1 Polynomials in One Variable

The study of systems of polynomial equations in many variables requires a good understanding of what can be said about one polynomial equation in one variable. The purpose of this lecture is to provide some basic tools on this matter. We shall consider the problem of how to compute and how to represent the zeros of a general polynomial of degree d in one variable x:

$$p(x) \quad = \quad a_d x^d + a_{d-1} x^{d-1} + \cdots + a_2 x^2 + a_1 x + a_0. \tag{1}$$

1.1 The Fundamental Theorem of Algebra

We begin by assuming that the coefficients a_i lie in the field $\mathbb{Q}$ of rational numbers, with $a_d \neq 0$, where the variable x ranges over the field $\mathbb{C}$ of complex numbers. Our starting point is the fact that $\mathbb{C}$ is algebraically closed.

Theorem 1. (Fundamental Theorem of Algebra) *The polynomial $p(x)$ has d roots, counting multiplicities, in the field $\mathbb{C}$ of complex numbers.*

If the degree d is four or less, then the roots are functions of the coefficients which can be expressed in terms of radicals. The command `solve` in `maple` will produce these familiar expressions for us:

```
> solve(a2 * x^2 +  a1 * x + a0, x );
```

$$\frac{-a1 + (a1^2 - 4\,a2\,a0)^{1/2}}{a2}, \quad \frac{-a1 - (a1^2 - 4\,a2\,a0)^{1/2}}{a2}$$

```
> lprint( solve(a3 * x^3 + a2 * x^2 +  a1 * x + a0, x )[1] );

1/6/a3*(36*a1*a2*a3-108*a0*a3^2-8*a2^3+12*3^(1/2)*(4*a1^3*a3
-a1^2*a2^2-18*a1*a2*a3*a0+27*a0^2*a3^2+4*a0*a2^3)^(1/2)*a3)
^(1/3)+2/3*(-3*a1*a3+a2^2)/a3/(36*a1*a2*a3-108*a0*a3^2-8*a2^3
+12*3^(1/2)*(4*a1^3*a3-a1^2*a2^2-18*a1*a2*a3*a0+27*a0^2*a3^2
+4*a0*a2^3)^(1/2)*a3)^(1/3)-1/3*a2/a3
```

The polynomial $p(x)$ has d distinct roots if and only if its *discriminant* is nonzero. Can you spot the discriminant of the cubic equation in the previous `maple` output ? In general, the discriminant is computed from the resultant of $p(x)$ and its first derivative $p'(x)$ as follows:

$$\mathrm{discr}_x(p(x)) \quad = \quad \frac{1}{a_d} \cdot \mathrm{res}_x(p(x), p'(x)).$$

This is an irreducible polynomial in the coefficients $a_0, a_1, \ldots, a_d$. It follows from Sylvester's matrix for the resultant that the discriminant is a homogeneous polynomial of degree $2d-2$. Here is the discriminant of a quartic:

```
> f :=  a4 * x^4 +  a3 * x^3 +  a2 * x^2 +  a1 * x +  a0 :
> lprint(resultant(f,diff(f,x),x)/a4);

-192*a4^2*a0^2*a3*a1-6*a4*a0*a3^2*a1^2+144*a4*a0^2*a2*a3^2
+144*a4^2*a0*a2*a1^2+18*a4*a3*a1^3*a2+a2^2*a3^2*a1^2
-4*a2^3*a3^2*a0+256*a4^3*a0^3-27*a4^2*a1^4-128*a4^2*a0^2*a2^2
-4*a3^3*a1^3+16*a4*a2^4*a0-4*a4*a2^3*a1^2-27*a3^4*a0^2
-80*a4*a3*a1*a2^2*a0+18*a3^3*a1*a2*a0
```

This sextic is the determinant of the following 7×7-matrix divided by `a4`:

```
> with(linalg):
> sylvester(f,diff(f,x),x);

[ a4      a3      a2      a1      a0      0       0 ]
[                                                   ]
[ 0       a4      a3      a2      a1      a0      0 ]
[                                                   ]
[ 0       0       a4      a3      a2      a1      a0]
[                                                   ]
[4 a4    3 a3    2 a2     a1      0       0       0 ]
[                                                   ]
[ 0      4 a4    3 a3    2 a2     a1      0       0 ]
[                                                   ]
[ 0       0      4 a4    3 a3    2 a2     a1      0 ]
[                                                   ]
[ 0       0       0      4 a4    3 a3    2 a2    a1]
```

Galois theory tells us that there is no general formula which expresses the roots of $p(x)$ in radicals if $d \geq 5$. For specific instances with d not too big, say $d \leq 10$, it is possible to compute the Galois group of $p(x)$ over $\mathbb{Q}$. Occasionally, one is lucky and the Galois group is solvable, in which case `maple` has a chance of finding the solution of $p(x) = 0$ in terms of radicals.

```
> f := x^6 + 3*x^5 + 6*x^4 + 7*x^3 + 5*x^2 + 2*x+1:
> galois(f);

"6T11", {"[2^3]S(3)", "2 wr S(3)", "2S_4(6)"}, "-", 48,

    {"(2 4 6)(1 3 5)", "(1 5)(2 4)", "(3 6)"}

> solve(f,x)[1];
                       1/2 1/3
1/12 (-6 (108 + 12 69   )

                       1/2 2/3     1/2                1/2 1/3 1/2
+ 6 I (3 (108 + 12 69   )  + 8 69  + 8 (108 + 12 69   )   )  )

              /             1/2 1/3
 + 72 )      /  (108 + 12 69   )
            /
```

The number 48 is the order of the Galois group and its name is `"6T11"`. Of course, the user now has to consult `help(galois)` in order to learn more.

1.2 Numerical Root Finding

In symbolic computation, we frequently consider a polynomial problem as solved if it has been reduced to finding the roots of one polynomial in one variable. Naturally, the latter problem can still be a very interesting and challenging one from the perspective of numerical analysis, especially if d gets very large or if the a_i are given by floating point approximations. In the problems studied in this course, however, the a_i are usually exact rational numbers and the degree d rarely exceeds 200. For numerical solving in this range, `maple` does reasonably well and `matlab` has no difficulty whatsoever.

```
> Digits := 6:
> f := x^200 - x^157 + 8 * x^101 - 23 * x^61 + 1:
> fsolve(f,x);
                                  .950624, 1.01796
```

This polynomial has only two real roots. To list the complex roots, we say:

```
> fsolve(f,x,complex);

-1.02820-.0686972 I, -1.02820+.0686972 I, -1.01767-.0190398 I,
-1.01767+.0190398 I, -1.01745-.118366 I, -1.01745 + .118366 I,
-1.00698-.204423 I, -1.00698+.204423 I, -1.00028 - .160348 I,
-1.00028+.160348 I, -.996734-.252681 I, -.996734 + .252681 I,
-.970912-.299748 I, -.970912+.299748 I, -.964269 - .336097 I,
ETC...ETC..
```

Our polynomial $p(x)$ is represented in `matlab` as the row vector of its coefficients $[a_d\, a_{d-1}\ \ldots\ a_2\, a_1\, a_0]$. For instance, the following two commands compute the three roots of the dense cubic $p(x) = 31x^3 + 23x^2 + 19x + 11$.

```
>> p = [31 23 19 11];
>> roots(p)
ans =
  -0.0486 + 0.7402i
  -0.0486 - 0.7402i
  -0.6448
```

Representing the sparse polynomial $p(x) = x^{200} - x^{157} + 8x^{101} - 23x^{61} + 1$ considered above requires introducing lots of zero coefficients:

```
>> p=[1 zeros(1,42) -1 zeros(1,55) 8 zeros(1,39) -23 zeros(1,60) 1]
>> roots(p)
ans =
  -1.0282 + 0.0687i
  -1.0282 - 0.0687i
  -1.0177 + 0.0190i
  -1.0177 - 0.0190i
  -1.0174 + 0.1184i
  -1.0174 - 0.1184i
ETC...ETC..
```

We note that convenient facilities are available for calling `matlab` inside of `maple` and for calling `maple` inside of `matlab`. We wish to encourage our readers to experiment with the passage of data between these two programs.

Some numerical methods for solving a univariate polynomial equation $p(x) = 0$ work by reducing this problem to computing the eigenvalues of the companion matrix of $p(x)$, which is defined as follows. Consider the quotient ring $V = \mathbb{Q}[x]/\langle p(x) \rangle$ modulo the ideal generated by the polynomial $p(x)$. The ring V is a d-dimensional $\mathbb{Q}$-vector space. Multiplication by the variable x defines a linear map from this vector space to itself.

$$\text{Times}_x \ : \ V \ \rightarrow \ V \, , \, f(x) \ \mapsto x \cdot f(x). \tag{2}$$

The *companion matrix* is the $d \times d$-matrix which represents the endomorphism Times_x with respect to the distinguished monomial basis $\{1, x, x^2, \ldots, x^{d-1}\}$ of V. Explicitly, the companion matrix of $p(x)$ looks like this:

$$\text{Times}_x \quad = \quad \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0/a_d \\ 1 & 0 & \cdots & 0 & -a_1/a_d \\ 0 & 1 & \cdots & 0 & -a_2/a_d \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_{d-1}/a_d \end{pmatrix} \tag{3}$$

Proposition 2. *The zeros of $p(x)$ are the eigenvalues of the matrix Times_x.*

Proof. Suppose that $f(x)$ is a polynomial in $\mathbb{C}[x]$ whose image in $V \otimes \mathbb{C} = \mathbb{C}[x]/\langle p(x) \rangle$ is an eigenvector of (2) with eigenvalue λ. Then $x \cdot f(x) = \lambda \cdot f(x)$ in the quotient ring, which means that $(x - \lambda) \cdot f(x)$ is a multiple of $p(x)$. Since $f(x)$ is not a multiple of $p(x)$, we conclude that λ is a root of $p(x)$ as desired. Conversely, if μ any root of $p(x)$ then the polynomial $f(x) = p(x)/(x - \mu)$ represents an eigenvector of (2) with eigenvalue μ. $\square$

Corollary 3. *The following statements about $p(x) \in \mathbb{Q}[x]$ are equivalent:*

- *The polynomial $p(x)$ is square-free, i.e., it has no multiple roots in $\mathbb{C}$.*
- *The companion matrix Times_x is diagonalizable.*
- *The ideal $\langle p(x) \rangle$ is a radical ideal in $\mathbb{Q}[x]$.*

We note that the set of multiple roots of $p(x)$ can be computed symbolically by forming greatest common divisor of $p(x)$ and its derivative:

$$q(x) \quad = \quad \gcd(p(x), p'(x)) \tag{4}$$

Thus the three conditions in the Corollary are equivalent to $q(x) = 1$. In general, we compute the radical of any ideal in $\mathbb{Q}[x]$ as follows:

$$\mathrm{Rad}\big(\langle p(x)\rangle\big) \quad = \quad \langle p(x)/q(x)\rangle \tag{5}$$

1.3 Real Roots

In this subsection we describe symbolic methods for computing information about the real roots of a univariate polynomial $p(x)$. In what follows, we assume that $p(x)$ is a squarefree polynomial. It is easy to achieve this by removing all multiplicities as in (4) and (5). The *Sturm sequence* of $p(x)$ is the following sequence of polynomials of decreasing degree:

$$p_0(x) := p(x),\ p_1(x) := p'(x),\ p_i(x) := -\mathrm{rem}(p_{i-2}(x), p_{i-1}(x)) \text{ for } i \geq 2.$$

Thus $p_i(x)$ is the negative of the remainder on division of $p_{i-2}(x)$ by $p_{i-1}(x)$. Let $p_m(x)$ be the last non-zero polynomial in this sequence.

Theorem 4. (Sturm's Theorem) *If $a < b$ in $\mathbb{R}$ and neither is a zero of $p(x)$ then the number of real zeros of $p(x)$ in the interval $[a, b]$ is the number of sign changes in the sequence $p_0(a), p_1(a), p_2(a), \ldots, p_m(a)$ minus the number of sign changes in the sequence $p_0(b), p_1(b), p_2(b), \ldots, p_m(b)$.*

We note that any zeros are ignored when counting the number of sign changes in a sequence of real numbers. For instance, a sequence of twelve number with signs $+, +, 0, +, -, -, 0, +, -, 0, -, 0$ has three sign changes.

If we wish to count all real roots of a polynomial $p(x)$ then we can apply Sturm's Theorem to $a = -\infty$ and $b = \infty$, which amounts to looking at the signs of the leading coefficients of the polynomials p_i in the Sturm sequence. Using bisection, one gets an efficient method for isolating the real roots by rational intervals. This method is conveniently implemented in `maple`:

```
> p := x^11-20*x^10+99*x^9-247*x^8+210*x^7-99*x^2+247*x-210:
> sturm(p,x,-INFINITY, INFINITY);
                                    3
```

```
> sturm(p,x,0,10);
                                    2
> sturm(p,x,5,10);
                                    0

> realroot(p,1/1000);
                  1101  551    1465  733    14509  7255
                [[----, ---], [----, ---], [-----, ----]]
                  1024  512    1024  512     1024   512

> fsolve(p);
                  1.075787072, 1.431630905, 14.16961992
```

Another important classical result on real roots is the following:

Theorem 5. (Déscartes' Rule of Sign) *The number of positive real roots of a polynomial is at most the number of sign changes in its coefficient sequence.*

For instance, the polynomial $p(x) = x^{200} - x^{157} + 8x^{101} - 23x^{61} + 1$, which was featured in Section 1.2, has four sign changes in its coefficient sequence. Hence it has at most four positive real roots. The true number is two.

Corollary 6. *A polynomial with m terms can have at most $2m-1$ real zeros.*

The bound in this corollary is optimal as the following example shows:

$$x \cdot \prod_{j=1}^{m-1} (x^2 - j)$$

All $2m - 1$ zeros of this polynomial are real, and its expansion has m terms.

1.4 Puiseux series

Suppose now that the coefficients a_i of our given polynomial are not rational numbers by they are rational functions $a_i(t)$ in another parameter t. Hence we wish to determine the zeros of a polynomial in $K[x]$ where $K = \mathbb{Q}(t)$.

$$p(t;x) \quad = \quad a_d(t)x^d + a_{d-1}(t)x^{d-1} + \cdots + a_2(t)x^2 + a_1(t)x + a_0(t). \quad (6)$$

The role of the ambient algebraically closed field containing K is now played by the field $\mathbb{C}(\!(t)\!)$ of *Puiseux series*. These are formal power series in t with

coefficients in $\mathbb{C}$ and having rational exponents, subject to the condition that the set of appearing exponents is bounded below and has a common denominator. The field $\mathbb{C}(\!(t)\!)$ is known to be algebraically closed.

Theorem 7. (Puiseux's Theorem) *The polynomial $p(t;x)$ has d roots, counting multiplicities, in the field of Puiseux series $\mathbb{C}(\!(t)\!)$.*

The proof of Puiseux's theorem is algorithmic, and, lucky for us, there is an implementation of this algorithm in `maple`. Here is how it works:

```
> with(algcurves):  p :=  x^2 +  x - t^3;
                                      2        3
                                p := x  + x - t
> puiseux(p,t=0,x,20);
      18       15      12      9    6    3
{-42 t   + 14 t   - 5 t   + 2 t  - t  + t ,
            18       15      12      9    6    3
      + 42 t   - 14 t   + 5 t   - 2 t  + t  - t  - 1 }
```

We note that this program generally does not compute all Puiseux series solutions but only enough to generate the splitting field of $p(t;x)$ over K.

```
> with(algcurves):  q :=  x^2 + t^4 * x - t:
> puiseux(q,t=0,x,20);
                              29/2        15/2        4    1/2
                  {- 1/128 t     + 1/8 t     - 1/2 t  + t   }
> S := solve(q,x):
> series(S[1],t,20);
            1/2        4        15/2          29/2      43/2
           t    - 1/2 t  + 1/8 t     - 1/128 t     + O(t    )
> series(S[2],t,20);
            1/2        4        15/2          29/2      43/2
          -t    - 1/2 t  - 1/8 t     + 1/128 t     + O(t    )
```

We shall explain how to compute the first term (lowest order in t) in each of the d Puiseux series solutions $x(t)$ to our equation $p(t;x) = 0$. Suppose that the i-th coefficient in (6) has the Laurent series expansion:

$$a_i(t) \quad = \quad c_i \cdot t^{A_i} + \text{higher terms in } t.$$

Each Puiseux series looks like

$$x(t) \quad = \quad \gamma \cdot t^{\tau} + \text{higher terms in } t.$$

We wish to characterize the possible pairs of numbers (τ, γ) in $\mathbb{Q} \times \mathbb{C}$ which allow the identity $p(t; x(t)) = 0$ to hold. This is done by first finding the possible values of τ. We ignore all higher terms and consider an equation

$$c_d \cdot t^{A_d + d\tau} + c_{d-1} \cdot t^{A_{d-1} + (d-1)\tau} + \cdots + c_1 \cdot t^{A_1 + \tau} + c_0 \cdot t^{A_0} = 0. \qquad (7)$$

This equation imposes the following piecewise-linear condition on τ:

$$\min\{A_d + d\tau, A_{d-1} + (d-1)\tau, A_2 + 2\tau, A_1 + \tau, A_0\} \quad \text{is attained twice.} \quad (8)$$

The crucial condition (8) will reappear in various guises later in these lectures. As an illustration consider the example $p(t; x) = x^2 + x - t^3$, where (8) reads

$$\min\{\, 0 + 2\tau,\, 0 + \tau,\, 3 \,\} \quad \text{is attained twice.}$$

The sentence means the following disjunction of linear inequality systems:

$$2\tau = \tau \leq 3 \quad \text{or} \quad 2\tau = 3 \leq \tau \quad \text{or} \quad 3 = \tau \leq 2\tau.$$

This disjunction is equivalent to

$$\tau = 0 \quad \text{or} \quad \tau = 3,$$

which gives us the lowest terms in the two Puiseux series produced by `maple`.

It is customary to phrase the procedure described above in terms of the *Newton polygon* of $p(t; x)$. This polygon is the convex hull in $\mathbb{R}^2$ of the points (i, A_i) for $i = 0, 1, \ldots, d$. The condition (8) is equivalent to saying that $-\tau$ equals the slope of an edge on the lower boundary of the Newton polygon.

1.5 Hypergeometric series

The method of Puiseux series can be extended to the case when the coefficients a_i are rational functions in several variables $t_1, \ldots, t_m$. The case $m = 1$ was discussed in the last section. We now examine the generic case when all $d + 1$ coefficients $a_0, \ldots, a_d$ in (1) are indeterminates. Each zero X of the polynomial in (1) is an algebraic function of $d + 1$ variables, written $X = X(a_0, \ldots, x_d)$. The following theorem due to Karl Mayr (1937) characterizes these functions by the differential equations which they satisfy.

