

# REAL ALGEBRAIC GEOMETRY

KENNETH R. DRIESSEL

## HERMITE'S ROOT COUNTING METHOD

*References:* In this section I mainly follow the books Basu, Pollack and Roy(2003) and Cox, Little and O'Shea(1991,1998) .

*Notation:* Throughout this section we use the following notation:  $K$  denotes a field with characteristic 0,  $R$  denotes a real closed extension of  $K$  and  $C$  denotes an algebraically closed extension of  $R$ . In particular, we have  $K \subseteq R \subseteq C$ .

The standard example is  $K := \mathbb{Q}$ ,  $R := \mathbb{R}$ ,  $C := \mathbb{C}$ . (Here is another choice for  $R$  and  $C$ : Take  $R$  to be the real algebraic numbers and take  $C$  to be the algebraic numbers.)

This section is divided into three subsections with titles "Introduction", "The univariate case" and " The multivariate case".

**Subsection: Introduction.** We consider the following

**Problem:** Given a finite set  $\mathcal{P}$  of polynomials and a polynomial  $Q$  in  $K[X_1, \dots, X_k]$ , determine the quantity

$$\#\{x \in \text{Zero}(\mathcal{P}, R^k) : Q(x) > 0\} - \#\{x \in \text{Zero}(\mathcal{P}, R^k) : Q(x) < 0\}$$

where  $\text{Zero}(\mathcal{P}, R^k) := \{x \in R^k : (\forall P \in \mathcal{P}) P(x) = 0\}$ .

The difference between the number of solutions  $\text{Zero}(\mathcal{P}, R^k)$  at which  $Q(x)$  is positive and the number at which  $Q(x)$  is negative is called the **Tarski query** or **Sturm query** of  $Q$  for  $\mathcal{P}$ . We shall use  $\text{TaQ}(Q, \mathcal{P})$  to denote this difference. If we can compute Tarski queries then we can compute many other useful numbers. Here are a few examples:

*Example:* The number of solutions of  $\mathcal{P}$  in  $R^k$ .

Let  $Q(X_1, \dots, X_k) := 1$  be the constant one polynomial. Then the number of solutions of  $\mathcal{P}$  in  $R^k$  equals the Tarski query of  $Q$  for  $\mathcal{P}$ ; in symbols,

$$\#\text{Zero}(\mathcal{P}, R^k) = \text{TaQ}(1, \mathcal{P}).$$

*Example:* The number of solutions of  $\mathcal{P}$  in  $R^k$  at which  $Q$  is positive.

Note

$$\begin{aligned} & \#\{x \in \text{Zero}(\mathcal{P}, R^k) : Q(x) > 0\} + \#\{x \in \text{Zero}(\mathcal{P}, R^k) : Q(x) < 0\} \\ &= \text{TaQ}(1, \mathcal{P}) - \text{TaQ}(1, \mathcal{P} \cup \{Q\}). \end{aligned}$$

Hence

$$\begin{aligned} & 2 \#\{x \in \text{Zero}(\mathcal{P}, R^k) : Q(x) > 0\} \\ &= \text{TaQ}(Q, \mathcal{P}) + \text{TaQ}(1, \mathcal{P}) - \text{TaQ}(1, \mathcal{P} \cup \{Q\}). \end{aligned}$$

*Example:* The approximate location of solutions in  $R^k$ .

We can use the polynomials which define spheres to approximate the location of solutions. For example, if  $k = 2$ , we can use

$$Q(X, Y) := r^2 - (X - a)^2 - (Y - b)^2$$

where  $a, b$  and  $r$  are parameters. Then

$$\#\{(x, y) \in \text{Zero}(\mathcal{P}, R^2) : Q(x, y) > 0\}$$

equals the number of solutions in the ball

$$\{(x, y) \in R^2 : (x - a)^2 + (y - b)^2 < r^2\}.$$

**Definition:** If the set  $\text{Zero}(\mathcal{P}, C^k)$  of solutions in  $C^k$  is finite then we say that  $\mathcal{P}$  is **zero-dimensional**.

We shall usually assume that we are in this zero-dimensional setting. Recall that we can algebraically determine if we are in this setting by means of the following result.

*Notation:* The ideal in  $K[X_1, \dots, X_k]$  generated by a set of polynomials  $\mathcal{P}$  in  $K[X_1, \dots, X_n]$  is denoted  $\text{Ideal}(\mathcal{P}, K)$

**Proposition 1. Finiteness.** Let  $A := K[X_1, \dots, X_k] / \text{Ideal}(\mathcal{P}, K)$ . Then

- The  $K$ -vector space  $A$  is finite dimensional if and only if  $\mathcal{P}$  is zero-dimensional.
- If the dimension of  $A$  is finite then the number of solutions of  $\mathcal{P}$  in  $C^k$  is less than or equal to this dimension.

