

Companion Matrices for Systems of Polynomial Equations

Alessandro Chiesa
©January 15, 2008

Abstract. This paper shows how to solve a system of polynomial equations known to have finitely many solutions by looking at the companion matrices of the corresponding ideal.

1. Introduction. Solving systems of polynomial equations is a subject of both theoretical and practical interest. It is of theoretical interest because the set of solutions forms a variety, which is a geometric object of primary importance in algebraic geometry. And it is of practical interest because such systems arise naturally in problems that impose some constraints on the possible solutions. Such situations occur in many fields in the sciences, most notably in robotics (see [5] and [1, Ch. 6]), error-correcting codes (see [3] and [2, Ch. 9]), and statistics (see [4]), where the use of algebraic geometry has provided very deep insights.

More formally, suppose we are given m polynomials f_1, \dots, f_m in n variables with coefficients in some field k . We are interested in the points p of k^n at which *all* of these polynomials vanish. We may then ask the two questions:

- (1) Can we determine for which f_i there are finitely many such points p ?
- (2) If so, can we find the p ?

In this paper, we study only systems with finitely many solutions, and we show how to answer affirmatively both of these questions using algebraic geometry.

More precisely, the polynomials of a system of equations define an ideal. Then we can consider the quotient ring of the ring of polynomials modulo the ideal. Multiplication by a polynomial is a linear map on this ring. In particular, multiplication by the i th coordinate function is. Since any linear map can be represented by a matrix, the i th coordinate function has a corresponding matrix, called the *i th companion matrix*. The main result asserts that the eigenvalues of the i th companion matrix are exactly the i th coordinates of the solutions to the system of polynomial equations when the system has finitely many solutions. These matrices provide an elegant and computationally advantageous way to find the solutions of the system of polynomial equations.

In Section 2, we introduce the basic definitions of algebraic geometry. In Section 3, we study the conditions that hold for a system of polynomial equations to have finitely many solutions. In Section 4 we study the properties of the linear map defined by multiplication by polynomials on the quotient ring of the ring of polynomials modulo an ideal. In Section 5, we describe the construction of the companion matrices. Finally, in Section 6, we illustrate the process with an example.

2. Background. In this section, we give the basic definitions of algebraic geometry that we need to study the solutions of systems of polynomial equations. For a more comprehensive introduction to algebraic geometry, see [1].

Let k denote a field, and $k[x_1, \dots, x_n]$ the polynomial ring in n variables with coefficients in k . Consider the system of polynomial equations

$$\begin{aligned} f_1(x_1, \dots, x_n) &= 0 \\ &\vdots \\ f_m(x_1, \dots, x_n) &= 0 \end{aligned}$$

where the f_i belong to $k[x_1, \dots, x_n]$. Its set of solutions is so important that we give it a name, and use special notation. Specifically, the set

$$\mathbf{V}(f_1, \dots, f_m) := \{(a_1, \dots, a_n) \in k^n : f_i(a_1, \dots, a_n) = 0 \text{ for all } 1 \leq i \leq m\}$$

is called the *variety* defined by f_1, \dots, f_m .

It is clear then that studying the solutions of $f_1 = \dots = f_m = 0$ is equivalent to studying the points of its variety $\mathbf{V}(f_1, \dots, f_m)$. Because of this equivalence, we may limit our attention to the study the variety.

Closely related to the geometric concept of variety is the algebraic concept of ideal. Ideals and varieties are the two main objects of study of algebraic geometry.

A subset I of $k[x_1, \dots, x_n]$ is called an *ideal* if it satisfies the following properties:

- (1) $0 \in I$;
- (2) if $f, g \in I$, then $f + g \in I$; and
- (3) if $f \in I$ and $h \in k[x_1, \dots, x_n]$, then $hf \in I$.

Given polynomials f_1, \dots, f_m , we can form an ideal with the m polynomials. In fact, it is easy to show that the set

$$\langle f_1, \dots, f_m \rangle := \left\{ \sum_{i=1}^m h_i f_i : h_1, \dots, h_m \in k[x_1, \dots, x_n] \right\}$$

is an ideal. Moreover, it is the smallest ideal containing the f_i . As is clear from the definition of $\langle f_1, \dots, f_m \rangle$, all of the elements in $\langle f_1, \dots, f_m \rangle$ are polynomial-linear combinations of the f_i . Therefore, we call $\{f_1, \dots, f_m\}$ a generating set of the ideal $\langle f_1, \dots, f_m \rangle$. Notice that an ideal I may have more than one generating set.

