Quite Inefficient Algorithms for Solving Systems of Polynomial Equations

Without Gröbner Bases!

Resultants, Primary Decomposition and Galois Groups

Lars Wallenborn   Jesko Hüttenhain

11/25/11 — 12/02/11

Contents

1 Resultants

We discuss some methods for solving zero-dimensional systems of polynomial equations without the aid of Gröbner bases. It will turn out that these methods also yield techniques to compute resultants.

In this section, let always $f_0, \dots, f_n \in \Bbbk[x_1, \dots, x_n]$ be polynomials over an algebraically closed field $\Bbbk = \overline{\Bbbk}$ of degrees $d_i := \deg(f_i)$. Let $I = (f_1, \dots, f_n)$. We assume that f_0 is homogeneous of degree $d_0 = 1$. We set $\mu := d_1 \cdots d_n$, $d := d_1 + \cdots + d_n - (n-1)$ and

$$\nu := \sum_{k=0}^{d} \binom{n+d}{d} - \mu = \binom{n+d+1}{d} - \mu.$$

A point $p \in \Bbbk^n$ is called a **solution** if $p \in Z(I)$. Finally, denote by

$$\overline{\cdot} : \Bbbk[x_1, \dots, x_n] \longrightarrow \Bbbk[x_1, \dots, x_n]/I =: A$$

the canonical projection. We denote by $L_f : A \to A$ the $\Bbbk$-automorphism of A which is given by multiplication with $\overline{f}$ for some $f \in \Bbbk[x_1, \dots, x_n]$.

1.1 Solving without Gröbner Bases

Theorem 1.1 (Bézout's Theorem).

(a). If there are only finitely many solutions, then their number, counted with multiplicity, is at most μ.

(b). For a generic choice of $f_1, \dots, f_n$, there are precisely μ solutions, each with multiplicity one. $\square$

Remark 1.2. *Here, "generic" means that for "almost every" choice of polynomials, the above holds. To properly define this "almost", one needs more algebraic geometry than we are willing to present here.*

Definition 1.3. *In the following, $\gamma = (\gamma_1, \dots, \gamma_n) \in \mathbb{N}^n$ denotes a multiindex. We partition the set $S = \{ x^\gamma \,:\, |\gamma| \le d \}$ as $S = S_0 \uplus \cdots \uplus S_n$, where*

$$S_i := \left\{ x^\gamma \in S \;\middle|\; \begin{array}{cc} \forall j < i : & d_j > \gamma_j \\ \textit{and} & d_i \le \gamma_i \end{array} \right\} \tag{1.1}$$

*for $i > 0$ and $S_0 := \{ x^\gamma \in S \mid \forall j : d_j > \gamma_j \}$. A monomial $x^\gamma \in S_i$ is said to be **reduced** if $d_j > \gamma_i$ for all $j > i$.*

We set $S_+ := S_1 \cup \cdots \cup S_n$. S_0 is the set of monomials of degree at most d which are not divisible by any of the $x_i^{d_i}$. Since S_0 will play a special role, we use x^α to denote its elements and x^β for elements in S_+.

Fact 1.4. *Behold:*

(a). If $x^\alpha \in S_0$, then $\deg x^\alpha \leq d-1$. Furthermore, $|S_0| = \mu$ and $|S_+| = \nu$.

(b). If $x^\beta \in S_i$ for $i > 0$, then $\deg\left(x^\beta \big/ x_i^{d_i}\right) \leq d - d_i$

Proof. Part (a) follows from $S_0 = \{\, x_1^{\gamma_1} \cdots x_n^{\gamma_n} \mid \forall i : 0 \leq \gamma_i \leq d_i - 1 \,\}$ and the observation that $d - 1 = \sum_{i=1}^n (d_i - 1)$. For part (b), use the explicit description (1.1). $\square$

Definition 1.5. *For $x^\gamma \in S_i$, define*

$$f_\gamma := \begin{cases} x^\gamma f_i \big/ x_i^{d_i} & ; \quad i \neq 0 \\ x^\gamma f_0 & ; \quad i = 0 \end{cases}$$

Fact 1.6. *We can write f_γ as a $\Bbbk$-linear combination of the $x^\gamma \in S$.*

Proof. For $\gamma \in S_0$, this follows from the assertions $d_0 = 1$ and Fact 1.4.(a). For $\gamma \in S_i$, this follows because $\deg(f_\gamma) \leq |\gamma| \leq d$. $\square$

Definition 1.7. *Write $S_0 = \{\, x^{\alpha_1}, \ldots, x^{\alpha_\mu} \,\}$ and $S_+ = \{\, x^{\beta_1}, \ldots, x^{\beta_\nu} \,\}$. The **Sylvester-type matrix** associated to our given data is the matrix $\mathcal{M}$ such that*

$$\mathcal{M} \cdot \begin{pmatrix} x^{\alpha_1} \\ \vdots \\ x^{\alpha_\mu} \\ x^{\beta_1} \\ \vdots \\ x^{\beta_\nu} \end{pmatrix} = \begin{pmatrix} f_{\alpha_1} \\ \vdots \\ f_{\alpha_\mu} \\ f_{\beta_1} \\ \vdots \\ f_{\beta_\nu} \end{pmatrix}. \tag{1.2}$$

Such a matrix exists by Fact 1.6. We write

$$\mathcal{M} = \begin{pmatrix} \mathcal{M}_{00} & \mathcal{M}_{01} \\ \mathcal{M}_{10} & \mathcal{M}_{11} \end{pmatrix}. \tag{1.3}$$

where $\mathcal{M}_{00}$ is a $\mu \times \mu$ square matrix and $\mathcal{M}_{11}$ is a $\nu \times \nu$ square matrix.

