Rational Univariate Representation

In [4] and [6], symbolic recipes apply linear algebra methods to zero-dimensional ideals. Recipes are also given in [1]. Our application is a 6R robot arm [3], [5]. Oversimplifying, we assume \( \mathbb{C} \) as number field.

1 Stickelberger’s Theorem

A Rational Univariate Representation (abbreviated by RUR) is

\[
R = \left\{ p_0(T) = 0, \ x_i = \frac{p_i(T)}{q(T)}, \ i = 1, 2, \ldots, n \right\}
\]

where \( p_0, p_1, p_2, \ldots, p_n, q \in \mathbb{C}[T] \). This set \( R \) represents the coordinates of the zeroes of a solution set \( V \), \(#V = D < \infty\), of some system of polynomials in \( \mathbb{C}[x] \). \( D \) is counted with multiplicities. The number of distinct zeroes in \( V \) is denoted by \( d \).

Let the linear form \( L(x) \) separate the zeroes: \( L(z_i) \neq L(z_j) \), for \( z_i, z_j \in V \) \( i \neq j \). Consider the multiplication map

\[
m_L : \mathbb{C}[x]/I(V) \to \mathbb{C}[x]/I(V) : h \mapsto (h \cdot L) \to g_{>}, r
\]

where \( \to g_{>}, r \) represents the normal form algorithm implemented by the division algorithm using some Gröbner basis \( g_{>}, r \). The existence of a separating linear form is demonstrated by the following lemma:

**Lemma 1.1** If \( V \) has \( d \) distinct zeroes, then at least one of the

\[
u_i(x_1, x_2, x_3, \ldots, x_n) = x_1 + i x_2 + i^2 x_3 + \cdots + i^{n-1} x_n, \text{ for } 0 \leq i \leq (n-1) \left( \begin{array}{c} d \\ 2 \end{array} \right)
\]

is separating, i.e.: \( \forall z_j, z_k \in V, j \neq k, u_i(z_j) \neq u_i(z_k) \).

**Proof.** For the pair \( (z_j, z_k) \), \( j \neq k \) of two distinct zeroes in \( V \) with components \( z_j = (z_{j1}, z_{j2}, \ldots, z_{jn}) \)
\( z_k = (z_{k1}, z_{k2}, \ldots, z_{kn}) \), consider the bad situation when \( u_i(z_j) = u_i(z_k) \), corresponding to

\[
p(t) = (z_{j1} - z_{k1}) + (z_{j2} - z_{k2}) t + \cdots + (z_{jn} - z_{kn}) t^{n-1}.
\]

Because \( z_j \neq z_k \), \( p \neq 0 \) and therefore \( p \) can have at most \( n - 1 \) zeroes. So for each pair of zeroes of \( V \) we have at most \( n - 1 \) bad choices for \( u_i \) and the number of pairs of zeroes is \( d(d-1)/2 \), yielding a total of at most \( (n-1)d(d-1)/2 \) nonseparating \( u_i \)'s. But the set of \( u_i \)'s consists of \((n-1)d(d-1)/2 + 1\) elements, so there is at least one separating \( u_i \). \( \square \)

If no variables separates the zeroes, then eliminating with any variable order will increase multiplicities.

**Theorem 1.1 (Stickelberger’s theorem)** The multiplication map \( m_L \) is a linear map with matrix \( M_L \). The eigenvalues of \( M_L \) give values for \( L(z) \), for all \( z \in V \), occurring with the same multiplicity \( \mu_z \). As a consequence of this theorem, we have that \( p_0(T) \) is the characteristic polynomial of \( M_L \):

\[
p_0(T) = \det(M_L - TI_D) = \prod_{z \in V} (T - L(z))^{\mu_z}.
\]

with \( \mu_z \) the multiplicity of the root.