*Proof.* See Cox, Little and O’Shea(1991,1998) or Basu, Pollack and Roy(2003) for a proof.  $\square$

*Notation:* Let  $A$  be a finite dimensional algebra over the field  $K$ . For  $f$  in  $A$ , let  $L(f) : A \rightarrow A$  be the linear **multiplication map** defined by  $L(f)(g) := fg$ . For  $q \in A$ , let  $B(q)$  denote the quadratic form on  $A$  defined by  $B(q)(f) := \text{Trace}(L(qf^2))$ . The notation used for this quadratic form is somewhat different in the polynomial setting: Let  $\mathcal{P}$  be a finite subset of  $K[X_1, \dots, X_n]$ , let  $A := K[X_1, \dots, X_n]/\text{Ideal}(\mathcal{P}, K)$  and let  $Q$  be an element of  $K[X_1, \dots, X_n]$ . Then  $\text{Herm}(\mathcal{P}, Q)$  denotes the quadratic form on  $A$  (assumed to be finite dimensional) defined by  $\text{Herm}(\mathcal{P}, Q)(f) := \text{Trace}(L(Qf^2))$  and this form is called the **Hermite quadratic form determined by  $\mathcal{P}$  and  $Q$** . We shall use the expressions  $\text{Rank}(\text{Herm}(\mathcal{P}, Q))$  and  $\text{Signature}(\text{Herm}(\mathcal{P}, Q))$  to denote the rank and signature of this quadratic form.

The following result provides an answer to the problem raised above.

**Proposition 2. Hermite’s root counting.** *Let  $\mathcal{P}$  be a finite set of polynomials which have a zero-dimensional solution set. Then the rank of the quadratic form  $\text{Herm}(\mathcal{P}, Q)$  equals the number of distinct roots of  $\mathcal{P}$  in  $C^k$  at which  $Q$  is different than zero; in symbols,*

$$\text{Rank}(\text{Herm}(\mathcal{P}, Q)) = \#\{x \in \text{Zero}(\mathcal{P}, C^k) : Q(x) \neq 0\}.$$

*The signature of the quadratic form  $\text{Herm}(\mathcal{P}, Q)$  equals the difference between the number of distinct roots of  $\mathcal{P}$  in  $R^k$  at which  $Q$  is positive and the number of distinct roots in  $R^k$  at which  $Q$  is negative; in symbols,*

$$\begin{aligned} \text{Signature}(\text{Herm}(\mathcal{P}, Q)) = \\ \#\{x \in \text{Zero}(\mathcal{P}, R^k) : Q(x) > 0\} - \#\{x \in \text{Zero}(\mathcal{P}, R^k) : Q(x) < 0\}. \end{aligned}$$

**\*\*\*TODO:** Perhaps insert a simple example here.

The following subsections provide proofs of this result. In particular, the next subsection provides a proof in the univariate case. And the following subsection deals with the multivariate case. The analyses of the two cases are similar. We use the following result (which we proved in an earlier section) in both cases.

**Proposition 3. Decomposition using idempotents.** *Let  $A$  be a commutative ring with identity. Let  $e_1, \dots, e_n$  be elements of  $A$  which satisfy the following conditions:*

- $e_1 + e_2 + \dots + e_n = 1$  and
- if  $i \neq j$  then  $e_i e_j = 0$ .

*Then*

- $e_i^2 = e_i$ ,
- $e_i A$  is a subring of  $A$  with identity, namely,  $e_i$ , and
- $A$  is the direct sum of the subrings  $e_i A$ :

$$A = e_1 A \oplus e_2 A \oplus \dots \oplus e_n A.$$

Furthermore, for  $f \in A$ ,

$$L(f) = L(e_1 f) \oplus \dots \oplus L(e_n f)$$

and, for  $q \in A$ ,

$$B(q) = B(e_1 q) \oplus \dots \oplus B(e_n q).$$

**Subsection: The univariate case.** In this subsection we shall consider the quotient algebra  $K[X]/I$  where  $I$  is an ideal in the univariate polynomial ring  $K[X]$ .

**Proposition 4. Decomposition of a univariate quotient algebra using idempotents.** *Let  $P$  be a polynomial in the ring  $C[X]$  with distinct roots  $x_1, \dots, x_n$ . Let  $\text{Ideal}(P, C)$  be the ideal generated by  $P$ . Then there exist elements  $e_1, \dots, e_n$  in the quotient ring  $C[X]/\text{Ideal}(P, C)$  such that*

- $e_1 + e_2 + \dots + e_n = 1$ ,
- if  $i \neq j$  then  $e_i e_j = 0$ ,
- $e_i^2 = e_i$ ,
- if  $i \neq j$  then  $e_i(x_j) = 0$
- $e_i(x_i) = 1$ ,
- $A = e_1 A \oplus e_2 A \oplus \dots \oplus e_n A$ , and
- $\dim(e_i A) = \mu_i$  where  $\mu_i$  is the multiplicity of  $x_i$  as a root of  $P$ .