Theorem 8. *The roots of the general equation of degree d are a basis for the solution space of the following system of linear partial differential equations:*

$$\frac{\partial^2 X}{\partial a_i \partial a_j} \;=\; \frac{\partial^2 X}{\partial a_k \partial a_l} \qquad \textit{whenever} \quad i+j = k+l, \tag{9}$$

$$\sum_{i=0}^{d} i a_i \frac{\partial X}{\partial a_i} \;=\; -X \qquad \textit{and} \qquad \sum_{i=0}^{d} a_i \frac{\partial X}{\partial a_i} \;=\; 0. \tag{10}$$

The meaning of the statement "are a basis for the solution space of" will be explained at the end of this section. Let us first replace this statement by "are solutions of" and prove the resulting weaker version of the theorem.

Proof. The two Euler equations (10) express the scaling invariance of the roots. They are gotten by applying the operator d/dt to the identities

$$\begin{aligned} X(a_0, ta_1, t^2 a_2, \ldots, t^{d-1} a_{d-1}, t^d a_d) \quad &= \quad \tfrac{1}{t} \cdot X(a_0, a_1, a_2, \ldots, a_{d-1}, a_d), \\ X(ta_0, ta_1, ta_2, \ldots, ta_{d-1}, ta_d) \quad &= \quad X(a_0, a_1, a_2, \ldots, a_{d-1}, a_d). \end{aligned}$$

To derive (9), we consider the first derivative $f'(x) = \sum_{i=1}^{d} i a_i x^{i-1}$ and the second derivative $f''(x) = \sum_{i=2}^{d} i(i-1) a_i x^{i-2}$. Note that $f'(X) \neq 0$, since $a_0, \ldots, a_d$ are indeterminates. Differentiating the defining identity $\sum_{i=0}^{d} a_i X(a_0, a_1, \ldots, a_d)^i = 0$ with respect to a_j, we get

$$X^j \,+\, f'(X) \cdot \frac{\partial X}{\partial a_j} \quad = \quad 0. \tag{11}$$

We next differentiate $\partial X / \partial a_j$ with respect to the indeterminate a_i:

$$\frac{\partial^2 X}{\partial a_i \partial a_j} \;=\; \frac{\partial}{\partial a_i}\Big(-\frac{X^j}{f'(X)}\Big) \;=\; \frac{\partial f'(X)}{\partial a_i} X^j f'(X)^{-2} - j X^{j-1} \frac{\partial X}{\partial a_i} f'(X)^{-1}. \tag{12}$$

Using (11) and the resulting identity $\frac{\partial f'(X)}{\partial a_i} = -\frac{f''(X)}{f'(X)} \cdot X^i + i X^{i-1}$, we can rewrite (12) as follows:

$$\frac{\partial^2 X}{\partial a_i \partial a_j} \quad = \quad -f''(X) X^{i+j} f'(X)^{-3} \;+\; (i+j) X^{i+j-1} f'(X)^{-2}.$$

This expression depends only on the sum of indices $i+j$. This proves (9). □

We check the validity of our differential system for the case $d = 2$ and we note that it characterizes the series expansions of the quadratic formula.

```
> X := solve(a0 + a1 * x + a2 * x^2, x)[1];
                                          2           1/2
                                 -a1 + (a1  - 4 a2 a0)
                       X := 1/2 ------------------------
                                          a2

> simplify(diff(diff(X,a0),a2) - diff(diff(X,a1),a1));
                                         0
> simplify( a1*diff(X,a1) + 2*a2*diff(X,a2) + X );
                                         0
> simplify(a0*diff(X,a0)+a1*diff(X,a1)+a2*diff(X,a2));
                                         0
> series(X,a1,4);
              1/2                                1/2
       (-a2 a0)           1             (-a2 a0)      2       4
      ----------- - 1/2 ---- a1 - 1/8 ----------- a1  + O(a1 )
          a2             a2                2
```

What do you get when you now say `series(X,a0,4)` or `series(X,a2,4)` ?

Writing series expansions for the solutions to the general equation of degree d has a long tradition in mathematics. In 1757 Johann Lambert expressed the roots of the trinomial equation $x^p + x + r$ as a *Gauss hypergeometric function* in the parameter r. Series expansions of more general algebraic functions were subsequently given by Euler, Chebyshev and Eisenstein, among others. The widely known poster "Solving the Quintic with Mathematica" published by Wolfram Research in 1994 gives a nice historical introduction to series solutions of the general equation of degree five:

$$a_5 x^5 + a_4 x^4 + a_3 x^3 + a_2 x^2 + a_1 x + a_0 \quad = \quad 0. \tag{13}$$

Mayr's Theorem can be used to write down all possible Puiseux series solutions to the general quintic (13). There are $16 = 2^{5-1}$ distinct expansions. For instance, here is one of the 16 expansions of the five roots:

$$X_1 = -\left[\tfrac{a_0}{a_1}\right], \qquad X_2 = -\left[\tfrac{a_1}{a_2}\right] + \left[\tfrac{a_0}{a_1}\right], \qquad X_3 = -\left[\tfrac{a_2}{a_3}\right] + \left[\tfrac{a_1}{a_2}\right],$$
$$X_4 = -\left[\tfrac{a_3}{a_4}\right] + \left[\tfrac{a_2}{a_3}\right], \qquad X_5 = -\left[\tfrac{a_4}{a_5}\right] + \left[\tfrac{a_3}{a_4}\right].$$

Each bracket is a series having the monomial in the bracket as its first term:

$$\begin{aligned}
\Big[\frac{a_0}{a_1}\Big] &= \tfrac{a_0}{a_1}+\tfrac{a_0^2a_2}{a_1^3}-\tfrac{a_0^3a_3}{a_1^4}+2\tfrac{a_0^3a_2^2}{a_1^5}+\tfrac{a_0^4a_4}{a_1^5}-5\tfrac{a_0^4a_2a_3}{a_1^6}-\tfrac{a_0^5a_5}{a_1^6}+\cdots\\
\Big[\frac{a_1}{a_2}\Big] &= \tfrac{a_1}{a_2}+\tfrac{a_1^2a_3}{a_2^3}-\tfrac{a_1^3a_4}{a_2^4}-3\tfrac{a_0a_1^2a_5}{a_2^4}+2\tfrac{a_1^3a_3^3}{a_2^5}+\tfrac{a_1^4a_5}{a_2^5}-5\tfrac{a_1^4a_3a_4}{a_2^6}+\cdots\\
\Big[\frac{a_2}{a_3}\Big] &= \tfrac{a_2}{a_3}-\tfrac{a_0a_5}{a_3^2}-\tfrac{a_1a_4}{a_3^2}+2\tfrac{a_1a_2a_5}{a_3^3}+\tfrac{a_2^2a_4}{a_3^3}-\tfrac{a_2^3a_5}{a_3^4}+2\tfrac{a_2^3a_4^2}{a_3^5}+\cdots\\
\Big[\frac{a_3}{a_4}\Big] &= \tfrac{a_3}{a_4}-\tfrac{a_2a_5}{a_4^2}+\tfrac{a_2^2a_5}{a_4^3}+\tfrac{a_1a_5^2}{a_4^3}-3\tfrac{a_2a_3a_5^2}{a_4^4}-\tfrac{a_0a_5^3}{a_4^4}+4\tfrac{a_1a_3a_5^3}{a_4^5}+\cdots\\
\Big[\frac{a_4}{a_5}\Big] &= \tfrac{a_4}{a_5}
\end{aligned}$$

The last bracket is just a single Laurent monomial. The other four brackets $\big[\frac{a_{i-1}}{a_i}\big]$ can easily be written as an explicit sum over $\mathbb{N}^4$. For instance,

$$\Big[\frac{a_0}{a_1}\Big] \;=\; \sum_{i,j,k,l\geq 0} \frac{(-1)^{2i+3j+4k+5l}\,(2i+3j+4k+5l)!}{i\,!\,j\,!\,k\,!\,l\,!\,(i+2j+3k+4l+1)!}\cdot\frac{a_0^{i+2j+3k+4l+1}a_2^ia_3^ja_4^ka_5^l}{a_1^{2i+3j+4k+5l+1}}$$

Each coefficient appearing in one of these series is integral. Therefore these five formulas for the roots work over any ring. The situation is different for the other 15 series expansions of the roots of the quintic (13). For instance, consider the expansions into positive powers in a_1, a_2, a_3, a_4. They are

$$X_\xi \;=\; \xi\cdot\Big[\frac{a_0^{1/5}}{a_5^{1/5}}\Big]+\frac{1}{5}\cdot\Big(\xi^2\Big[\frac{a_1}{a_0^{3/5}a_5^{2/5}}\Big]+\xi^3\Big[\frac{a_2}{a_0^{2/5}a_5^{3/5}}\Big]+\xi^4\Big[\frac{a_3}{a_0^{1/5}a_5^{4/5}}\Big]-\Big[\frac{a_4}{a_5}\Big]\Big)$$

where ξ runs over the five complex roots of the equation $\xi^5=-1$, and

$$\begin{aligned}
\Big[\frac{a_0^{1/5}}{a_5^{1/5}}\Big] &= \tfrac{a_0^{1/5}}{a_5^{1/5}}-\tfrac{1}{25}\tfrac{a_1a_4}{a_0^{4/5}a_5^{6/5}}-\tfrac{1}{25}\tfrac{a_2a_3}{a_0^{4/5}a_5^{6/5}}+\tfrac{2}{125}\tfrac{a_1^2a_3}{a_0^{9/5}a_5^{6/5}}+\tfrac{3}{125}\tfrac{a_2a_4^2}{a_0^{4/5}a_5^{11/5}}+\cdots\\
\Big[\frac{a_1}{a_0^{3/5}a_5^{2/5}}\Big] &= \tfrac{a_1}{a_0^{3/5}a_5^{2/5}}-\tfrac{1}{5}\tfrac{a_3^2}{a_0^{3/5}a_5^{7/5}}-\tfrac{2}{5}\tfrac{a_2a_4}{a_0^{3/5}a_5^{7/5}}+\tfrac{7}{25}\tfrac{a_3a_4^2}{a_0^{3/5}a_5^{12/5}}+\tfrac{6}{25}\tfrac{a_1a_2a_3}{a_0^{8/5}a_5^{7/5}}+\cdots\\
\Big[\frac{a_2}{a_0^{2/5}a_5^{3/5}}\Big] &= \tfrac{a_2}{a_0^{2/5}a_5^{3/5}}-\tfrac{1}{5}\tfrac{a_1^2}{a_0^{7/5}a_5^{3/5}}-\tfrac{3}{5}\tfrac{a_3a_4}{a_0^{2/5}a_5^{8/5}}+\tfrac{6}{25}\tfrac{a_1a_2a_4}{a_0^{7/5}a_5^{8/5}}+\tfrac{3}{25}\tfrac{a_1a_3^2}{a_0^{7/5}a_5^{8/5}}+\cdots\\
\Big[\frac{a_3}{a_0^{1/5}a_5^{4/5}}\Big] &= \tfrac{a_3}{a_0^{1/5}a_5^{4/5}}-\tfrac{1}{5}\tfrac{a_1a_2}{a_0^{6/5}a_5^{4/5}}-\tfrac{2}{5}\tfrac{a_4^2}{a_0^{1/5}a_5^{9/5}}+\tfrac{1}{25}\tfrac{a_1^3}{a_0^{11/5}a_5^{4/5}}+\tfrac{4}{25}\tfrac{a_1a_3a_4}{a_0^{6/5}a_5^{9/5}}+\cdots
\end{aligned}$$

Each of these four series can be expressed as an explicit sum over the lattice points in a 4-dimensional polyhedron. The general formula can be found

in Theorem 3.2 of Sturmfels (2000). That reference gives all 2^{n-1} distinct Puiseux series expansions of the solution of the general equation of degree d.

The system (9)-(10) is a special case of the hypergeometric differential equations discussed in (Saito, Sturmfels and Takayama, 1999). More precisely, it is the Gel'fand-Kapranov-Zelevinsky system with parameters $\binom{-1}{0}$ associated with the integer matrix

$$\mathcal{A} \quad = \quad \begin{pmatrix} 0 & 1 & 2 & 3 & \cdots & n-1 & n \\ 1 & 1 & 1 & 1 & \cdots & 1 & 1 \end{pmatrix}.$$

We abbreviate the derivation $\frac{\partial}{\partial a_i}$ by the symbol ∂_i and we consider the ideal generated by the operators (10) in the commutative polynomial ring $\mathbb{Q}[\partial_0, \partial_1, \ldots, \partial_d]$. This is the ideal of the 2×2-minors of the matrix

$$\begin{pmatrix} \partial_0 & \partial_1 & \partial_2 & \cdots & \partial_{d-1} \\ \partial_1 & \partial_2 & \partial_3 & \cdots & \partial_d \end{pmatrix}.$$

This ideal defines a projective curve of degree d, namely, the *rational normal curve*, and from this it follows that our system (9)-(10) is *holonomic of rank* d This means the following: Let $(a_0, \ldots, a_d)$ be any point in $\mathbb{C}^{d+1}$ such that the discriminant of $p(x)$ is non-zero, and let $\mathcal{U}$ be a small open ball around that point. Then the set of holomorphic functions on $\mathcal{U}$ which are solutions to (9)-(10) is a complex vector space of dimension d. Theorem 8 states that the d roots of $p(x) = 0$ form a distinguished basis for that vector space.

1.6 Exercises

(1) Describe the Jordan canonical form of the companion matrix Times_x. What are the generalized eigenvectors of the endomorphism (2) ?

(2) We define a unique cubic polynomial $p(x)$ by four interpolation conditions $p(x_i) = y_i$ for $i = 0, 1, 2, 3$. The discriminant of $p(x)$ is a rational function in $x_0, x_1, x_2, x_3, y_0, y_1, y_2, y_3$. What is the denominator of this rational function, and how many terms does the numerator have ?

(3) Create a symmetric 50×50-matrix whose entries are random integers between -10 and 10 and compute the eigenvalues of your matrix.

(4) For which complex parameters α is the following system solvable ?

$$x^d - \alpha \quad = \quad x^3 + x + 1 \quad = \quad 0.$$

(5) Consider the set of all $65,536$ polynomials of degree 15 whose coefficients are $+1$ or -1. Answer the following questions about this set:

   (a) Which polynomial has largest discriminant ?

   (b) Which polynomial has the smallest number of complex roots ?

   (c) Which polynomial has the complex root of largest absolute value ?

   (d) Which polynomial has the most real roots ?

(6) Give a necessary and sufficient condition for quartic equation

$$a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 \quad = \quad 0$$

to have exactly two real roots. We expect a condition which is a Boolean combination of polynomial inequalities involving a_0, a_1, a_2, a_3, a_4.

(7) Describe an algebraic algorithm for deciding whether a polynomial $p(x)$ has a complex root of absolute value one.

(8) Compute all five Puiseux series solutions $x(t)$ of the quintic equation

$$x^5 + t \cdot x^4 + t^3 \cdot x^3 + t^6 \cdot x^2 + t^{10} \cdot x + t^{15} \quad = \quad 0$$

What is the coefficient of t^n in each of the five series ?

(9) Fix two real symmetric $n \times n$-matrices A and B. Consider the set of points (x, y) in the plane $\mathbb{R}^2$ such that all eigenvalues of the matrix $xA + yB$ are non-negative. Show that this set is closed and convex. Does every closed convex semi-algebraic subset of $\mathbb{R}^2$ arise in this way ?

(10) Let α and β be integers and consider the following system of linear differential equations for an unknown function $X(a_0, a_1, a_2)$:

$$\begin{aligned} \partial^2 X/\partial a_0 \partial a_2 \;&=\; \partial^2 X/\partial a_1^2 \\ a_1 \tfrac{\partial X}{\partial a_1} + 2a_2 \tfrac{\partial X}{\partial a_1} &= \alpha \cdot X \\ a_0 \tfrac{\partial X}{\partial a_0} + a_1 \tfrac{\partial X}{\partial a_1} + a_2 \tfrac{\partial X}{\partial a_2} &= \beta \cdot X \end{aligned}$$

For which values of α and β do (non-zero) polynomial solutions exist ? Same question for rational solutions and algebraic solutions.

2 Gröbner Bases of Zero-Dimensional Ideals

Suppose we are given polynomials $f_1, \ldots, f_m$ in $\mathbb{Q}[x_1, \ldots, x_n]$ which are known to have only finitely many common zeros in $\mathbb{C}^n$. Then $I = \langle f_1, \ldots, f_m \rangle$, the ideal generated by these polynomials, is zero-dimensional. In this section we demonstrate how Gröbner bases can be used to compute the zeros of I.

2.1 Computing standard monomials and the radical

Let $\prec$ be a term order on the polynomial ring $S = \mathbb{Q}[x_1, \ldots, x_n]$. Every ideal I in S has a unique reduced Gröbner basis $\mathcal{G}$ with respect to $\prec$. The leading terms of the polynomials in $\mathcal{G}$ generate the initial monomial ideal $\mathrm{in}_\prec(I)$. Let $\mathcal{B} = \mathcal{B}_\prec(I)$ denote the set of all monomials $x^u = x_1^{u_1} x_2^{u_2} \cdots x_n^{u_n}$ which do not lie in $\mathrm{in}_\prec(I)$. These are the *standard monomials* of I with respect to $\prec$. Every polynomial f in S can be written uniquely as a $\mathbb{Q}$-linear combination of $\mathcal{B}$ modulo I, using the division algorithm with respect to the Gröbner basis $\mathcal{G}$. We write $\mathcal{V}(I) \subset \mathbb{C}^n$ for the complex variety defined by the ideal I.