An ideal is called *radical* if it satisfies the following property:

$$\text{if } f^m \in I \text{ then } f \in I.$$

If an ideal I is not radical, then we can extend it to the smallest ideal containing I that satisfies the above property. We define the set

$$\sqrt{I} := \{h : h^m \in I \text{ for some integer } m\}$$

to be the *radical* of I . It is easy to show that \sqrt{I} is indeed a radical ideal containing I , and it is the smallest such ideal. Furthermore, $\sqrt{I} = I$ if and only if I is radical.

The Hilbert Nullstellensatz is a theorem that relates ideals and varieties. For its proof, see [1, Ch. 4].

Theorem 2-1 (The Hilbert Nullstellensatz). *Let k be an algebraically closed field. If a polynomial $f \in k[x_1, \dots, x_n]$ vanishes everywhere on the variety V of an ideal I of $k[x_1, \dots, x_n]$, then there exists an integer $m \geq 1$ such that $f^m \in I$.*

The Nullstellensatz tells us that the polynomials that vanish on the variety V of an ideal I are the polynomials in \sqrt{I} , that is, the polynomials in I plus some other polynomials each with a power in I .

In order to simplify the notation for monomials, we let

$$x^\alpha := x^{\alpha_1} \cdots x^{\alpha_n}$$

where $\alpha := (\alpha_1, \dots, \alpha_n)$ is an n -tuple of nonnegative integers. With this notation, any polynomial $f \in k[x_1, \dots, x_n]$ can be written as follows:

$$f = \sum_{\alpha} a_{\alpha} x^{\alpha} \text{ with } a_{\alpha} \in k.$$

A *monomial ordering* $>$ on $k[x_1, \dots, x_n]$ is a relation on the set of monomials x^α satisfying the following properties:

- (1) $>$ is a total ordering on the set of monomials;
- (2) if $x^\alpha > x^\beta$, then $x^\alpha x^\gamma > x^\beta x^\gamma$ for all monomials x^γ ; and
- (3) $>$ is a well ordering on the set of monomials.

For a fixed monomial order, a polynomial has a unique leading term. We denote the leading term of f by $LT(f)$. Furthermore, for any nonempty ideal I , we can form the set of leading terms of polynomials of I

$$LT(I) := \{cx^\alpha : \text{there exists } f \in I \text{ with } LT(f) = cx^\alpha\}.$$

Then we can define $\langle LT(I) \rangle$ to be the ideal generated by the elements of $LT(I)$.

Among all generating sets $G = \{g_1, \dots, g_t\}$ of an ideal I , there are some that have the additional property

$$\langle LT(I) \rangle = \langle LT(g_1), \dots, LT(g_t) \rangle. \quad (2-1)$$

Equation (2-1) says that the leading terms of the elements in G generate $\langle LT(I) \rangle$. We call such a generating set a *Gröbner basis* for I .

In this paper, we do not discuss Gröbner bases in detail, but we do need to know that given a polynomial f , we can calculate a remainder upon division by one. In fact, for a fixed monomial order, given a polynomial $f \in k[x_1, \dots, x_n]$ and any ordered m -tuple of polynomials $F := (f_1, \dots, f_m)$, we can write f in the form

$$f = a_1 f_1 + \cdots + a_m f_m + r$$

where $a_i, r \in k[x_1, \dots, x_n]$. Furthermore, r is either 0 or a k -linear combination of monomials none of which is divisible by any of $LT(f_1), \dots, LT(f_m)$. The polynomial r is called the *remainder* and denoted \bar{f}^F .

When we divide a polynomial f by a Gröbner basis G , we have the additional property that the remainder \bar{f}^G is unique, and it has the property

$$\bar{f}^G = \bar{g}^G \text{ if and only if } f - g \in I. \quad (2-2)$$

For a more detailed discussion of Gröbner bases, see [2, Ch. 2].