For  $f \in A$ , let  $L(f)$  denote the linear multiplication map on  $A$  defined by  $L(f)(g) := fg$ . Then

$$L(f) = L(e_1 f) \oplus \dots \oplus L(e_n f).$$

For  $q \in A$  let  $B(q)$  be the bilinear function defined on  $A$  by  $B(q)(g, h) := \text{Trace}(L(qgh))$ . Then

$$B(q) = B(e_1 q) \oplus \dots \oplus B(e_n q).$$

**Definition:** The element  $e_i$  is called the **idempotent associated with the root  $x_i$** .

*Proof.* We have

$$P(X) = (X - x_1)^{\mu_1} \dots (X - x_n)^{\mu_n}.$$

If  $n = 1$  we can simply take  $e_1$  to be the constant 1 polynomial. The assertions of the proposition are then obvious.

So we assume that  $n \geq 2$ . For  $i = 1, \dots, n$ , let  $S_i(X)$  be the Lagrange interpolating polynomial which is defined to be the following product:

$$S_i(X) := \prod \left\{ \frac{X - x_j}{x_i - x_j} : j \neq i \right\}.$$

Note  $S_i(x_i) = 1$  and, for  $j \neq i$ ,  $S_i(x_j) = 0$ . For  $i = 1, \dots, n$ , let  $T_i := S_i^\mu$  where  $\mu$  is the maximum of the multiplicities  $\mu_1, \dots, \mu_n$ . Note that  $T_i(x_i) = 1$  and, for  $i \neq j$ , the polynomial  $P$  divides  $T_i T_j$ . Since the  $T_i$  are relatively prime, there are polynomials  $U_i$  such that

$$U_1 T_1 + \dots + U_n T_n = 1.$$

Set  $e_i := U_i T_i + \text{Ideal}(P)$ . It is easy to see that these elements of  $A$  satisfy the assertions of the proposition. Also note that the assertion that  $e_i$  is an idempotent follows from the first two assertions concerning the  $e_i$  as in the general idempotent decomposition result for rings. This general result also provides a proof of the decomposition assertion given here.

TODO: Find a nice proof of the assertion relating dimensions and multiplicities. □

\*\*\*TODO: Include the alternative proof of the last result.

**Example:** This example illustrates the constructions of the last proof. Let  $P(X) := X(X - 1)$ . Note  $A := \mathbb{Q}[X]/\text{Ideal}(P, \mathbb{Q})$  has dimension two. The roots of  $P$  are 0 and 1. We shall use the roots as indices. The Lagrange interpolating polynomials are:

- $S_0(X) := 1 - X$  and
- $S_1(X) := X$ .

We can take  $T_0 := S_0$  and  $T_1 := S_1$ . Note:

$$(1 - X) + X = 1.$$

Consequently we take  $e_0(X) := 1 - X$  and  $e_1(X) := X$ . Then  $e_0 + e_1 = 1$ ,  $e_0 e_1 = 0$ ,  $e_0(0) = 1$ ,  $e_0(1) = 0$ ,  $e_1(0) = 0$  and  $e_1(1) = 1$ . Here is the multiplication table for  $A$  for the basis  $e_1, e_2$ :

·	e <sub>0</sub>	e <sub>1</sub>
e <sub>0</sub>	e <sub>0</sub>	0
e <sub>1</sub>	0	e <sub>1</sub>

We have  $e_0 A = e_0 \text{Span}(e_0, e_1) = \text{Span}(e_0)$  and  $e_1 A = e_1 \text{Span}(e_1, e_0) = \text{Span}(e_1)$ .

**Example:** Here is another example which illustrates the constructions of the last proof. Let  $P(X) := X^2(X - 1)$ . Note  $A := \mathbb{Q}[X]/\text{Ideal}(P, \mathbb{Q})$  has dimension three. The roots of  $P$  are 0 and 1. We shall use the roots as indices. The Lagrange interpolating polynomials are:

- $S_0(X) := 1 - X$  and
- $S_1(X) := X$ .

