2.

a. Compute $D_0(u_0,-1,0)$ and $D_0(u_0,0,-1)$.
b. Find the roots of these polynomials numerically. Hint: Try the Maple command `fsolve`. In general, `fsolve` should be used with the `complex` option, though in this case it's not necessary since the roots are real.
c. What does this say about the coordinates of the solutions of the equations $x_1^2 + x_2^2 = 10$, $x_1^2 + x_1x_2 + 2x_2^2 = 16$? Can you figure out what the solutions are?

As this exercise illustrates, the univariate polynomials we get from the u-resultant enable us to find the individual coordinates of the solutions, but they don't tell us how to match them up. One method for doing this (based on [CM]) will be explained in Exercise 7 at the end of the section. We should also mention that a different u-resultant method for computing solutions is given in [Can2].

All of the u-resultant methods make strong genericity assumptions on the polynomials $f_0, \ldots, f_n$. In practice, one doesn't know in advance if a given system of equations is generic. Here are some of the things that can go wrong when trying to apply the above methods to non-generic equations:

- There might be solutions at infinity. This problem can be avoided by making a generic linear change of coordinates.
- If too many coefficients are zero, it might be necessary to use the sparse resultants of Chapter 7.
- The equations (5.1) might have infinitely many solutions. In the language of algebraic geometry, the projective variety $\mathbf{V}(F_1, \ldots, F_n)$ might have components of positive dimension, together with some isolated solutions. One is still interested in the isolated solutions, and techniques for finding them are described in Section 4 of [Can1].
- The denominator D_0' in the resultant formula (5.10) might vanish. When this happens, one can use the *generalized characteristic polynomials* described in §4 to avoid this difficulty. See Section 4.1 of [CM] for details.
- Distinct solutions might have the same x_i-coordinate for some i. The polynomial giving the x_i-coordinates would have multiple roots, which are computationally unstable. This problem can be avoided with a generic change of coordinates. See Section 4.2 of [CM] for an example.

Also, Chapter 4 will give versions of (5.12) and Proposition (5.8) for the case when $f_1 = \cdots = f_n = 0$ has solutions of multiplicity > 1.

Hidden Variables

One of the better known resultant techniques for solving equations is the *hidden variable* method. The basic idea is to regard one of variables as a constant and then take a resultant. To illustrate how this works, consider the affine equations we get from Exercise 2 by setting $x_0 = 1$:

$$\text{(5.13)} \qquad \begin{aligned} f_1 &= x_1^2 + x_2^2 - 10 = 0 \\ f_2 &= x_1^2 + x_1x_2 + 2x_2^2 - 16 = 0. \end{aligned}$$

If we regard x_2 as a constant, we can use the resultant of §1 to obtain

$$\mathrm{Res}(f_1, f_2) = 2x_2^4 - 22x_2^2 + 36 = 2(x_2 - 3)(x_2 + 3)(x_2 - \sqrt{2})(x_2 + \sqrt{2}).$$

The resultant is a polynomial in x_2, and its roots are *precisely* the x_2-coordinates of the solutions of the equations (as we found in Exercise 2).

To generalize this example, we first review the affine form of the resultant. Given $n+1$ homogeneous polynomials $G_0, \ldots, G_n$ of degrees $d_0, \ldots, d_n$ in $n+1$ variables $x_0, \ldots, x_n$, we get $\mathrm{Res}_{d_0,\ldots,d_n}(G_0, \ldots, G_n)$. Setting $x_0 = 1$ gives

$$g_i(x_1, \ldots, x_n) = G_i(1, x_1, \ldots, x_n),$$

and since the g_i and G_i have the same coefficients, we can write the resultant as $\mathrm{Res}_{d_0,\ldots,d_1}(g_0, \ldots, g_n)$. Thus, $n+1$ polynomials $g_0, \ldots, g_n$ in n variables $x_1, \ldots, x_n$ have a resultant. It follows that from the affine point of view, forming a resultant requires that *the number of polynomials be one more than the number of variables.*

Now, suppose we have n polynomials $f_1, \ldots, f_n$ of degrees $d_1, \ldots, d_n$ in n variables $x_1, \ldots, x_n$. In terms of resultants, we have the wrong numbers of equations and variables. One solution is to add a new polynomial, which leads to the u-resultant. Here, we will pursue the other alternative, which is to get rid of one of the variables. The basic idea is what we did above: we *hide* a variable, say x_n, by regarding it as a constant. This gives n polynomials $f_1, \ldots, f_n$ in $n-1$ variables $x_1, \ldots, x_{n-1}$, which allows us to form their resultant. We will write this resultant as

$$\mathrm{Res}^{x_n}_{d_1,\ldots,d_n}(f_1, \ldots, f_n). \tag{5.14}$$

The superscript x_n reminds us that we are regarding x_n as constant. Since the resultant is a polynomial in the coefficients of the f_i, (5.14) is a polynomial in x_n.

We can now state the main result of the hidden variable technique.

(5.15) Proposition. *Generically,* $\mathrm{Res}^{x_n}_{d_1,\ldots,d_n}(f_1, \ldots, f_n)$ *is a polynomial in* x_n *whose roots are the* x_n*-coordinates of the solutions of (5.3).*

Proof. The basic strategy of the proof is that by (5.12), we already know a polynomial whose roots are the x_n-coordinates of the solutions, namely

$$\mathrm{Res}_{1,d_1,\ldots,d_n}(u_0 - x_n, f_1, \ldots, f_n).$$

We will prove the theorem by showing that this polynomial is the same as the hidden variable resultant (5.14). However, (5.14) is a polynomial in x_n, while $\mathrm{Res}(u_0 - x_n, f_1, \ldots, f_n)$ is a polynomial in u_0. To compare these two polynomials, we will write

$$\mathrm{Res}^{x_n=u_0}_{d_1,\ldots,d_n}(f_1, \ldots, f_n)$$

to mean the polynomial obtained from (5.14) by the substitution $x_n = u_0$. Using this notation, the theorem will follow once we show that

$$\mathrm{Res}^{x_n=u_0}_{d_1,\ldots,d_n}(f_1, \ldots, f_n) = \pm\mathrm{Res}_{1,d_1,\ldots,d_n}(u_0 - x_n, f_1, \ldots, f_n).$$

We will prove this equality by applying Theorem (3.4) separately to the two resultants in this equation.