Proposition 9. *The variety $\mathcal{V}(I)$ is finite if and only if the set $\mathcal{B}$ is finite, and the cardinality of $\mathcal{B}$ equals the cardinality of $\mathcal{V}(I)$, counting multiplicities.*

Consider an example with three variables denoted $S = \mathbb{Q}[x, y, z]$:

$$I \quad = \quad \langle\, (x-y)^3 - z^2,\, (z-x)^3 - y^2,\, (y-z)^3 - x^2 \,\rangle. \tag{14}$$

The following `Macaulay2` computation verifies that I is zero-dimensional:

```
i1 : S = QQ[x,y,z];
i2 : I = ideal( (x-y)^3-z^2, (z-x)^3-y^2, (y-z)^3-x^2 );
o2 : Ideal of S

i3 : dim I, degree I
o3 = (0, 14)

i4 : gb I

o4 = | y2z-1/2xz2-yz2+1/2z3+13/60x2-1/12y2+7/60z2
     x2z-xz2-1/2yz2+1/2z3+1/12x2-13/60y2-7/60z2
     y3-3y2z+3yz2-z3-x2
     xy2-2x2z-3y2z+3xz2+4yz2-3z3-7/6x2+5/6y2-1/6z2
```

```
    x2y-xy2-x2z+y2z+xz2-yz2+1/3x2+1/3y2+1/3z2
    x3-3x2y+3xy2-3y2z+3yz2-z3-x2-z2
    z4+1/5xz2-1/5yz2+2/25z2
    yz3-z4-13/20xz2-3/20yz2+3/10z3+2/75x2-4/75y2-7/300z2
    xz3-2yz3+z4+29/20xz2+19/20yz2-9/10z3-8/75x2+2/15y2+7/300z2
    xyz2-3/2y2z2+xz3+yz3-3/2z4+y2z-1/2xz2
              -7/10yz2+1/5z3+13/60x2-1/12y2-1/12z2|

i5 : toString (x^10 % I)

o5 = -4/15625*x*z^2+4/15625*z^3-559/1171875*x^2
     -94/1171875*y^2+26/1171875*z^2

i6 : R = S/I;  basis R

o7 = | 1 x x2 xy xyz xz xz2 y y2 yz yz2 z z2 z3 |
              1       14
o7 : Matrix R  <--- R
```

The output `o4` gives the reduced Gröbner basis for I with respect to the reverse lexicographic term order with $x > y > z$. We see in `o7` that there are 14 standard monomials. In `o5` we compute the expansion of x^{10} in this basis of S/I. We conclude that the number of complex zeros of I is at most 14.

If I is a zero-dimensional ideal in $S = \mathbb{Q}[x_1, \ldots, x_n]$ then the elimination ideal $I \cap \mathbb{Q}[x_i]$ is non-zero for all $i = 1, 2, \ldots, n$. Let $p_i(x_i)$ denote the generator of $I \cap \mathbb{Q}[x_i]$. The univariate polynomial p_i can be gotten by Gröbner basis for I with respect to an elimination term order. Another method is to use an arbitrary Gröbner basis compute the normal form of successive powers of x_i until they first become linearly dependent.

We denote the square-free part of the polynomial $p_i(x_i)$ by

$$p_{i,red}(x_i) \quad = \quad p_i(x_i)/gcd(p_i(x_i), p_i'(x_i)).$$

Theorem 10. *A zero-dimensional ideal I is radical if and only if the n elimination ideals $I \cap \mathbb{Q}[x_i]$ are radical. Moreover, the radical of I equals*

$$\mathrm{Rad}(I) \quad = \quad I + \langle\, p_{1,red},\, p_{2,red},\, \ldots, p_{n,red} \,\rangle.$$

Our example in (14) is symmetric with respect to the variables, so that

$$I \cap \mathbb{Q}[x] \,=\, \langle p(x) \rangle, \quad I \cap \mathbb{Q}[y] \,=\, \langle p(y) \rangle, \quad I \cap \mathbb{Q}[z] \,=\, \langle p(z) \rangle.$$

The common generator of the elimination ideals is a polynomial of degree 8:

$$p(x) \quad = \quad x^8 + \frac{6}{25}x^6 + \frac{17}{625}x^4 + \frac{8}{15625}x^2$$

This polynomial is not squarefree. Its squarefree part equals

$$p_{red}(x) \quad = \quad x^7 + \frac{6}{25}x^5 + \frac{17}{625}x^3 + \frac{8}{15625}x.$$

Hence our ideal I is not radical. Using Theorem 10, we compute its radical:

$$\begin{aligned}
\mathrm{Rad}(I) \quad &= \quad I + \langle p_{red}(x), p_{red}(y), p_{red}(z) \rangle \\
&= \quad \langle \underline{x} - 5/2y^2 - 1/2y + 5/2z^2 - 1/2z, \\
&\quad \underline{y} + 3125/8z^6 + 625/4z^5 + 375/4z^4 + 125/4z^3 + 65/8z^2 + 3z, \\
&\quad \underline{z^7} + 6/25z^5 + 17/625z^3 + 8/15625z \rangle.
\end{aligned}$$

The three given generators form a lexicographic Gröbner basis. We see that $\mathcal{V}(I)$ has cardinality seven. The only real root is the origin. The other six zeros of I in $\mathbb{C}^3$ are not real. They are gotten by cyclically shifting

$$\begin{aligned}
(x, y, z) \;&=\; \big(-0.14233 - 0.35878i,\, 0.14233 - 0.35878i,\, 0.15188i\,\big) \\
\text{and} \quad (x, y, z) \;&=\; \big(-0.14233 + 0.35878i,\, 0.14233 + 0.35878i,\, -0.15188i\big).
\end{aligned}$$

2.2 Localizing and removing known zeros

In the example above, the origin is a zero of multiplicity 8. and it would have made sense to remove this distinguished zero right from the beginning. In this section we explain how to do this and how the number 8 could have been derived a priori. Let I be a zero-dimensional ideal in $S = \mathbb{Q}[x_1, \ldots, x_n]$ and $p = (p_1, \ldots, p_n)$ any point with coordinates in $\mathbb{Q}$. We consider the associated *maximal ideal*

$$M \quad = \quad \langle x_1 - p_1, x_2 - p_2, \ldots, x_n - p_n \rangle \quad \subset \quad S.$$

The *ideal quotient* of I by M is defined as

$$\big(I \,:\, M\big) \quad = \quad \big\{\, f \in S \;:\; f \cdot M \,\subseteq\, I \big\}.$$

We can iterate this process to get the increasing sequence of ideals

$$I \,\subseteq\, (I : M) \,\subseteq\, (I : M^2) \,\subseteq\, (I : M^3) \,\subseteq\, \cdots$$

This sequence stabilizes with an ideal called the *saturation*

$$\big(I : M^\infty\big) \quad = \quad \big\{ f \in S \,:\, \exists\, m \in \mathbb{N} \,:\, f^m \cdot M \,\subseteq\, I \big\}.$$

Proposition 11. *The variety of* $(I : M^\infty)$ *equals* $\mathcal{V}(I) \backslash \{p\}$.

Here is how we compute the ideal quotient and the saturation in `Macaulay 2`. We demonstrate this for the ideal in the previous section and $p = (0, 0, 0)$:

```
i1 : R = QQ[x,y,z];
i2 : I = ideal( (x-y)^3-z^2, (z-x)^3-y^2,  (y-z)^3-x^2 );
i3 : M = ideal( x , y, z );

i4 : gb (I : M)

o4 = | y2z-1/2xz2-yz2+1/2z3+13/60x2-1/12y2+7/60z2
       xyz+3/4xz2+3/4yz2+1/20x2-1/20y2 x2z-xz2-1/2yz2+ ....

i5 : gb saturate(I,M)

o5 = | z2+1/5x-1/5y+2/25 y2-1/5x+1/5z+2/25
       xy+xz+yz+1/25 x2+1/5y-1/5z+2/25 |

i6 :  degree I, degree (I:M), degree (I:M^2), degree(I:M^3)

o6 = (14, 13, 10, 7)

i7 : degree (I : M^4), degree (I : M^5), degree (I : M^6)

o7 = (6, 6, 6)
```

In this example, the fourth ideal quotient $(I : M^4)$ equals the saturation $(I : M^\infty) =$ `saturate(I,M)`. Since $p = (0, 0, 0)$ is a zero of high multiplicity, namely eight, it would be interesting to further explore the local ring S_p/I_p. This is an 8-dimensional $\mathbb{Q}$-vector space which tells the *scheme structure* at p, meaning the manner in which those eight points pile on top of one another.

The following general method can be used to compute the local ring at an isolated zero of any polynomial system. Form the ideal quotient

$$J \quad = \quad \big(I : (I : M^\infty)\big). \tag{15}$$

Proposition 12. *The ring S/J is isomorphic to the local ring S_p/I_p under the natural map $x_i \mapsto x_i$. In particular, the multiplicity of p as a zero of I equals the number of standard monomials for any Gröbner basis of J.*

In our example, the local ideal J is particularly simple and the multiplicity eight is obvious. Here is how the `Macaulay 2` session continues:

```
i8 : J = ( I : saturate(I,M) )

              2   2   2
o8 = ideal (z , y , x )

i9 : degree J

o9 = 8
```

Propositions 11 and 12 provide a decomposition of the given ideal:

$$I \quad = \quad J \quad \cap \quad (I : M^\infty). \tag{16}$$

Here J is the iterated ideal quotient in (15). This ideal is primary to the maximal ideal M, that is, $\mathrm{Rad}(J) = M$. We can now iterate by applying this process to the ideal $(I : M^\infty)$, and this will eventually lead to the *primary decomposition* of I. We shall return to this topic in later lectures.

For the ideal in our example, the decomposition (16) is already the primary decomposition when working over the field of rational numbers. It equals

$$\begin{aligned} &\langle\, (x-y)^3 - z^2,\, (z-x)^3 - y^2,\, (y-z)^3 - x^2 \,\rangle \quad = \\ &\langle\, x^2\,,\, y^2\,,\, z^2 \,\rangle \quad \cap \quad \langle\, \underline{z^2} + \tfrac{1}{5}x - \tfrac{1}{5}y + \tfrac{2}{25},\, \underline{y^2} - \tfrac{1}{5}x + \tfrac{1}{5}z + \tfrac{2}{25}, \\ &\qquad\qquad \underline{x^2} + \tfrac{1}{5}y - \tfrac{1}{5}z + \tfrac{2}{25},\, \underline{xy} + xz + yz + \tfrac{1}{25} \,\rangle \end{aligned}$$

Note that the second ideal is maximal and hence prime in $\mathbb{Q}[x, y, z]$. The given generators are a Gröbner basis with leading terms underlined.

2.3 Companion matrices

Let I be a zero-dimensional ideal in $S = \mathbb{Q}[x_1, \ldots, x_n]$, and suppose that the $\mathbb{Q}$-vectorspace S/I has dimension d. In this section we assume that some

Gröbner basis of I is known. Let $\mathcal{B}$ denote the associated monomial basis for S/I. Multiplication by any of the variables x_i defines an endomorphism

$$S/I \;\rightarrow\; S/I\,,\; f \;\mapsto x_i \cdot f \tag{17}$$

We write T_i for the $d \times d$-matrix over $\mathbb{Q}$ which represents the linear map (17) with respect to the basis $\mathcal{B}$. The rows and columns of T_i are indexed by the monomials in $\mathcal{B}$. If $x^u, x^v \in \mathcal{B}$ then the entry of T_i in row x^u and column x^v is the coefficient of x^u in the normal form of $x_i \cdot x^v$. We call T_i the *i-th companion matrix* of the ideal I. It follows directly from the definition that the companion matrices commute pairwise:

$$T_i \cdot T_j \quad = \quad T_j \cdot T_i \qquad \text{for } 1 \leq i < j \leq n.$$

The matrices T_i generate a commutative subalgebra of the non-commutative ring of $d \times d$-matrices, and this subalgebra is isomorphic to our ring

$$\mathbb{Q}[T_1, \ldots, T_n] \quad \simeq \quad S/I\,, \quad T_i \;\mapsto\; x_i.$$

Theorem 13. *The complex zeros of the ideal I are the vectors of joint eigenvalues of the companion matrices $T_1, \ldots, T_n$, that is,*

$$\mathcal{V}(I) \quad = \quad \big\{\, (\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^n \;:\; \exists\, v \in \mathbb{C}^n \;\; \forall\, i \;:\; T_i \cdot v \,=\, \lambda_i \cdot v \,\big\}.$$

Proof. Suppose that v is a non-zero complex vector such that $T_i \cdot v = \lambda_i \cdot v$ for all i. Then, for any polynomial $p \in S$,

$$p(T_1, \ldots, T_n) \cdot v \quad = \quad p(\lambda_1, \ldots, \lambda_n) \cdot v.$$

If p is in the ideal I then $p(T_1, \ldots, T_n)$ is the zero matrix and we conclude that $p(\lambda_1, \ldots, \lambda_n) = 0$. Hence $\mathcal{V}(I)$ contains the set on the right hand side. We prove the converse under the hypothesis that I is a radical ideal. (The general case is left to the reader). Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be any zero of I. There exists a polynomial $q \in S \otimes \mathbb{C}$ such that $p(\lambda) = 1$ and p vanishes at all points in $\mathcal{V}(I) \backslash \{\lambda\}$. Then $x_i \cdot q = \lambda_i \cdot q$ holds on $\mathcal{V}(I)$, hence $(x_i - \lambda_i) \cdot q$ lies in the radical ideal I. Let v be the non-zero vector representing the element q of $S/I \otimes \mathbb{C}$. Then v is a joint eigenvector with joint eigenvalue λ. □

If I is a radical ideal then we can form a square invertible matrix V whose columns are the eigenvectors v described above. Then $V^{-1} \cdot T_i \cdot V$ is a diagonal matrix whose entries are the i-th coordinates of all the zeros of I.

Corollary 14. *The companion matrices $T_1, \ldots, T_n$ can be simultaneously diagonalized if I is a radical ideal.*

As an example consider the Gröbner basis given at the end of the last section. The given ideal is a prime ideal in $\mathbb{Q}[x, y, z]$ having degree $d = 6$. We determine the three companion matrices `Tx`, `Ty` and `Tz` using `maple`:

```
> with(Groebner):

> GB := [z^2+1/5*x-1/5*y+2/25, y^2-1/5*x+1/5*z+2/25,
>        x*y+x*z+y*z+1/25, x^2+1/5*y-1/5*z+2/25]:

> B :=  [1, x, y, z, x*z, y*z]:

> for v in [x,y,z] do
>  T := array([],1..6,1..6):
>  for j from 1 to 6 do
>    p := normalf( v*B[j], GB, tdeg(x,y,z)):
>    for i from 1 to 6 do
>     T[i,j] := coeff(coeff(coeff(p,x,degree(B[i],x)),y,
>                 degree(B[i],y)),z,degree(B[i],z)):
>  od:
> od:

> print(cat(T,v),T);
> od:
```

$$
\mathtt{Tx},\ \begin{bmatrix} -\frac{2}{25} & -\frac{1}{25} & 0 & -\frac{2}{125} & 0 \\ 0 & 0 & 0 & -\frac{1}{25} & \frac{1}{25} \\ 0 & -\frac{1}{5} & 0 & 0 & \frac{1}{25} & \frac{1}{25} \\ 0 & 0 & 0 & 0 & -\frac{1}{25} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{2}{25} \end{bmatrix}
$$

```
    [0    1/5     0     0     --     1/25]
    [                       25         ]
    [                                  ]
    [0     0      -1    1     0       0  ]
    [                                  ]
    [0     0      -1    0    -1/5     0  ]
```

```
    [     -1     -2                      ]
    [0    --     --     0     0      2/125]
    [     25     25                      ]
    [                                    ]
    [0    0     1/5     0    1/25    1/25 ]
    [                                    ]
    [                                -1  ]
    [1    0      0      0    1/25     --  ]
Ty, [                                25  ]
    [                                    ]
    [                                -2  ]
    [0    0     -1/5    0    1/25     --  ]
    [                                25  ]
    [                                    ]
    [0    -1     0      0     0       1/5 ]

    [0    -1     0      1     0        0  ]
```

```
    [                -2              -1  ]
    [0    0    0     --     1/125    --- ]
    [                25              125 ]
    [                                    ]
    [                          -2        ]
    [0    0    0    -1/5     --      1/25]
    [                          25        ]
    [                                    ]
    [                                -2  ]
Tz, [0    0    0    1/5     1/25     --  ]
```

```
[                                                  25 ]
[                                                     ]
[                                        -1        -1 ]
[1     0     0     0                     --        -- ]
[                                        25        25 ]
[                                                     ]
[0     1     0     0                   -1/5      1/5 ]

[0     0     1     0                   -1/5      1/5 ]
```

The matrices `Tx`, `Ty` and `Tz` commute pairwise and they can be simultaneously diagonalized. The entries on the diagonal are the six complex zeros. We invite the reader to compute the common basis of eigenvectors using `matlab`.

2.4 The trace form

In this section we explain how to compute the number of real roots of a zero-dimensional ideal which is presented to us by a Gröbner basis as before. Fix any other polynomial $h \in S$ and consider the following bilinear form on our vector space $S/I \simeq \mathbb{Q}^d$. This is called the *trace form for* h:

$$B_h \,:\, S/I \times S/I \,\mapsto\, \mathbb{Q},\ (f,g) \,\mapsto\, \text{trace}\big((f \cdot g \cdot h)(T_1, T_2, \ldots, T_n)\big).$$

We represent the quadratic form T_h by a symmetric $d \times d$-matrix over $\mathbb{Q}$ with respect to the basis $\mathcal{B}$. If $x^u, x^v \in \mathcal{B}$ then the entry of B_h in row x^u and column x^v is the sum of the diagonal entries in the $d \times d$-matrix gotten by substituting the companion matrices T_i for the variables x_i in the polynomial $x^{u+v} \cdot h$. This rational number can be computed by summing, over all $x^w \in \mathcal{B}$, the coefficient of x^w in the normal form of $x^{u+v+w} \cdot h$ modulo I.

Since the matrix B_h is symmetric, all of its eigenvalues are real numbers. The *signature* of B_h is the number of positive eigenvalues of B_h minus the number of negative eigenvalues of B_h. It turns out that this number is always non-negative for symmetric matrices of the special form B_h. In the following theorem, multiple real zeros of I are counted only once.