An ideal I induces an equivalence relation among the elements of $k[x_1, \dots, x_n]$. We define the *class* of $f \in k[x_1, \dots, x_n]$ to be the set

$$[f] := \{g \in k[x_1, \dots, x_n] : f - g \in I\}.$$

We also call $[f] \in A$ the *coset* of f with respect to I . The definition of $[f]$ shows that it has the following property:

$$[f] = [g] \text{ if and only if } f - g \in I. \quad (2-3)$$

Addition and *multiplication* on the equivalence classes can be defined in terms of elements of the classes according to the following relations:

$$[f] + [g] := [f + g] \text{ and } [f] \cdot [g] := [f \cdot g]. \quad (2-4)$$

It is easy to show that the set of classes is a ring. We define the *quotient ring*

$$A := k[x_1, \dots, x_n]/I := \{[f] : f \in k[x_1, \dots, x_n]\}$$

to be the set of all equivalence classes.

We are now ready to study the conditions that have to hold for a variety to be finite.

3. Finite Varieties. In this section, we prove properties about finite varieties. Most of this material is covered in [2, pp. 230–236]. We first need to introduce the concept of standard monomials of an ideal I .

Equations (2-2) and (2-3) show that there is a one-to-one correspondence between the remainders \bar{f}^G upon division by a Gröbner basis G and the cosets $[f]$ in the quotient ring A . Furthermore, \bar{f}^G is a k -linear combination of the monomials

$$x^\alpha \notin \langle LT(I) \rangle. \quad (3-1)$$

Any monomial x^α satisfying the relation (3-1) is called a *standard monomial* of I .

Thus the monomials not in $\langle LT(I) \rangle$ form the set B of standard monomials of I :

$$B := \{x^\alpha : x^\alpha \notin \langle LT(I) \rangle\}.$$

In the new notation, \bar{f}^G is a k -linear combination of elements of B . Since \bar{f}^G corresponds to $[f] \in A$, we have the following lemma relating the quotient ring A to the standard monomials B of I .

Lemma 3-1 (Macaulay). *Let $I \subset k[x_1, \dots, x_n]$ be an ideal. Then the quotient ring $k[x_1, \dots, x_n]/I$ is isomorphic as a k -vector space to the k -linear span of the set B of standard monomials.*

Proof: Let $S := \text{Span}(B)$ be the k -linear span of the standard monomials of I . Consider the natural map $\psi: A \rightarrow S$ that sends the coset $[f]$ to the remainder \bar{f}^G . From the discussion above, ψ is bijective; so it remains to show that ψ preserves the vector-space operations.

It is an easy exercise to show that, for $c \in k$, the following two properties hold:

$$\overline{f + g}^G = \bar{f}^G + \bar{g}^G \text{ and } \overline{c \cdot f}^G = c \cdot \bar{f}^G.$$

Therefore, we may write

$$\begin{aligned} \psi([f] + [g]) &= \psi([f + g]) = \overline{f + g}^G = \bar{f}^G + \bar{g}^G = \psi([f]) + \psi([g]) \text{ and} \\ \psi([c] \cdot [f]) &= \psi([c \cdot f]) = \overline{c \cdot f}^G = c \cdot \bar{f}^G = \psi([c]) \cdot \psi([f]). \end{aligned}$$

Thus ψ preserves the vector-space operations. □

Lemma 3-1 implies the equivalence of conditions (2) and (3) in the following theorem.

Theorem 3-2 (Finiteness Theorem). *The following conditions are equivalent:*

- (1) *the variety $\mathbf{V}(I)$ is finite;*
- (2) *the set B of standard monomials is finite; and*
- (3) *the \mathbb{C} -vector space A is finite dimensional.*

Proof: Assume (1). Say that $\mathbf{V}(I) = \{p_1, \dots, p_m\}$ where $p_j = (a_{1j}, \dots, a_{nj})$ for $j = 1, \dots, m$. For $i = 1, \dots, n$, consider the polynomial $h_i = \prod_{j=1}^m (x_i - a_{ij})$. By construction, h_i vanishes everywhere on $\mathbf{V}(I)$. By the Nullstellensatz, Theorem 2-1, some power h_i^l is in I . Clearly $LT(h_i) = x_i^{ml}$. Hence x_i^{ml} is in $\langle LT(I) \rangle$. Therefore, the set B of standard monomials contains at most $(ml)^n$ elements. Thus (1) implies (2).