Beginning with $\mathrm{Res}(u_0 - x_n, f_1, \ldots, f_n)$, first recall that it equals the homogeneous resultant $\mathrm{Res}(u_0x_0 - x_n, F_1, \ldots, F_n)$ via (5.2). Since u_0 is a coefficient, we will work over the field $\mathbb{C}(u_0)$ of rational functions in u_0. Then, adapting Theorem (3.4) to the situation of (5.2) (see Exercise 8), we see that $\mathrm{Res}(u_0x_0 - x_n, F_1, \ldots, F_n)$ equals

$$(5.16) \qquad \mathrm{Res}_{1,d_1,\ldots,d_{n-1}}(-x_n, \overline{F}_1, \ldots, \overline{F}_{n-1})^{d_n} \det(m_{f_n}),$$

where $-x_n, \overline{F}_1, \ldots, \overline{F}_{n-1}$ are obtained from $u_0x_0 - x_n, F_1, \ldots, F_{n-1}$ by setting $x_0 = 0$, and $m_{f_n} : A \to A$ is multiplication by f_n in the ring

$$A = \mathbb{C}(u)[x_1, \ldots, x_n]/\langle u - x_n, f_1, \ldots, f_n\rangle.$$

Next, consider $\mathrm{Res}^{x_n=u_0}(f_1, \ldots, f_n)$, and observe that if we define

$$\hat{f}_i(x_1, \ldots, x_{n-1}) = f_i(x_1, \ldots, x_{n-1}, u_0),$$

then $\mathrm{Res}^{x_n=u_0}(f_1, \ldots, f_n) = \mathrm{Res}(\hat{f}_1, \ldots, \hat{f}_n)$. If we apply Theorem (3.4) to the latter resultant, we see that it equals

$$(5.17) \qquad \mathrm{Res}_{d_1,\ldots,d_{n-1}}(\widetilde{F}_1, \ldots, \widetilde{F}_{n-1})^{d_n} \det(m_{\hat{f}_n}),$$

where $\widetilde{F}_i$ is obtained from $\hat{f}_i$ by first homogenizing with respect to x_0 and then setting $x_0 = 0$, and $m_{\hat{f}_n} : \widehat{A} \to \widehat{A}$ is multiplication by $\hat{f}_n$ in

$$\widehat{A} = \mathbb{C}(u_0)[x_1, \ldots, x_{n-1}]/\langle \hat{f}_1, \ldots, \hat{f}_n\rangle.$$

To show that (5.16) and (5.17) are equal, we first examine (5.17). We claim that if f_i homogenizes to F_i, then $\widetilde{F}_i$ in (5.17) is given by

$$(5.18) \qquad \widetilde{F}_i(x_1, \ldots, x_{n-1}) = F_i(0, x_1, \ldots, x_{n-1}, 0).$$

To prove this, take a term of F_i, say

$$c\, x_0^{a_0} \cdots x_n^{a_n}, \quad a_0 + \cdots + a_n = d_i.$$

Since $x_0 = 1$ gives f_i and $x_n = u_0$ then gives $\hat{f}_i$, the corresponding term in $\hat{f}_i$ is

$$c\, 1^{a_0} x_1^{a_1} \cdots x_{n-1}^{a_{n-1}} u_0^{a_n} = c u_0^{a_n} \cdot x_1^{a_1} \cdots x_{n-1}^{a_{n-1}}.$$

When homogenizing $\hat{f}_i$ with respect to x_0, we want a term of total degree d_i in $x_0, \ldots, x_{n-1}$. Since $cu_0^{a_n}$ is a constant, we get

$$cu_0^{a_n} \cdot x_0^{a_0+a_n} x_1^{a_1} \cdots x_{n-1}^{a_{n-1}} = c \cdot x_0^{a_0} \cdots x_{n-1}^{a_{n-1}} (u_0x_0)^{a_n}.$$

It follows that the homogenization of $\hat{f}_i$ is $F_i(x_0, \ldots, x_{n-1}, u_0x_0)$, and since $\widetilde{F}_i$ is obtained by setting $x_0 = 0$ in this polynomial, we get (5.18).

Once we know (5.18), Exercise 12 of §3 shows that

$$\mathrm{Res}_{1,d_1,\ldots,d_{n-1}}(-x_n, \overline{F}_1, \ldots, \overline{F}_{n-1}) = \pm\mathrm{Res}_{d_1,\ldots,d_{n-1}}(\widetilde{F}_1, \ldots, \widetilde{F}_{n-1})$$

since $\overline{F}_i(x_1, \ldots, x_n) = F_i(0, x_1, \ldots, x_n)$. Also, the ring homomorphism

$$\mathbb{C}(u_0)[x_1, \ldots, x_n] \to \mathbb{C}(u_0)[x_1, \ldots, x_{n-1}]$$

defined by $x_n \mapsto u_0$ carries f_i to $\hat{f}_i$. It follows that this homomorphism induces a ring isomorphism $A \cong \widehat{A}$ (you will check the details of this in Exercise 8). Moreover, multiplication by f_n and $\hat{f}_n$ give a diagram

$$\begin{array}{ccc} A & \cong & \widehat{A} \\ {\scriptstyle m_{f_n}}\big\downarrow & & \big\downarrow{\scriptstyle m_{\hat{f}_n}} \\ A & \cong & \widehat{A} \end{array} \tag{5.19}$$

In Exercise 8, you will show that going across and down gives the same map $A \to \widehat{A}$ as going down and across (we say that (5.19) is a *commutative diagram*). From here, it is easy to show that $\det(m_{f_n}) = \det(m_{\hat{f}_n})$, and it follows that (5.16) and (5.17) are equal. □

The advantage of the hidden variable method is that it involves resultants with fewer equations and variables than the u-resultant. For example, when dealing with the equations $f_1 = f_2 = 0$ from (5.13), the u-resultant $\mathrm{Res}_{1,2,2}(f_0, f_1, f_2)$ uses the 10×10 matrix from Exercise 2, while $\mathrm{Res}^{x_2}_{2,2}(f_1, f_2)$ only requires a 4×4 matrix.