Theorem 15. *The signature of the trace form B_h equals the number of real roots p of I with $h(p) > 0$ minus the number of real roots p of I with $h(p) < 0$.*

The special case when $h = 1$ is used to count all real roots:

Corollary 16. *The number of real roots of I equals the signature of B_1.*

We compute the symmetric 6×6-matrix B_1 for the case of the polynomial system whose companion matrices were determined in the previous section.

```
> with(linalg): with(Groebner):

> GB := [z^2+1/5*x-1/5*y+2/25, y^2-1/5*x+1/5*z+2/25,
>        x*y+x*z+y*z+1/25, x^2+1/5*y-1/5*z+2/25]:
> B :=  [1, x, y, z, x*z, y*z]:

> B1 := array([],1..6,1..6):
> for j from 1 to 6 do
> for i from 1 to 6 do
> B1[i,j] :=  0:
> for k from 1 to 6 do
> B1[i,j] := B1[i,j] +  coeff(coeff(coeff(
> normalf(B[i]*B[j]*B[k], GB, tdeg(x,y,z)),x,
> degree(B[k],x)), y, degree(B[k],y)),z, degree(B[k],z)):
> od:
> od:
> od:

> print(B1);
                  [                                  -2      -2 ]
                  [6      0      0       0           --      -- ]
                  [                                  25      25 ]
                  [                                             ]
                  [      -12     -2      -2      -2             ]
                  [0     ---     --      --      --      0      ]
                  [      25      25      25      25             ]
                  [                                             ]
                  [      -2     -12      -2                     ]
                  [0     --     ---      --      0      2/25    ]
                  [      25     25       25                     ]
                  [                                             ]
                  [      -2      -2     -12             -2      ]
                  [0     --      --     ---     2/25     --      ]
```

```
                  [      25      25     25                25 ]
                  [                                        ]
                  [-2    -2                    34       -16 ]
                  [--    --      0     2/25    ---      --- ]
                  [25    25                    625      625 ]
                  [                                        ]
                  [-2                   -2     -16      34  ]
                  [--     0     2/25     --     ---      --- ]
                  [25                   25     625      625 ]
```

```
> charpoly(B1,z);
```

```
 6   2918  5   117312  4   1157248  3   625664   2
z  - ---- z  - ------ z  - ------- z  - ------- z
      625       15625       390625     9765625

     4380672        32768
  + -------- z  -  ------
    48828125       9765625
```

```
> fsolve(%);
```

```
-.6400000, -.4371281, -.4145023, .04115916, .1171281, 6.002143
```

Here the matrix B_1 has three positive eigenvalues and three negative eigenvalues, so the trace form has signature zero. This confirms our earlier finding that these equations have no real zeros. We note that we can read off the signature of B_1 directly from the characteristic polynomial. Namely, the characteristic polynomial has three sign changes in its coefficient sequence. Using the following result, which appears in Exercise 5 on page 67 of (Cox, Little & O'Shea, 1998), we infer that there are three positive real eigenvalues and this implies that the signature of B_1 is zero.

Lemma 17. *The number of positive eigenvalues of a real symmetric matrix equals the number of sign changes in the coefficient sequence of its characteristic polynomial.*

It is instructive to examine the trace form for the case of one polynomial

in one variable. Consider the principal ideal

$$I \quad = \quad \langle\, a_d x^d + a_{d-1}x^{d-1} + \cdots + a_2 x^2 + a_1 x + a_0 \,\rangle \quad \subset \quad S = \mathbb{Q}[x].$$

We consider the traces of successive powers of the companion matrix:

$$b_i \quad := \quad \mathrm{trace}\big(\mathrm{Times}_x^i\big) \quad = \quad \sum_{u \in \mathcal{V}(I)} u^i.$$

Thus b_i is a Laurent polynomial of degree zero in $a_0, \ldots, a_d$, which is essentially the familiar Newton relation between elementary symmetric functions and power sum symmetric functions. The trace form is given by the matrix

$$B_1 \quad = \quad \begin{pmatrix} b_0 & b_1 & b_2 & \cdots & b_{d-1} \\ b_1 & b_2 & b_3 & \cdots & b_d \\ b_2 & b_3 & b_4 & \cdots & b_{d+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{d-1} & b_d & b_{d+1} & \cdots & b_{2d-2} \end{pmatrix} \tag{18}$$

Thus the number of real zeros of I is the signature of this Hankel matrix. For instance, for $d = 4$ the entries in the 4×4-Hankel matrix B_1 are

$$\begin{aligned}
b_0 &= 4 \\
b_1 &= \tfrac{-a_3}{a_4} \\
b_2 &= \tfrac{-2a_4a_2+a_3^2}{a_4^2} \\
b_3 &= \tfrac{-3a_4^2a_1+3a_4a_3a_2-a_3^3}{a_4^3} \\
b_4 &= \tfrac{-4a_4^3a_0+4a_4^2a_3a_1+2a_4^2a_2^2-4a_4a_3^2a_2+a_3^4}{a_4^4} \\
b_5 &= \tfrac{-5a_4^3a_3a_0-5a_4^3a_2a_1+5a_4^2a_3^2a_1+5a_4^2a_3a_2^2-5a_4a_3^3a_2+a_3^5}{a_4^5} \\
b_6 &= \tfrac{-6a_4^4a_2a_0-3a_4^4a_1^2+6a_4^3a_3^2a_0+12a_4^3a_3a_2a_1+2a_4^3a_2^3-6a_4^2a_3^3a_1-9a_4^2a_3^2a_2^2+6a_4a_3^4a_2-a_3^6}{a_4^6},
\end{aligned}$$

and the characteristic polynomial of the 4×4-matrix B_1 equals

$$\begin{aligned}
& x^4 \quad + \quad (-b_0 - b_2 - b_4 - b_6) \cdot x^3 \\
+ \quad & (b_0b_2 + b_0b_4 + b_0b_6 - b_5^2 - b_1^2 - b_2^2 + b_2b_4 + b_2b_6 - 2b_3^2 - b_4^2 + b_4b_6) \cdot x^2 \\
+ \quad & (b_0b_5^2 - b_0b_2b_4 - b_0b_2b_6 + b_0b_3^2 + b_0b_4^2 - b_0b_4b_6 + b_5^2b_2 - 2b_5b_2b_3 - 2b_5b_3b_4 + b_1^2b_4 \\
& \quad + b_1^2b_6 - 2b_1b_2b_3 - 2b_1b_3b_4 + b_2^3 + b_2^2b_6 + b_2b_3^2 - b_2b_4b_6 + b_3^2b_4 + b_3^2b_6 + b_4^3) \cdot x \\
- \quad & b_0b_5^2b_2 + 2b_0b_5b_3b_4 + b_0b_2b_4b_6 - b_0b_3^2b_6 - b_0b_4^3 + b_5^2b_1^2 - 2b_5b_1b_2b_4 - 2b_5b_1b_3^2 \\
& \quad + 2b_5b_2^2b_3 - b_1^2b_4b_6 + 2b_1b_2b_3b_6 + 2b_1b_3b_4^2 - b_2^3b_6 + b_2^2b_4^2 - 3b_2b_3^2b_4 + b_3^4
\end{aligned}$$

By considering sign alternations among these expressions in $b_0, b_1, \ldots, b_6$, we get explicit conditions for the general quartic to have zero, one, two, three, or four real roots respectively. These are *semialgebraic conditions*. This means the conditions are Boolean combinations of polynomial inequalities in the five indeterminates a_0, a_1, a_2, a_3, a_4. In particular, all four zeros of the general quartic are real if and only if the trace form of positive definite. Recall that a symmetric matrix is positive definite if and only if its principal minors are positive. Hence the quartic has four real roots if and only if

$$\begin{array}{c} b_0 > 0 \text{ and } b_0 b_2 - b_1^2 > 0 \text{ and } b_0 b_2 b_4 - b_0 b_3^2 - b_1^2 b_4 + 2 b_1 b_2 b_3 - b_2^3 > 0 \text{ and } \\ 2 b_0 b_5 b_3 b_4 - b_0 b_5^2 b_2 + b_0 b_2 b_4 b_6 - b_0 b_3^2 b_6 - b_0 b_4^3 + b_5^2 b_1^2 - 2 b_5 b_1 b_2 b_4 - 2 b_5 b_1 b_3^2 \\ + 2 b_5 b_2^2 b_3 - b_1^2 b_4 b_6 + 2 b_1 b_2 b_3 b_6 + 2 b_1 b_3 b_4^2 - b_2^3 b_6 + b_2^2 b_4^2 - 3 b_2 b_3^2 b_4 + b_3^4 \ > \ 0. \end{array}$$

The last polynomial is the determinant of B_1. It equals the discriminant of the quartic (displayed in `maple` at the beginning of Lecture 1) divided by a_4^6.

2.5 Exercises

(1) Let $A = (a_{ij})$ be a non-singular $n \times n$-matrix whose entries are positive integers. How many complex solutions do the following equations have:

$$\prod_{j=1}^{n} x^{a_{1j}} \ = \ \prod_{j=1}^{n} x^{a_{2j}} \ = \ \cdots \ = \ \prod_{j=1}^{n} x^{a_{nj}} \ = \ 1.$$

(2) Pick a random homogeneous cubic polynomial in four variables. Compute the 27 lines on the cubic surface defined by your polynomial.

(3) Given d arbitrary rational numbers $a_0, a_1, \ldots, a_{d-1}$, consider the system of d polynomial equations in d unknowns $z_1, z_2, \ldots, z_d$ given by setting

$$x^d + a_{d-1} x^{d-1} \cdots + a_1 x + a_0 \ = \ (x - z_1)(x - z_2) \cdots (x - z_d).$$

Describe the primary decomposition of this ideal in $\mathbb{Q}[z_1, z_1, \ldots, z_d]$. How can you use this to find the Galois group of the given polynomial ?

(4) For any two positive integers m, n, find an explicit radical ideal I in $\mathbb{Q}[x_1, \ldots, x_n]$ and a term order $\prec$ such that $in_\prec(I) \ = \ \langle x_1, x_2, \ldots, x_n \rangle^m$.

(5) Fix the monomial ideal $M = \langle x, y \rangle = \langle x^3, x^2y, xy^2, y^3 \rangle$ and compute its companion matrices T_x, T_y. Describe all polynomial ideals in $\mathbb{Q}[x, y]$ which are within distance $\epsilon = 0.0001$ from M, in the sense that the companion matrices are ϵ-close to T_x, T_y in your favorite matrix norm.

(6) Does every zero-dimensional ideal in $\mathbb{Q}[x, y]$ have a radical ideal in all of its ϵ-neighborhoods ? How about zero-dimensional ideals in $\mathbb{Q}[x, y, z]$?

(7) How many distinct real vectors $(x, y, z) \in \mathbb{R}^3$ satisfy the equations

$$x^3 + z = 2y^2, \quad y^3 + x = 2z^2, \quad z^3 + y = 2x^2 \quad ?$$

(8) Pick eight random points in the real projective plane. Compute the 12 nodal cubic curves passing through your points. Can you find eight points such that all 12 cubic polynomials have real coefficients ?

(9) Consider a quintic polynomial in two variables, for instance,

$$\begin{aligned} f \quad = \quad & 5y^5 + 19y^4x + 36y^3x^2 + 34y^2x^3 + 16yx^4 + 3x^5 \\ & +6y^4 + 4y^3x + 6y^2x^2 + 4yx^3 + x^4 + 10y^3 + 10y^2 + 5y + 1. \end{aligned}$$

Determine the irreducible factor of f in $\mathbb{R}[x, y]$, and also in $\mathbb{C}[x, y]$.

(10) Consider a polynomial system which has infinitely many complex zeros but only finitely many of them have all their coordinates distinct. How would you compute those zeros with distinct coordinates ?

(11) Does there exist a Laurent polynomial in $\mathbb{C}[t, t^{-1}]$ of the form

$$f \quad = \quad t^{-4} + x_3t^{-3} + x_2t^{-2} + x_1t^{-1} \; + \; y_1t + y_2t^2 + y_3t^3 + t^4$$

such that the powers f^2, f^3, f^4, f^5, f^6 and f^7 all have zero constant term ? Can you find such a Laurent polynomial with real coefficients ? What if we also require that the constant term of t^8 is zero ?

(12) A well-studied problem in number theory is to find rational points on elliptic curves. Given an ideal $I \subset \mathbb{Q}[x_1, \ldots, x_n]$ how can you decide whether $\mathcal{V}(I)$ is an elliptic curve, and, in the affirmative case, which computer program would you use to look for points in $\mathcal{V}(I) \cap \mathbb{Q}^n$?

3 Bernstein's Theorem and Fewnomials

The Gröbner basis methods described in the previous lecture apply to arbitrary systems of polynomial equations. They are so general that they are frequently not the best choice when dealing with specific classes polynomial systems. A situation encountered in many applications is a system of n sparse polynomial equations in n variables which have finitely many roots. Algebraically, this situation is special because we are dealing with a complete intersection, and sparsity allows us to use polyhedral techniques for counting and computing the zeros. This lecture gives a gentle introduction to sparse polynomial systems by explaining some basic techniques for $n = 2$.

3.1 From Bézout's Theorem to Bernstein's Theorem

A polynomial in two unknowns looks like

$$f(x,y) \quad = \quad a_1 x^{u_1} y^{v_1} + a_2 x^{u_2} y^{v_2} + \cdots + a_m x^{u_m} y^{v_m}, \tag{19}$$

where the exponents u_i and v_i are non-negative integers and the coefficients a_i are non-zero rationals. Its *total degree* $deg(f)$ is the maximum of the numbers $u_1 + v_1, \ldots, u_m + v_m$. The following theorem gives an upper bound on the number of common complex zeros of two polynomials in two unknowns.

Theorem 18. (Bézout's Theorem) *Consider two polynomial equations in two unknowns: $g(x,y) = h(x,y) = 0$. If this system has only finitely many zeros $(x,y) \in \mathbb{C}^2$, then the number of zeros is at most $deg(g) \cdot deg(h)$.*

Bézout's Theorem is best possible in the sense that almost all polynomial systems have $deg(g) \cdot deg(h)$ distinct solutions. An explicit example is gotten by taking g and h as products of linear polynomials $u_1 x + u_2 y + u_3$. More precisely, there exists a polynomial in the coefficients of g and h such that whenever this polynomial is non-zero then f and g have the expected number of zeros. The first exercise below concerns finding such a polynomial.

A drawback of Bézout's Theorem is that it yields little information for polynomials that are sparse. For example, consider the two polynomials

$$g(x,y) \;=\; a_1 + a_2 x + a_3 xy + a_4 y\,, \quad h(x,y) \;=\; b_1 + b_2 x^2 y + b_3 x y^2. \tag{20}$$

These two polynomials have precisely four distinct zeros $(x,y) \in \mathbb{C}^2$ for generic choices of coefficients a_i and b_j. Here "generic" means that a certain

polynomial in the coefficients a_i, b_j, called the *discriminant*, should be non-zero. The discriminant of the system (20) is the following expression

$$\begin{gathered}
4a_1^7a_3b_2^3b_3^3 + a_1^6a_2^2b_2^2b_3^4 - 2a_1^6a_2a_4b_2^3b_3^3 + a_1^6a_4^2b_2^4b_3^2 + 22a_1^5a_2a_3^2b_1b_2^2b_3^3 \\
+22a_1^5a_3^2a_4b_1b_2^3b_3^2 + 22a_1^4a_2^3a_3b_1b_2b_3^4 + 18a_1a_2a_3a_4^5b_1^2b_2^4 - 30a_1^4a_2a_3a_4^2b_1b_2^3b_3^2 \\
+a_1^4a_3^4b_1^2b_2^2b_3^2 + 22a_1^4a_3a_4^3b_1b_2^4b_3 + 4a_1^3a_2^5b_1b_3^5 - 14a_1^3a_2^4a_4b_1b_2b_3^4 \\
+10a_1^3a_2^3a_4^2b_1b_2^2b_3^3 + 22a_1^3a_2^2a_3^3b_1^2b_2b_3^3 + 10a_1^3a_2^2a_4^3b_1b_2^3b_3^2 + 116a_1^3a_2a_3^3a_4b_1^2b_2^2b_3^2 \\
-14a_1^3a_2a_4^4b_1b_2^4b_3 + 22a_1^3a_3^3a_4^2b_1^2b_2^3b_3 + 4a_1^3a_4^5b_1b_2^5 + a_1^2a_2^4a_3^2b_1^2b_3^4 \\
+94a_1^2a_2^3a_3^2a_4b_1^2b_2b_3^3 - 318a_1^2a_2^2a_3^2a_4^2b_1^2b_2^2b_3^2 + 396a_1a_2^3a_3a_4^3b_1^2b_2^2b_3^2 + a_1^2a_3^2a_4^4b_1^2b_2^4 \\
+94a_1^2a_2a_3^2a_4^3b_1^2b_2^3b_3 + 4a_1^2a_2a_3^5b_1^3b_2b_3^2 + 4a_1^2a_3^5a_4b_1^3b_2^2b_3 + 18a_1a_2^5a_3a_4b_1^2b_3^4 \\
-216a_1a_2^4a_3a_4^2b_1^2b_2b_3^3 + 96a_1a_2^2a_3^4a_4b_1^3b_2b_3^2 - 216a_1a_2^2a_3a_4^4b_1^2b_2^3b_3 - 27a_2^6a_4^2b_1^2b_3^4 \\
-30a_1^4a_2^2a_3a_4b_1b_2^2b_3^3 + 96a_1a_2a_3^4a_4^2b_1^3b_2^2b_3 + 108a_2^5a_3^4b_1^2b_2b_3^3 \\
+4a_2^4a_3^3a_4b_1^3b_3^3 - 162a_2^4a_4^4b_1^2b_2^2b_3^2 - 132a_2^3a_3^3a_4^2b_1^3b_2b_3^2 + 108a_2^3a_4^5b_1^2b_2^3b_3 \\
-132a_2^2a_3^3a_4^3b_1^3b_2^2b_3 - 27a_2^2a_4^6b_1^2b_2^4 + 16a_2a_3^6a_4b_1^4b_2b_3 + 4a_2a_3^3a_4^4b_1^3b_2^3
\end{gathered}$$

If this polynomial of degree 14 is non-zero, then the system (20) has four distinct complex zeros. This discriminant is computed in `maple` as follows.

```
g := a1 + a2 * x + a3 * x*y + a4 * y;
h := b1 + b2 * x^2 * y + b3 * x * y^2;
R := resultant(g,h,x):
S :=  factor( resultant(R,diff(R,y),y) ):
discriminant := op( nops(S), S);
```

Bezout's Theorem would predict $deg(g) \cdot deg(h) = 6$ common complex zeros for the equations in (20). Indeed, in projective geometry we would expect the cubic curve $\{g = 0\}$ and the quadratic curve $\{h = 0\}$ to intersect in six points. But these particular curves never intersect in more than four points in $\mathbb{C}^2$. How come ? To understand why the number is four and not six, we need to associate convex polygons with our given polynomials.

Convex polytopes have been studied since the earliest days of mathematics. We shall see that they are very useful for analyzing and solving polynomial equations. A *polytope* is a subset of $\mathbb{R}^n$ which is the convex hull of a finite set of points. A familiar example is the convex hull of $\{(0,0,0),(0,1,0),(0,0,1),(0,1,1),(1,0,0),(1,1,0),(1,0,1),\ (1,1,1)\}$ in $\mathbb{R}^3$; this is the regular 3-cube. A d-dimensional polytope has many *faces*, which

are again polytopes of various dimensions between 0 and $d-1$. The 0-dimensional faces are called *vertices*, the 1-dimensional faces are called *edges*, and the $(d-1)$-dimensional faces are called *facets*. For instance, the cube has 8 vertices, 12 edges and 6 facets. If $d = 2$ then the edges coincide with the facets. A 2-dimensional polytope is called a *polygon*.