Remark 1.8. *For a generic choice of $f_1, \ldots, f_n$, the matrix $\mathcal{M}_{11}$ is invertible. Let us understand why. If we let $\mathcal{M}_{11} = (\lambda_{ij})_{i,j}$ and consider the equality*

$$\mathcal{M}_{11} \cdot \begin{pmatrix} x^{\beta_1} \\ \vdots \\ x^{\beta_\nu} \end{pmatrix} = \begin{pmatrix} f_{\beta_1} \\ \vdots \\ f_{\beta_\nu} \end{pmatrix},$$

this means nothing more than

$$f_{\beta_i} = \sum_{k=1}^{\nu} \lambda_{ik} \cdot x^{\beta_k}.$$

A generic choice of the f_i means a generic choice of coefficients λ_{ik}, and for almost every such choice, the vectors $(\lambda_{i1}, \dots, \lambda_{i\nu})$ for $1 \leq i \leq \nu$ are linearly independent.

Hence, the following Assumption 1.9 is justified:

Assumption 1.9. *We henceforth assume $\mathcal{M}_{11}$ to be invertible. Note that this implies that A is a zero-dimensional ring, i.e. a finite $\Bbbk$-vectorspace.*

Definition 1.10. *We define*

$$\widetilde{\mathcal{M}} := \widetilde{\mathcal{M}}(f_0) := \mathcal{M}_{00} - \mathcal{M}_{01}\mathcal{M}_{11}^{-1}\mathcal{M}_{10}. \tag{1.4}$$

Furthermore, we define two maps $\phi : \Bbbk^n \to \Bbbk^\mu$ and $\psi : \Bbbk^n \to \Bbbk^\nu$ where

$$\phi(p) := \begin{pmatrix} p^{\alpha_1} \\ \vdots \\ p^{\alpha_\mu} \end{pmatrix} \qquad\qquad \psi(p) := \begin{pmatrix} p^{\beta_1} \\ \vdots \\ p^{\beta_\nu} \end{pmatrix}$$

In other words, the maps are induced by the monomials in S_0 and S_+, respectively.

Theorem 1.11. *For all solutions p, the vector $\phi(p)$ is an eigenvector of $\widetilde{\mathcal{M}}$ with eigenvalue $f_0(p)$. Furthermore, for a generic choice of f_0, the set $\phi(Z(I))$ is linearly independent.*

Proof. Let $p \in Z(I)$. Then,

$$\begin{pmatrix} \mathcal{M}_{00} & \mathcal{M}_{01} \\ \mathcal{M}_{10} & \mathcal{M}_{11} \end{pmatrix} \cdot \begin{pmatrix} \phi(p) \\ \psi(p) \end{pmatrix} = \mathcal{M} \cdot \begin{pmatrix} \phi(p) \\ \psi(p) \end{pmatrix} = \begin{pmatrix} f_{\alpha_1}(p) \\ \vdots \\ f_{\alpha_\mu}(p) \\ f_{\beta_1}(p) \\ \vdots \\ f_{\beta_\nu}(p) \end{pmatrix} = \begin{pmatrix} (x^{\alpha_1} f_0)(p) \\ \vdots \\ (x^{\alpha_\mu} f_0)(p) \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} f_0(p) \cdot \phi(p) \\ \mathbf{0} \end{pmatrix}$$

by (1.2). This gives us the two identities

$$\mathcal{M}_{00} \cdot \phi(p) + \mathcal{M}_{01} \cdot \psi(p) = f_0(p) \cdot \phi(p) \tag{1.5}$$

$$\mathcal{M}_{10} \cdot \phi(p) + \mathcal{M}_{11} \cdot \psi(p) = 0 \tag{1.6}$$

which we can use to conclude

$$\begin{aligned}
\widetilde{\mathcal{M}} \cdot \phi(p) &= \mathcal{M}_{00} \cdot \phi(p) - \mathcal{M}_{01}\mathcal{M}_{11}^{-1}\mathcal{M}_{10} \cdot \phi(p) && \text{by (1.4)} \\
&= \mathcal{M}_{00} \cdot \phi(p) + \mathcal{M}_{01}\mathcal{M}_{11}^{-1}\mathcal{M}_{11} \cdot \psi(p) && \text{by (1.6)} \\
&= \mathcal{M}_{00} \cdot \phi(p) + \mathcal{M}_{01} \cdot \psi(p) \\
&= f_0(p) \cdot \phi(p) && \text{by (1.5)}
\end{aligned}$$

Finally, we note that for a generic choice, f_0 takes distinct values at all the $p \in Z(I)$, therefore the eigenvalues are distinct and the corresponding eigenvectors linearly independent. □

Theorem 1.12. *For generic f_0, the set $\overline{S}_0$ is a $\Bbbk$-basis of A. Furthermore, $\widetilde{\mathcal{M}}$ is the matrix corresponding to L_{f_0} with respect to the basis $\overline{S}_0$.*

Proof. By Bézout (Theorem 1.1), A has dimension μ over $\Bbbk$. Since this is also the cardinality of S_0, the first part of the theorem will follow once we show that the $\overline{x}^\alpha$ are linearly independent. Assume that

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

Evaluating this equation at a solution p gives

$$c_1 \cdot p^{\alpha_1} + \cdots + c_\mu \cdot p^{\alpha_\mu} = 0.$$

Let $Z(I) = \{ p_1, \ldots, p_\mu \}$ and define the square matrix $P := (p_j^{\alpha_i})_{ij}$. Since

$$(c_1, \ldots, c_\mu) \cdot P = \mathbf{0}$$

and P is invertible by Theorem 1.11, we conclude $c_i = 0$ for all i, which proves that $\overline{S}_0$ is linearly independent.