Lemma 3-1 implies the equivalence of conditions (2) and (3).

Finally, assume that the \mathbb{C} -vector space A is finite dimensional with dimension m . Then, for $i = 1, \dots, n$, the $m+1$ cosets $[x_i]^0, \dots, [x_i]^m$ are linearly dependent. So there exist $c_0, \dots, c_m \in \mathbb{C}$, not all zero, such that

$$\sum_{j=0}^m c_j [x_i]^j = [0].$$

But the polynomial $\sum_{j=0}^m c_j x_i^j$ has finitely many roots. Therefore, only finitely many complex numbers may appear as the i th coordinate of points in $\mathbf{V}(I)$. Since i is arbitrary, we have shown that (3) implies (1). \square

If an ideal I satisfies any of the conditions in Theorem 3-2, then I is said to be *zero-dimensional*. So I is zero-dimensional if and only if its variety $\mathbf{V}(I)$ is finite. In this paper, we are only interested in systems of polynomial equations that have finitely many zeros; so for the rest of this paper, we concentrate on proving properties of zero-dimensional ideals.

Next, we prove a lemma that allows us to give an upper bound on the number of points in the variety $\mathbf{V}(I)$.

Lemma 3-3. *Let $S = \{p_1, \dots, p_m\}$ be a finite subset of \mathbb{C}^n . Then there exist polynomials $g_i \in \mathbb{C}[x_1, \dots, x_n]$ for $i = 1, \dots, m$ such that $g_i(p_j) = 0$ if $i \neq j$ and $g_i(p_i) = 1$.*

Proof: We show how to construct g_1 . The construction of the other polynomials is similar.

Since the p_i are distinct, p_1 and p_i for $i \geq 2$ must differ at some coordinate, say the j_i th coordinate. Let a_{j_i} be the j_i th coordinate of p_1 , and b_{j_i} be the j_i th coordinate of p_i . Then consider the polynomial

$$f_i := \frac{x_{j_i} - b_{j_i}}{a_{j_i} - p_{j_i}}.$$

The polynomial f_i satisfies $f_i(p_1) = 1$ and $f_i(p_i) = 0$ for $i \geq 2$. The polynomial given by $g_1 := f_2 \cdot f_3 \cdots f_m$ satisfies the desired properties. \square

The following theorem bounds the number of points in $\mathbf{V}(I)$ whenever I is zero-dimensional.

Theorem 3-4. *Let I be a zero-dimensional ideal in $\mathbb{C}[x_1, \dots, x_n]$. Then the number of points in $\mathbf{V}(I)$ is at most $\dim_{\mathbb{C}}(A)$. Equality occurs if and only if I is a radical ideal.*

Proof: We show first that the number of points of $\mathbf{V}(I)$ is bounded above by the dimension of A . Since I is zero dimensional, $\mathbf{V}(I)$ is finite, say $\mathbf{V}(I) = \{p_1, \dots, p_m\}$. By Lemma 3-3, we can find polynomials $g_i \in \mathbb{C}[x_1, \dots, x_n]$ for $i = 1, \dots, m$ such that $g_i(p_j) = 0$ if $i \neq j$ and $g_i(p_i) = 1$.

To prove the statement, it is sufficient to show that the cosets $[g_1], \dots, [g_m] \in A$ are linearly independent in A . Suppose that, for some $a_i \in \mathbb{C}$, we have that

$$\sum_{i=1}^m a_i [g_i] = [0] \text{ in } A.$$

If we let $g = \sum_{i=1}^m a_i g_i$, then g must be in the ideal I , so that g vanishes everywhere on $\mathbf{V}(I)$. Therefore, we may write

$$0 = g(p_j) = \sum_{i=1}^m a_i g_i(p_j) = 0 + a_j g_j(p_j) = a_j$$

for $j = 1, \dots, m$. Hence a_1, \dots, a_m vanish, so that $[g_i]$ are linearly independent in A .