We can take  $T_0 := S_0$  and  $T_1 := S_1^2$ . Note:

$$(1 + X)(1 - X) + X^2 = 1.$$

Consequently we take  $e_0(X) := (1 + X)(1 - X)$  and  $e_1(X) := X^2$ . Then  $e_0 + e_1 = 1$ ,  $e_0 e_1 = 0$ ,  $e_0(0) = 1$ ,  $e_0(1) = 0$ ,  $e_1(0) = 0$  and  $e_1(1) = 1$ . Here is

the multiplication table for  $A$  for the basis  $1, X, X^2$ :

$\cdot$		$1$	$X$	$X^2$
$1$		$1$	$X$	$X^2$
$X$		$X$	$X^2$	$X^2$
$X^2$		$X^2$	$X^2$	$X^2$

We compute  $e_0A = \text{Span}(e_0, Xe_0)$  and  $e_1A = \text{Span}(e_1)$ . Here is the multiplication table for  $A$  for the basis  $e_0, Xe_0, e_1$ :

$\cdot$		$e_0$	$Xe_0$	$e_1$
$e_0$		$e_0$	$Xe_0$	$0$
$Xe_0$		$Xe_0$	$0$	$0$
$e_1$		$0$	$0$	$e_1$

**Proposition 5. Eigenvalues of a multiplication map.** *Let  $P$  be a polynomial in  $K[X]$  with roots  $x_1, \dots, x_n$  in  $C$ . Let  $A := K[X]/\text{Ideal}(P, K)$  and let  $f$  be an element of  $A$ . Then the linear multiplication map  $L(f)$  on  $A$  has eigenvalues  $f(x_i)$  for  $i = 1, \dots, n$ . Furthermore, the multiplicity of  $f(x_i)$  as a root of the characteristic polynomial of  $L(f)$  equals the multiplicity of  $x_i$  as a root of the polynomial  $P$ .*

*Proof.* Let  $e_1, \dots, e_n$  be idempotents for  $A$  associated with the roots of  $P$ . We consider the elements  $f_i := e_i(f - f(x_i))$  of  $A$ . Since  $f_i$  vanishes at all of the roots of  $P$ , there exists a natural number  $m$  such that  $P$  divides  $f_i^m$ . It follows that the linear map  $L(f_i) : e_iA \rightarrow e_iA$  is nilpotent. Hence  $L(f_i)$  has a unique eigenvalue  $0$  with multiplicity  $\dim(e_iA) = \mu_i$ . It follows that  $L(e_i f)$  has exactly one eigenvalue  $f(x_i)$  with multiplicity  $\mu_i$ . We can use the decomposition proposition to complete the proof.  $\square$

**Proposition 6. Stickelberger's theorem in the univariate case.** *Let  $P$  be an polynomial in  $K[X]$  with roots  $x_1, \dots, x_n$  in  $C$  with multiplicities  $\mu_1, \dots, \mu_n$  respectively. Let  $A := K[X]/\text{Ideal}(P, K)$ , let  $f$  be an element of  $A$  and let  $L(f) := A \rightarrow A : g \mapsto fg$  be the multiplication map associated with  $f$ . Then*

$$\text{Trace}(L(f)) = \sum_{i=1}^n \mu_i f(x_i).$$

Note that this trace is an element of  $K$ . In particular, note that the sum is a symmetric function of the roots of  $P$ .

*Proof.* This result follows immediately from the result concerning the eigenvalues of  $L(f)$ . In particular, there is a choice of basis for  $C[X]/\text{Ideal}(P, C)$  for which the matrix of  $L(f)$  is diagonal.  $\square$

**Example:** Let

$$P(X) := (X^2 + 1)(X - 2)^2(X + 2) = X^5 - 2X^4 - 3X^3 + 6X^2 - 4X + 8.$$

Let  $A := \mathbb{Q}[X]/\text{Ideal}(P, \mathbb{Q})$ . We take the polynomials  $1, X, X^2, X^3, X^4$  as an ordered basis of  $A$ . We consider the multiplication map  $L(X)$ . Note

$X \cdot X^i = X^{i+1}$  for  $i = 0, 1, 2, 3$  and  $X \cdot X^4 \equiv 2X^4 + 3X^3 - 6X^2 + 4X - 8$ . Hence the matrix of  $L(X)$  with respect to the given basis is

$$M := \begin{pmatrix} 0 & 0 & 0 & 0 & -8 \\ 1 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & -6 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

This matrix is the companion matrix of the polynomial  $P$ . The eigenvalues of  $L(X)$  are the roots of  $P$ :

$$x_1 = -2, x_2 = 2, x_3 = i, x_4 = -i$$

with multiplicities:

$$\mu_1 = 1, \mu_2 = 2, \mu_3 = 1, \mu_4 = 1.$$

Note  $\text{Trace}(L(X)) = 2 = -2 + 2 + 2 + i - i$ .

If  $f = f_0 + f_1X + f_2X^2 + f_3X^3 + f_4X^4$  is a general element of  $A$  then  $L(f) = f_0I + f_1L(X) + f_2L(X)^2 + f_3L(X)^3 + f_4L(X)^4$  and the matrix of  $L(f)$  is  $f_0I + f_1M + f_2M^2 + f_3M^3 + f_4M^4$ . The eigenvalues of  $L(f)$  are  $f(x_i)$  with multiplicities  $\mu_i$ .