The trace of \( M_L \) is

\[
\text{trace}(M_L) = \sum_{z \in V} \mu_z L(z).
\]

The determinant of \( M_L \) is

\[
\text{det}(M_L) = \prod_{z \in V} L(z)^{\mu_z}.
\]
2 The Elbow Manipulator

The application in this section is based on [3], but following the original notation of [5].

The elbow manipulator is a spatial robot arm with three links, of lengths \( L_2, L_3, \) and \( L_4 \). Abbreviating the sines and cosines of the angles \( \theta_i \) respectively by \( s_i = \sin(\theta_i) \) and \( c_i = \cos(\theta_i) \), for \( i = 1, 2, \ldots, 6 \), the successive transformations from one coordinate frame to the next are described by

\[
\begin{pmatrix}
c_1 & 0 & s_1 & 0 \\
s_1 & 0 & -c_1 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
c_2 & -s_2 & 0 & c_2 L_2 \\
s_2 & 0 & c_2 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
c_3 & -s_3 & 0 & c_3 L_3 \\
s_3 & 0 & c_3 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]

where the matrix at the right of (8) represents the position and orientation of the hand of the robot: \( p = (p_x, p_y, p_z) \) is the position of the origin at the hand, and the tree unit vectors \( n = (n_x, n_y, n_z) \), \( o = (o_x, o_y, o_z) \) and \( a = (a_x, a_y, a_z) \) are respectively called the normal, orientation, and approach vector, related to each other by the cross product \( n = o \times a \). In addition to (8), we have the usual relations between the angles, given by

\[
c_i^2 + s_i^2 = 1, \quad i = 1, 2, \ldots, 6. \tag{9}
\]

Taking one as the value for the three lengths \( L_2, L_3, \) and \( L_4 \), in [3], the following examples for the matrix at the right of (8) are proposed:

\[
\begin{pmatrix}
1 & 0 & 0 & 4 \\
0 & 1 & 0 & 8 \\
0 & 0 & 1 & 12 \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
\frac{2}{3} & \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\
\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} & \frac{2}{3} \\
\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
\frac{3}{3} & \frac{6}{3} & \frac{2}{3} & \frac{1}{3} \\
\frac{3}{3} & \frac{6}{3} & \frac{2}{3} & \frac{1}{3} \\
\frac{3}{3} & \frac{6}{3} & \frac{2}{3} & \frac{1}{3} \\
0 & 0 & 0 & 1
\end{pmatrix} \tag{10}
\]

For these five choices the number of complex solutions equals 8 for each instance, while the number of real solutions equals 0, 4, 8, and 8 respectively.

3 Removing Multiplicities

The \textit{radical} of an ideal \( I \), denoted by \( \sqrt{I} \), consists of all polynomial in \( I \) with removed multiplicities. For example, if \( I = \langle x^2 \rangle \), then \( \sqrt{I} = \langle x \rangle \). The quotient ring \( A = \mathbb{C}[x]/(I) \) is a finite dimensional vector space. For any \( h \) in the quotient ring, we define the \textit{h-trace bilinear form} (TrB):

\[
\text{TrB}_h : A \times A \to \mathbb{C} : (f, g) \mapsto \text{trace}(L_{fgh}) \tag{11}
\]

and the \textit{Hermite quadratic form} \( Q_h \):

\[
Q_h : A \to \mathbb{C} : f \mapsto \text{trace}(L_{f^2h}). \tag{12}
\]

The kernel of \( Q_1 \) is

\[
\ker(Q_1) = \{ f \in A \mid \forall g \in A : \text{TrB}_1(f, g) = 0 \}. \tag{13}
\]

\textbf{Theorem 3.1} \( p \in \sqrt{I} \iff p \in \ker(Q_1) \)

The proof is based on the formula for the trace:

\[
\text{TrB}_1(f, g) = \sum_{z \in V} \mu_z f(z)g(z) = 0, \quad \forall g \in A. \tag{14}
\]

For \( h = 1 \), we write \( \text{TrB}_1 \) as \( \text{TrM} \).
4 Newton Identities

A polynomial \( p \) in one variable \( x \) defined by its roots \( x_1, x_2, x_3, \) and \( x_4 \), written as a monic polynomial:

\[
p(x) = (x - x_1)(x - x_2)(x - x_3)(x - x_4) = x^4 - (x_1 + x_2 + x_3 + x_4)x^3 + (x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4)x^2 - (x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4)x + x_1x_2x_3x_4.
\]

The polynomials \( e_1, e_2, e_3, \) and \( e_4 \) are the elementary symmetric polynomials.