In general, we can compute $\mathrm{Res}^{x_n}(f_1, \ldots, f_n)$ by Theorem (4.9), and as with the u-resultant, we can again ignore the denominator. More precisely, if we write

$$\mathrm{Res}^{x_n}_{d_1,\ldots,d_n}(f_1, \ldots, f_n) = \pm \frac{\widehat{D}_0}{\widehat{D}'_0}, \tag{5.20}$$

then $\widehat{D}'_0$ doesn't involve x_n. The proof of this result is a nice application of Proposition (4.6), and the details can be found in Exercise 10 at the end of the section. Thus, when using the hidden variable method, it suffices to use the numerator $\widehat{D}_0$—when $f_1, \ldots, f_n$ are generic, its roots give the x_n-coordinates of the affine equations (5.3).

Of course, there is nothing special about hiding x_n—we can hide any of the variables in the same way, so that the hidden variable method can be used to find the x_i-coordinates of the solutions for any i. One limitation of this method is that it only gives the individual coordinates of the solution points and doesn't tell us how they match up.

Exercise 5. Consider the affine equations

$$\begin{aligned} f_1 &= x_1^2 + x_2^2 + x_3^2 - 3 \\ f_2 &= x_1^2 + x_3^2 - 2 \\ f_3 &= x_1^2 + x_2^2 - 2x_3. \end{aligned}$$

a. If we compute the u-resultant with $f_0 = u_0 + u_1x_1 + u_2x_2 + u_3x_3$, show that Theorem (4.9) expresses $\mathrm{Res}_{1,2,2,2}(f_0, f_1, f_2, f_3)$ as a quotient of determinants of sizes 35×35 and 15×15 respectively.
b. If we hide x_3, show that $\mathrm{Res}^{x_3}_{2,2,2}(f_1, f_2, f_3)$ is a quotient of determinants of sizes 15×15 and 3×3 respectively.
c. Hiding x_3 as in part b, use (2.8) to express $\mathrm{Res}^{x_3}_{2,2,2}(f_1, f_2, f_3)$ as the determinant of a 6×6 matrix, and show that up to a constant, the resultant is $(x_3^2 + 2x_3 - 3)^4$. Explain the significance of the exponent 4. Hint: You will need to regard x_3 as a constant and homogenize the f_i with respect to x_0. Then (2.8) will be easy to apply.

The last part of Exercise 5 illustrates how formulas such as (2.8) allow us, in special cases, to represent a resultant as a *single* determinant of relatively small size. This can reduce dramatically the amount of computation involved and explains the continuing interest in finding determinant formulas for resultants (see, for example, [DD], [SZ], and [WZ]).

Additional Exercises for §5

Exercise 6. In the text, we claimed that generically, the solutions of n affine equations $f_1 = \cdots = f_n = 0$ have multiplicity one. This exercise will prove this result. Assume as usual that the f_i come from homogeneous polynomials F_i of degree d_i by setting $x_0 = 1$. We will also use the following fact from multiplicity theory: if $F_1 = \cdots = F_n = 0$ has finitely many solutions and p is a solution such that the gradient vectors

$$\nabla F_i(p) = \Big(\frac{\partial F_i}{\partial x_0}(p), \ldots, \frac{\partial F_i}{\partial x_n}(p)\Big), \quad 1 \le i \le n$$

are linearly independent, then p is a solution of multiplicity one.

a. Consider the affine space $\mathbb{C}^M$ consisting of all possible coefficients of the F_i. As in the discussion at the end of §2, the coordinates of $\mathbb{C}^M$ are $c_{i,\alpha}$, where for fixed i, the $c_{i,\alpha}$ are the coefficients of F_i. Now consider the set $W \subset \mathbb{C}^M \times \mathbb{P}^n \times \mathbb{P}^{n-1}$ defined by

$$\begin{aligned} W = \{(c_{i,\alpha}, p, a_1, \ldots, a_n) \in \mathbb{C}^M \times \mathbb{P}^n \times \mathbb{P}^{n-1} : \ & p \text{ is a} \\ & \text{nontrivial solution of } F_0 = \cdots = F_n = 0 \text{ and} \\ & a_1\nabla F_1(p) + \cdots + a_n\nabla F_n(p) = 0\}. \end{aligned}$$

Under the projection map $\pi : \mathbb{C}^M \times \mathbb{P}^n \times \mathbb{P}^{n-1} \to \mathbb{C}^M$, explain why a generalization of the Projective Extension Theorem from §2 would imply that $\pi(W) \subset \mathbb{C}^M$ is a variety.

b. Show that $\pi(W) \subset \mathbb{C}^M$ is a proper variety, i.e., find $F_1, \ldots, F_n$ such that $(F_1, \ldots, F_n) \in \mathbb{C}^M \setminus \pi(W)$. Hint: Let $F_i = \Pi_{j=1}^{d_i}(x_i - jx_0)$ for $1 \leq i \leq n$.

c. By parts a and b, we can find a nonzero polynomial G in the coefficients of the F_i such that G vanishes on $\pi(W)$. Then consider $G \cdot \mathrm{Res}(\overline{F}_1, \ldots, \overline{F}_n)$. We can regard this as a polynomial in the coefficients of the f_i. Prove that if this polynomial is nonvanishing at $f_1, \ldots, f_n$, then the equations $f_0 = \cdots = f_n = 0$ have $d_1 \cdots d_n$ many solutions in $\mathbb{C}^n$, all of which have multiplicity one. Hint: Use Theorem (5.5).

Exercise 7. As we saw in (5.12), we can find the x_n-coordinates of the solutions using $\mathrm{Res}(u - x_n, f_1, \ldots, f_n)$, and in general, the x_i-coordinates can be found by replacing $u - x_n$ by $u - x_i$ in the resultant. In this exercise, we will describe the method given in [CM] for matching up coordinates to get the solutions. We begin by assuming that we've found the x_1- and x_2-coordinates of the solutions. To match up these two coordinates, let α and β be randomly chosen numbers, and consider the resultant

$$R_{1,2}(u) = \mathrm{Res}_{1,d_1,\ldots,d_n}(u - (\alpha x_1 + \beta x_2), f_1, \ldots, f_n).$$

a. Use (5.9) to show that

$$R_{1,2}(u) = C' \prod_{i=1}^{d_1 \cdots d_n} \big(u - (\alpha a_{i1} + \beta a_{i2})\big),$$

where C' is a nonzero constant and, as in (5.9), the solutions are $p_i = (a_{i1}, \ldots, a_{in})$.