Let M be the coordinate matrix of L_{f_0} in the basis $\overline{S}_0$. Clearly, $M\phi(p) = f_0(p)\phi(p)$ for every solution p. But from Theorem 1.11 we know that $f_0(p)\phi(p) = \widetilde{\mathcal{M}}\phi(p)$ and therefore, $M\phi(p) = \widetilde{\mathcal{M}}\phi(p)$ for every solution p. We know that the μ different $\phi(p)$ are linearly independent, so they form a basis. □

1.1.1 Multiplication Matrices

By setting $f_0 = x_i$ in Theorem 1.12, we get that the matrix of multiplication by x_i is $\widetilde{\mathcal{M}}(x_i)$. However, it is possible to compute all of these maps simultaneously by using $f_0 = u_1x_1 + \cdots + u_nx_n$, where $u_1, \ldots, u_n$ are varibles. Thus, by means of

$$\widetilde{\mathcal{M}}(f_0) = u_1\widetilde{\mathcal{M}}(x_1) + \cdots + u_n\widetilde{\mathcal{M}}(x_n),$$

we can compute all multiplication matrices at once.

1.1.2 Solving via Multivariate Factorization

As above, suppose that $f_0 = u_1x_1+\cdots+u_nx_n$ where $u_1, \ldots, u_n$ are variables. In this case, $\det(\widetilde{\mathcal{M}}(f_0))$ becomes a polynomial in $\Bbbk\,[u_1, \ldots, u_n]$. The results of this section imply that the eigenvalues of $\widetilde{\mathcal{M}}(f_0)$ are $f_0(Z(I))$. Since all of the eigenspaces have dimension 1,

$$\det(\widetilde{\mathcal{M}}) = \prod_{p\in Z(I)} f_0(p) = \prod_{(p_1,\ldots,p_n)\in Z(I)} (u_1p_1 + \cdots + u_np_n). \tag{1.7}$$

By factoring $\det(\widetilde{\mathcal{M}}(f_0))$ into irreducibles in $\Bbbk\,[u_1, \ldots, u_n]$, we get all solutions.

1.1.3 Ideal Membership

For a given $f \in \Bbbk\,[x_1, \ldots, x_n]$ we want to decide whether $f \in I$ or not. Using the above $\widetilde{\mathcal{M}}(x_i)$, we can solve this problem by the following

Fact 1.13. $f \in I \Leftrightarrow f(\widetilde{\mathcal{M}}(x_1), \ldots, \widetilde{\mathcal{M}}(x_n)) = \mathbf{0} \in \Bbbk^{\mu\times\mu}$.

Proof. We know that $f \in I$ if and only if $\overline{f} = 0$, i.e. f becomes zero in A. This is equivalent to saying that L_f is the zero map. If we write $L_i := L_{x_i}$ and $f = \sum_\lambda a_\lambda x^\lambda$ then

$$\begin{aligned} L_f &= \Big(y \mapsto \sum\nolimits_\lambda a_\lambda x^\lambda \cdot y\Big) \\ &= \sum\nolimits_\lambda a_\lambda \prod\nolimits_{i=1}^{n} (y \mapsto x_i y)^{\circ\lambda_i} \\ &= \sum\nolimits_\lambda a_\lambda L^{\circ\lambda} \\ &= f(L_1, \ldots, L_n), \end{aligned}$$

Thus, $\widetilde{\mathcal{M}}(f) = f(\widetilde{\mathcal{M}}(x_1), \ldots, \widetilde{\mathcal{M}}(x_n))$ is the zero matrix if and only if $f \in I$. □

1.2 Multivariate Resultants

We now set up notation for this section. Let f_{ij} be the homogeneous component of f_i in degree j. Let $F_0, \ldots, F_n \in \Bbbk[x_0, \ldots, x_n]$ arise from f_i by homogenization in a new variable x_0. This means $F_i := x_0^{d_i} f_i\left(\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}\right)$, $F_0 = f_0$ and $F_i = \sum_{j=0}^{d_i} f_{ij} x_0^{d_i - j}$. We say that $p = [p_0 : \ldots : p_n] \in \mathbb{P}^n$ is a **solution at infinity** if $p_0 = 0$ and $F_i(p) = 0$ for all i. We set $G_i := f_{i,d_i}$ for all i.

Definition 1.14. *Let $F_0, \ldots, F_n \in \Bbbk\,[x_0, \ldots, x_n]$ be homogeneous polynomials. The* ***resultant*** $\operatorname{res}(F_0, \ldots, F_n)$ *is a polynomial in their coefficients which is the zero polynomial if and only if the F_i have a common projective root. We denote the same polynomial by* $\operatorname{res}(f_0, \ldots, f_n)$ *if F_i is the homogenization of f_i by x_0.*

Fact 1.15. *There exist solutions at infinity if and only if* $\operatorname{res}(G_1, \ldots, G_n) = 0$.

Proof. Note that $F_i(0, x_1, \ldots, x_n) = G_i$. Thus, there exists a solution at infinity if and only if the homogeneous polynomials G_i share a common root. □

Theorem 1.16. *If there are no solutions at infinity and M is a matrix corresponding to L_{f_0}, then*

$$\operatorname{res}(f_0, \ldots, f_n) = \operatorname{res}(G_1, \ldots, G_n) \cdot \det(M).$$

Furthermore, if we denote by $\chi_M(T)$ the characteristic polynomial of M in an indeterminante T,

$$\operatorname{res}(T - f_0, f_1, \ldots, f_n) = \operatorname{res}(G_1, \ldots, G_n) \cdot \chi_M(T).$$

Metaproof. A proof for these equalities can be traced back to [Jou91].

Definition 1.17. *Recall Definition 1.7. We define the matrix $\mathcal{M}'$ to be the matrix that arises from $\mathcal{M}$ by deleting all rows and columns corresponding to reduced monomials.*

Proposition 1.18. *It is* $\operatorname{res}(f_0, \ldots, f_n) \cdot \det(\mathcal{M}') = \det(\mathcal{M})$.

Metaproof. This is [CLO98, Theorem 4.9]. □

Corollary 1.19. *It is* $\operatorname{res}(G_1, \ldots, G_n) \cdot \det(\mathcal{M}') = \det(\mathcal{M}_{11})$.

