TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 359, Number 7, July 2007, Pages 3071–3083
S 0002-9947(07)04123-2
Article electronically published on January 4, 2007

AN IDEAL SEPARATING EXTENSION OF AFFINE SPACE

PAUL S. PEDERSEN

Abstract. In affine space the set of solutions to a system of polynomial equations does not uniquely determine the system. We extend affine space so that the solutions (in the extension) to a system of equations uniquely determines the system.

1. Statement of the problem

In particular, for field R let elements of $R[x] = R[x_1, ..., x_n]$ act on the set of power series $R[[T]] = R[[T_1, ..., T_n]]$ by way of the linear extension of the action $x^k(T^m) = T^{m-k}$ for $m - k$ in N^n and $x^k(T^m) = 0$ for $m - k \notin N^n$. For the ideal $I \in \Im$ (the set of all ideals in $R[x]$) let $\mathbf{N}(I) = \{f(T) \in R[[T]] \mid P(x)f(T) = 0 \forall P(x) \in I\}$ (which we call the "generalized solutions" of I). We also set $Z(I) = \{r \in R^n \mid P(r) = 0 \forall P(x) \in I\}$. The injection $\rho : R^n \to R[[T]]$ given by $\rho(r) = \rho(r_1, ..., r_n) = \sum r^k T^k$ is such that $\rho(Z(I)) \subset \mathbf{N}(I)$ and so ρ maps the affine solutions to a system of equations into the generalized solutions of that system. Let $\mathcal{N}$ be the set of all $\mathbf{N}(I)$ for $I \in \Im$. We show that $\mathbf{N}$ is a one-to-one order reversing bijection from $\Im$ to $\mathcal{N}$ which implies that the generalized solutions to a system of equations uniquely determines the system.

2. Introduction

Let $\Im$ denote the set of all ideals in the polynomial ring $R[x] \equiv R[x_1, x_2, ..., x_n]$ where R is a field, let

$$Z(I) \equiv \{r = (r_1, ..., r_n) \in R^n | P(r) = 0 \text{ for all } P(x) \in I\}$$

be the *algebraic set* corresponding to any $I \in \Im$, and let

$$\zeta \equiv \{Z(I) | I \in \Im\}$$

be the set of all the algebraic sets in R^n. We will say that R^n "separates" a set of ideals $\hat{\Im} \subset \Im$ if we have $I_1 = I_2$ whenever we have $Z(I_1) = Z(I_2)$ for $I_1, I_2 \in \hat{\Im}$.

While working on problems in invariant theory, Hilbert proved the powerful theorem (called Nullstellensatz) that $rad(\Im)$ is separated in R^n when R is an *algebraically closed* field (where $rad(I) \equiv \{g = g(x) \in R[x] | \exists m \in N$ so that $g^m \in I\}$ is the radical of the ideal I and where $rad(\Im)$ is the set of all radical ideals).

When $n = 1$ the radical ideals are of the form $\langle P(x) \rangle$ where $P(x)$ has distinct roots. Hilbert's Nullstellensatz implies that $\langle P(x) \rangle$ has as many solutions as the degree of $P(x)$. Consequently, Hilbert's Nullstellensatz theorem has been called an n-dimensional generalization of the fundamental theorem of algebra.

Received by the editors April 24, 2003 and, in revised form, March 9, 2005.
2000 *Mathematics Subject Classification.* Primary 14xx, 13xx.

A consequence of Hilbert's Nullstellensatz is that when R is algebraically closed there is a one to one order reversing correpondence between the lattice of radical ideals and the lattice of algebraic sets of R^n (both partially ordered by inclusion).