b. A random choice of α and β will ensure that for solutions p_i, p_j, p_k, we have $\alpha a_{i1} + \beta a_{j2} \neq \alpha a_{k1} + \beta a_{k2}$ except when $p_i = p_j = p_k$. Conclude that the only way the condition

$$\alpha \cdot (\text{an } x_1\text{-coordinate}) + \beta \cdot (\text{an } x_2\text{-coordinate}) = \text{root of } R_{1,2}(u)$$

can hold is when the x_1-coordinate and x_2-coordinate come from the same solution.

c. Explain how we can now find the first two coordinates of the solutions.

d. Explain how a random choice of α, β, γ will enable us to construct a polynomial $R_{1,2,3}(u)$ which will tell us how to match up the x_3-coordinates with the two coordinates already found.

e. In the affine equations $f_1 = f_2 = 0$ coming from (5.13), compute $\mathrm{Res}(u - x_1, f_1, f_2)$, $\mathrm{Res}(u - x_2, f_1, f_2)$ and (in the notation of part a) $R_{1,2}(u)$, using $\alpha = 1$ and $\beta = 2$. Find the roots of these polynomials numerically and explain how this gives the solutions of our equations. Hint: Try the Maple command `fsolve`. In general, `fsolve` should be used with the `complex` option, though in this case it's not necessary since the roots are real.

Exercise 8. This exercise is concerned with Proposition (5.15).

a. Explain what Theorem (3.4) looks like if we use (5.2) instead of (3.3), and apply this to (5.16), (5.17) and Proposition (5.8).
b. Show carefully that the the ring homomorphism

$$\mathbb{C}(u)[x_1, \ldots, x_n] \longrightarrow \mathbb{C}(u)[x_1, \ldots, x_{n-1}]$$

defined by $x_n \mapsto u$ carries f_i to $\hat{f}_i$ and induces a ring isomorphism $A \cong \widehat{A}$.
c. Show that the diagram (5.19) is commutative and use it to prove that $\det(m_{f_n}) = \det(m_{\hat{f}_n})$.

Exercise 9. In this exercise, you will develop a homogeneous version of the hidden variable method. Suppose that we have homogeneous polynomials $F_1, \ldots, F_n$ in $x_0, \ldots, x_n$ such that

$$f_i(x_1, \ldots, x_n) = F_i(1, x_1, \ldots, x_n).$$

We assume that F_i has degree d_i, so that f_i has degree at most d_i. Also define

$$\hat{f}_i(x_1, \ldots, x_{n-1}) = f_i(x_1, \ldots, x_{n-1}, u).$$

As we saw in the proof of Proposition (5.15), the hidden variable resultant can be regarded as the affine resultant $\mathrm{Res}_{d_1,\ldots,d_n}(\hat{f}_1, \ldots, \hat{f}_n)$. To get a homogeneous resultant, we homogenize $\hat{f}_i$ with respect to x_0 to get a homogeneous polynomial $\widehat{F}_i(x_0, \ldots, x_{n-1})$ of degree d_i. Then

$$\mathrm{Res}_{d_1,\ldots,d_n}(\hat{f}_1, \ldots, \hat{f}_n) = \mathrm{Res}_{d_1,\ldots,d_n}(\widehat{F}_1, \ldots, \widehat{F}_n).$$

a. Prove that

$$\widehat{F}_i(x_0, \ldots, x_{n-1}) = F_i(x_0, x_1, \ldots, x_0 u).$$

Hint: This is done in the proof of Proposition (5.15).
b. Explain how part a leads to a purely homogeneous construction of the hidden variable resultant. This resultant is a polynomial in u.
c. State a purely homogeneous version of Proposition (5.15) and explain how it follows from the affine version stated in the text. Also explain why the roots of the hidden variable resultant are a_n/a_0 as $p = (a_0, \ldots, a_n)$ varies over all homogeneous solutions of $F_1 = \cdots = F_n = 0$ in $\mathbb{P}^n$.

Exercise 10. In (5.20), we expressed the hidden variable resultant as a quotient of two determinants $\pm\widehat{D}_0/\widehat{D}_0'$. If we think of this resultant as a polynomial in u, then use Proposition (4.6) to prove that the denominator $\widehat{D}_0'$ does *not* involve u. This will imply that the numerator $\widehat{D}_0$ can be regarded as the hidden variable resultant. Hint: By the previous exercise, we can write the hidden variable resultant as $\mathrm{Res}(\widehat{F}_1, \ldots, \widehat{F}_n)$. Also note that Proposition (4.6) assumed that x_n is last, while here $\widehat{D}_0$ and $\widehat{D}'_0$ mean that x_0 is taken last. Thus, applying Proposition (4.6) to the $\widehat{F}_i$ means setting $x_0 = 0$ in $\widehat{F}_i$. Then use part a of Exercise 9 to explain why u disappears from the scene.

Exercise 11. Suppose that $f_1, \ldots, f_n$ are polynomials of total degrees $d_1, \ldots, d_n$ in $k[x_1, \ldots, x_n]$.

a. Use Theorem (2.10) of Chapter 2 to prove that the ideal $\langle f_1, \ldots, f_n\rangle$ is radical for $f_1, \ldots, f_n$ generic. Hint: Use the notion of generic discussed in Exercise 6.
b. Explain why Exercise 16 of Chapter 2, §4, describes a *lex* Gröbner basis (assuming x_n is the last variable) for the ideal $\langle f_1, \ldots, f_n\rangle$ when the f_i are generic.

§6 Solving Equations via Eigenvalues and Eigenvectors

In Chapter 2, we learned that solving the equations $f_1 = \cdots = f_n = 0$ can be reduced to an eigenvalue problem. We did this as follows. The monomials not divisible by the leading terms of a Gröbner basis G for $\langle f_1, \ldots, f_n\rangle$ give a basis for the quotient ring

$$A = \mathbb{C}[x_1, \ldots, x_n]/\langle f_1, \ldots, f_n\rangle. \tag{6.1}$$

(see §2 of Chapter 2). Using this basis, we find the matrix of a multiplication map m_{f_0} by taking a basis element x^α and computing the remainder of $x^\alpha f_0$ on division by G (see §4 of Chapter 2). Once we have this matrix, its eigenvalues are the values $f_0(p)$ for $p \in \mathbf{V}(f_1, \ldots, f_n)$ by Theorem (4.5) of Chapter 2. In particular, the eigenvalues of the matrix for m_{x_i} are the x_i-coordinates of the solution points.