Proof. If $\det(\mathcal{M}_{11}) \neq 0$, we can write

$$\begin{aligned}\det(\widetilde{\mathcal{M}}) \cdot \det(\mathcal{M}_{11}) &= \det \begin{pmatrix} \widetilde{\mathcal{M}} & \mathbf{0} \\ \mathcal{M}_{10} & \mathcal{M}_{11} \end{pmatrix} = \det \begin{pmatrix} I & -\mathcal{M}_{01}\mathcal{M}_{11}^{-1} \\ \mathbf{0} & I \end{pmatrix} \cdot \det \begin{pmatrix} \mathcal{M}_{00} & \mathcal{M}_{01} \\ \mathcal{M}_{10} & \mathcal{M}_{11} \end{pmatrix} \\ &= \det(\mathcal{M}).\end{aligned}$$

Hence by Proposition 1.18, Theorem 1.16 and Theorem 1.12 (in this order),

$$\begin{aligned}\det(\widetilde{\mathcal{M}}) \cdot \det(\mathcal{M}_{11}) &= \det(\mathcal{M}) \\ &= \operatorname{res}(f_0, \ldots, f_n) \cdot \det(\mathcal{M}') \\ &= \operatorname{res}(G_1, \ldots, G_n) \cdot \det(L_{f_0}) \cdot \det(\mathcal{M}') \\ &= \operatorname{res}(G_1, \ldots, G_n) \cdot \det(\widetilde{\mathcal{M}}) \cdot \det(\mathcal{M}').\end{aligned}$$

when $f_1, \ldots, f_n$ are sufficiently generic. Cancelling $\det(\widetilde{\mathcal{M}})$, which is generically nonzero, we conclude that

$$\det(\mathcal{M}_{11}) = \operatorname{res}(G_1, \ldots, G_n) \cdot \det(\mathcal{M}')$$

holds for almost all choices of coefficients of the f_i.

Since both sides of this equation are polynomial in the coefficients of the f_i, it means that the equality holds in general. □

2 Primary Ideal Decomposition

In this section, I the ideal generated by the polynomials $f_1, \ldots, f_s \in \Bbbk[x_1, \ldots, x_n]$ over any field $\Bbbk$. We denote by $\mathbb{K} := \overline{\Bbbk}$ the algebraic closure of $\Bbbk$. We then write

$$\overline{\cdot} : \Bbbk[x_1, \ldots, x_n] \twoheadrightarrow \Bbbk[x_1, \ldots, x_n]/I =: A$$

for the canonical projection. Assume that A is of dimension zero, i.e. $V := Z(I) \subseteq \mathbb{K}^n$ is finite.

We denote by $n_f(p)$ (resp. $m_f(p)$) the multiplicity of $(T - f(p))$ in the characteristic (resp. minimal) polynomial of L_f, where $L_f : A \to A$ is the multiplication by $\overline{f}$ for some polynomial f. When there is no risk of confusion, we write $n(p)$ instead of $n_f(p)$ and equivalently, $m(p)$ instead of $m_f(p)$.

2.1 Preliminaries

Definition 2.1. *An ideal $J \subseteq R$ of a ring is said to be **primary** if $fg \in I$ implies that $f \in I$ or $g^k \in I$ for some $k \in \mathbb{N}$. If J is primary, then $\sqrt{J}$ is a prime ideal.*

Definition 2.2. *A **primary decomposition** of an ideal J is a decomposition of the form $J = J_1 \cap \cdots \cap J_r$ with J_i primary. By [Eis94, Theorem 3.10], every ideal of a Noetherian ring has a primary decomposition. We say that the decomposition is **minimal** if r is minimal. In this case, the ideals $P_i := \sqrt{J_i}$ are prime ideals which are minimal over J.*

Fact 2.3. *Assume that I and J are primary ideals of a domain R. If $\sqrt{I} = \sqrt{J}$, then $I \cap J$ is primary.*

Proof. Let $fg \in I \cap J \subseteq I$. We can assume that $f \notin I \cap J$. Hence, let us assume that $f \notin I$. This means $g^r \in I$ for some r. But this means $g \in \sqrt{I} = \sqrt{J}$, so $g^s \in J$ for some s. We set $k := \max(r, s)$ and obtain $g^k \in I \cap J$. □

Fact 2.4. *Let P be a maximal ideal in a domain R. Then, for any $h \notin P$, there exists an element $g \in R$ such that $1 + gh \in P$.*

Proof. Since $h \notin P$, it becomes a unit in $K := R/P$. Choosing any element $g \in R$ which is mapped to $-h^{-1}$ under $R \twoheadrightarrow R/P$ yields $1 + gh = 0$ in K, i.e. $1 + gh \in P$. □

Lemma 2.5. *The ideal I has a minimal primary decomposition $I = I_1 \cap \cdots \cap I_r$. For any such decomposition, the $P_i := \sqrt{I_i}$ are distinct maximal ideals. Furthermore, $I_i \not\subset \bigcup_{j \neq i} P_j$ and for any $g \in I_i \setminus \bigcup_{j \neq i} P_j$, it is $I_i = I + (g)$.*

Proof. We already established that such a primary decomposition always exists. Note that P_i is a minimal prime ideal over I, so $\dim(I_i) = \dim(P_i) = \dim(I) = 0$. Since P_i is zero-dimensional and prime, it is maximal. If $P_i = P_j$ for some $i \neq j$, then $I_i \cap I_j$ is primary by Fact 2.3, contradicting minimality of the decompostion. If $I_i \subseteq \bigcup_{j \neq i} P_j$, then $I_i \subseteq P_j$ for some $j \neq i$ by Prime Avoidance (see [Eis94, Lemma 3.3]). This means $P_i \subseteq P_j$ and hence, $P_i = P_j$ by maximality, which is absurd.