In this paper we describe an extension of affine space having enough points to separate all ideals. In fact we show that to separate all ideals it is sufficient to consider an extension of affine space which only separates the zero ideals (those having $Z(I) = \{0\}$). An added feature of the method is that the field need *not be algebraically closed.*

To do this we use the R linear space of power series $R[[T]] = R[[T_1, T_2, ..., T_n]]$ on which we allow the elements of $R[x]$ to operate by the linear extension of the action (for $k, j \in N^n$) given by $x^j(T^k) = T^{k-j}$ if $k - j \in N^n$ and by $x^j(T^k) = 0$ if $k - j \notin N^n$. We use

$$\mathbf{N}(I) = \{f(T) \in R[[T]] | P(x)f(T) = 0 \forall P(x) \in I\} \tag{1}$$

to denote the *algebraic nullspace* associated with the ideal I and we let $\mathcal{N}$ denote the set of all algebraic nullspaces of $R[[T]]$.

We *embed* R^n in $R[[T]]$ by way of the injection

$$\rho(r) = \rho(r_1, ..., r_n) = \sum_{k \in N^n} r^k T^k \tag{2}$$

where $T^k = {T_1}^{k_1} ... {T_n}^{k_n}$. It is easily checked that

$$\rho(Z(I)) \subset \mathbf{N}(I).$$

We prove that for any ideals in $\Im$

$$I_1 \supset I_2 \text{ if and only if } \mathbf{N}(I_1) \subset \mathbf{N}(I_2). \tag{3}$$

In other words the lattice of ideals in $R[x]$ is in one to one order reversing correspondance with the lattice of algebraic nullspaces of $R[[T]]$. As a consequence we have that $R[[T]]$ separates the ideals in $\Im$.

For a set $V \subset R[[T]]$ let $\mathbf{I}(V)$ denote the smallest ideal in $\Im$ containing V. Another consequence of (3) is the theorem

$$\mathbf{I}(\mathbf{N}(I)) = I \text{ for all } I \in \Im. \tag{4}$$

As other corollaries to (3) we show that

$$\mathbf{N}(I_1 \cap I_2) = \mathbf{N}(I_1) + \mathbf{N}(I_2),$$
$$\mathbf{N}(I_1 + I_2) = \mathbf{N}(I_1) \cap \mathbf{N}(I_2),$$
$$\mathbf{I}(\mathbf{N}(I_1) + \mathbf{N}(I_2)) = \mathbf{I}(\mathbf{N}(I_1)) \cap \mathbf{I}(\mathbf{N}(I_2)),$$

and

$$\mathbf{I}(\mathbf{N}(I_1) \cap \mathbf{N}(I_2)) = \mathbf{I}(\mathbf{N}(I_1)) + \mathbf{I}(\mathbf{N}(I_2))$$

where $I_i, i = 1, 2$, are arbitrary ideals in $R[x]$.

Our proof of (3) will rely on a construction which gives an explicit countable *basis* for $\mathbf{N}(I)$ for any given $I \in \Im$. We refer to this set as a basis because it has the property that any element of $\mathbf{N}(I)$ can be written as a *unique countable* R linear combination of its elements.

We now give a rough outline of how we prove our theorems (in particular the Nullspace Basis Theorem). Let $R[\alpha] = R[\alpha_1, ..., \alpha_n] \cong \frac{R[y]}{I(y)}$, which is to say that $R[\alpha]$ is isomorphic to the residue class ring $\frac{R[y]}{I(y)}$ where $\alpha_i \leftrightarrow [y_i]$. Using *lexicographic ordering* on N^n we *totally order* the monomials in $R[\alpha]$. Using

these monomials and this total ordering we define an R ordered vector space basis B for $R[\alpha]$ as follows: $\alpha^k \in B$ if α^k *cannot* be written as an R linear combination of monomials *preceding* it in the lexicographic ordering. So there exists a subset J of N^n so that $B = \{\alpha^k | k \in J\}$ is a vector space basis for $R[\alpha]$ with the property that every monomial in $R[\alpha]$ which is not in B can be written as a unique, linear combination of monomials *preceding* it in B.