Consider the polynomial $f(x,y)$ in (19). Each term $x^{u_i}y^{v_i}$ appearing in $f(x,y)$ can be regarded as a lattice point (u_i, v_i) in the plane $\mathbb{R}^2$. The convex hull of all these points is called the *Newton polygon* of $f(x,y)$. In symbols,

$$New(f) \quad := \quad \text{conv}\big\{\,(u_1,v_1),\,(u_2,v_2),\ldots,(u_m,v_m)\big\}$$

This is a polygon in $\mathbb{R}^2$ having at most m vertices. More generally, every polynomial in n unknowns gives rise to a *Newton polytope* in $\mathbb{R}^n$.

Our running example in this lecture is the the pair of polynomials in (20). The Newton polygon of the polynomial $g(x,y)$ is a quadrangle, and the Newton polygon of $h(x,y)$ is a triangle. If P and Q are any two polygons in the plane, then their *Minkowski sum* is the polygon

$$P+Q \quad := \quad \big\{\,p+q\,:\,p\in P,\,q\in Q\,\big\}.$$

Note that each edge of $P+Q$ is parallel to an edge of P or an edge of Q.

The geometric operation of taking the Minkowski sum of polytopes mirrors the algebraic operation of multiplying polynomials. More precisely, the Newton polytope of a product of two polynomials equals the Minkowski sum of two given Newton polytopes:

$$New(g\cdot h) \quad = \quad New(g) \,+\, New(h).$$

If P and Q are any two polygons then we define their *mixed area* as

$$\mathcal{M}(P,Q) \quad := \quad area(P+Q) \,-\, area(P) \,-\, area(Q).$$

For instance, the mixed area of the two Newton polygons in (20) equals

$$\mathcal{M}(P,Q) \;=\; \mathcal{M}(New(g),New(h)) \;=\; \frac{13}{2} \,-\, 1 - \frac{3}{2} \;=\; 4.$$

The correctness of this computation can be seen in the following diagram:

`Figure here: Mixed subdivision`

This figure shows a subdivision of $P+Q$ into five pieces: a translate of P, a translate of Q and three parallelograms. The mixed area is the sum of the areas of the three parallelograms, which is four. This number coincides with the number of common zeros of g and h. This is not an accident, but it is an instance of a general theorem due to David Bernstein (1975). We abbreviate $\mathbb{C}^* := \mathbb{C}\backslash\{0\}$. The set $(\mathbb{C}^*)^2$ of pairs (x, y) with $x \neq 0$ and $y \neq 0$ is a group under multiplication, called the *two-dimensional algebraic torus*.

Theorem 19. (Bernstein's Theorem)
If g and h are two generic bivariate polynomials, then the number of solutions of $g(x, y) = h(x, y) = 0$ in $(\mathbb{C}^*)^2$ equals the mixed area $\mathcal{M}(New(g), New(h))$.

Actually, this assertion is valid for *Laurent polynomials*, which means that the exponents in our polynomials (19) can be any integers, possibly negative. Bernstein's Theorem implies the following combinatorial fact about lattice polygons. If P and Q are lattice polygons (i.e., the vertices of P and Q have integer coordinates), then $\mathcal{M}(P, Q)$ is a non-negative integer.

We remark that Bézout's Theorem follows as a special case from Bernstein's Theorem. Namely, if g and h a general polynomials of degree d and e respectively, then their Newton polygons are the triangles

$$\begin{aligned} P \;:=\; New(g) \;&=\; \mathrm{conv}\{(0,0),(0,d),(d,0)\}\,, \\ Q \;:=\; New(h) \;&=\; \mathrm{conv}\{(0,0),(0,e),(e,0)\}\,, \\ P+Q \;:=\; New(g\cdot h) \;&=\; \mathrm{conv}\{(0,0),(0,d+e),(d+e,0)\}. \end{aligned}$$

The areas of these triangles are $d^2/2$, $e^2/2$, $(d+e)^2/2$, and hence

$$\mathcal{M}(P,Q) \quad = \quad \frac{(d+e)^2}{2} - \frac{d^2}{2} - \frac{e^2}{2} \quad = \quad d \cdot e.$$

Hence two general plane curves of degree d and e meet in $d \cdot e$ points.

We shall present a proof of Bernstein's Theorem. This proof is algorithmic in the sense that it tells us how to approximate all the zeros numerically. The steps in this proof from the foundation for the method of polyhedral homotopies for solving polynomial systems. This is an active area of research, with lots of exciting progress by work of T.Y. Li, Jan Verschelde and others.

We proceed in three steps. The first deals with an easy special case.

3.2 Zero-dimensional binomial systems

A *binomial* is a polynomial with two terms. We first prove Theorem 1.1 in the case when g and h are binomials. After multiplying or dividing both binomials by suitable scalars and powers of the variables, we may assume that our given equations are

$$g \quad = \quad x^{a_1}y^{b_1} - c_1 \qquad \text{and} \qquad h \quad = \quad x^{a_2}y^{b_2} - c_2, \tag{21}$$

where a_1, a_2, b_1, b_2 are integers (possibly negative) and c_1, c_2 are non-zero complex numbers. Note that multiplying the given equations by a (Laurent) monomial changes neither the number of zeros in $(\mathbb{C}^*)^2$ nor the mixed area of their Newton polygons

To solve the equations $g = h = 0$, we compute an invertible integer 2×2-matrix $U = (u_{ij}) \in SL_2(\mathbb{Z})$ such that

$$\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \cdot \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \quad = \quad \begin{pmatrix} r_1 & r_3 \\ 0 & r_2 \end{pmatrix}.$$

This is accomplished using the *Hermite normal form* algorithm of integer linear algebra. The invertible matrix U triangularizes our system of equations:

$$\begin{array}{cc} & g = h = 0 \\ \iff & x^{a_1}y^{b_1} = c_1 \quad \text{and} \quad x^{a_2}y^{b_2} = c_2 \\ \iff & (x^{a_1}y^{b_1})^{u_{11}}(x^{a_2}y^{b_2})^{u_{12}} = c_1^{u_{11}}c_2^{u_{12}} \text{ and } (x^{a_1}y^{b_1})^{u_{21}}(x^{a_2}y^{b_2})^{u_{22}} = c_1^{u_{21}}c_2^{u_{22}} \\ \iff & x^{r_1}y^{r_3} = c_1^{u_{11}}c_2^{u_{12}} \quad \text{and} \quad y^{r_2} = c_1^{u_{21}}c_2^{u_{22}}. \end{array}$$

This triangularized system has precisely $r_1 r_2$ distinct non-zero complex solutions. These can be expressed in terms of radicals in the coefficients c_1 and c_2. The number of solutions equals

$$r_1 r_2 \;=\; \det \begin{pmatrix} r_1 & r_3 \\ 0 & r_2 \end{pmatrix} \;=\; \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \quad = \quad \text{area}(New(g) + New(h)).$$

This equals the mixed area $\mathcal{M}(New(g), New(h))$, since the two Newton polygons are just segments, so that $area(New(g)) = area(New(h)) = 0$. This proves Bernstein's Theorem for binomials. Moreover, it gives a simple algorithm for finding all zeros in this case.

The method described here clearly works also for n binomial equations in n variables, in which case we are to compute the Hermite normal form of an

integer $n \times n$-matrix. We note that the Hermite normal form computation is similar but not identical to the computation of a lexicographic Gröbner basis. We illustrate this in `maple` for a system with $n = 3$ having 20 zeros:

```
> with(Groebner): with(linalg):
> gbasis([
>              x^3 * y^5  * z^7   -  c1,
>             x^11 * y^13 * z^17  -  c2,
>             x^19 * y^23 * z^29  -  c3],        plex(x,y,z));

    13  3     8  10    15   2  2     9  8    6  3       4  7
[-c2  c1  + c3  z  , c2   c1  y  - c3  z , c2 c1  x - c3  z  y]

> ihermite( array([
> [ 3,  5,  7  ],
> [ 11, 13, 17 ],
> [ 19, 23, 29 ] ]));
                                  [1    1     5]
                                  [             ]
                                  [0    2     2]
                                  [             ]
                                  [0    0    10]
```

3.3 Introducing a toric deformation

We introduce a new indeterminate t, and we multiply each monomial of g and each monomial of h by a power of t. What we want is the solutions to this system for $t = 1$, but what we will do instead is to analyze it for t in neighborhood of 0. For instance, our system (20) gets replaced by

$$\begin{aligned} g_t(x,y) &= a_1 t^{\nu_1} + a_2 x t^{\nu_2} + a_3 xy t^{\nu_3} + a_4 y t^{\nu_4} \\ h_t(x,y) &= b_1 t^{\omega_1} + b_2 x^2 y t^{\omega_2} + b_3 x y^2 t^{\omega_3} \end{aligned}$$

We require that the integers ν_i and ω_j are "sufficiently generic" in a sense to be made precise below. The system $g_t = h_t = 0$ can be interpreted as a bivariate system which depends on a parameter t. Its zeros $(x(t), y(t))$ depend on that parameter. They define the branches of an *algebraic function* $t \mapsto (x(t), y(t))$. Our goal is to identify the branches.

In a neighborhood of the origin in the complex plane, each branch of our algebraic function can be written as follows:

$$\begin{aligned} x(t) &= x_0 \cdot t^u + \text{higher order terms in } t, \\ y(t) &= y_0 \cdot t^v + \text{higher order terms in } t, \end{aligned}$$

where x_0, y_0 are non-zero complex numbers and u, v are rational numbers. To determine the exponents u and v we substitute $x = x(t)$ and $y = y(t)$ into the equations $g_t(x, y) = h_t(x, y) = 0$. In our example this gives

$$\begin{aligned} g_t\big(x(t), y(t)\big) &= a_1 t^{\nu_1} + a_2 x_0 t^{u+\nu_2} + a_3 x_0 y_0 t^{u+v+\nu_3} + a_4 y_0 t^{v+\nu_4} + \cdots, \\ h_t\big(x(t), y(t)\big) &= b_1 t^{\omega_1} + b_2 x_0^2 y_0 t^{2u+v+\omega_2} + b_3 x_0 y_0^2 t^{u+2v+\omega_3} + \cdots. \end{aligned}$$

In order for (1.6) to be a root the term of lowest order must vanish. Since x_0 and y_0 are chosen to be non-zero, this is possible only if the lowest order in t is attained by at least two different terms. This implies the following two piecewise-linear equations for the indeterminate vector $(u, v) \in \mathbb{Q}^2$:

$$\begin{aligned} &\min\big\{\, \nu_1,\, u+\nu_2,\, u+v+\nu_3,\, v+\nu_4 \,\big\} \quad \text{is attained twice} \\ &\min\big\{\, \omega_1,\, 2u+v+\omega_2,\, u+2v+\omega_3 \,\big\} \quad \text{is attained twice.} \end{aligned}$$

As in Lecture 1, each of these translates into a disjunction of linear equations and inequalities. For instance, the second "min-equation" translates into

$$\begin{aligned} & \omega_1 = 2u+v+\omega_2 \geq u+2v+\omega_3 \\ \text{or} \quad & \omega_1 = u+2v+\omega_3 \geq 2u+v+\omega_2 \\ \text{or} \quad & 2u+v+\omega_2 = u+2v+\omega_3 \geq \omega_1 \end{aligned}$$

It is now easy to state what we mean by the ν_i and ω_j being *sufficiently generic*. It means that "Min" is attained twice but not thrice. More precisely, at every solution (u, v) of the two piecewise-linear equations, precisely two of the linear forms attain the minimum value in each of the two equations.

One issue in the algorithm for Bernstein's Theorem is to chose powers of t that are small but yet generic. In our example, the choice $\nu_1 = \nu_2 = \nu_3 = \nu_4 = \omega_3 = 0$, $\omega_1 = \omega_2 = 1$ is generic. Here the two polynomial equations are

$$g_t(x, y) = a_1 + a_2 x + a_3 xy + a_4 y, \qquad h_t(x, y) = b_1 t + b_2 x^2 y t + b_3 x y^2,$$

and the corresponding two piecewise-linear equations are

$$\min\big\{\, 0,\, u,\, u+v,\, v \,\big\} \quad \text{and} \quad \min\big\{\, 1,\, 2u+v+1,\, u+2v \,\big\} \quad \text{are attained twice.}$$

This system has precisely three solutions:

$$(u,v) \quad \in \quad \big\{\,(1,0),\,(0,1/2),\,(-1,0)\,\big\}.$$

For each of these pairs (u,v), we now obtain a binomial system $g'(x_0,y_0) = h'(x_0,y_0)$ which expresses the fact that the lowest terms in $g_t\big(x(t),y(t)\big)$ and $h_t\big(x(t),y(t)\big)$ do indeed vanish. The three binomial systems are

- $g'(x_0,y_0) = a_1 + a_4 y_0$ and $h'(x_0,y_0) = b_1 + b_3 x_0 y_0^2$ for $(u,v) = (1,0)$.
- $g'(x_0,y_0) = a_1 + a_2 x_0$ and $h'(x_0,y_0) = b_1 + b_3 x_0 y_0^2$ for $(u,v) = (0,1/2)$.
- $g'(x_0,y_0) = a_2 x_0 + a_3 x_0 y_0$ and $h'(x_0,y_0) = b_2 x_0^2 y_0 + b_3 x_0 y_0^2$ for $(u,v) = (-1,0)$.

These binomial systems have one, two and one root respectively. For instance, the unique Puiseux series solution for $(u,v) = (1,0)$ has

$$x_0 \,=\, -a_4^2 b_1 / a_1^2 b_3 \qquad \text{and} \qquad y_0 \,=\, -a_1/a_4.$$

Hence our algebraic function has a total number of four branches. If one wishes more information about the four branches, one can now compute further terms in the Puiseux expansions of these branches. For instance,

$$\begin{aligned} x(t) \quad &= \quad -\tfrac{a_4^2 b_1}{a_1^2 b_3}\cdot t \quad + \quad 2\cdot \tfrac{a_4^3 b_1^2 (a_1 a_3 - a_2 a_4)}{a_1^5 b_3^2}\cdot t^2 \\ &\qquad + \quad \tfrac{a_4^4 b_1^2 (a_1^3 a_4 b_2 - 5 a_1^2 a_3^2 b_1 + 12 a_1 a_2 a_3 a_4 b_1 - 7 a_2^2 a_4^2 b_1)}{a_1^8 b_3^8}\cdot t^3 \quad + \quad \cdots \\ y(t) \quad &= \quad -\tfrac{a_1}{a_4} \quad + \quad \tfrac{b_1(a_1 a_3 - a_2 a_4)}{a_1^2 b_3}\cdot t \quad + \quad \tfrac{a_4 b_1^2 (a_1 a_3 - a_2 a_4)(a_1 a_3 - 2 a_2 a_4)}{a_1^5 b_3^2}\cdot t^2 \,+\, \cdots. \end{aligned}$$

For details on computing multivariate Puiseux series see (McDonald 1995).

3.4 Mixed subdivisions of Newton polytopes

We fix a generic toric deformation $g_t = h_t = 0$ of our equations. In this section we introduce a polyhedral technique for solving the associated piecewise linear equation and, in order to prove Bernstein's Theorem, we show that the total number of branches equals the mixed area of the Newton polygons.

Let us now think of g_t and h_t as Laurent polynomials in three variables (x,y,t) whose zero set is a curve in $(\mathbb{C}^*)^3$. The *Newton polytopes* of these trivariate polynomials are the following two polytopes in $\mathbb{R}^3$:

$$\begin{aligned} P \quad &:= \quad \operatorname{conv}\big\{(0,0,\nu_1),(1,0,\nu_2),(1,1,\nu_3),(0,1,\nu_4)\big\} \\ \text{and} \quad Q \quad &:= \quad \operatorname{conv}\big\{(0,0,\omega_1),(2,1,\omega_2),(1,2,\omega_3)\big\}. \end{aligned}$$

The Minkowski sum $P+Q$ is a polytope in $\mathbb{R}^3$. By a *facet* of $P+Q$ we mean a two-dimensional face. A facet F of $P+Q$ is a *lower facet* if there is a vector $(u,v) \in \mathbb{R}^2$ such that $(u,v,1)$ is an inward pointing normal vector to $P+Q$ at F. Our genericity conditions for the integers ν_i and ω_j is equivalent to:

(1) The Minkowski sum $P+Q$ is a 3-dimensional polytope.

(2) Every lower facet of $P+Q$ has the form F_1+F_2 where either

   (a) F_1 is a vertex of P and F_2 is a facet of Q, or

   (b) F_1 is an edge of P and F_2 is an edge of Q, or

   (c) F_1 is a facet of P and F_2 is a vertex of Q.

As an example consider our lifting from before, $\nu_1 = \nu_2 = \nu_3 = \nu_4 = \omega_3 = 0$ and $\omega_1 = \omega_2 = 1$. It meets the requirements (1) and (2). The polytope P is a quadrangle and Q is triangle. But they lie in non-parallel planes in $\mathbb{R}^3$. Their Minkowski sum $P+Q$ is a 3-dimensional polytope with 10 vertices:

`Figure here: The 3-dimensional polytope P+Q`

The union of all lower facets of $P+Q$ is called the *lower hull* of the polytope $P+Q$. Algebraically speaking, the lower hull is the subset of all points in $P+Q$ at which some linear functional of the form $(x_1,x_2,x_3) \mapsto ux_1+vx_2+x_3$ attains its minimum. Geometrically speaking, the lower hull is that part of the boundary of $P+Q$ which is visible from below. Let $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ denote the projection onto the first two coordinates. Then

$$\pi(P) = New(g), \;\; \pi(Q) = New(h), \quad \text{and} \quad \pi(P+Q) = New(g)+New(h).$$

The map π restricts to a bijection from the lower hull onto $New(g)+New(h)$. The set of polygons $\Delta := \{\pi(F) : F \text{ lower facet of } P+Q\}$ defines a subdivision of $New(g)+New(h)$. A subdivision Δ constructed by this process, for some choice of ν_i and ω_j, is called a *mixed subdivision* of the given Newton polygons. The polygons $\pi(F)$ are the *cells* of the mixed subdivision Δ.

Every cell of a mixed subdivision Δ has the form F_1+F_2 where either

(a) $F_1 = \{(u_i,v_i)\}$ where $x^{u_i}y^{v_i}$ appears in g and F_2 is the projection of a facet of Q, or

(b) F_1 is the projection of an edge of P and F_2 is the projection of an edge of Q, or

(c) F_1 is the projection of a facet of P and $F_2 = \{(u_i, v_i)\}$ where $x^{u_i}y^{v_i}$ appears in h.

The cells of type (b) are called the *mixed cells* of Δ.