The amazing fact is that we can do all of this using resultants! We first show how to find a basis for the quotient ring.

(6.2) Theorem. *If $f_1, \ldots, f_n$ are generic polynomials of total degree $d_1, \ldots, d_n$, then the cosets of the monomials*

$$x_1^{a_1} \cdots x_n^{a_n}, \ \textit{where } 0 \le a_i \le d_i - 1 \textit{ for } i = 1, \ldots, n$$

form a basis of the ring A of (6.1).

Proof. Note that these monomials are *precisely* the monomials obtained from S_0 in (5.11) by setting $x_0 = 1$. As we will see, this is no accident.

By $f_1, \ldots, f_n$ generic, we mean that there are no solutions at ∞, that all solutions have multiplicity one, and that the matrix M_{11} which appears below is invertible.

Our proof will follow [ER] (see [PS1] for a different proof). There are $d_1 \cdots d_n$ monomials $x_1^{a_1} \cdots x_n^{a_n}$ with $0 \le a_i \le d_i - 1$. Since this is the dimension of A in the generic case by Theorem (5.5), it suffices to show that the cosets of these polynomials are linearly independent.

To prove this, we will use resultants. However, we have the wrong number of polynomials: since $f_1, \ldots, f_n$ are not homogeneous, we need $n+1$ polynomials in order to form a resultant. Hence we will add the polynomial $f_0 = u_0 + u_1 x_1 + \cdots + u_n x_n$, where $u_0, \ldots, u_n$ are independent variables. This gives the resultant $\mathrm{Res}_{1,d_1,\ldots,d_n}(f_0, \ldots, f_n)$, which we recognize as the u-resultant. By (5.10), this resultant is the quotient D_0/D_0', where $D_0 = \det(M_0)$ and M_0 is the matrix coming from the equations (4.1).

We first need to review in detail how the matrix M_0 is constructed. Although we did this in (4.1), our present situation is different in two ways: first, (4.1) ordered the variables so that x_n was last, while here, we want x_0 to be last, and second, (4.1) dealt with homogeneous polynomials, while here we have dehomogenized by setting $x_0 = 1$. Let's see what changes this makes.

As before, we begin in the homogeneous situation and consider monomials $x^\gamma = x_0^{a_0} \cdots x_n^{a_n}$ of total degree $d = 1 + d_1 + \cdots + d_n - n$ (remember that the resultant is $\mathrm{Res}_{1,d_1,\ldots,d_n}$). Since we want to think of x_0 as last, we divide these monomials into $n+1$ disjoint sets as follows:

$$
\begin{aligned}
S_n &= \{x^\gamma : |\gamma| = d,\ x_n^{d_n} \text{ divides } x^\gamma\}\\
S_{n-1} &= \{x^\gamma : |\gamma| = d,\ x_n^{d_n} \text{ doesn't divide } x^\gamma \text{ but } x_{n-1}^{d_{n-1}} \text{ does}\}\\
&\ \ \vdots\\
S_0 &= \{x^\gamma : |\gamma| = d,\ x_n^{d_n}, \ldots, x_1^{d_1} \text{ don't divide } x^\gamma \text{ but } x_0 \text{ does}\}
\end{aligned}
$$

(remember that $d_0 = 1$ in this case). You should check that S_0 is precisely as described in (5.11). The next step is to dehomogenize the elements of S_i by setting $x_0 = 1$. If we denote the resulting set of monomials as S_i', then $S_0' \cup S_1' \cup \cdots \cup S_n'$ consists of *all* monomials of total degree $\le d$ in $x_1, \ldots, x_n$. Furthermore, we see that S_0' consists of the $d_1 \cdots d_n$ monomials in the statement of the theorem.

Because of our emphasis on S_0', we will use x^α to denote elements of S_0' and x^β to denote elements of $S_1' \cup \cdots \cup S_n'$. Then observe that

$$
\begin{aligned}
&\text{if } x^\alpha \in S_0', \text{ then } x^\alpha \text{ has degree} \le d-1,\\
&\text{if } x^\beta \in S_i',\ i > 0, \text{ then } x^\alpha/x_i^{d_i} \text{ has degree} \le d - d_i.
\end{aligned}
$$

Then consider the equations:

$$\begin{aligned}
x^\alpha\, f_0 &= 0 \quad \text{for all } x^\alpha \in S'_0 \\
(x^\beta/x_1^{d_1})\, f_1 &= 0 \quad \text{for all } x^\beta \in S'_1 \\
&\ \ \vdots \\
(x^\beta/x_n^{d_n})\, f_n &= 0 \quad \text{for all } x^\beta \in S'_n.
\end{aligned}$$

Since the $x^\alpha f_0$ and $x^\beta/x_i^{d_i} f_i$ have total degree $\leq d$, we can write these polynomials as linear combinations of the x^α and x^β. We will order these monomials so that the elements $x^\alpha \in S'_0$ come first, followed by the elements $x^\beta \in S'_1 \cup \cdots \cup S'_n$. This gives a square matrix M_0 such that

$$M_0 \begin{pmatrix} x^{\alpha_1} \\ x^{\alpha_2} \\ \vdots \\ x^{\beta_1} \\ x^{\beta_2} \\ \vdots \end{pmatrix} = \begin{pmatrix} x^{\alpha_1}\, f_0 \\ x^{\alpha_2}\, f_0 \\ \vdots \\ x^{\beta_1}/x_1^{d_1}\, f_1 \\ x^{\beta_2}/x_1^{d_1}\, f_1 \\ \vdots \end{pmatrix},$$

where, in the column on the left, the first two elements of S'_0 and the first two elements of S'_1 are listed explicitly. This should make it clear what the whole column looks like. The situation is similar for the column on the right.