Finally, let $g \in I_i \setminus \bigcup_{j \neq i} P_j$. Certainly, $I + (g) \subseteq I_i$. Now for every $j \neq i$, note that $g \notin P_j$ and so we can choose a $h_j \in R$ with the property that $1 + gh_j \in P_j$, by Fact 2.4. We choose $m \in \mathbb{N}$ such that, for all j, $(1 + gh_j)^m \in I_j$. Expanding the product

$$\prod_{j \neq i} (1 + gh_j)^m \in \prod_{j \neq i} I_j \subseteq \bigcap_{j \neq i} I_j,$$

we conclude $1 + gh \in \bigcap_{j \neq i} I_j$ for a particular $h \in R$. Given any $a \in I_i$,

$$a(1 + gh) \in I_i \cap \bigcap_{j \neq i} I_j = I$$

and therefore $a = (1 + gh)a + g(-ha) \in I + (g)$ as desired. □

In the following, we describe a method to compute certain $g_i \in I_i \setminus \bigcup_{j \neq i} P_j$ which yield the ideals I_i by Lemma 2.5. We can do this, again, without using Gröbner bases.

2.2 Rational Case

Recall that a solution $p \in V$ is **$\Bbbk$-rational** if $p \in \Bbbk^n$. For this subsection, we assume that all solutions are $\Bbbk$-rational.

Proposition/Definition 2.6. *$I = \bigcap_{p \in V} I_p$ is a minimal primary decomposition with*

$$I_p := \{\, f \in \Bbbk\,[x_1, \ldots, x_n] \mid \exists g \in \Bbbk\,[x_1, \ldots, x_n] : gf \in I \text{ and } g(p) \neq 0 \,\}.$$

In this case, $\sqrt{I_p}$ is the maximal ideal $\mathfrak{m}_p := (x_1 - p_1, \ldots, x_n - p_n)$ where the p_i are the coordinates of p.

Remark. In fact, in the zero-dimensional case, the minimal primary decomposition is unique, but we will not give a proof for this.

Proof. The assumption that all solutions are $\Bbbk$-rational means that we can assume $\Bbbk = \mathbb{K}$ is algebraically closed.

We show $\bigcap_{p \in V} I_p \subseteq I$. Pick any f such that $f \in I_p$ for all p. Choose polynomials g_p such that $g_p(p) \neq 0$ and $g_p f \in I$. We have already seen that there exist idempotents $e_p \in A$, i.e. $e_p(q) = \delta_{pq}$. The function $g := \sum_{p \in V} e_p g_p$ satisfies $g(p) = g_p(p) \neq 0$ for all $p \in V$. Let $h := \sum_{p \in V} \frac{e_p}{g(p)}$, then $(1 - gh)$ vanishes on all of V and thus, $(1 - gh)^k \in I$ for some k, by the Hilbert Nullstellensatz. Multiplying this out gives $1 - gh' \in I$ for some polynomial h'. Since $gf \in I$, we know $gfh' \in I$. Since $f - fgh' = f(1 - gh') \in I$, we conclude $f \in I$.

To show that $\sqrt{I_p} = \mathfrak{m}_p$, it will suffice to show that there exists a natural number $k \in \mathbb{N}$ such that $(x_i - p_i)^k \in I_p$ because it implies $\mathfrak{m}_p \subseteq \sqrt{I_p}$ and the statement then follows by maximality of $\mathfrak{m}_p$. Set

$$g_i := \prod\nolimits_{\substack{q \in V \\ q_i \neq p_i}} (x_i - q_i)$$

Then, $g_i \cdot (x_i - p_i)$ vanihses on all of V. Hence, $g_i^k (x_i - p_i)^k \in I$ for some $k \in \mathbb{N}$. Since $g_i^k(p) \neq 0$, we know $(x_i - p_i)^k \in I_p$ by definition. This also proves that I_p is primary.

We are left to show minimality of the decomposition. Assume that $I = \bigcap_{i=1}^r I_i$ is minimal. Let $P_i := \sqrt{I_i}$. It is a maximal ideal with corresponding point $p_i \in \mathbb{K}^n$. Then,

$$V = Z(I) = Z\left(\sqrt{I}\right) = Z\left(\bigcap\nolimits_i P_i\right) \subseteq Z\left(\prod\nolimits_i P_i\right) = \bigcup\nolimits_{i=1}^r Z(P_i) = \{ p_1, \ldots, p_r \}$$

implies $|V| \leq r$ and we are done. We used well-known facts about algebraic sets, see [Har06, Propositions 1.1 and 1.2], for instance. □

Proposition 2.7. *If $f \in \Bbbk[x_1, \ldots, x_n]$ takes distinct values at all solutions, then*

$$\forall p \in V : \quad I_p = I + \left((f - f(p))^{m(p)} \right).$$

Proof. Pick $p \in V$ and set $g := (f - f(p))^{m(p)}$. By Lemma 2.5 and Proposition 2.6, it suffices to show that $g \in I_p$ and $g \notin \mathfrak{m}_q$ for all $q \neq p$. The latter condition is equivalent to $g(q) \neq 0$, which follows because $f(q) \neq f(p)$ by assumption. To prove that $g \in I_p$, let

$$h := \prod_{q \neq p} (f - f(q))^{m(q)}.$$

Denote by μ the minimal polynomial of the multiplication map L_f. Then, by definition $gh = \mu(f)$. However, the Cayley-Hamilton Theorem [Eis94, Theorem 4.3] says that $\mu(L_f)$ is the zero map on A. Applied to 1, we obtain

$$0 = \mu(L_f)(1) = \mu(\overline{f}) = \overline{\mu(f)}.$$

Hence, $gh \in I \subseteq I_p$. Since $h(p) = \prod_{q \neq p} (f(p) - f(q))^{m(q)} \neq 0$ by our assumption on f, we know $h \notin \mathfrak{m}_p$. Thus, no power of h can be contained in I_p. Since I_p is primary, this means $g \in I_p$. □

Example 2.8. *Consider the case $\Bbbk = \mathbb{Q}$ and*

$$\begin{aligned} f_1 &:= x^2 + 2y(y-1) \\ f_2 &:= xy(y-1) \\ f_3 &:= y(y^2 - 2y + 1) \end{aligned}$$

First, note that $p = (0, 0)$ and $q = (0, 1)$ are $\mathbb{Q}$-rational solutions. Since y takes different values at these, we can used $f = y$ in Proposition 2.7.