**Subsection: The zero-dimensional multi-variate case.** In this subsection we consider a quotient algebra  $K[X_1, \dots, X_k]/I$  where  $I$  is an ideal in the polynomial ring  $K[X_1, \dots, X_k]$ . We considered the univariate case earlier. Here we consider the multivariate case. Of course, the multivariate case includes the univariate case but the analysis of the multivariate case is somewhat more complicated.

**Definition:** Let  $V$  be a vector space over field  $K$ . Let  $Z := \{z_1, \dots, z_n\}$  be a finite subset of  $V$ . Let  $V' := (V \rightarrow K)$  denote the space of linear functionals on  $V$ ; in other words,  $V'$  denotes the dual space of  $V$ . Then an element  $a$  of  $V'$  is **separating for  $Z$**  if  $a$  has distinct values at the distinct elements of  $Z$ ; in symbols,  $i \neq j \implies a(z_i) \neq a(z_j)$ .

**Proposition 7. Separating a finite subset of a vector space.** *Let  $V$  be a finite dimensional vector space over a field  $K$  with characteristic zero and let  $Z$  be a finite subset of  $V$ . Then there is a separating linear functional for  $Z$ .*

*Proof.* This result is obvious since  $V'$  has infinitely many elements. Nevertheless, here is a “constructive” proof.

Let  $k$  be the dimension of  $V$ . Let  $v_1, \dots, v_k$  be a basis for  $V$  and let  $w_1, \dots, w_k$  be a dual basis for  $V'$ ; in other words,  $w_i(v_j) = \delta_{ij}$ , where  $\delta$  is the Kronecker symbol. Let  $n$  be the number of elements in  $Z$ . For  $0 \leq i \leq (k-1)\binom{n}{2}$ , let

$$a_i := w_1 + iw_2 + i^2w_3 + \dots + i^{k-1}w_k.$$

Note there are  $(k-1)\binom{n}{2} + 1$  linear functionals in this list. For distinct  $x$  and  $y$  in  $Z$  let

$$E(x, y) := \{a_i : a_i(x) = a_i(y)\}.$$

In other words, the set  $E(x, y)$  is the set of  $a_i$  which do not separate  $x$  and  $y$ . Note that  $a_i(x) = a_i(y)$  iff

$$w_1(x-y) + iw_2(x-y) + \dots + i^{k-1}w_k = 0.$$

Since the polynomial

$$P(T) := w_1(x-y) + w_2(x-y)T + \dots + w_k T^{k-1}$$

has at most  $k-1$  roots, we see that  $E(x, y)$  has at most  $k-1$  elements.

Finally, since the number of distinct pairs in  $Z$  is  $\binom{n}{2}$ , we have

$$\# \cup \{E(x, y) : x \in Z \wedge y \in Z \wedge x \neq y\} \leq (k-1)\binom{n}{2}.$$

□

**Definition:** Let  $\mathcal{P}$  be a zero-dimensional set of polynomials in  $K[X_1, \dots, X_k]$ . Then an element  $a$  in  $A := K[X_1, \dots, X_k]/\text{Ideal}(\mathcal{P}, K)$  is **separating for  $\mathcal{P}$**  if  $a$  has distinct values at the distinct roots of  $\mathcal{P}$  in  $C^k$ .

The next proposition says that there is a linear polynomial which is separating for  $\mathcal{P}$  if the number of roots of  $\mathcal{P}$  in  $C^k$  is finite.

**Proposition 8.** *Let  $\mathcal{P} \subseteq K[X_1, \dots, X_k]$  be a set of polynomials with a finite solution set  $\text{Zero}(\mathcal{P}, C^k)$ . If this solution set has  $n$  elements then there is an integer between 0 and  $(k-1)\binom{n}{2}$  such that the linear polynomial*

$$a_i := X_1 + iX_2 + i^2X_3 + \dots + i^{k-1}X_k$$

*is separating for  $\mathcal{P}$ .*

*Proof.* Note that each of the  $a_i$  may be regarded as an element the dual space  $(C^k)'$ . Apply the separating result for vector spaces. See also the proof of that result. □

**Proposition 9. Independence of powers of a separating element.**

*Let  $A := K[X_1, \dots, X_k]/\text{Ideal}(\mathcal{P}, K)$  and assume  $n := \#\text{Zero}(\mathcal{P}, C^k)$  is finite. If  $a \in A$  is a separating element then the elements  $1, a, a^2, \dots, a^{n-1}$  are linearly independent in  $A$ .*