Lemma 20. *Let Δ be any mixed subdivision for g and h. Then the sum of the areas of the mixed cells in Δ equals the mixed area $\mathcal{M}(New(g), New(h))$.*

Proof. Let γ and δ be arbitrary positive reals and consider the polytope $\gamma P + \delta Q$ in $\mathbb{R}^3$. Its projection into the plane $\mathbb{R}^2$ equals

$$\pi(\gamma P + \delta Q) \quad = \quad \gamma\pi(P) + \delta\pi(Q) \quad = \quad \gamma \cdot New(g) \,+\, \delta \cdot New(h).$$

Let $A(\gamma, \delta)$ denote the area of this polygon. This polygon can be subdivided into cells $\gamma F_1 + \delta F_2$ where $F_1 + F_2$ runs over all cells of Δ. Note that $area(\gamma F_1 + \delta F_2)$ equals $\delta^2 \cdot area(F_1 + F_2)$ if $F_1 + F_2$ is a cell of type (a), $\gamma\delta \cdot area(F_1 + F_2)$ if it is a mixed cell, and $\gamma^2 \cdot area(F_1 + F_2)$ if is has type (c). The sum of these areas equals $A(\gamma, \delta)$. Therefore $A(\gamma, \delta) = A_{(a)} \cdot \delta^2 + A_{(b)} \cdot \gamma\delta + A_{(c)} \cdot \gamma^2$, where $A_{(b)}$ is the sum of the areas of the mixed cells in Δ. We conclude $A_{(b)} = A(1,1) - A(1,0) - A(0,1) = \mathcal{M}(New(g), New(h))$. □

The following lemma makes the connection with the previous section.

Lemma 21. *A pair $(u, v) \in \mathbb{Q}^2$ solves the piecewise-linear min-equations if and only if $(u, v, 1)$ is the normal vector to a mixed lower facet of $P + Q$.*

This implies that the valid choices of (u, v) are in bijection with the mixed cells in the mixed subdivision Δ. Each mixed cell of Δ is expressed uniquely as the Minkowski sum of a Newton segment $New(g')$ and a Newton segment $New(h')$, where g' is a binomial consisting of two terms of g, and h' is a binomial consisting of two terms of h. Thus each mixed cell in Δ can be identified with a system of two binomial equations $g'(x, y) = h'(x, y) = 0$. In this situation we can rewrite our system as follows:

$$\begin{aligned} g_t(x(t), y(t)) \quad &= \quad g'(x_0, y_0) \cdot t^a \;+ \text{ higher order terms in } t, \\ h_t(x(t), y(t)) \quad &= \quad h'(x_0, y_0) \cdot t^b \;+ \text{ higher order terms in } t, \end{aligned}$$

where a and b suitable rational numbers. This implies the following lemma.

Lemma 22. *Let (u, v) as in Lemma 21. The corresponding choices of $(x_0, y_0) \in (\mathbb{C}^*)^2$ are the solutions of the binomial system $g'(x_0, y_0) = h'(x_0, y_0) = 0$.*

We are now prepared to complete the proof of Bernstein's Theorem. This is done by showing that the equations $g_t(x, y) = h_t(x, y) = 0$ have $\mathcal{M}(\text{New}(g), \text{New}(h))$ many distinct isolated solutions in $(K^*)^2$ where $K = \mathbb{C}(\!(t)\!)$ is the algebraically closed field of Puiseux series.

By Section 3.2, the number of roots $(x_0, y_0) \in (\mathbb{C}^*)^2$ of the binomial system in Lemma 22 coincides with the area of the mixed cell $\text{New}(g') + \text{New}(h')$. Each of these roots provides the leading coefficients in a Puiseux series solution $(x(t), y(t))$ to our equations. Conversely, by Lemma 21 every series solution arises from some mixed cell of Δ. We conclude that the number of series solutions equals the sum of these areas over all mixed cells in Δ. By Lemma 20, this quantity coincides with the mixed area $\mathcal{M}(\text{New}(f), \text{New}(g))$. General facts from algebraic geometry guarantee that the same number of roots is attained for almost all choices of coefficients, and that we can descend from the field K to the complex numbers $\mathbb{C}$ under the substitution $t = 1$. $\square$

Our proof of Bernstein's Theorem gives rise to a numerical algorithm for finding of all roots of a sparse system of polynomial equations. This algorithm belongs to the general class of *numerical continuation* methods (Allgower & Georg, 1990), which are sometimes also called *homotopy methods* (Drexler 1978). The idea is to trace each of the branches of the algebraic curve $(x(t), y(t))$ between $t = 0$ and $t = 1$. We have shown that the number of branches equals the mixed area. Our constructions give sufficient information about the Puiseux series so that we can approximate $(x(t), y(t))$ for any t in a small neighborhood of zero. Using numerical continuation, it is now possible to approximate $(x(1), y(1))$. We return to this topic in a later lecture.

3.5 Khovanskii's Theorem on Fewnomials

Polynomial equations arise in many mathematical models in science and engineering. In such applications one is typically interested in solutions over the real numbers $\mathbb{R}$ instead of the complex numbers $\mathbb{C}$. This study of real roots of polynomial systems is considerably more difficult than the study of complex roots. Even the most basic questions remain unanswered to-date. Let us start out with a very concrete such question:

Question 23. *What is the maximum number of isolated real roots of any system of two polynomial equations in two variables each having four terms ?*

The polynomial equations considered here look like

$$\begin{aligned} f(x,y) &= a_1 x^{u_1} y^{v_1} + a_2 x^{u_2} y^{v_2} + a_3 x^{u_3} y^{v_3} + a_4 x^{u_4} y^{v_4}, \\ g(x,y) &= b_1 x^{\tilde{u}_1} y^{\tilde{v}_1} + b_2 x^{\tilde{u}_2} y^{\tilde{v}_2} + b_3 x^{\tilde{u}_3} y^{\tilde{v}_3} + b_4 x^{\tilde{u}_4} y^{\tilde{v}_4}. \end{aligned}$$

where a_i, b_j are arbitrary real numbers and $u_i, v_j, \tilde{u}_i, \tilde{v}_j$ are arbitrary integers. To stay consistent with our earlier discussion, we shall count only solutions (x, y) in $(\mathbb{R}^*)^2$, that is, we require that both x and y are non-zero reals.

There is an obvious lower bound for the number Question 23: *thirty-six*. It is easy to write down a system of the above form that has 36 real roots:

$$f(x) = (x^2-1)(x^2-2)(x^2-3) \quad \text{and} \quad g(y) = (y^2-1)(y^2-2)(y^2-3).$$

Each of the polynomials f and g depends on one variable only, and it has 6 non-zero real roots in that variable. Therefore the system $f(x) = g(y) = 0$ has 36 distinct isolated roots in $(\mathbb{R}^*)^2$. Note also that the expansions of f and g have exactly four terms each, as required.

A priori it is not clear whether Question 23 even makes sense: why should such a maximum exist ? It certainly does not exist if we consider complex zeros, because one can get arbitrarily many complex zeros by increasing the degrees of the equations. The point is that such an unbounded increase of roots is impossible over the real numbers. This was proved by Khovanskii (1980). He found a bound on the number of real roots which does not depend on the degrees of the given equations. We state the version for positive roots.

Theorem 24. (Khovanskii's Theorem) *Consider n polynomials in n variables involving m distinct monomials in total. The number of isolated roots in the positive orthant $(\mathbb{R}_+)^n$ of any such system is at most $2^{\binom{m}{2}} \cdot (n+1)^m$.*

The basic idea behind the proof of Khovanskii's Theorem is to establish the following more general result. We consider systems of n equations which can be expressed as polynomial functions in at most m monomials in $\mathbf{x} = (x_1, \ldots, x_n)$. If we abbreviate the i-th such monomial by $\mathbf{x}^{\mathbf{a}_i} := x_1^{a_{i1}} x_2^{a_{i2}} \cdots x_n^{a_{in}}$, then we can write our n polynomials as

$$F_i\big(\mathbf{x}^{\mathbf{a}_1}, \mathbf{x}^{\mathbf{a}_2}, \ldots, \mathbf{x}^{\mathbf{a}_m}\big) = 0 \qquad (i = 1, 2, \ldots, n)$$

We claim that the number of real zeros in the positive orthant is at most

$$2^{\binom{m}{2}} \cdot \big(1 + \sum_{i=1}^{n} deg(F_i)\big)^m \cdot \prod_{i=1}^{d} deg(F_i).$$

Theorem 2.3 concerns the case where $deg(F_i) = 1$ for all i.

We proceed by induction on $m - n$. If $m = n$ then (2.3) is expressed in n monomials in n unknowns. By a multiplicative change of variables

$$x_i \quad \mapsto \quad z_1^{u_{i1}} z_2^{u_{i2}} \cdots z_n^{u_{in}}$$

we can transform our d monomials into the n coordinate functions $z_1, \ldots, z_n$. (Here the u_{ij} can be rational numbers, since all roots under consideration are positive reals.) Our assertion follows from Bezout's Theorem, which states that the number of isolated complex roots is at most the product of the degrees of the equations.

Now suppose $m > n$. We introduce a new variable t, and we multiply one of the given monomials by t. For instance, we may do this to the first monomial and set

$$G_i(t, x_1, \ldots, x_n) \quad := \quad F_i\left(\mathbf{x}^{\mathbf{a}_1} \cdot t\,,\, \mathbf{x}^{\mathbf{a}_2}\,,\, \ldots\,,\, \mathbf{x}^{\mathbf{a}_m}\right) \qquad (i = 1, 2, \ldots, n)$$

This is a system of equations in $\mathbf{x}$ depending on the parameter t. We study the behavior of its positive real roots as t moves from 0 to 1. At $t = 0$ we have a system involving one monomial less, so the induction hypothesis provides a bound on the number of roots. Along our trail from 0 to 1 we encounter some bifurcation points at which two new roots are born. Hence the number of roots at $t = 1$ is at most twice the number of bifurcation points plus the number of roots of $t = 0$.

Each bifurcation point corresponds to a root $(\mathbf{x}, t)$ of the augmented system

$$J(t, \mathbf{x}) \;=\; G_1(t, \mathbf{x}) \;=\; \cdots \;=\; G_n(t, \mathbf{x}) \;=\; 0, \qquad (2.4)$$

where $J(t, \mathbf{x})$ denotes the *toric Jacobian*:

$$J(t, x_1, \ldots, x_m) \quad = \quad det\left(x_i \cdot \frac{\partial}{\partial x_j} G_j(t, \mathbf{x})\right)_{1 \le i,j \le m}.$$

Now, the punch line is that each of the $n + 1$ equations in (2.4) – including the Jacobian – can be expressed in terms of only m monomials

$\mathbf{x}^{\mathbf{a}_1} \cdot t$, $\mathbf{x}^{\mathbf{a}_2}$, $\cdots$, $\mathbf{x}^{\mathbf{a}_m}$. Therefore we can bound the number of bifurcation points by the induction hypothesis, and we are done.

This was only to give the flavor of how Theorem 2.3 is proved. There are combinatorial and topological fine points which need most careful attention. The reader will find the complete proof in (Khovanskii 1980), in (Khovanskii 1991) or in (Benedetti & Risler 1990).

Khovanskii's Theorem implies an upper bound for the root count suggested in Question 23. After multiplying one of the given equations by a suitable monomial, we may assume that our system has seven distinct monomials. Substituting $n = 2$ and $m = 7$ into Khovanskii's formula, we see that there are at most $2^{\binom{7}{2}} \cdot (2+1)^7 = 4,586,471,424$ roots in the positive quadrant. By summing over all four quadrants, we conclude that the maximum in Question 23 lies between 36 and $18,345,885,696 = 2^2 \cdot 2^{\binom{7}{2}} \cdot (2+1)^7$. The gap between 36 and 18,345,885,696 is frustratingly large. Experts agree that the truth should be closer to the lower bound than to the upper bound, but at the moment nobody knows the exact value. Could it be 36 ?

The original motivation for Khovanskii's work was the following conjecture from the 1970's due to Kouchnirenko. *Consider any system of n polynomial equations in n unknown, where the i-th equation has at most m_i terms. The number of isolated real roots in $(\mathbb{R}_+)^n$ of such a system is at most* $(m_1-1)(m_2-1)\cdots(m_d-1)$. This number is attained by equations in distinct variables, as was demonstrated by our example with $d = 2, m_1 = m_2 = 4$ which has $(m_1 - 1)(m_2 - 1) = 16$ real zeros.

Remarkably, Kouchnirenko's conjecture remained open for many years after Khovanskii had developed his theory of fewnomials which includes the above theorem. Only two years ago, Bertrand Haas (2000) found the following counterexample to Kouchnirenko's conjecture in the case $d = 2, m_1 = m_2 = 4$. Proving the following proposition from scratch is a nice challenge.

Proposition 25. (Haas) *The two equations*

$$x^{108} + 1.1y^{54} - 1.1y \quad = \quad y^{108} + 1.1x^{54} - 1.1x \quad = \quad 0$$

have five distinct strictly positive solutions $(x, y) \in (\mathbb{R}_+)^2$.

It was proved by Li, Rojas and Wang (2001) that the lower bound provided by Haas' example coincides with the upper bound for two trinomials.

Theorem 26. (Li, Rojas and Wang) *A system of two trinomials*

$$\begin{aligned} f(x,y) &= a_1x^{u_1}y^{v_1} + a_2x^{u_2}y^{v_2} + a_3x^{u_3}y^{v_3}, \\ g(x,y) &= b_1x^{\tilde{u}_1}y^{\tilde{v}_1} + b_2x^{\tilde{u}_2}y^{\tilde{v}_2} + b_3x^{\tilde{u}_3}y^{\tilde{v}_3}, \end{aligned}$$

with $a_i, b_j \in \mathbb{R}$ *and* $u_i, v_j, \tilde{u}_i, \tilde{v}_j \in \mathbb{Z}$ *has at most five positive real zeros.*

3.6 Exercises

(1) Consider the intersection of a general conic and a general cubic curve

$$\begin{aligned} a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + a_6 &= 0 \\ b_1x^3+b_2x^2y+b_3xy^2+b_4y^3+b_5x^2+b_6xy+b_7y^2+b_8x+b_9y+b_{10} &= 0 \end{aligned}$$

Compute an explicit polynomial in the unknowns a_i, b_j such that equations have six distinct solutions whenever your polynomial is non-zero.

(2) Draw the Newton polytope of the following polynomial

$$f(x_1,x_2,x_3,x_4) = (x_1-x_2)(x_1-x_3)(x_1-x_4)(x_2-x_3)(x_2-x_4)(x_3-x_4).$$

(3) For general $\alpha_i, \beta_j \in \mathbb{Q}$, how many vectors $(x,y) \in (\mathbb{C}^*)^2$ satisfy

$$\alpha_1x^3y + \alpha_2xy^3 = \alpha_3x + \alpha_4y \quad \text{and} \quad \beta_1x^2y^2 + \beta_2xy = \beta_3x^2 + \beta_4y^2 \ ?$$

Can your bound be attained with all real vectors $(x,y) \in (\mathbb{R}^*)^2$?

(4) Find the first three terms in each of the four Puiseux series solutions $(x(t), y(t))$ of the two equations

$$\begin{aligned} t^2x^2 + t^5xy + t^{11}y^2 + t^{17}x + t^{23}y + t^{31} &= 0 \\ t^3x^2 + t^7xy + t^{13}y^2 + t^{19}x + t^{29}y + t^{37} &= 0 \end{aligned}$$

(5) State and prove Bernstein's Theorem for n equations in n variables.

(6) Bernstein's Theorem can be used in reverse, namely, we can calculate the mixed volume of n polytopes by counting the number of zeros in $(\mathbb{C}^*)^n$ of a sparse system of polynomial equations. Pick your favorite three distinct three-dimensional lattice polytopes in $\mathbb{R}^3$ and compute their mixed volume with this method using `Macaulay 2`.

(7) Show that Kouchnirenko's Conjecture is true for $d = 2$ and $m_1 = 2$.

(8) Prove Proposition 25. Please use any computer program of your choice.

(9) Can Haas' example be modified to show that the answer to Question 23 is strictly larger than 36 ?

4 Resultants

Elimination theory deals with the problem of eliminating one or more variables from a system of polynomial equations, thus reducing the given problem to a smaller problem in fewer variables. For instance, if we wish to solve

$$a_0 + a_1 x + a_2 x^2 \quad = \quad b_0 + b_1 x + b_2 x^2 \quad = \quad 0,$$

with $a_2 \neq 0$ and $b_2 \neq 0$ then we can eliminate the variable x to get

$$a_0^2 b_2^2 - a_0 a_1 b_1 b_2 - 2 a_0 a_2 b_0 b_2 + a_0 a_2 b_1^2 + a_1^2 b_0 b_2 - a_1 a_2 b_0 b_1 + a_2^2 b_0^2 \;=\; 0. \quad (22)$$

This polynomial of degree 4 is the *resultant*. It vanishes which vanishes if and only if the given quadratic polynomials have a common complex root x. The resultant (22) has the following three determinantal representations:

$$\begin{vmatrix} a_0 & a_1 & a_2 & 0 \\ 0 & a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 & 0 \\ 0 & b_0 & b_1 & b_2 \end{vmatrix} \quad = \quad - \begin{vmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \\ [01] & [02] & 0 \end{vmatrix} \quad = \quad - \begin{vmatrix} [01] & [02] \\ [02] & [12] \end{vmatrix} \quad (23)$$

where $[ij] = a_i b_j - a_j b_i$. Our aim in this section is to discuss such formulas.

The computation of resultants is an important tool for solving polynomial systems. It is particularly well suited for eliminating all but one variable from a system of n polynomials in n unknowns which has finitely many solutions.

4.1 The univariate resultant

Consider two general polynomials in one variable of degrees d and e:

$$\begin{aligned} f \quad &= \quad a_0 + a_1 x + a_2 x^2 + \cdots + a_{d-1} x^{d-1} + a_d x^d, \\ g \quad &= \quad b_0 + b_1 x + b_2 x^2 + \cdots + b_{e-1} x^{e-1} + b_e x^e. \end{aligned}$$

Theorem 27. *There exists a unique (up to sign) irreducible polynomial* Res *in* $\mathbb{Z}[a_0, a_1, \ldots, a_d, b_0, b_1, \ldots, b_d]$ *which vanishes whenever the polynomials* $f(x)$ *and* $g(x)$ *have a common zero.*

Here and throughout this section "common zeros" may lie in any algebraically closed field (say, $\mathbb{C}$) which contains the field to which we specialize the coefficients a_i and b_j of the given polynomials (say, $\mathbb{Q}$). Note that a polynomial with integer coefficients being "irreducible" implies that the coefficients are relatively prime. The resultant $\mathrm{Res} = \mathrm{Res}_x(f, g)$ can be expressed as the determinant of the *Sylvester matrix*

$$\mathrm{Res}_x(f,g) \quad = \quad \begin{vmatrix} a_0 & & & & b_0 & & & \\ a_1 & a_0 & & & b_1 & b_0 & & \\ & a_1 & \ddots & & & b_1 & \ddots & \\ \vdots & & \ddots & a_0 & \vdots & & \ddots & b_0 \\ & \vdots & & a_1 & & \vdots & & b_1 \\ a_d & & & & b_e & & & \\ & a_d & & \vdots & & b_e & & \vdots \\ & & \ddots & & & & \ddots & \\ & & & a_d & & & & b_e \end{vmatrix} \tag{24}$$

where the blank spaces are filled with zeroes. See the left formula in (22).