We state without proof that the minimal polynomial of L_y is $\mu(T) = T(T-1)^2$. It follows that the primary components are

$$\begin{aligned} I_p &= (f_1, f_2, f_3, y) = \left(x^2, y\right) \\ I_q &= \left(f_1, f_2, f_3, (y-1)^2\right) \\ &= \left(x^2 + 2(y-1), x(y-1), (y-1)^2\right). \end{aligned}$$

The following, weaker statement can be proven analogously:

Corollary 2.9. *If $f \in \Bbbk[x_1, \dots, x_n]$ takes distinct values at all solutions, then*

$$\forall p \in V: \quad I_p = I + \left((f - f(p))^{n(p)}\right). \qquad \square$$

2.3 General Case

We are now in the general setting where not all solutions must be $\Bbbk$-rational. However, we do assume that $\Bbbk$ is a perfect field. Let $I = \bigcap_{i=1}^r I_i$ be a minimal primary decomposition of I over $\Bbbk$. Define $V_i := V(I_i) \subseteq \mathbb{K}^n$. Let $\mathbb{L} \subseteq \mathbb{K}$ be the smallest field such that $V \subseteq \mathbb{L}^n$, i.e.

$$\mathbb{L} := \Bbbk[\{\, \lambda \mid \exists p = (p_1, \dots, p_n) \in V : \exists i : p_i = \lambda \,\}]$$

We set $G := \mathrm{Gal}(\mathbb{L}/\Bbbk)$. Note that this is a finite group.

Definition 2.10. *For any $\sigma \in G$ and $p = (p_1, \dots, p_n) \in \mathbb{L}^n$, write*

$$\sigma(p) := (\sigma(p_1), \dots, \sigma(p_n)).$$

For $p \in V$, we note that $f_i(\sigma(p)) = \sigma(f_i(p)) = 0$, so the Galois group acts on V in the above way.

Proposition 2.11. *G acts transitively on V_i.*

Proof. Let $p \in V_i$ be a point corresponding to a maximal ideal $\mathfrak{m}_p \subset \mathbb{L}[x_1, \dots, x_n]$. Note that for any p, the ideal $\mathfrak{m}_p \cap \Bbbk[x_1, \dots, x_n]$ is a prime ideal containing I_i, therefore equal to $\mathfrak{p} := \sqrt{I_i}$. Assume that $\mathfrak{m}_p \neq \sigma(\mathfrak{m}_q)$ for any $q \in V_i$. By the Chinese Remainder Theorem (see [Bos06, 2.3, Satz 12]), there exists an $h \in \mathbb{L}[x_1, \dots, x_n]$ such that

$$h \equiv 0 \pmod{\mathfrak{m}_p} \qquad \text{and} \qquad \forall q \in V_i \setminus \{\, p \,\} : \forall \sigma \in G : \quad h \equiv 1 \pmod{\sigma(\mathfrak{m}_q)}.$$

Then, by the well-known fact [Bos06, 4.7, Satz 4] about norms and since $\Bbbk$ is perfect,

$$g := \prod_{\sigma \in G} \sigma(h) = N_{\mathbb{L}/\Bbbk}(h) \in \Bbbk$$

and since $\mathrm{id} \in G$, this means $g \in \Bbbk \cap \mathfrak{m}_p = \mathfrak{p}$. On the other hand, we can pick any $q \in V_i \setminus \{ p \}$ and see that $h \notin \sigma(\mathfrak{m}_q)$ for any $\sigma \in G$, hence $\sigma(h) \notin \mathfrak{m}_q$. Consequently, $g \notin \Bbbk \cap \mathfrak{m}_q = \mathfrak{p}$ is a contradiction. □

Fact 2.12. *$V_i \cap V_j = \emptyset$ for $i \neq j$.*

Proof. Assume the converse. Then, there exists a maximal ideal $P \subseteq \mathbb{L}[x_1, \ldots, x_n]$ which contains both I_i and I_j. Then, $P' := P \cap \Bbbk[x_1, \ldots, x_n]$ is a prime ideal which contains both I_i and I_j, contradicting minimality by $\sqrt{I_i} = P' = \sqrt{I_j}$. □

Theorem 2.13. *Let $\chi \in \Bbbk[T]$ be the characteristic polynomial of the multiplication map L_f for some polynomial f, which takes distinct values at all solutions. Then,*

$$\chi = \prod_{i=1}^{r} \chi_i^{k_i} \qquad \textit{for} \qquad \chi_i := \prod_{p \in V_i} (T - f(p))$$

is an irreducible factorization and the χ_i are distinct. The minimal primary decomposition $I = I_1 \cap \cdots \cap I_r$ satisfies $I_i = I + \left(\chi_i(f)^{k_i}\right)$.

Proof. We can write χ over $\mathbb{L}$ as

$$\chi = \prod_{p \in V} (T - f(p))^{n(p)} = \prod_{i=1}^{r} \prod_{p \in V_i} (T - f(p))^{n(p)}$$

Observe that χ has coefficients in $\Bbbk$ and so,

$$\begin{aligned}\prod_{i=1}^{r} \prod_{p \in V_i} (T - f(p))^{n(p)} = \chi = \sigma(\chi) &= \sigma\left(\prod_{i=1}^{r} \prod_{p \in V_i} (T - f(p))^{n(p)}\right) \\ &= \prod_{i=1}^{r} \prod_{p \in V_i} (T - f(\sigma(p)))^{n(p)}\end{aligned}$$

for all $\sigma \in G$. By Proposition 2.11, for any $p, q \in V_i$, we can find a $\sigma \in G$ with $\sigma(p) = q$ and conclude that $n(p) = n(q) =: k_i$ only depends on i. We are now going to show that χ_i is irreducible and has coefficients in $\Bbbk$. The latter follows because $\chi_i = \sigma(\chi_i)$ for all $\sigma \in G$ and this means that all coefficients of χ_i are in $\mathbb{L}^G = \Bbbk$. Irreducibility now follows from [Bos06, 4.3, Satz 1].