For $p \in \mathbf{V}(f_1, \ldots, f_n)$, we have $f_1(p) = \cdots = f_n(p) = 0$. Thus, evaluating the above equation at p yields

$$M_0 \begin{pmatrix} p^{\alpha_1} \\ p^{\alpha_2} \\ \vdots \\ p^{\beta_1} \\ p^{\beta_2} \\ \vdots \end{pmatrix} = \begin{pmatrix} p^{\alpha_1}\, f_0(p) \\ p^{\alpha_2}\, f_0(p) \\ \vdots \\ 0 \\ 0 \\ \vdots \end{pmatrix}.$$

To simplify notation, we let $\mathbf{p}^\alpha$ be the column vector $(p^{\alpha_1}, p^{\alpha_2}, \ldots)^T$ given by evaluating all monomials in S'_0 at p (and T means transpose). Similarly, we let $\mathbf{p}^\beta$ be the column vector $(p^{\beta_1}, p^{\beta_2}, \ldots)^T$ given by evaluating all monomials in $S'_1 \cup \cdots \cup S'_n$ at p. With this notation, we can rewrite the above equation more compactly as

$$M_0 \begin{pmatrix} \mathbf{p}^\alpha \\ \mathbf{p}^\beta \end{pmatrix} = \begin{pmatrix} f_0(p)\, \mathbf{p}^\alpha \\ \mathbf{0} \end{pmatrix}. \tag{6.3}$$

The next step is to partition M_0 so that the rows and columns of M_0 corresponding to elements of S'_0 lie in the upper left hand corner. This means writing M_0 in the form

$$M_0 = \begin{pmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{pmatrix},$$

where M_{00} is a $\mu \times \mu$ matrix for $\mu = d_1 \cdots d_n$, and M_{11} is also a square matrix. With this notation, (6.3) can be written

$$\begin{pmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{pmatrix} \begin{pmatrix} \mathbf{p}^\alpha \\ \mathbf{p}^\beta \end{pmatrix} = \begin{pmatrix} f_0(p)\,\mathbf{p}^\alpha \\ \mathbf{0} \end{pmatrix}. \tag{6.4}$$

By Lemma 4.4 of [Emi1], M_{11} is invertible for most choices of $f_1, \ldots, f_n$. Note that this condition is generic since it is given by $\det(M_{11}) \neq 0$ and $\det(M_{11})$ is a polynomial in the coefficients of the f_i. Hence, for generic $f_1, \ldots, f_n$, we can define the $\mu \times \mu$ matrix

$$\widetilde{M} = M_{00} - M_{01}M_{11}^{-1}M_{10}. \tag{6.5}$$

Note that the entries of $\widetilde{M}$ are polynomials in $u_0, \ldots, u_n$ since these variables only appear in M_{00} and M_{01}. If we multiply each side of (6.4) on the left by the matrix

$$\begin{pmatrix} I & -M_{01}M_{11}^{-1} \\ 0 & I \end{pmatrix},$$

then an easy computation gives

$$\begin{pmatrix} \widetilde{M} & 0 \\ M_{10} & M_{11} \end{pmatrix} \begin{pmatrix} \mathbf{p}^\alpha \\ \mathbf{p}^\beta \end{pmatrix} = \begin{pmatrix} f_0(p)\,\mathbf{p}^\alpha \\ \mathbf{0} \end{pmatrix}.$$

This implies

$$\widetilde{M}\,\mathbf{p}^\alpha = f_0(p)\,\mathbf{p}^\alpha, \tag{6.6}$$

so that for each $p \in \mathbf{V}(f_1, \ldots, f_n)$, $f_0(p)$ is an eigenvalue of $\widetilde{M}$ with $\mathbf{p}^\alpha$ as the corresponding eigenvector. Since $f_0 = u_0 + u_1x_1 + \cdots + u_nx_n$, the eigenvalues $f_0(p)$ are distinct for $p \in \mathbf{V}(f_1, \ldots, f_n)$. Standard linear algebra implies that the corresponding eigenvectors $\mathbf{p}^\alpha$ are linearly independent.

We can now prove the theorem. Write the elements of S_0' as $x^{\alpha_1}, \ldots, x^{\alpha_\mu}$, where as usual $\mu = d_1 \cdots d_n$, and recall that we need only show that the cosets $[x^{\alpha_1}], \ldots, [x^{\alpha_\mu}]$ are linearly independent in the quotient ring A. So suppose we have a linear relation among these cosets, say

$$c_1[x^{\alpha_1}] + \cdots + c_\mu[x^{\alpha_\mu}] = 0.$$

Evaluating this equation at $p \in \mathbf{V}(f_1, \ldots, f_n)$ makes sense by Exercise 12 of Chapter 2, §4 and implies that $c_1p^{\alpha_1} + \cdots + c_\mu p^{\alpha_\mu} = 0$. In the generic case, $\mathbf{V}(f_1, \ldots, f_n)$ has $\mu = d_1 \cdots d_n$ points $p_1, \ldots, p_\mu$, which gives μ equations

$$\begin{aligned} c_1p_1^{\alpha_1} + \cdots + c_\mu p_1^{\alpha_\mu} &= 0 \\ &\vdots \\ c_1p_\mu^{\alpha_1} + \cdots + c_\mu p_\mu^{\alpha_\mu} &= 0. \end{aligned}$$

In the matrix of these equations, the ith row is $(p_i^{\alpha_1}, \ldots, p_i^{\alpha_\mu})$, which in the notation used above, is the transpose of the column vector $\mathbf{p}_i^\alpha$ obtained by evaluating the monomials in S_0' at p_i. The discussion following (6.6) showed that the vectors $\mathbf{p}_i^\alpha$ are linearly independent. Thus the rows are linearly independent, so $c_1 = \cdots = c_\mu = 0$. We conclude that the cosets $[x^{\alpha_1}], \ldots, [x^{\alpha_\mu}]$ are linearly independent. □

Now that we know a basis for the quotient ring A, our next task it to find the matrix of the multiplication map m_{f_0} relative to this basis. Fortunately, this is easy since we already know the matrix!

(6.7) Theorem. *Let $f_1, \ldots, f_n$ be generic polynomials of total degrees $d_1, \ldots, d_n$, and let $f_0 = u_0 + u_1 x_1 + \cdots + u_n x_n$. Using the basis of $A = \mathbb{C}[x_1, \ldots, x_n]/\langle f_1, \ldots, f_n\rangle$ from Theorem (6.2), the matrix of the multiplication map m_{f_0} is the* **transpose** *of the matrix*

$$\widetilde{M} = M_{00} - M_{01} M_{11}^{-1} M_{10}$$

from (6.5).