There are many other useful formulas for the resultant. For instance, suppose that the roots of f are $\xi_1, \ldots, \xi_d$ and the roots of g are $\eta_1, \ldots, \eta_e$. Then we have the following product formulas:

$$\mathrm{Res}_x(f,g) \;=\; a_0^e b_0^d \prod_{i=1}^{d}\prod_{j=1}^{e}(\xi_i - \eta_j) \;=\; a_0^e \prod_{i=1}^{d} g(\xi_i) \;=\; (-1)^{de} b_0^d \prod_{j=1}^{e} f(\eta_j).$$

¿From this we conclude the following proposition.

Proposition 28. *If* C_f *and* C_g *are the companion matrices of* f *and* g *then*

$$\mathrm{Res}_x(f,g) \quad = \quad a_0^e \cdot \det\big(g(C_f)\big) \;=\; (-1)^{de} b_0^d \cdot \det\big(f(C_g)\big).$$

If f and g are polynomials of the same degree $d = e$, then the following method for computing the resultant is often used in practise. Compute the

following polynomial in two variables, which is called the *Bézoutian*

$$B(x,y) \quad = \quad \frac{f(x)g(y)-f(y)g(x)}{x-y} \quad = \quad \sum_{i,j=0}^{d-1} c_{ij} x^i y^j.$$

Form the symmetric $d \times d$-matrix $C = (c_{ij})$. Its entries c_{ij} are sums of brackets $[kl] = a_k b_l - a_l b_k$. The case $d = 2$ appears in (22) on the right.

Theorem 29. (Bézout resultant) *The determinant of C equals $\pm \mathrm{Res}_x(f,g)$.*

Proof. The resultant $Res_x(f,g)$ is an irreducible polynomial of degree $2d$ in $a_0, \ldots, a_d, b_0, \ldots, b_d$. The determinant of C is also a polynomial of degree $2d$. We will show that the zero set of $Res_x(f,g)$ is contained in the zero set of $det(C)$. This implies that the two polynomials are equal up to a constant. Looking at leading terms one finds the constant to be either 1 or -1.

If $(a_0, \ldots, a_d, b_0, \ldots, b_d)$ is in the zero set of $Res_x(f,g)$ then the system $f = g = 0$ has a complex solution x_0. Then $B(x_0, y)$ is identically zero as a polynomial in y. This implies that the non-zero complex vector $(1, x_0, x_0^2, \ldots, x_0^{m-1})$ lies in the kernel of C, and therefore $det(C) = 0$. □

The 3×3-determinants in the middle of (22) shows that one can also use mixtures of Bézout matrices and Sylvester matrices. Such hybrid formulas for resultants are very important in higher-dimensional problems as we shall see below. Let us first show three simple applications of the univariate resultant.

Example. *(Intersecting two algebraic curves in the real plane)*
Consider two polynomials in two variables, say,

$$f \;=\; x^4 + y^4 - 1 \quad \text{and} \quad g \;=\; x^5 y^2 - 4x^3 y^3 + x^2 y^5 - 1.$$

We wish to compute the intersection of the curves $\{f = 0\}$ and $\{g = 0\}$ in the real plane $\mathbb{R}^2$, that is, all points $(x,y) \in \mathbb{R}^2$ with $f(x,y) = g(x,y) = 0$. To this end we evaluate the resultant with respect to one of the variables,

$$\begin{aligned} \mathrm{Res}_x(f,g) \quad = \quad & 2y^{28} - 16y^{27} + 32y^{26} + 249y^{24} + 48y^{23} - 128y^{22} + 4y^{21} \\ & -757y^{20} - 112y^{19} + 192y^{18} - 12y^{17} + 758y^{16} + 144y^{15} - 126y^{14} \\ & +28y^{13} - 251y^{12} - 64y^{11} + 30y^{10} - 36y^9 - y^8 + 16y^5 + 1. \end{aligned}$$

This is an irreducible polynomial in $\mathbb{Q}[y]$. It has precisely four real roots

$$y = -0.9242097, \quad y = -0.5974290, \quad y = 0.7211134, \quad y = 0.9665063.$$

Hence the two curves have four intersection points, with these y-coordinates. By the symmetry in f and g, the same values are also the possible x-coordinates. By trying out (numerically) all 16 conceivable x-y-combinations, we find that the following four pairs are the real solutions to our equations:

$$(x,y) = (-0.9242, 0.7211), \quad (x,y) = (0.7211, -0.9242),$$
$$(x,y) = (-0.5974, 0.9665), \quad (x,y) = (0.9665, -0.5974).$$

Example. *(Implicitization of a rational curve in the plane)*
Consider a plane curve which is given to us parametrically:

$$\mathcal{C} \quad = \quad \left\{ \left(\frac{a(t)}{b(t)}, \frac{c(t)}{d(t)}\right) \in \mathbb{R}^2 \quad : \quad t \in \mathbb{R} \right\},$$

where $a(t), b(t), c(t), d(t)$ are polynomials in $\mathbb{Q}[t]$. The goal is to find the unique irreducible polynomial $f \in \mathbb{Q}[x,y]$ which vanishes on $\mathcal{C}$. We may find f by the general Gröbner basis approach explained in (Cox, Little & O'Shea). It is more efficient, however, to use the following formula:

$$f(x,y) \quad = \quad \mathrm{Res}_t\big(b(t) \cdot x - a(t),\, d(t) \cdot y - c(t) \big).$$

Here is an explicit example in `maple` of a rational curve of degree six:

```
> a := t^3 - 1: b := t^2 - 5:
> c := t^4 - 3: d := t^3 - 7:
> f := resultant(b*x-a,d*y-c,t);
                                          2        2          2
f := 26 - 16 x - 162 y + 18 x y + 36 x  - 704 x  y + 324 y

                    2         2  2        3
          + 378 x y   + 870 x  y  - 226 x  y

           3        4        3  2        4           3
    + 440 x  - 484 x  + 758 x  y  - 308 x  y - 540 x y

          2  3       3  3       4  2        3
     - 450 x  y  - 76 x  y  + 76 x  y  - 216 y
```

Example. *(Computation with algebraic numbers)*
Let α and β be algebraic numbers over $\mathbb{Q}$. They are represented by their

minimal polynomials $f, g \in \mathbb{Q}[x]$. These are the unique (up to scaling) irreducible polynomials satisfying $f(\alpha) = 0$ and $g(\beta) = 0$. Our problem is to find the minimal polynomials p and q for their sum $\alpha + \beta$ and their product $\alpha \cdot \beta$ respectively. The answer is given by the following two formulas

$$p(z) \;=\; \mathrm{Res}_x\big(\, f(x),\, g(z-x)\,\big) \quad \text{and} \quad q(z) \;=\; \mathrm{Res}_x\big(\, f(x),\, g(z/x)\cdot x^{deg(g)}\,\big).$$

It is easy to check the identities $p(\alpha+\beta) = 0$ and $q(\alpha\cdot\beta) = 0$. As an example we consider two algebraic numbers given in terms of radicals:

$$\alpha \;=\; \sqrt[5]{2}, \qquad \beta \;=\; \sqrt[3]{-7/2 - 1/18\sqrt{3981}} \;+\; \sqrt[3]{-7/2 + 1/18\sqrt{3981}}.$$

Their minimal polynomials are $\alpha^5 - 2$ and $\beta^3+\beta+7$ respectively. Using the above formulas, we find that the minimal polynomial for their sum $\alpha+\beta$ is

$$\begin{aligned} p(z) \;\;=\;\; & z^{15} + 5\,z^{13} + 35\,z^{12} + 10\,z^{11} + 134\,z^{10} + 500\,z^9 + 240\,z^8 + 2735\,z^7 \\ & +3530z^6 + 1273z^5 - 6355z^4 + 12695z^3 + 1320z^2 + 22405z + 16167, \end{aligned}$$

and the minimal polynomial for their product $\alpha\cdot\beta$ equals

$$q(z) \;\;=\;\; z^{15} - 70\,z^{10} + 984\,z^5 + 134456.$$

4.2 The classical multivariate resultant

Consider a system of n homogeneous polynomials in n indeterminates

$$f_1(x_1,\ldots,x_n) \;\;=\;\; \cdots \;\;=\;\; f_n(x_1,\ldots,x_n) \;\;=\;\; 0. \tag{25}$$

We assume that the i-th equation is homogeneous of degree $d_i > 0$, that is,

$$f_i \;\;=\;\; \sum_{j_1+\cdots+j_n=d_i} c^{(i)}_{j_1,\ldots,j_n}\, x_1^{j_1}\cdots x_n^{j_n},$$

where the sum is over all $\binom{n+d_i-1}{d_i}$ monomials of degree d_i in $x_1,\ldots,x_n$. Note that the zero vector $(0,0,\ldots,0)$ is always a solution of (25). Our question is to determine under which condition there is a non-zero solution. As a simple example we consider the case of linear equations $(n = 3, d_1 = d_2 = d_3 = 1)$:

$$\begin{array}{rcccccc} f_1 & = & c^1_{100}x_1 & + & c^1_{010}x_2 & + & c^1_{001}x_3 \;\;=\;\; 0 \\ f_2 & = & c^2_{100}x_1 & + & c^2_{010}x_2 & + & c^2_{001}x_3 \;\;=\;\; 0 \\ f_3 & = & c^3_{100}x_1 & + & c^3_{010}x_2 & + & c^3_{001}x_3 \;\;=\;\; 0. \end{array}$$

This system a non-zero solution if and only if the determinant is zero:

$$\det \begin{pmatrix} c^1_{100} & c^1_{010} & c^1_{001} \\ c^2_{100} & c^2_{010} & c^2_{001} \\ c^3_{100} & c^3_{010} & c^3_{001} \end{pmatrix} \quad = \quad 0.$$

Returning to the general case, we regard each coefficient $c^{(i)}_{j_1,\ldots,j_n}$ of each polynomial f_i as an unknown, and we write $\mathbb{Z}[c]$ for the ring of polynomials with integer coefficients in these variables. The total number of variables in $\mathbb{Z}[c]$ equals $N = \sum_{i=1}^{n} \binom{n+d_i-1}{d_i}$. For instance, the 3×3-determinant in the example above may be regarded as a cubic polynomial in $\mathbb{Z}[c]$. The following theorem characterizes the classical multivariate resultant $\mathrm{Res} = \mathrm{Res}_{d_1 \cdots d_n}$.

Theorem 30. *Fix positive degrees $d_1, \ldots, d_n$. There exists a unique (up to sign) irreducible polynomial $\mathrm{Res} \in \mathbb{Z}[c]$ which has the following properties:*

(a) Res *vanishes under specializing the $c^{(i)}_{j_1 \ldots, j_n}$ to rational numbers if and only if the corresponding equations (25) have a non-zero solution in $\mathbb{C}^n$.*

(b) Res *is irreducible, even when regarded as a polynomial in $\mathbb{C}[c]$.*

(c) Res *is homogeneous of degree $d_1 \cdots d_{i-1} \cdot d_{i+1} \cdots d_n$ in the coefficients $(c^{(i)}_a : |a| = d_i)$ of the polynomial f_i, for each fixed $i \in \{1, \ldots, n\}$.*

We sketch a proof of Theorem 30. It uses results from algebraic geometry.

Proof. The elements of $\mathbb{C}[u]$ are polynomial functions on the affine space $\mathbb{C}^N$. We regard $x = (x_1, \ldots, x_n)$ as homogeneous coordinates for the complex projective space P^{n-1}. Thus (u, x) are the coordinates on the product variety $\mathbb{C}^N \times P^{n-1}$. Let $\mathcal{I}$ denote the subvariety of $\mathbb{C}^N \times P^{n-1}$ defined by the equations

$$\sum_{j_1+\cdots+j_n=d_i} c^{(i)}_{j_1,\ldots,j_n} x_1^{j_1} \cdots x_n^{j_n} \quad = \quad 0 \qquad \text{for } \; i = 1, 2, \ldots, n.$$

Note that $\mathcal{I}$ is defined over $\mathbb{Q}$. Consider the projection $\phi : \mathbb{C}^N \times P^{n-1} \to P^{n-1}$, $(u, x) \mapsto x$. Then $\phi(\mathcal{I}) = P^{n-1}$. The preimage $\phi^{-1}(x)$ of any point $x \in P^{n-1}$ can be identified with the set $\{\, u \in \mathbb{C}^N \,:\, (u, x) \in \mathcal{I} \,\}$. This is a linear subspace of codimension n in $\mathbb{C}^N$. To this situation we apply (Shafarevich 1977, §I.6.3, Theorem 8) to conclude that the variety $\mathcal{I}$ is a closed and irreducible of codimension n in $\mathbb{C}^N \times P^{n-1}$. Hence $dim(\mathcal{I}) = N-1$.

Consider the projection $\psi : \mathbb{C}^N \times P^{n-1} \rightarrow \mathbb{C}^N$, $(u, x) \mapsto u$. It follows from the *Main Theorem of Elimination Theory*, (Eisenbud 1994, Theorem 14.1) that $\psi(\mathcal{I})$ is an irreducible subvariety of $\mathbb{C}^N$ which is defined over $\mathbb{Q}$ as well. Every point c in $\mathbb{C}^N$ can be identified with a particular polynomial system $f_1 = \cdots = f_n = 0$. That system has a nonzero root if and only if c lies in the subvariety $\psi(\mathcal{I})$. For every such c we have

$$dim(\psi(\mathcal{I})) \quad \leq \quad dim(\mathcal{I}) \;=\; N-1 \quad \leq \quad dim(\psi^{-1}(c)) + dim(\psi(\mathcal{I}))$$

The two inequalities in follow respectively from parts (2) and (1) of Theorem 7 in Section I.6.3 of (Shafarevich 1977). We now choose $c = (f_1, \ldots, f_n)$ as follows. Let $f_1, \ldots, f_{n-1}$ be any equations as in (25) which have only finitely many zeros in P^{n-1}. Then choose f_n which vanishes at exactly one of these zeros, say $y \in P^{n-1}$. Hence $\psi^{-1}(c) = \{(c, y)\}$, a zero-dimensional variety. For this particular choice of c both inequalities hold with equality. This implies $dim(\psi(\mathcal{I})) = N - 1$.

We have shown that the image of $\mathcal{I}$ under ψ is an irreducible hypersurface in $\mathbb{C}^N$, which is defined over $\mathbb{Z}$. Hence there exists an irreducible polynomial $Res \in \mathbb{Z}[c]$, unique up to sign, whose zero set equals $\psi(\mathcal{I})$. By construction, this polynomial $Res(u)$ satisfies properties (a) and (b) of Theorem 30.

Part (c) of the theorem is derived from Bézout's Theorem. □

Various determinantal formulas are known for the multivariate resultant. The most useful formulas are mixtures of Bézout matrices and Sylvester matrices like the expression in the middle of (23). Exact division-free formulas of this kind are available for $n \leq 4$. We discuss such formulas for $n = 3$.

The first non-trivial case is $d_1 = d_2 = d_3 = 2$. Here the problem is to eliminate two variables x and y from a system of three quadratic forms

$$\begin{aligned} F &= a_0x^2 + a_1xy + a_2y^2 + a_3xz + a_4yz + a_5z^2, \\ G &= b_0x^2 + b_1xy + b_2y^2 + b_3xz + b_4yz + b_5z^2, \\ H &= c_0x^2 + c_1xy + c_2y^2 + c_3xz + c_4yz + c_5z^2. \end{aligned}$$

To do this, we first compute their *Jacobian determinant*

$$J \quad := \quad \det \begin{pmatrix} \partial F/\partial x & \partial F/\partial y & \partial F/\partial z \\ \partial G/\partial x & \partial G/\partial y & \partial G/\partial z \\ \partial H/\partial x & \partial H/\partial y & \partial H/\partial z \end{pmatrix}.$$

We next compute the partial derivatives of J. They are quadratic as well:

$$\begin{aligned}
\partial J/\partial x &= u_0x^2 + u_1xy + u_2y^2 + u_3xz + u_4yz + u_5z^2, \\
\partial J/\partial y &= v_0x^2 + v_1xy + v_2y^2 + v_3xz + v_4yz + v_5z^2, \\
\partial J/\partial z &= w_0x^2 + w_1xy + w_2y^2 + w_3xz + w_4yz + w_5z^2.
\end{aligned}$$

Each coefficient u_i, v_j or w_k is a polynomial of degree 3 in the original coefficients a_i, b_j, c_k. The resultant of F, G and H coincides with the following 6×6-determinant:

$$\mathrm{Res}_{2,2,2} \quad = \quad \det \begin{pmatrix}
a_0 & b_0 & c_0 & u_0 & v_0 & w_0 \\
a_1 & b_1 & c_1 & u_1 & v_1 & w_1 \\
a_2 & b_2 & c_2 & u_2 & v_2 & w_2 \\
a_3 & b_3 & c_3 & u_3 & v_3 & w_3 \\
a_4 & b_4 & c_4 & u_4 & v_4 & w_4 \\
a_5 & b_5 & c_5 & u_5 & v_5 & w_5
\end{pmatrix} \tag{26}$$

This is a homogeneous polynomial of degree 12 in the 18 unknowns $a_0, a_1, \ldots, a_5, b_0, b_1, \ldots, b_5, c_0, c_1, \ldots, c_5$. The full expansion of Res has $21,894$ terms.

In a typical application of $\mathrm{Res}_{2,2,2}$, the coefficients a_i, b_j, c_k will themselves be polynomials in another variable t. Then the resultant is a polynomial in t which represents the projection of the desired solutions onto the t-axis.