Substituting the roots \( x_1, x_2, x_3, \) and \( x_4 \) into \( p(x) = x^4 - e_1x^3 + e_2x^2 - e_3x + e_4 \) gives

\[
0 = p(x_1) = x_1^4 - e_1x_1^3 + e_2x_1^2 - e_3x_1 + e_4,
\]

\[
0 = p(x_2) = x_2^4 - e_1x_2^3 + e_2x_2^2 - e_3x_2 + e_4,
\]

\[
0 = p(x_3) = x_3^4 - e_1x_3^3 + e_2x_3^2 - e_3x_3 + e_4,
\]

\[
0 = p(x_4) = x_4^4 - e_1x_4^3 + e_2x_4^2 - e_3x_4 + e_4.
\]

Adding the four relations above leads to

\[
0 = s_4 - e_1s_3 + e_2s_2 - e_3s_1 + 4e_4,
\]

where \( s_1, s_2, s_3, \) and \( s_4 \) are the power sums:

\[
s_1 = x_1 + x_2 + x_3 + x_4,
\]

\[
s_2 = x_1^2 + x_2^2 + x_3^2 + x_4^2,
\]

\[
s_3 = x_1^3 + x_2^3 + x_3^3 + x_4^3,
\]

\[
s_4 = x_1^4 + x_2^4 + x_3^4 + x_4^4.
\]

So we can express the power sums in terms of the elementary symmetric polynomials:

\[
s_1 = e_1,
\]

\[
s_2 = e_1s_1 - 2e_2,
\]

\[
s_3 = e_1s_2 - e_2s_1 + 3e_3,
\]

\[
s_4 = e_1s_3 - e_2s_2 + e_3s_1 - 4e_4.
\]

The relations above allow the derivations of the power sums from the coefficients of a monic polynomial. Writing the Newton identities in another way:

\[
e_1 = s_1,
\]

\[
2e_2 = e_1s_1 - s_2,
\]

\[
3e_3 = e_1s_2 - e_2s_1 + s_3,
\]

\[
4e_4 = e_1s_3 - e_2s_2 + e_3s_1 - s_4,
\]

we see that, given the power sums of the roots, we can derive the coefficients of the monic polynomial that vanishes at those roots.
5 Application of the Newton Identities

We follow [4] and [6] and we derive an expression for the characteristic polynomial $p_0$. The trace of $L^i$ is

$$s_i = \sum_{z \in V} \mu_z L^i(z). \tag{38}$$

If

$$p_0(T) = \sum_{i=0}^{D} b_i T^{D-i}, \quad D = \#V, \quad b_i \in \mathbb{C}[x],$$

then

$$\frac{p'_0(T)}{p_0(T)} = \sum_{z \in V} \frac{\mu_z}{T - L(z)} = \sum_{j \geq 0} \frac{\text{trace}(L^j)}{T^{j+1}}. \tag{40}$$

Then

$$p'_0(T) = \sum_{i=0}^{D-1} \sum_{j=0}^{D-i-1} \text{trace}(L^j) b_i T^{D-i-j-1}. \tag{41}$$

Identification of the formula obtained above with the derivative of (39) gives Newton’s formula:

$$(D - i)b_i = \sum_{j=0}^{i} \text{trace}(L^j) b_{i-j}. \tag{42}$$

This leads to a linear system to find the coefficients $b_i$ of $p_0$.

After the computation of the characteristic polynomial $p_0(T)$ of the multiplication map $m_L$ where $L$ is separating for the $d$ solutions in $V$.