Proof. Let $M_{f_0} = (m_{ij})$ be the matrix of m_{f_0} relative to the basis $[x^{\alpha_1}], \ldots, [x^{\alpha_\mu}]$ of A from Theorem (6.2), where $\mu = d_1 \cdots d_n$. The proof of Proposition (4.7) of Chapter 2 shows that for $p \in \mathbf{V}(f_1, \ldots, f_n)$, we have

$$f_0(p)(p^{\alpha_1}, \ldots, p^{\alpha_\mu}) = (p^{\alpha_1}, \ldots, p^{\alpha_\mu})\, M_{f_0}.$$

Letting $\mathbf{p}^\alpha$ denote the column vector $(p^{\alpha_1}, \ldots, p^{\alpha_\mu})^T$ as in the previous proof, we can take the transpose of each side of this equation to obtain

$$\begin{aligned} f_0(p)\, \mathbf{p}^\alpha &= \big(f_0(p)(p^{\alpha_1}, \ldots, p^{\alpha_\mu})\big)^T \\ &= \big((p^{\alpha_1}, \ldots, p^{\alpha_\mu})\, M_{f_0}\big)^T \\ &= (M_{f_0})^T\, \mathbf{p}^\alpha, \end{aligned}$$

where $(M_{f_0})^T$ is the transpose of M_{f_0}. Comparing this to (6.6), we get

$$(M_{f_0})^T\, \mathbf{p}^\alpha = \widetilde{M}\, \mathbf{p}^\alpha$$

for all $p \in \mathbf{V}(f_1, \ldots, f_n)$. Since $f_1, \ldots, f_n$ are generic, we have μ points $p \in \mathbf{V}(f_1, \ldots, f_n)$, and the proof of Theorem (6.2) shows that the corresponding eigenvectors $\mathbf{p}^\alpha$ are linearly independent. This implies $(M_{f_0})^T = \widetilde{M}$, and then $M_{f_0} = \widetilde{M}^T$ follows easily. □

Since $f_0 = u_0 + u_1 x_1 + \cdots + u_n x_n$, Corollary (4.3) of Chapter 2 implies

$$M_{f_0} = u_0\, I + u_1\, M_{x_1} + \cdots + u_n\, M_{x_n},$$

where M_{x_i} is the matrix of m_{x_i} relative to the basis of Theorem (6.2). By Theorem (6.7), it follows that if we write

$$\widetilde{M} = u_0\, I + u_1\, \widetilde{M}_1 + \cdots + u_n\, \widetilde{M}_n, \tag{6.8}$$

where each $\widetilde{M}_i$ has constant entries, then $M_{f_0} = \widetilde{M}^T$ implies that $M_{x_i} = (\widetilde{M}_i)^T$ for all i. Thus $\widetilde{M}$ *simultaneously* computes the matrices of the n multiplication maps $m_{x_1}, \ldots, m_{x_n}$.

Exercise 1. For the equations

$$\begin{aligned} f_1 &= x_1^2 + x_2^2 - 10 = 0 \\ f_2 &= x_1^2 + x_1x_2 + 2x_2^2 - 16 = 0 \end{aligned}$$

(this is the affine version of Exercise 2 of §5), show that $\widetilde{M}$ is the matrix

$$\widetilde{M} = \begin{pmatrix} u_0 & u_1 & u_2 & 0 \\ 4u_1 & u_0 & 0 & u_1 + u_2 \\ 6u_2 & 0 & u_0 & u_1 - u_2 \\ 0 & 3u_1 + 3u_2 & 2u_1 - 2u_2 & u_0 \end{pmatrix}.$$

Use this to determine the matrices M_{x_1} and M_{x_2}. What is the basis of $\mathbb{C}[x_1, x_2]/\langle f_1, f_2\rangle$ in this case? Hint: The matrix M_0 of Exercise 2 of §5 is already partitioned into the appropriate submatrices.

Now that we have the matrices M_{x_i}, we can find the x_i-coordinates of the solutions of (5.3) using the eigenvalue methods mentioned in Chapter 2 (see especially the discussion following Corollary (4.6)). This still leaves the problem of finding how the coordinates match up. We will follow Chapter 2 and show how the left eigenvectors of M_{f_0}, or equivalently, the right eigenvectors of $\widetilde{M} = (M_{f_0})^T$, give the solutions of our equations.

Since $\widetilde{M}$ involves the variables $u_0, \ldots, u_n$, we need to specialize them before we can use numerical methods for finding eigenvectors. Let

$$f_0' = c_0 + c_1x_1 + \cdots + c_nx_n,$$

where $c_0, \ldots, c_n$ are constants chosen so that the values $f_0'(p)$ are distinct for $p \in \mathbf{V}(f_1, \ldots, f_n)$. In practice, this can be achieved by making a random choice of $c_0, \ldots, c_n$. If we let $\widetilde{M}'$ be the matrix obtained from $\widetilde{M}$ by letting $u_i = c_i$, then (6.6) shows that $\mathbf{p}^\alpha$ is a right eigenvector for $\widetilde{M}'$ with eigenvalue $f_0'(p)$. Since we have $\mu = d_1 \cdots d_n$ distinct eigenvalues in a vector space of the same dimension, the corresponding eigenspaces all have dimension 1.

To find the solutions, suppose that we've used a standard numerical method to find an eigenvector v of $\widetilde{M}'$. Since the eigenspaces all have dimension 1, it follows that $v = c\,\mathbf{p}^\alpha$ for some solution $p \in \mathbf{V}(f_1, \ldots, f_n)$ and nonzero constant c. This means that whenever x^α is a monomial in S_0', the corresponding coordinate of v is cp^α. The following exercise shows how to reconstruct p from the coordinates of the eigenvector v.

Exercise 2. As above, let $p = (a_1, \ldots, a_n) \in \mathbf{V}(f_1, \ldots, f_n)$ and let v be an eigenvector of $\widetilde{M}'$ with eigenvalue $f_0'(p)$. This exercise will explain how to recover p from v when $d_1, \ldots, d_n$ are all > 1, and Exercise 5 at the end of the section will explore what happens when some of the degrees equal 1.

a. Show that $1, x_1, \ldots, x_n \in S_0'$, and conclude that for some $c \neq 0$, the numbers $c, ca_1, \ldots, ca_n$ are among the coordinates of v.

b. Prove that a_j can be computed from the coordinates of v by the formula

$$a_j = \frac{ca_j}{c} \qquad \text{for } j = 1, \ldots, n.$$

This shows that the solution p can be easily found using ratios of certain coordinates of the eigenvector v.