*Proof.* Assume  $0 = c_0 + c_1a + \dots + c_{n-1}a^{n-1}$  in  $A$  where the  $c_i$  are in  $K$ . Consider the polynomial  $b := c_0 + c_1a + \dots + c_{n-1}a^{n-1}$  in  $K[X_1, \dots, X_k]$ . Note that  $b \in \text{Ideal}(\mathcal{P}, K)$ . Let  $x_1, \dots, x_n$  be the distinct roots of  $\mathcal{P}$  in  $C^k$ . Note that for every  $x$  in this list we have  $b(x) = 0$ . It follows that the univariate polynomial  $p(T) := c_0 + c_1T + \dots + c_{n-1}T^{n-1}$  has  $n$  distinct roots (namely,  $a(x_1), a(x_2), \dots, a(x_{n-1})$ ). Hence  $p(T)$  must be the zero polynomial, that is,  $c_0 = c_1 = \dots = c_{n-1} = 0$ . □

In the proof of the following decomposition result we use two versions of Hilbert's Nullstellensatz. In particular, we use the following well-known results.

**Proposition 10. Weak form of Hilbert's Nullstellensatz.** *Let  $\mathcal{P} := \{P_1, \dots, P_s\}$  be a subset of  $K[X_1, \dots, X_k]$ . Then  $\text{Zero}(\mathcal{P}, C^k) = \emptyset$  if and only if there exist  $A_1, \dots, A_s$  in  $K[X_1, \dots, X_k]$  such that  $A_1P_1 + \dots + A_sP_s = 1$ .*

**Proposition 11. Hilbert's Nullstellensatz.** *Let  $\mathcal{P}$  be a finite subset of  $K[X_1, \dots, X_k]$ . If  $P \in K[X_1, \dots, X_k]$  vanishes on  $\text{Zero}(\mathcal{P}, C^k)$  then there is a natural number  $m$  such that  $P^m$  is in the ideal  $\text{Ideal}(\mathcal{P}, K)$  generated by  $\mathcal{P}$ .*

**Proposition 12. Decomposition using idempotents.** *Let  $\bar{A}$  denote the algebra  $C[X_1, \dots, X_k]/\text{Ideal}(\mathcal{P}, C)$ . Assume that this algebra is finite dimensional (equivalently, that the set  $\text{Zero}(\mathcal{P}, C^k)$  is finite). Then there is a map  $e : \text{Zero}(\mathcal{P}, C^k) \rightarrow \bar{A}$  with the following properties:*

- $\sum \{e_x : x \in \text{Zero}(\mathcal{P}, C^k)\} = 1$ .
- For all  $x, y$  in  $\text{Zero}(\mathcal{P}, C^k)$ ,  $x \neq y \implies e_x e_y = 0$
- For all  $x$  in  $\text{Zero}(\mathcal{P}, C^k)$ ,  $e_x(x) = 1$ .
- For all  $x, y$  in  $\text{Zero}(\mathcal{P}, C^k)$ ,  $x \neq y \implies e_x(y) = 0$ .
- $e_x^2 = e_x$
- $\bar{A} = \bigoplus \{e_x \bar{A} : x \in \text{Zero}(\mathcal{P}, C^k)\}$ .

*Proof.* To simplify notation let  $Z := \text{Zero}(\mathcal{P}, C^k)$

*Claim:* There is a map  $s : Z \rightarrow \bar{A}$  with the following properties: For every  $x \in Z$ ,  $s_x(x) = 1$  and for all  $x, y \in Z$ ,  $x \neq y \implies s_x(y) = 0$ .

We may assume the  $X_1$  is separating. (If it is not then perform an appropriate linear change of coordinates.) We define  $s_x$  by means of the Lagrange interpolation polynomial:

$$s_x := \prod \left\{ \frac{X_1 - y_1}{x_1 - y_1} : y \neq x \wedge y \in Z \right\}.$$

It is easy to see that the map  $s$  defined in this way satisfies the assertions of the claim.

*Claim:* There is a map  $t : Z \rightarrow \bar{A}$  with the following properties: For every  $x \in Z$ ,  $t_x(x) = 1$  and for all  $x, y \in Z$ ,  $x \neq y \implies t_x t_y = 0$ .

Consider any  $x$  and  $y$  in  $Z$  with  $x \neq y$ . Note that for all  $z \in Z$ ,  $s_x s_y(z) = 0$ . Hence, by Hilbert's Nullstellensatz, there is a natural number  $m(y)$  such that  $(s_x s_y)^{m(y)} = 0$ . Let  $n(x)$  denote the maximum of these natural numbers  $m(y)$ ; in symbols,  $n(x) := \max\{m(y) : y \neq x \wedge y \in Z\}$ . Take  $t_x = s_x^{n(x)}$ . It is easy to see that the map  $t$  satisfies the assertions of the claim.