We now proceed similar to the proof of Proposition 2.7. By Lemma 2.5, it suffices to show that $g := \chi_i^{k_i} \in I_i$ and $g \notin P_j = \sqrt{I_j}$ for all $j \neq i$. Note that $g(q) \neq 0$ for all $q \in V_j$ because f takes distinct values on all solutions by assumption. Hence,

$$g \notin \mathfrak{m}_q \cap \Bbbk[x_1, \ldots, x_n] = P_j.$$

To prove that $g \in I_i$, let

$$h := \prod_{j \neq i} \chi_j^{k_j}.$$

Then, by definition $gh = \chi$. However, the Cayley-Hamilton Theorem [Eis94, Theorem 4.3] says that $\chi(L_f)$ is the zero map on A. Applied to 1, we obtain

$$0 = \chi(L_f)(1) = \chi(\overline{f}) = \overline{\chi(f)}.$$

Hence, $gh \in I \subseteq I_i$. Since $h(p) \neq 0$ by our assumption on f, we know $h \notin \mathfrak{m}_p$. Thus, no power of h can be contained in I_i, since $\sqrt{I_i} = \mathfrak{m}_p \cap \Bbbk[x_1, \ldots, x_n]$. Since I_i is primary, this means $g \in I_i$. $\square$

Remark 2.14. *Provided that we can find an appropriate f, Theorem 2.13 yields an algorithm for computing primary decompositions. If we choose a random, homogeneous linear polynomoial f, then it is suitable with very high probability. However, if we do not want to end up with a probabilitstic algorithm, we need a certificate for f to take distinct values at all solutions.*

- *If I is radical, then f takes distinct values at the solutions if and only if the characteristic polynomial χ of L_f has distinct roots. Hence, we only need to compute $\gcd(\chi, \chi')$ by [Bos06, 3.6, Lemma 1].*
- *If I is not radical, we can simply compute its radical and proceed as before.*

We can therefore choose a random f and check numerically if it is a suitable choice for Theorem 2.13.

Algorithm 2.15 (Primary Ideal Decomposition).

(a). Pick $f_0 = t_1 x_1 + \cdots + t_n x_n$ such that all eigenspaces of L_{f_0} have dimension one.

(b). Calculate the irreducible factorization of the characteristic polynomial of L_{f_0}.

(c). Use Theorem 2.13 to calculate generators for a minimal primary decomposition.

3 Galois Groups

Let $\mathbb{k}$ be an infinite field with algebraic closure $\mathbb{K} := \overline{\mathbb{k}}$ and consider a monic polynomial with distinct roots

$$f = \sum_{i=0}^{n} (-1)^i \cdot c_i \cdot T^{n-i} \in \mathbb{k}[T].$$

Definition 3.1. *The* ***elementary symmetric polynomials*** $\sigma_0, \ldots, \sigma_n \in \mathbb{k}[x_1, \ldots, x_n]$ *are defined by* $\sigma_0 = 1$ *and the identity*

$$\prod_{i=1}^{n} (T - x_i) = \sum_{i=0}^{n} (-1)^i \sigma_i T^{n-i}.$$

For the rest of this section, we also set $f_i := \sigma_i - c_i$ *and consider the associated algebra* $A = \mathbb{k}[x_1, \ldots, x_n]/I$, *where* $I = (f_1, \ldots, f_n)$. *We write* $s_i := \overline{\sigma_i} = \overline{c_i}$.

3.1 The Universal Property

Fact 3.2. *The polynomial* f *splits completely over* A.

Proof. We claim $f = \prod_i (T - s_i)$. Indeed,

$$f(s_j) = \sum_{i=0}^{n} (-1)^i c_i \overline{c_j}^{n-i} = \sum_{i=0}^{n} (-1)^i c_i \overline{\sigma_j}^{n-i} = \prod_{i=0}^{n} (\sigma_j - \sigma_i) = 0.$$

□

Proposition 3.3. *Suppose that* R *is a* $\mathbb{k}$*-algebra such that* $f = \prod_{i=1}^{n} (T - \alpha_i)$ *for certain* $\alpha_1, \ldots, \alpha_n \in R$. *Then, there exists a homomorphism* $\phi : A \to R$ *of* $\mathbb{k}$*-algebras which maps* $\phi(s_i) = \alpha_i$.

Proof. We define a map $\psi : \mathbb{k}[x_1, \ldots, x_n] \to R$ by $\psi(x_i) := \alpha_i$. We have to verify that $\ker(\psi) \supseteq I$ because this yields an induced map $\phi : A \to R$. Since

$$\begin{aligned} f(T) &= \prod_{i=1}^{n} (T - \psi(x_i)) = \psi\Big(\prod\nolimits_{i=1}^{n} (T - x_i)\Big) \\ &= \psi\Big(\sum\nolimits_{i=0}^{n} (-1)^i \sigma_i T^{n-i}\Big) = \sum_{i=0}^{n} (-1)^i \psi(\sigma_i) T^{n-i}, \end{aligned}$$

we know $\psi(\sigma_i) = c_i = \psi(c_i)$ by comparing coefficients, so $\psi(f_i) = \psi(\sigma_i - c_i) = 0$. □

3.2 The Dimension of A

Proposition 3.4. *There are $|Z(I)| = n!$ solutions and each of them has multiplicity 1. The coordinates of each solution are the roots of f in $\mathbb{K}$. The symmetric group acts on the solutions by permuting coordinates. In particular, $\dim_{\Bbbk}(A) = n!$.*