Suppose I is a radical ideal. To prove equality, it suffices to show that $[g_1], \dots, [g_m]$ span A . Consider any $[g] \in A$. Set $a_i = g(p_i)$ and let $h := g - \sum_{i=1}^m a_i f_i$ where f_i are the same f_i considered above. It is easy to show that $h(p_j) = 0$ for all j , so that h vanishes everywhere on V . By the Hilbert Nullstellensatz, we know that there exists an integer $m \geq 1$ such that $h^m \in I$. But I is radical, so that $h \in I$. Hence $[h] = [0]$ in A , so that $[g] = \sum_{i=1}^m a_i [f_i]$.

Conversely, suppose that m is the dimension of A . Define the ring

$$B := k[x_1, \dots, x_n] / \sqrt{I}$$

to be the ring of all equivalence classes with respect to \sqrt{I} . Consider the natural map $\psi: A \rightarrow B$ that maps the coset $[f]$ of I in A to the coset $[f]$ of \sqrt{I} in B . It is clear that ψ is surjective. By hypothesis, A has dimension m . However, we just proved that the dimension of B is the number of points in $\mathbf{V}(\sqrt{I})$, so that $\mathbf{V}(\sqrt{I}) = \mathbf{V}(I) = m$. Hence, ψ is a bijection. Thus, $\sqrt{I} = I$, as desired. \square

We now proceed to develop the theory that allows us to find the points of a finite variety $\mathbf{V}(I)$.

4. A Linear Map on A . Throughout this section, I denotes a zero-dimensional ideal.

Multiplication by an element $f \in \mathbb{C}[x_1, \dots, x_n]$ defines a linear map from A to itself. That is, we have the map $m_f: A \rightarrow A$ given by

$$m_f([g]) := [f] \cdot [g] := [fg] \in A$$

for any $f \in \mathbb{C}[x_1, \dots, x_n]$.

Also, $[f] = [g]$ if and only if $f - g \in I$. Thus, since elements in the same coset $[f]$ define the same map m_f , we know that $m_f = m_g$ if and only if $f - g \in I$. So we may take f to be \bar{f}^G , since \bar{f}^G is the representative element of the coset $[f]$ in A . We can also represent m_f as a $d \times d$ matrix with respect to some basis.

Let us pick the set B of standard monomials as a basis. The set B is finite since the ideal I is zero-dimensional by the definition of zero-dimensional ideal given after

Theorem 3-2. Furthermore, we can carry out computations in the quotient ring A by finding the multiplication table for the elements of the basis B . Hence, if we index the rows and columns of m_f by the monomials in B , then, for $x^u, x^v \in B$, the entry of m_f in row x^u and column x^v is the coefficient of x^u in $\overline{f \cdot x^v}$.

Let $M_{d \times d}(\mathbb{C})$ be the ring of $d \times d$ matrices with complex entries. Let

$$\phi: \mathbb{C}[x_1, \dots, x_n] \rightarrow M_{d \times d}(\mathbb{C})$$

be the map that sends f to its corresponding matrix m_f . We have the following proposition.

Proposition 4-1. *The map ϕ is a ring homomorphism.*

Proof: A ring homomorphism must preserve the algebraic structure. First, $\phi(1) = I_d$ follows directly from the definition of m_f : the matrix defined by the class $[1]$ is the $d \times d$ identity matrix, so that ϕ carries the multiplicative identity of $\mathbb{C}[x_1, \dots, x_n]$ to the multiplicative identity of $M_{d \times d}(\mathbb{C})$. Furthermore, the two relations in (2-4) imply that

$$\begin{aligned} \phi(f + g) &= m_{f+g} = m_f + m_g = \phi(f) + \phi(g) \text{ and} \\ \phi(f \cdot g) &= m_{f \cdot g} = m_f \cdot m_g = \phi(f) \cdot \phi(g). \end{aligned}$$

Thus ϕ preserves addition and multiplication. □

As a consequence of the fact that ϕ is a ring homomorphism, we obtain the following property that we use to prove the main theorem of Section 5.

Proposition 4-2. *Let $f \in \mathbb{C}[x_1, \dots, x_n]$ and $h \in \mathbb{C}[t]$. Then*

$$m_{h(f)} = h(m_f).$$

Furthermore, $h(m_f) = 0$ if and only if $h([f]) = [0]$.