Exercise 3. For the equations $f_1 = f_2 = 0$ of Exercise 1, consider the matrix $\widetilde{M}'$ coming from $(u_0, u_1, u_2, u_3) = (0, 1, 0, 0)$. In the notation of (6.8), this means $\widetilde{M}' = \widetilde{M}_1 = (M_{x_1})^T$. Compute the eigenvectors of this matrix and use Exercise 2 to determine the solutions of $f_1 = f_2 = 0$.

While the right eigenvectors of $\widetilde{M}$ relate to the solutions of $f_1 = \cdots = f_n = 0$, the left eigenvectors give a nice answer to the *interpolation problem*. This was worked out in detail in Exercise 17 of Chapter 2, §4, which applies without change to the case at hand. See Exercise 6 at the end of this section for an example.

Eigenvalue methods can also be applied to the hidden variable resultants discussed earlier in this section. We will discuss this very briefly. In Proposition (5.15), we showed that the x_n-coordinates of the solutions of the equations $f_1 = \cdots = f_n = 0$ could be found using the resultant $\mathrm{Res}^{x_n}_{d_1,\ldots,d_n}(f_1, \ldots, f_n)$ obtained by regarding x_n as a constant. As we learned in (5.20),

$$\mathrm{Res}^{x_n}_{d_1,\ldots,d_n}(f_1, \ldots, f_n) = \pm \frac{\widehat{D}_0}{\widehat{D}_0'},$$

and if $\widehat{M}_0$ is the corresponding matrix (so that $\widehat{D}_0 = \det(\widehat{M}_0)$), one could ask about the eigenvalues and eigenvectors of $\widehat{M}_0$. It turns out that this is not quite the right question to ask. Rather, since $\widehat{M}_0$ depends on the variable x_n, we write the matrix as

$$\widehat{M}_0 = A_0 + x_n A_1 + \cdots + x_n^l A_l, \tag{6.9}$$

where each A_i has constant entries and $A_l \neq 0$. Suppose that $\widehat{M}_0$ and the A_i are $m \times m$ matrices. If A_l is invertible, then we can define the *generalized companion matrix*

$$C = \begin{pmatrix} 0 & I_m & 0 & \cdots & 0 \\ 0 & 0 & I_m & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_m \\ -A_l^{-1}A_0 & -A_l^{-1}A_1 & -A_l^{-1}A_2 & \cdots & -A_l^{-1}A_{l-1} \end{pmatrix},$$

where I_m is the $m \times m$ identity matrix. Then the correct question to pose concerns the eigenvalues and eigenvectors of C. One can show that the eigenvalues of the generalized companion matrix are precisely the roots of the polynomial $\widehat{D}_0 = \det(\widehat{M}_0)$, and the corresponding eigenvectors have a nice interpretation as well. Further details of this technique can be found in [Man2] and [Man3].

Finally, we should say a few words about how eigenvalue and eigenvector methods behave in the non-generic case. As in the discussion of u-resultants in §5, there are many things which can go wrong. All of the problems listed earlier are still present when dealing with eigenvalues and eigenvectors, and there are two new difficulties which can occur:

- In working with the matrix M_0 as in the proof of Theorem (6.2), it can happen that M_{11} is not invertible, so that $\widetilde{M} = M_{00} - M_{01}M_{11}^{-1}M_{10}$ doesn't make sense.
- In working with the matrix $\widehat{M}_0$ as in (6.9), it can happen that the leading term A_l is not invertible, so that the generalized companion matrix C doesn't make sense.

Techniques for avoiding both of these problems are described in [Emi2], [Man1], [Man2], and [Man3].

Exercise 4. Express the 6×6 matrix of part c of Exercise 5 of §5 in the form $A_0 + x_3A_1 + x_3^2A_2$ and show that A_2 is *not* invertible.

The idea of solving equations by a combination of eigenvalue/eigenvector methods and resultants goes back to the work of Auzinger and Stetter [AS]. This has now become an active area of research, not only for the resultants discussed here (see [BMP], [Man3], [Mou1] and [Ste], for example) but also for the *sparse resultants* to be introduced in Chapter 7. Also, we will say more about multiplication maps in §2 of Chapter 4.

Additional Exercises for §6

Exercise 5. This exercise will explain how to recover the solution $p = (a_1, \ldots, a_n)$ from an eigenvector v of the matrix $\widetilde{M}'$ in the case when some of the degrees $d_1, \ldots, d_n$ are equal to 1. Suppose for instance that $d_i = 1$. This means that $x_i \notin S_0'$, so that the ith coordinate a_i of the solution p doesn't appear in the eigenvector $\mathbf{p}^\alpha$. The idea is that the matrix M_{x_i} (which we know by Theorem (6.7)) has all of the information we need. Let $c_1, \ldots, c_\mu$ be the entries of the column of M_{x_i} corresponding to $1 \in S_0'$.

a. Prove that $[x_i] = c_1[x^{\alpha_1}] + \cdots + c_\mu[x^{\alpha_\mu}]$ in A, where $S_0' = \{x^{\alpha_1}, \ldots, x^{\alpha_\mu}\}$.

b. Prove that $a_i = c_1\mathbf{p}^{\alpha_1} + \cdots + c_\mu\mathbf{p}^{\alpha_\mu}$.

It follows that if we have an eigenvector v as in the discussion preceding Exercise 2, it is now straightforward to recover *all* coordinates of the solution p.

Exercise 6. The equations $f_1 = f_2 = 0$ from Exercise 1 have solutions p_1, p_2, p_3, p_4 (they are listed in projective form in Exercise 2 of §5). Apply Exercise 17 of Chapter 2, §4, to find the polynomials g_1, g_2, g_3, g_4 such that $g_i(p_j) = 1$ if $i = j$ and 0 otherwise. Then use this to write down explicitly a polynomial h which takes preassigned values $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ at the points p_1, p_2, p_3, p_4. Hint: Since the x_1-coordinates are distinct, it suffices to find the eigenvectors of M_{x_1}. Exercise 1 will be useful.