We complete the proof of the proposition as follows. Note that the  $t_x$  are relatively prime. Hence, by the weak form of Hilbert's Nullstellensatz, there are polynomials  $u_x$  such that  $1 = \sum \{u_x t_x : x \in Z\}$ . Let  $e_x := u_x t_x$ . It is easy to see that these elements of  $\bar{A}$  satisfy the assertions of the proposition. (In particular, that these  $e_x$  are idempotents, follows from their

orthogonality as in the abstract decomposition result. And the direct sum decomposition also follows as in that abstract result.)  $\square$

**Proposition 13. Eigenvalues of multiplication maps.** *Let  $\mathcal{P}$  be a finite set of polynomials in  $K[X_1, \dots, X_k]$ . Assume that the algebra  $A := K[X_1, \dots, X_k]/\text{Ideal}(\mathcal{P}, K)$  is finite dimensional. Let  $f$  be an element of  $A$ . Then the eigenvalues of the linear multiplication map  $L(f) : A \rightarrow A$  are the values  $f(x)$ ,  $x \in \text{Zero}(\mathcal{P}, C^k)$ , with multiplicities  $\mu(x) := \dim(e_x \bar{A})$ , where  $e$  and  $\bar{A}$  are as in the last decomposition result.*

*Proof.* Note that  $e_x(f - f(x))$  vanishes on the set of solutions  $\text{Zero}(\mathcal{P}, C^k)$ . By Hilbert's Nullstellensatz,  $(e_x(f - f(x)))^m = 0$  in  $\bar{A}$  for some natural number  $m$ . Hence  $L(e_x(f - f(x)))$  is a nilpotent linear map with eigenvalue 0 and multiplicity  $\mu(x)$ . It follows that

$$L(e_x f) = L(f(x)I(e_x \bar{A})) = \dim(e_x \bar{A})f(x)$$

where  $I(e_x \bar{A})$  is the identity map on  $e_x \bar{A}$ .  $\square$

**Proposition 14. Stickelberger.** *Let  $A := K[X_1, \dots, X_k]/\text{Ideal}(\mathcal{P}, K)$  where  $\mathcal{P}$  is a finite set of polynomials in  $K[X_1, \dots, X_k]$ . Let  $f$  be an element of  $A$  and let  $L(f) : A \rightarrow A$  be the multiplication map determined by  $f$ . Let  $e_x$  for  $x \in \text{Zero}(\mathcal{P}, C^k)$  be orthogonal idempotents for  $\bar{A}$ . Then*

$$\text{Trace } L(f) = \sum \{\mu(x)f(x) : x \in \text{Zero}(\mathcal{P}, C^k)\}$$

where  $\mu(x) = \dim(e_x \bar{A})$ .

*Proof.* This result follows immediately from the eigenvalue result. In particular, recall that the trace of a linear map equals the sum of its eigenvalues.  $\square$

**Proposition 15. Hermite's root counting.** *Let  $\mathcal{P}$  be a finite set of polynomials which have a zero-dimensional solution set. Then the rank of the quadratic form  $\text{Herm}(\mathcal{P}, Q)$  equals the number of distinct roots in  $C^k$  of  $\mathcal{P}$  at which  $Q$  is different than zero; in symbols,*

$$\text{Rank}(\text{Herm}(\mathcal{P}, Q)) = \#\{x \in \text{Zero}(\mathcal{P}, C^k) : Q(x) \neq 0\}.$$

*The signature of the quadratic form  $\text{Herm}(\mathcal{P}, Q)$  equals the difference between the number of distinct roots in  $R^k$  of  $\mathcal{P}$  at which  $Q$  is positive and the number of distinct roots in  $R^k$  at which  $Q$  is negative; in symbols,*

$$\begin{aligned} \text{Signature}(\text{Herm}(\mathcal{P}, Q)) = \\ \#\{x \in \text{Zero}(\mathcal{P}, R^k) : Q(x) > 0\} - \#\{x \in \text{Zero}(\mathcal{P}, R^k) : Q(x) < 0\}. \end{aligned}$$

*Proof.* Let  $a$  be a separating element of  $A$ . Recall that  $1, a, a^2, \dots, a^{n-1}$  are independent in  $A$ . Let  $\omega_1 := 1, \omega_2 := a, \dots, \omega_{n-1} := a^{n-1}, \omega_{n+1}, \dots, \omega_N$  be a basis of the  $K$  vector space  $A$ . By Stickelberger's result we have, for

$f = \sum_{j=1}^N f_j \omega_j$  in  $A$ ,

$$\begin{aligned} \text{Herm}(\mathcal{P}, Q)(f) &= \text{Trace} L(Qf^2) \\ &= \sum \{ \mu(x) Q(x) (\sum f_j \omega_j(x))^2 : x \in \text{Zero}(\mathcal{P}, C^k) \}. \end{aligned}$$