Proof: Let $h(t) = \sum_{i=0}^m c_i t^i$. If we plug f in for t in $h(t)$, then the expression given by $h(f) = \sum_{i=0}^m c_i f^i$ is a polynomial in $\mathbb{C}[x_1, \dots, x_n]$. Similarly, if we plug m_f in for t in $h(t)$, then $h(m_f)$ is the zero map. From the definition of m_f , we know that m_f operates by multiplication by the coset $[f]$. Thus that $m_{h(f)} = h(m_f)$.

Assume $h(m_f) = 0$. Then the discussion in the previous paragraph says that $m_{h(f)}$ is the zero map. Therefore, $[h(f)] \cdot [g] = [0]$ for any polynomial $g \in \mathbb{C}[t]$. If we choose $g = 1$, then we deduce that $[h(f)] = [0]$. But $[h(f)] = h([f])$, so that $h([f]) = [0]$. Conversely, assume that $h([f]) = [0]$. Then the map $m_{h(f)}$ is the zero map. Using the relation proved in the previous paragraph, we obtain $h(m_f) = 0$. □

Let M be a $d \times d$ matrix with complex entries. Then define I_M to be the set of all polynomials $h_M \in \mathbb{C}[t]$ such that $h(M) = 0$. It is easy to show that I_M is an ideal of $\mathbb{C}[t]$. Since $\mathbb{C}[t]$ is a principal ideal domain, we know that

$$I_M = \langle h_M \rangle$$

for some polynomial h_M . We call h_M the *minimal polynomial* of M .

In Section 5, we show how to use the properties we just proved about the linear map m_f to find the points of $\mathbf{V}(I)$.

5. Companion Matrices. In this section, we show how to find the points of $\mathbf{V}(I)$ where $I \subset \mathbb{C}[x_1, \dots, x_n]$ is again a zero-dimensional ideal. To do so, we first prove a stronger result that relates the values of a polynomial $f \in \mathbb{C}[x_1, \dots, x_n]$ at the points of $\mathbf{V}(I)$ to the eigenvalues of the linear map m_f that f defines on A . Then linear algebra tells us that the eigenvalues of m_f are in turn related to the zeros of the minimal polynomial h_f of m_f . Recall that A was defined in Section 2 as the quotient ring $k[x_1, \dots, x_n]/I$.

For standard theorems in linear algebra, please refer to any text book on the subject, such as that by Strang [6]. A more extensive treatment of companion matrices, together with examples of other applications of companion matrices, can be found in the short book by Sturmfels [7, Ch. 2].

Combining the results we proved in Section 4, we obtain the following theorem.

Theorem 5-1. *Preserve the conditions above. Let $\lambda \in \mathbb{C}$. Then the following statements are equivalent:*

- (1) λ is a zero of h_f ;
- (2) λ is an eigenvalue of m_f ; and
- (3) $f(p) = \lambda$ for some $p \in \mathbf{V}(I)$.

Proof: The Cayley–Hamilton Theorem has the corollary that the minimal polynomial h_f of m_f divides the characteristic polynomial of m_f . Therefore, if (1) holds, then λ must be an eigenvalue of m_f ; thus (1) implies (2).

It can also be shown that the eigenvalues of m_f appear as zeros of its minimal polynomial h_f , proving that (2) implies (1).

Assume (2). Then $m_f \cdot [z] = \lambda[z]$ for some eigenvector $[z] \neq 0$. We know $\mathbf{V}(I)$ is finite since I is zero dimensional, so say $\mathbf{V}(I) = \{p_1, \dots, p_m\}$. Suppose by way of contradiction that $f(p_i) \neq \lambda$ for $i = 1, \dots, m$.