For any $v \in A$, we define

$$g_L(v, T) = \sum_{z \in V} \mu_z v(z) \prod_{y \in V(p_0)} (T - y). \tag{43}$$

Since

$$g_L(v, L(z)) = \sum_{z \in V} \mu_z v(z) \prod_{y \in V(p_0)} (L(z) - y)$$

$$g_L(1, L(z)) = \frac{\sum_{z \in V} \mu_z \prod_{y \in V(p_0)} (L(z) - y)}{\prod_{y \in V(p_0)} (L(z) - y)} = v(z), \tag{44}$$

as $v$ becomes a coordinate $x_i$ we have:

$$x_i = \frac{g_L(x_i, T)}{g_L(1, T)}, \quad i = 1, 2, \ldots, n. \tag{45}$$

Note

$$g_L(1, T) = \frac{p'_0(T)}{\text{GCD}(p'_0(T), p_0(T))}. \tag{46}$$

If all roots occur with multiplicity one, the denominator $g_L(1, T)$ is just the derivative of $p_0(T)$. To compute $g_L(v, T)$, we define $\tilde{p}_0(T) = p_0(T) / \text{GCD}(p'_0(T), p_0(T))$, and

$$g_L(v, T) = \sum_{j=0}^{d-1} \sum_{k=0}^{d-j-1} \text{trace}(v L^j) a_i T^{d-j-k-1}, \quad \tilde{p}_0 = \sum_{i=0}^{d} a_i T^{d-i}. \tag{47}$$

Then we set $q(T) = g_L(1, T)$ and $p_i(T) = g_L(x_i, T)$ and obtain a representation for (1).
6 Algorithms to Compute a RUR

Algorithms to compute a rational univariate representation can be found in [4] and [6]. As example, taken from [4], consider the katsura3 system:

\[
\begin{align*}
f(x) = \begin{cases} 
2x_1^2 + 2x_2^2 + 2x_3^2 + x_4^2 - x_4 = 0 \\
x_1x_2 + 2x_2x_3 + 2x_3x_4 - x_3 = 0 \\
x_1x_3 + x_3^2 + 2x_2x_4 - x_2 = 0 \\
x_1 + 2x_2 + 2x_3 + x_4 - 1 = 0.
\end{cases}
\end{align*}
\]

Using \(x_4 > x_3 > x_2 > x_1\) in the computation of a Gröbner basis with lexicographical term order gives a univariate polynomial in \(x_1\). Since all solutions of that polynomial in \(x_1\) are distinct, we say that the variable \(x_1\) is separating. We can use this polynomial in \(x_1\) thus as the polynomial \(p_0\) in the rational univariate representation:

\[
p_0(T) = 128304T^8 - 93312T^7 + 15552T^6 + 3144T^4 - 1120T^4 + 367T^3 + 15T^2 - T
\]

\[
q(T) = 7185024T^7 - 4572288T^6 + 653184T^5 + 110040T^4 - 31360T^3 + 756T^2 + 210T - 7
\]

\[
p_2(T) = 699840T^7 - 449712T^6 + 74808T^5 + 1956T^4 - 1308T^3 + 174T^2 - 18T
\]

\[
p_3(T) = 303264T^7 - 14928T^6 + 113544T^5 - 9840T^4 + 3000T^3 + 564T^2 - 12T
\]

\[
p_4(T) = 3872448T^7 - 2607552T^6 + 408528T^5 + 63088T^4 - 20224T^3 + 540T^2 + 172T - 7
\]

The polynomial \(p_1(T)\) is omitted since for this case, the roots of \(p_0(T)\) give the values for \(x_1\).

There are two advantages of using a RUR over the shape lemma:
1. the size of the coefficients is usually much less;
2. one does not need a Gröbner basis with respect to the lexicographical term order.

In [4], Recipe VII gives an algorithm to compute a rational univariate representation of a zero dimensional ideal, given a Gröbner basis.