Proof. Let $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ be the distinct roots of f in the algebraic closure of $\Bbbk$. The point $p = (\beta_1, \ldots, \beta_n) \in \mathbb{K}^n$ satisfies $\sigma_i(p) = c_i$ for all i if and only if

$$\prod_{i=1}^{n}(T - \beta_i) = \prod_{i=1}^{n}(T - x_i(p)) = \sum_{i=0}^{n}(-1)^i \sigma_i(p) T^{n-i} = \sum_{i=0}^{n}(-1)^i c_i T^{n-i} = f.$$

Any point with coordinates equal to some permutation of the α_i is therefore a solution. By Bézout (Theorem 1.1) and since $\deg(f_i) = i$, there are at most $n!$ many solutions counted with multiplicity, so each solution has multiplicity 1 and $\dim_{\Bbbk}(A) = n!$. □

Remark 3.5. *We note that the symmetric group S_n acts on $\Bbbk[x_1, \ldots, x_n]$ by permuting the variables, i.e. $\pi(x_i) := x_{\pi(i)}$ for $\pi \in S_n$. Since the elementary symmetric polynomials are invariant under this action, so are the f_i. Thus, we get an induced action $S_n \times A \to A$.*

3.3 The Emergence of Splitting Fields

Let $f_0 \in \Bbbk[x_1, \ldots, x_n]$ be a linear polynomial which takes distinct values at all $p \in Z(I)$ and let $\chi \in \Bbbk[T]$ be the characteristic polynomial of the map L_{f_0}. Since $\Bbbk$ is infinite, note that we can always find such an f_0. In fact, a generic homogeneous linear polynomial satisfies this condition.

Fact 3.6. *The eigenspaces of L_{f_0} are all one-dimensional. In particular, χ is the minimal polynomial of L_{f_0}.*

Proof. This follows because $\chi = \prod_{p \in V}(T - f_0(p))$. □

Lemma 3.7. *There is an algebra isomorphism $\Bbbk[T]/(\chi) \cong A$.*

Proof. Consider the projection $\pi : \Bbbk[T] \to A$ defined by $h \mapsto \overline{h(f_0)}$. We know that

$$h \in \ker(\pi) \quad \Longleftrightarrow \quad h(f_0) \in I \quad \Longleftrightarrow \quad h(L_{f_0}) = 0$$

The minimal polynomial μ of L_{f_0} is the nonzero polynomial of smallest degree with $\mu(L_{f_0}) = 0$. Hence, $\ker(\pi)$ is generated by μ and we get an injective morphism

$$\Bbbk[T]/(\mu) \hookrightarrow A.$$

By [Bos06, 3.2, Satz 6], Fact 3.6 and Proposition 3.4,

$$\dim_{\Bbbk}(\Bbbk[T]/(\mu)) = \deg(\mu) = \deg(\chi) = n! = \dim_{\Bbbk}(A).$$ □

Definition 3.8. *Let $\chi = \prod_{i=1}^{r} \chi_i$ be the irreducible factors of χ. They are distinct by Fact 3.6 and we define $\Bbbk_i := \Bbbk[T]/(\chi_i)$. Observe that $A \cong \Bbbk[T]/(\chi) \cong \prod_{i=1}^{r} \Bbbk_i$ by Lemma 3.7 and the Chinese Remainder Theorem.*

Remark 3.9. *For everyone who has lost track, let us understand what a permutation $\pi \in S_n$ does on an element of A, understood as the residue class of a polynomial $h \in \Bbbk[T]$. The isomorphism $\Bbbk[T]/(\chi) \cong A$ constructed in Lemma 3.7 is induced by mapping h to $h(f_0)$. Hence, applying π to h means to permute the variables of $h(f_0)$ and taking its residue class.*

Fact 3.10. *Let $\pi \in S_n$. For all i, there exists some j with $\pi(\Bbbk_i) = \Bbbk_j$.*

Proof. Since π induces an automorphism, $\pi(\Bbbk_i) \cap \pi(\Bbbk_j) = \pi(\Bbbk_i \cap \Bbbk_j) = \{ 0 \}$, therefore we know $\prod_i \Bbbk_i \cong \pi(\prod_i \Bbbk_i) = \prod_i \pi(\Bbbk_i)$ and the statement follows. $\square$

We state the following theoretical result without proof:

Proposition 3.11. *The symmetric group S_n acts transitively on the set $\{ \Bbbk_1, \ldots, \Bbbk_r \}$. Furthermore we have isomorphisms*

$$\operatorname{Gal}(\Bbbk_i/\Bbbk) \cong G_i := \{ \pi \in S_n \mid \pi(\Bbbk_i) = \Bbbk_i \}.$$

We now obtain a quite inefficient algorithm to compute the Galois group of f:

Algorithm 3.12 (Calculating Galois Groups).

(a). *Use Algorithm 2.15 to compute polynomials χ_i such that $I = \bigcap_{i=1}^{r} I_i$ is the minimal primary decomposition and $I_i = I + (\chi_i(f_0))$.*

(b). *Using the method of section 1.1.3[1], calculate*

$$\operatorname{Gal}(\Bbbk_i/\Bbbk) = \{ \pi \in S_n \mid \pi(I_i) = I_i \} = \{ \pi \in S_n \mid \pi(\chi_i) \in I_i \}.$$

[1]Mostly to avoid the devilish Gröbner bases.

References

[Bos06] Siegfried Bosch. *Algebra (Sixth Edition)*. Springer, Berlin, 2006.

[CLO98] D. Cox, J. Little, and D. O'Shea. *Using algebraic geometry*. Springer, New York, 1998. Volume 185 of Graduate Texts in Mathematics.

[Eis94] David Eisenbud. *Commutative Algebra with a View Toward Algebraic Geometry*. Springer, New York, 1994.

[Har06] Robin Hartshorne. *Algebraic Geometry*. Springer, New York, 2006.

[Jou91] J.-P. Jouanolou. Le formalisme du résultant. *Adv. Math.*, 90:117–263, 1991.