Let  $x_1, \dots, x_n$  be the elements of  $\text{Zero}(\mathcal{P}, C^k)$ . Let

$$\Gamma := \begin{pmatrix} 1 & a(x_1) & \dots & a(x_1)^{n-1} & \omega(x_1) & \dots & \omega_N(x_1) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & a(x_n) & \dots & a(x_n)^{n-1} & \omega(x_n) & \dots & \omega_N(x_n) \end{pmatrix}.$$

Let  $D := \text{diag}(\mu_1 Q(x_1), \dots, \mu_n Q(x_n))$  where  $\mu_i := \mu(x_i)$ . Let  $\mathbf{f}$  denote the column vector  $(f_1, \dots, f_N)^T$ . Note that

$$\text{Herm}(\mathcal{P}, Q) = \mathbf{f}^T \Gamma^T D \Gamma \mathbf{f}.$$

*Claim:*  $\text{Rank}(\Gamma) = \#\{x \in \text{Zero}(\mathcal{P}, C^k) : Q(x) \neq 0\}$ .

Note that the first  $n$  columns of  $\Gamma$  form a Vandermonde matrix. Since  $a$  is separating the values  $a(x_i)$  are distinct. Hence by Vandermonde's well-known theorem the first  $n$  columns of  $\Gamma$  form a matrix with rank  $n$ . Hence  $\Gamma$  is a matrix with rank  $n$ .

Using Stickelberger's theorem again we have

$$\begin{aligned} \text{Herm}(\mathcal{P}, Q)(f) &= \sum \{ \mu(y) Q(y) (\sum_{j=1}^N f_j \omega_j(y))^2 : y \in \text{Zero}(\mathcal{P}, R^k) \} \\ &\quad + \sum \{ \mu(z) Q(z) (\sum_{j=1}^N f_j \omega_j(z))^2 : z \in \text{Zero}(\mathcal{P}, C^k) \wedge z \notin \text{Zero}(\mathcal{P}, R^k) \}. \end{aligned}$$

The next claim implies that the second summation (which involves the complex solutions which are not real) contributes zero to the signature of  $\text{Herm}(\mathcal{P}, Q)$ . In other words, the proof of that claim will finish the proof of Hermite's root counting result.

In the rest of this proof I use terminology usually used for real and complex numbers. Recall that, since  $R$  is a real closed field, we have  $C = R[i]$  where  $i$  is a root of  $X^2 + 1$ . I shall say that elements of  $R$  are "real" and the elements of  $C$  are "imaginary", etc.

Consider any  $z$  which satisfies  $z \in \text{Zero}(\mathcal{P}, C^k)$  and  $z \notin \text{Zero}(\mathcal{P}, R^k)$  and  $Q(z) \neq 0$ .

*Claim:* The following quantity can be written as the difference of two squares of real linear forms:

$$\mu(z) Q(z) (\sum f_j \omega_j(z))^2 + \mu(\bar{z}) Q(\bar{z}) (\sum f_j \omega_j(\bar{z}))^2.$$

(I omit the limits on the summations to simplify notation.)

Let  $s_j$  and  $t_j$  be the real and imaginary parts of  $\omega_j(z)$ : in other words,  $\omega_j(z) = s_j + it_j$  where  $s_j$  and  $t_j$  are real. Note that  $\omega_j(\bar{z}) = s_j - it_j$  since

the conjugation map  $z \mapsto \bar{z}$  is an isomorphism. Choose real numbers  $a$  and  $b$  so that  $\mu(z)Q(z) = (a + bi)^2$ . Note that  $\mu(\bar{z})Q(\bar{z}) = (a - bi)^2$ . Set

$$(1) \quad L_1 := \sum (as_j - bt_j)f_j$$

$$(2) \quad L_2 := \sum (at_j + bs_j)f_j.$$

Note that these are real linear forms (that is, they have coefficients in  $R$ ). Note that

$$\begin{aligned} L_1 + iL_2 &= \sum (as_j - bt_j + i(at_j + bs_j))f_j \\ &= (a + bi) \sum f_j \omega_j(z) \end{aligned}$$

since  $(a + bi)\omega_j(z) = (a + bi)(s_j + t_j i)$ . The following computation finishes the proof of the claim:

$$\begin{aligned} \mu(z)Q(z)(\sum f_j \omega_j(z))^2 + \mu(\bar{z})Q(\bar{z})(\sum f_j \omega_j(\bar{z}))^2 & \\ &= (a + bi)^2 (\sum f_j \omega_j(z))^2 + (a - bi)^2 (\sum f_j \omega_j(\bar{z}))^2 \\ &= (L_1 + iL_2)^2 + (L_1 - iL_2)^2 \\ &= L_1^2 + 2iL_1L_2 - L_2^2 + L_1^2 - 2iL_1L_2 - L_2^2 \\ &= 2L_1^2 - 2L_2^2. \end{aligned}$$

□