**Algorithm 6.1 Rational Univariate Representation**

Input: \(G>\), a Gröbner basis for an ideal \(I\) with term order \(>\), \(#V(I) = D < \infty\).

Output: a rational univariate representation for \(V(\sqrt{T})\).

compute \(N>\), the basis vector for the quotient ring;
let \(D = \#N> = \#V(I)\), counted with multiplicities;
compute \(\text{TrM}\) and deduce \(d = \#\text{distinct zeroes}\);
choose a separating element \(u\) as one of the \(u_i\)'s;
compute for \(m\) from 1 to \(D\): \(\text{trace}(u^m)\) and use \(u\) to form \(p_0(T)\);
compute \(\bar{p}_0\) for \(\sqrt{T}\), if \(\deg(\bar{p}_0) < d\) then choose another \(u\);
for \(j\) from 1 to \(D\)
  for \(i\) from 0 to \(d\)
    compute \(\text{trace}(x_j u^i)\) and deduce \(g_u(x_j, T)\);
set \(q(T) = g_u(1, T)\) and \(p_i = g_u(x_i, T), i = 1, 2, \ldots, n\).

In Maple, we compute a rational univariate representation as follows:

\[
\begin{align*}
> v := x[4],x[3],x[2],x[1];
> Groebner[Basis](f,plex(v));
> Groebner[RationalUnivariateRepresentation](f,v,output=factored);
\]
7 Exercises

1. Construct an example of a solution set $V$ in two variables and $\#V > 1$ where all except for one choice of the $u_i$’s fail to be separating.

2. Use a lexicographic term order to compute a Gröbner basis for the system in (48). How many decimal places does the largest coefficient in this basis have? Compare with the size of the coefficients in (49).

3. Use Maple’s Groebner[RationalUnivariateRepresentation] on the example (48).

4. Create a Maple worksheet to define the polynomial system in (8), for general choices of the position. Solve the system using the choices in (10), either by Maple or by Sage.Singular.

5. Consider the system (from [4]):

$$f(x) = \begin{cases} 
24x_1x_2 - x_1^2 - x_2^2 - x_1^2x_2^2 - 13 = 0 \\
24x_2x_3 - x_2^2 - x_3^2 - x_2^2x_3^2 - 13 = 0 \\
24x_3x_1 - x_3^2 - x_1^2 - x_3^2x_1^2 - 13 = 0.
\end{cases} \tag{50}$$

(a) Solve the system to verify that none of the variables is separating.

(b) Find a separating element of the form as $u_i$ in (3) an run through the steps of Algorithm 6.1. Use a Maple worksheet or a Sage notebook to guide the computations.

(c) Use the builtin commands in Maple or Sage to compute a rational univariate representation. Compare the output with the outcome of the step-by-step execution of Algorithm 6.1.

6. Consider the following modification (suggested in [2]) of the cyclic 5-roots problem:

$$f(x) = \begin{cases} 
x_1 + x_2 + x_3 + x_4 + x_5 = 0 \\
x_1x_2 + x_2x_3 + x_3x_4 + x_4x_5 + x_5x_1 = 0 \\
x_1x_2x_3 + x_2x_3x_4 + x_3x_4x_5 + x_4x_5x_1 + x_5x_1x_2 = 0 \\
x_2x_3x_4 + x_2x_3x_4x_5 + x_3x_4x_5x_1 + x_4x_5x_1x_2 + x_5x_1x_2x_3 = 0 \\
x_1x_2x_3x_4x_5 - 1 = 0.
\end{cases} \tag{51}$$

where the monomial $x_1x_2x_3x_4$ in the original cyclic 5-roots system is replaced by $x_2x_3x_4$. Compute a Gröbner basis with the grded lexicographical order to determine the number of roots of this modified cyclic 5-roots system. Compute a rational univariate representation for this system. Compare the size of the coefficients between the lexicographical Gröbner basis (shape lemma) and the RUR.

References