Consider now the more general case of three ternary forms f, g, h of the same degree $d = d_1 = d_2 = d_3$. The following determinantal formula for their resultant was known to Sylvester. We know from part (c) of Theorem 30 that $\mathrm{Res}_{d,d,d}$ is a homogeneous polynomial of degree $3d^2$ in $3\binom{d+2}{2}$ unknowns. We shall express $\mathrm{Res}_{d,d,d}$ as the determinant of a square matrix of size

$$\binom{2d}{2} \quad = \quad \binom{d}{2} + \binom{d}{2} + \binom{d}{2} + \binom{d+1}{2}.$$

We write $S_e = \mathbb{Q}[x,y,z]_e$ for the $\binom{e+2}{2}$-dimensional vector space of ternary forms of degree e. Our matrix represents a linear map of the following form

$$\begin{aligned}
S_{d-2} \otimes S_{d-2} \otimes S_{d-2} \otimes S_{d-1} \quad &\rightarrow \quad S_{2d-2} \\
(a, b, c, u) \quad &\mapsto \quad a \cdot f + b \cdot g + c \cdot h + \delta(u),
\end{aligned}$$

where δ is a linear map from S_{d-1} to S_{2d-2} to be described next. We shall define δ by specifying its image on any monomial $x^iy^jz^k$ with $i+j+k = d-1$.

For any such monomial, we chose arbitrary representations

$$\begin{aligned} f &= x^{i+1} \cdot P_x + y^{j+1} \cdot P_y + z^{k+1} \cdot P_z \\ g &= x^{i+1} \cdot Q_x + y^{j+1} \cdot Q_y + z^{k+1} \cdot Q_z \\ h &= x^{i+1} \cdot R_x + y^{j+1} \cdot R_y + z^{k+1} \cdot R_z, \end{aligned}$$

where P_x, Q_x, R_x are homogeneous of degree $d-i-1$, P_y, Q_y, R_y are homogeneous of degree $d-j-1$, and P_z, Q_z, R_z are homogeneous of degree $d-k-1$. Then we define

$$\delta\big(x^i y^j z^k\big) \quad = \quad \det \begin{pmatrix} P_x & P_y & P_z \\ Q_x & Q & Q_z \\ R_x & R_y & R_z \end{pmatrix}.$$

Note that this determinant is indeed a ternary form of degree

$$(d-i-1)+(d-j-1)+(d-k-1) \quad = \quad 3d-3-(i+j+k) \quad = \quad 2d-2.$$

DISCUSS THE CASE d = 3 IN DETAIL WITH MAPLE CODE

4.3 The sparse resultant

Most systems of polynomial equations encountered in real world applications are *sparse* in the sense that only few monomials appear with non-zero coefficient. The classical multivariate resultant is not well suited to this situation. As an example consider the following system of three quadratic equations:

$$f \,=\, a_0 x + a_1 y + a_2 xy, \quad g \,=\, b_0 + b_1 xy + b_2 y^2, \quad h \,=\, c_0 + c_1 xy + c_2 x^2.$$

If we substitute the coefficients of f, g and h into the resultant $\mathrm{Res}_{2,2,2}$ in (26) then the resulting expression vanishes identically. This is consistent with Theorem 30 because the corresponding homogeneous equations

$$F \,=\, a_0 xz + a_1 yz + a_2 xy, \quad G \,=\, b_0 z^2 + b_1 xy + b_2 y^2, \quad H \,=\, c_0 z^2 + c_1 xy + c_2 y^2$$

always have the common root $(1:0:0)$, regardless of what the coefficients a_i, b_j, c_k are. In other words, the three given quadrics always intersect in the projective plane. But they generally do not intersect in the affine plane $\mathbb{C}^2$. In order for this to happen, the following polynomial in the coefficients must vanish:

$$\begin{aligned}
&\underline{a_1^2 b_2 b_1^2 c_0^2 c_1} - 2a_1^2 b_2 b_1 b_0 c_0 c_1^2 + a_1^2 b_2 b_0^2 c_1^3 - a_1^2 b_1^3 c_0^2 c_2 + 2a_1^2 b_1^2 b_0 c_0 c_1 c_2 \\
&-a_1^2 b_1 b_0^2 c_1^2 c_2 - 2a_1 a_0 b_2^2 b_1 c_0^2 c_1 + 2a_1 a_0 b_2^2 b_0 c_0 c_1^2 + 2a_1 a_0 b_2 b_1^2 c_0^2 c_2 \\
&-2a_1 a_0 b_2 b_0^2 c_1^2 c_2 - 2a_1 a_0 b_1^2 b_0 c_0 c_2^2 + 2a_1 a_0 b_1 b_0^2 c_1 c_2^2 + a_0^2 b_2^3 c_0^2 c_1 - a_0^2 b_2^2 b_1 c_0^2 c_2 \\
&-2a_0^2 b_2^2 b_0 c_0 c_1 c_2 + 2a_0^2 b_2 b_1 b_0 c_0 c_2^2 + a_0^2 b_2 b_0^2 c_1 c_2^2 - a_0^2 b_1 b_0^2 c_2^3 - a_2^2 b_2^2 b_1 c_0^3 \\
&+a_2^2 b_2^2 b_0 c_0^2 c_1 + 2a_2^2 b_2 b_1 b_0 c_0^2 c_2 - 2a_2^2 b_2 b_0^2 c_0 c_1 c_2 - a_2^2 b_1 b_0^2 c_0 c_2^2 + a_2^2 b_0^3 c_1 c_2^2.
\end{aligned}$$

The expression is the *sparse resultant* of f, g and h. This resultant is custom-tailored to the specific monomials appearing in the given input equations.

In this section we introduce the set-up of "sparse elimination theory". In particular, we present the precise definition of the sparse resultant. Let $\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_n$ be finite subsets of $\mathbb{Z}^n$. Set $m_i := \#(\mathcal{A}_i)$. Consider a system of $n+1$ Laurent polynomials in n variables $x = (x_1, \ldots, x_n)$ of the form

$$f_i(x) \quad = \quad \sum_{a \in \mathcal{A}_i} c_{ia}\, x^a \qquad\qquad (i = 0, 1, \ldots, n).$$

Here $x^a = x_1^{a_1} \cdots x_n^{a_n}$ for $a = (a_1, \ldots, a_n) \in \mathbb{Z}^n$. We say that $\mathcal{A}_i$ is the *support* of the polynomial $f_i(x)$. In the example above, $n = 2$, $m_1 = m_2 = m_3 = 3$, $\mathcal{A}_0 = \{\,(1,0),\,(0,1),\,(1,1)\,\}$ and $\mathcal{A}_1 = \mathcal{A}_2 = \{\,(0,0),\,(1,1),\,(0,2)\}$. For any subset $J \subseteq \{0, \ldots, n\}$ consider the affine lattice spanned by $\sum_{j \in J} \mathcal{A}_j$,

$$\mathcal{L}_J \quad := \quad \Big\{ \sum_{j \in J} \lambda_j a^{(j)} \;\mid\; a^{(j)} \in \mathcal{A}_j,\, \lambda_j \in \mathbb{Z} \text{ for all } j \in J \text{ and } \sum_{j \in J} \lambda_j \,=\, 1 \Big\}.$$

We may assume that $\mathcal{L}_{\{0,1,\ldots,n\}} = \mathbb{Z}^n$. Let $rank(J)$ denote the rank of the lattice $\mathcal{L}_J$. A subcollection of supports $\{\mathcal{A}_i\}_{i \in I}$ is said to be *essential* if

$$rank(I) \,=\, \#(I) - 1 \quad \text{and} \quad rank(J) \,\geq\, \#(J) \;\text{ for each proper subset } J \text{ of } I.$$

The vector of all coefficients c_{ia} appearing in $f_0, f_1, \ldots, f_n$ represents a point in the product of complex projective spaces $P^{m_0-1} \times \cdots \times P^{m_n-1}$. Let Z denote the subset of those systems (4.3) which have a solution x in $(\mathbb{C}^*)^n$, where $\mathbb{C}^* := \mathbb{C} \backslash \{0\}$. Let $\bar{Z}$ be the closure of Z in $P^{m_0-1} \times \cdots \times P^{m_n-1}$.

Lemma 31. *The projective variety $\bar{Z}$ is irreducible and defined over $\mathbb{Q}$.*

It is possible that $\bar{Z}$ is not a hypersurface but has codimension ≥ 2. This is where the condition of the supports being essential comes in. It is known that the codimension of $\bar{Z}$ in $P^{m_0-1} \times \cdots \times P^{m_n-1}$ equals the maximum of the numbers $\#(I) - rank(I)$, where I runs over all subsets of $\{0, 1, \ldots, n\}$.

We now define the *sparse resultant* Res. If $codim(\bar{Z}) = 1$ then Res is the unique (up to sign) irreducible polynomial in $\mathbb{Z}[\ldots, c_{ia}, \ldots]$ which vanishes on the hypersurface $\bar{Z}$. If $codim(\bar{Z}) \geq 2$ then Res is defined to be the constant 1. We have the following result. Theorem 32 is a generalization of Theorem 30, in the same way that Bernstein's Theorem generalizes Bézout's Theorem.

Theorem 32. *Suppose that $\{\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_n\}$ is essential, and let Q_i denote the convex hull of $\mathcal{A}_i$. For all $i \in \{0, \ldots, n\}$ the degree of* Res *in the i-th group of variables $\{c_{ia},\ a \in \mathcal{A}_i\}$ is a positive integer, equal to the mixed volume*

$$\mathcal{M}(Q_0, \ldots, Q_{i-1}, Q_{i+1} \ldots, Q_n) \ = \sum_{J \subseteq \{0,\ldots,i-1,i+1\ldots,n\}} (-1)^{\#(J)} \cdot \mathrm{vol}\,\big(\sum_{j \in J} Q_j\big).$$

We refer to (Gel'fand, Kapranov & Zelevinsky 1994) and (Pedersen & Sturmfels 1993) for proofs and details. The latter paper contains the following combinatorial criterion for the existence of a non-trivial sparse resultant. Note that, if each $\mathcal{A}_i$ is n-dimensional, then $I = \{0, 1, \ldots, n\}$ is essential.

Corollary 33. *The variety $\bar{Z}$ has codimension 1 if and only if there exists a unique subset $\{\mathcal{A}_i\}_{i \in I}$ which is essential. In this case the sparse resultant* Res *coincides with the sparse resultant of the equations $\{f_i : i \in I\}$.*

Example. For the linear system

$$c_{00}x + c_{01}y \quad = \quad c_{10}x + c_{11}y \quad = \quad c_{20}x + c_{21}y + c_{22} \quad = \quad 0.$$

the variety $\bar{Z}$ has codimension 1 in the coefficient space $P^1 \times P^1 \times P^2$. The unique essential subset consists of the first two equations. Hence the sparse resultant of this system is *not* the 3×3-determinant (which would be reducible). The sparse resultant is the 2×2-determinant $\mathrm{Res} = c_{00}c_{11} - c_{10}c_{01}$.

We illustrate Theorem 32 for our little system $\{f, g, h\}$. Clearly, the triple of support sets $\{\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3\}$ is essential, since all three *Newton polygons* $Q_i = conv(\mathcal{A}_i)$ are triangles. The mixed volume of two polygons equals

$$\mathcal{M}(Q_i, Q_j) \quad = \quad area(Q_i + Q_j) - area(Q_i) - area(Q_j).$$

In our example the triangles Q_2 and Q_3 coincide, and we have

$$area(Q_1) = 1/2,\ \ area(Q_2) = 1,\ \ area(Q_1 + Q_2) = 9/2,\ \ area(Q_2 + Q_3) = 4.$$

This implies

$$\mathcal{M}(Q_1, Q_2) \ = \ \mathcal{M}(Q_1, Q_3) \ = \ 3 \quad \text{and} \quad \mathcal{M}(Q_2, Q_3) \ = \ 2.4.6$$

This explains why the sparse resultant (4.2) is quadratic in (a_0, a_1, a_2) and homogeneous of degree 3 in (b_0, b_1, b_2) and in (c_0, c_1, c_2) respectively.

One of the central problems in elimination theory is to find "nice" determinantal formulas for resultants. The best one can hope for is a *Sylvester-type formula*, that is, a square matrix whose non-zero entries are the coefficients of the given equation and whose determinant equals precisely the resultant. The archetypical example of such a formula is (24). Sylvester-type formulas do not exist in general, even for the classical multivariate resultant.

If a Sylvester-type formula is not available or too hard to find, the next best thing is to construct a "reasonably small" square matrix whose determinant is a non-zero multiple of the resultant under consideration. For the sparse resultant such a construction was given in (Canny and Emiris 1995) and (Sturmfels 1994). A Canny-Emiris matrix for our example is

$$\begin{array}{c} \\ yf \\ y^2f \\ xy^2f \\ y^2g \\ xy^2g \\ yg \\ xyg \\ xy^2h \\ yh \\ xyh \end{array}
\begin{array}{c}
\begin{array}{cccccccccc} y^2 & y^3 & xy^3 & y^4 & xy^4 & xy^2 & x^2y^2 & x^2y^3 & y & xy \end{array} \\
\left(\begin{array}{cccccccccc}
a_1 & 0 & 0 & 0 & 0 & a_2 & 0 & 0 & 0 & a_0 \\
0 & a_1 & a_2 & 0 & 0 & a_0 & 0 & 0 & 0 & 0 \\
0 & 0 & a_1 & 0 & 0 & 0 & a_0 & a_2 & 0 & 0 \\
b_0 & 0 & b_1 & b_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & b_2 & b_0 & 0 & b_1 & 0 & 0 \\
0 & b_2 & 0 & 0 & 0 & b_1 & 0 & 0 & b_0 & 0 \\
0 & 0 & b_2 & 0 & 0 & 0 & b_1 & 0 & 0 & b_0 \\
0 & 0 & 0 & 0 & c_2 & c_0 & 0 & c_1 & 0 & 0 \\
0 & c_2 & 0 & 0 & 0 & c_1 & 0 & 0 & c_0 & 0 \\
0 & 0 & c_2 & 0 & 0 & 0 & c_1 & 0 & 0 & c_0
\end{array}\right)
\end{array}$$

The determinant of this matrix equals a_1b_2 times the sparse resultant (4.2).

The structure of this 10×10-matrix can be understood as follows. Form the product fgh and expand it into monomials in x and y. A certain combinatorial rule selects 10 out of the 15 monomials appearing in fgh. The columns are indexed by these 10 monomials. Suppose the i-th column is

indexed by the monomial $x^j y^k$, Next there is a second combinatorial rule which selects a monomial multiple of one of the input equations f, g or h such that this multiple contains $x^i y^j$ in its expansion. The i-th row is indexed by that polynomial. Finally the (i, j)-entry contains the coefficient of the j-th column monomial in the i-th row polynomial. This construction implies that the matrix has non-zero entries along the main diagonal. The two combinatorial rules mentioned in the previous paragraph are based on the geometric construction of a *mixed subdivision of the Newton polytypes*.

The main difficulty overcome by the Canny-Emiris formula is this: If one sets up a matrix like the one above just by "playing around" then most likely its determinant will vanish (try it), unless there is a good reason why it shouldn't vanish. Now the key idea is this: a big unknown polynomial (such as Res) will be non-zero if one can ensure that its initial monomial (with respect to some term order) is non-zero.

Consider the lexicographic term order induced by the variable ordering $a_1 > a_0 > a_2 > b_2 > b_1 > b_0 > c_0 > c_1 > c_2$. The 24 monomials of Res are listed in this order above. Consider all $10!$ permutations contribute a (possible) non-zero term to the expansion of the determinant of the Canny-Emiris matrix. There will undoubtedly be some cancellation. However, the unique largest monomial (in the above term order) appears only once, namely, on the main diagonal. This guarantees that the determinant is a non-zero polynomial. Note that the product of the diagonal elements in equals $a_1 b_2$ times the underlined leading monomial.

An explicit combinatorial construction for all possible initial monomials (with respect to any term order) of the sparse resultant is given in (Sturmfels 1993). It is shown there that for any such initial monomial there exists a Canny-Emiris matrix which has that monomial on its main diagonal.

4.4 The unmixed sparse resultant

In this section we consider the important special case when the given Laurent polynomials $f_0, f_1, \ldots, f_n$ all have the same support:

$$\mathcal{A} \quad := \quad \mathcal{A}_0 \;=\; \mathcal{A}_1 \;=\; \cdots \;=\; \mathcal{A}_n \quad \subset \quad \mathbb{Z}^n.$$

In this situation, the sparse resultant Res is the *Chow form* of the projective toric variety $X_{\mathcal{A}}$ which parametrically given by the vector of monomials $\big(x^a \,:\, a \in \mathcal{A}\big)$. Chow forms play a central role in elimination theory, and it

is of great importance to find determinantal formulas for Chow forms of frequently appearing projective varieties. Significant progress in this direction has been made in the recent work of Eisenbud, Floystad, Schreyer on exterior syzygies and the Bernstein-Bernstein-Beilinson correspondence. Khetan (2002) has applied these techniques to give an explicit determinantal formula of mixed Bézout-Sylvester type for the Chow form of any toric surface or toric threefold. This provides a very practical technique for eliminating two variables from three equations or three variables from four equations.

We describe Khetan's formula for an example. Consider the following unmixed system of three equations in two unknowns:

$$\begin{aligned} f &= a_1 + a_2 x + a_3 y + a_4 xy + a_5 x^2 y + a_6 xy^2, \\ g &= b_1 + b_2 x + b_3 y + b_4 xy + b_5 x^2 y + b_6 xy^2, \\ h &= c_1 + c_2 x + c_3 y + c_4 xy + c_5 x^2 y + c_6 xy^2. \end{aligned}$$

The common Newton polygon of f, g and h is a pentagon of normalized area 5. It defines a toric surface of degree 5 in projective 5-space. The sparse unmixed resultant $\mathrm{Res} = \mathrm{Res}(f, g, h)$ is the Chow form of this surface. It can be written as a homogeneous polynomial of degree 5 in the brackets

$$[ijk] \quad = \quad \begin{pmatrix} a_i & a_j & a_k \\ b_i & b_j & b_k \\ c_i & c_j & c_k \end{pmatrix}.$$

Hence Res is a polynomial of degree 15 in the 18 unknowns $a_1, a_2, \ldots, c_6$. It equals the determinant of the following 9×9-matrix

$$\begin{pmatrix} 0 & -[124] & 0 & [234] & [235] & [236] & a_1 & b_1 & c_1 \\ 0 & -[125] & 0 & 0 & 0 & 0 & a_2 & b_2 & c_2 \\ 0 & -[126] & 0 & -[146] & -[156]-[345] & -[346] & a_3 & b_3 & c_3 \\ 0 & 0 & 0 & [345]-[156]-[246] & -[256] & -[356] & a_4 & b_4 & c_4 \\ 0 & 0 & 0 & -[256] & 0 & 0 & a_5 & b_5 & c_5 \\ 0 & 0 & 0 & -[356] & -[456] & 0 & a_6 & b_6 & c_6 \end{pmatrix}$$

4.5 Exercises

PUT YOUR FAVORITE EXERCISES ABOUT RESULTANTS HERE