Let $g := f - \lambda$, so that $g(p_i) \neq 0$ for all i . Lemma 3-3 asserts that we can find polynomials g_i such that $g_i(p_j) = 0$ if $i \neq j$, and $g_i(p_i) = 1$. Construct the polynomial $g' = \sum_{i=1}^m \frac{1}{g(p_i)} g_i$. Then $g'g = 1$ for all points on $\mathbf{V}(I)$, so that $1 - g'g \in \mathbf{I}(\mathbf{V}(I))$. The Nullstellensatz tells us that some power $(1 - gg')^l$ is in I . Expanding, we can write

$$(1 - gg')^l = 1 - \tilde{g}g \in I,$$

for some $\tilde{g} \in \mathbb{C}[x_1, \dots, x_n]$. Taking cosets, we find that $[1] = [\tilde{g}][g]$, so that g is invertible in A . However, $m_f \cdot [z] = [f][z] = \lambda[z]$, so that $[g][z] = [f - \lambda][z] = 0$. If we multiply by \tilde{g} on both sides, we get that the eigenvector $[z]$ is zero, a contradiction. Hence, we must have $f(p) = \lambda$ for some $p \in \mathbf{V}(I)$. Thus (2) implies (3).

Finally, assume $f(p) = \lambda$ for some $p \in \mathbf{V}(I)$. Proposition 4-2 says that $h_f([f]) = [0]$, since h_f is the minimal polynomial of m_f . Then by the definition of A , we have that $h_f(f) \in I$. Hence, by the definition of variety of an ideal, $h_f(f)$ vanishes everywhere on $\mathbf{V}(I)$. Thus, $h_f(\lambda) = h_f(f(p)) = 0$, proving that (3) implies (1), and completing the proof of the theorem. \square

For the special case $f = x_i$, we set $T_i := m_f$ the i th companion matrix of the ideal I . More importantly, this special case tells us what the points in $\mathbf{V}(I)$ are. Thus, we have the following corollary.

Corollary 5-2 (Stickelberger’s Theorem). *Let $I \subset \mathbb{C}[x_1, \dots, x_n]$ be a zero-dimensional ideal. Then the eigenvalues of the i th companion matrix T_i are the x_i -coordinates of the points of $\mathbf{V}(I)$.*

Moreover, if I is radical, then we have a stronger statement than Stickelberger's Theorem.

Theorem 5-3. *The companion matrices T_1, \dots, T_n can be simultaneously diagonalized if I is radical.*

Proof: Since I is zero dimensional, $\mathbf{V}(I)$ is finite; say $\mathbf{V}(I) = \{p_1, \dots, p_m\}$.

Suppose I is radical. Consider any point $p = (\lambda_1, \dots, \lambda_n)$ in $\mathbf{V}(I)$. By Lemma 3-3, there exists a polynomial g such that $g(p) = 1$ and g vanishes everywhere else on $\mathbf{V}(I)$. Hence, the relation $x_i \cdot g = \lambda_i \cdot g$ is everywhere true on $\mathbf{V}(I)$. So the polynomial $(x_i - \lambda_i)g$ vanishes everywhere on $\mathbf{V}(I)$. The Hilbert Nullstellensatz says then that $(x_i - \lambda_i)g \in \sqrt{I} = I$. Then $[g]$ is a joint eigenvector of the companion matrices T_1, \dots, T_n .

Let V be the matrix constructed from all eigenvectors $[g]$ corresponding to each point p in $\mathbf{V}(I)$. The $[g]$ form a full set of distinct eigenvectors. Then $V^{-1}T_iV$ is a diagonal matrix whose entries are the i th coordinates of all the zeroes of I . \square

The computational advantage of Theorem 5-3 is that, if we diagonalize all the T_i at the same time, then the eigenvalues of the T_i simply appear on the diagonal, and we can read them off. By Stickelberger's Theorem, those eigenvalues are the x_i -coordinates of the points in $\mathbf{V}(I)$.

6. An Example. We now illustrate by way of example the method that we have described. We carry out all computations in the computer algebra system Maple. Consider the following polynomials in $\mathbf{C}[x, y]$:

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

The solutions of the system $f_1 = f_2 = 0$ are the points of the variety $\mathbf{V}(I)$ where $I = \langle f_1, f_2 \rangle$.

We load the packages `Ore_algebra` and `Groebner` that are needed to compute remainders of polynomials upon division by a set of polynomials and a Gröbner basis of an ideal, respectively. First, we find the Gröbner basis GB of I with the command `Basis`. A monomial ordering has to be specified; in our case we use a lexicographic ordering with $x > y$. Then we compute by hand the set of standard monomials B by noticing that

$$\langle LT(I) \rangle = \langle x^2, xy, y^3 \rangle.$$

We must have then $B = \{1, x, y, y^2\}$. The standard monomials form a basis for A , so we can compute the companion matrices T_x and T_y corresponding to multiplication by x and y on A : if x^α and x^β are in B , then the entry of T_x (or T_y) in row x^α and x^β is the coefficient of x^α in the remainder $\overline{x \cdot x^\beta}^G$ (or $\overline{y \cdot x^\beta}^G$). In Maple the remainder is computed via the command `NormalForm`. Finally, we diagonalize both T_x and T_y with a common set of eigenvectors. Then, the coordinates of the points in $\mathbf{V}(I)$ appear on the main diagonal of the matrices. The computations show that

$$\mathbf{V}(I) = \{(1, 0), (-1/2, 3/4), (-1, 0), (0, 1)\}.$$

The solutions may be checked with the Maple command `solve` on the original polynomials f_1 and f_2 .

Here is the transcript of the session:

```
> with(Ore_algebra):
> with(Groebner):
> A := poly_algebra(x,y):
> LEX := MonomialOrder(A, plex(x,y)):
> POLS := [x^2 + y -1, 2*x^2 - x*y + 2*y^2 + 1]:
> GB := Basis(POLS,LEX);

      GB := [3y - 7y^2 + 4y^3, xy - 2y^2 + 2y, x^2 + y - 1]
```

```
> B := [1,x,y,y^2]:
> for v in [x,y] do
> T:=array([],1..4,1..4):
> for j from 1 to 4 do
> p:= NormalForm(v*B[j],GB,LEX,PolynomialRing([x,y])):
> for i from 1 to 4 do
> T[i,j] := coeff(coeff(p,x,degree(B[i],x)),y,degree(B[i],y)):
> od:
> od:
> print(cat(T,v),T);
> od:
```

$$Tx, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & -2 & -\frac{3}{2} \\ 0 & 0 & 2 & \frac{3}{2} \end{bmatrix}$$

$$Ty, \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -2 & 0 & -\frac{3}{4} \\ 0 & 2 & 1 & \frac{7}{4} \end{bmatrix}$$

```

> Tx := Matrix([[0, 1, 0, 0], [1, 0, 0, 0], [0, -1, -2, -3/2],
> [0, 0, 2, 3/2]]):
> Ty := Matrix([[0, 0, 0, 0], [0, 0, 0, 0], [1, -2, 0, -3/4],
> [0, 2, 1, 7/4]]):
> P := Eigenvectors(Tx)[2];

```

$$P := \begin{bmatrix} \frac{-3}{4} & 0 & \frac{1}{4} & 0 \\ \frac{-3}{4} & 0 & \frac{-1}{4} & 0 \\ \frac{-1}{4} & -1 & \frac{-5}{4} & \frac{-3}{4} \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

```

> S := MatrixInverse(P):
> L1 := MatrixMatrixMultiply(S, MatrixMatrixMultiply(Tx, P)):
> L2 := MatrixMatrixMultiply(S, MatrixMatrixMultiply(Ty, P)):
> print(L1) : print(L2);

```

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{-1}{2} & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{3}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

REFERENCES

- [1] Cox, D., Little, J., and O’Shea, D., “Ideals, Varieties, and Algorithms,” Springer, Third Edition, 2007.
- [2] Cox, D., Little, J., and O’Shea, D., “Using Algebraic Geometry,” Springer, Second Edition, 2005.

- [3] Fitzgerald, J., and Lax, R. F., *Decoding affine variety codes using Gröbner bases*, Des. Codes Cryptogr. 13 (1998), no. 2, 147–158.
- [4] Holliday, T., Pistone, G., Riccomagno, E., and Wynn, H. P., *The application of computational algebraic geometry to the analysis of designed experiments: a case study*, Comput. Statist. 14 (1999), no. 2, 213–231.
- [5] Paul, R., “Robot Manipulators: Mathematics, Programming and Control,” MIT Press, 1981.
- [6] Strang, G., “Introduction to Linear Algebra,” Wellesley-Cambridge Press, 2003.
- [7] Sturmfels, B., “Solving Systems of Polynomial Equations,” CBMS Conference no. 97 (2002: Texas A and M University), Amer. Math. Soc., 2002.