After forming

$$w(T) = \sum_{k \in N^n} T^k \otimes \alpha^k$$

we replace every $\alpha^k \notin B$ in $w(T)$ by its unique R linear combination of preceding basis elements. Collecting terms we get an expression of the form

$$\sum_{\alpha^k \in B} w_k(T) \otimes \alpha^k$$

where the $w_k(T)$ are real formal power series in $T_1, ..., T_n$. We will show that every element of $\mathbf{N}(I)$ can be written as a *unique*, countable, R linear combination of the $w_k(T)$. Additionally, we show how one can compute the coefficients of this linear combination. In particular, the elements of $Z(I)$ are easily expressed as linear combinations of the $w_k(T)$.

In section 10 we give an example of a pair of inequivalent finite dimensional ideals having the same radical, having the same Hilbert Characteristic Function, and having identical $Z(I)$ with identical arithmetic multiplicity. From this example we conclude that affine space cannot separate all ideals even if we include multiplicity information.

As used in this paper, $R[[T]]$ can be thought of as the dual space of $R[x]$. Although he did not consider "generalized solutions", Macaulay in his famous book [1] used the dual space in the form of inverse systems to give (amongst other things) an algebraic characterization of intersection multiplicity for finite dimensional ideals.

3. Notation

The relation $\equiv$ will be used when making definitions. N will represent the non-negative integers, N_1 will denote the nonzero natural numbers, $\delta_{k,m} \in \{0, 1\}$ is defined for all $k, m \in N^n$ and $\delta_{k,m} \equiv 1$ if $k = m$ and $\delta_{k,m} \equiv 0$ if $k \neq m$. For $k \in N^n$,$|k| \equiv \sum_{1 \leq i \leq n} k_i$. For the set of objects $u_m, m \in N$, $Span_R\{u_m | m \in N\} \equiv \{\sum_{m \in N} a_m u_m | a_m \in R\}$. For $1 \leq i \leq n$, $\epsilon_i \equiv (0, ..., 0, 1, 0, ...0)$, the i^{th} standard basis vector. $R[T] \equiv R[T_1, ..., T_n]$ is the set of polynomials (with coefficients from R) in $(T_1, ..., T_n)$. $R[[T]] \equiv R[[T_1, ...T_n]]$ is the set of formal power series (with coefficients from R) in $(T_1, ..., T_n)$. Similarly, $R[\beta][[T]]$ is the set of power series with coefficients from the ring $R[\beta]$. For symbols $x_1, ..., x_n$, $\Im$ is the set of all ideals in $R[x]$ and $\Im_0$ will be the set of zero ideals (those only having 0 for their solution set). For $I \in \Im$, $Z(I) = \{r \in R^N | P(r) = 0 \forall P(x) \in I\}$. B will be a lexicographically totally ordered basis for the algebra $R[\alpha]$ and $\tilde{B}$ is its complement in the set $\{\alpha^k | k \in N^n\}$. $J \subset N^n$ is the set of exponents of α which lie in B and $\tilde{J}$ is its complement in N^n. In addition to its usual meaning over the reals, $<$ will be used to denote (compatible) total orders on the elements of N^n and on the monomials of $R[\alpha]$. $\leq\leq$ will be used for a partial order on N^n and on the monomials of $R[\alpha]$.

4. Properties of R linear spaces

In this section we discuss the consequences of having elements of $R[x]$ act as linear operators on an R linear space. We then describe conditions which must be met for an R linear space to separate the *zero* ideals (those having $Z(I) = \{0\}$). For the ideal I of $R[x]$ whose elements act on R linear space U we will use the notation $\mathbf{N}_U(I) = \{f \in U | P(x)f = 0 \forall P(x) \in I\}$.

Theorem 1. *Let U be any R linear space on which the elements of $R[x] = R[x_1, ..., x_n]$ act as R linear operators. Then*

(i) $\mathbf{N}_U(R[x]) = \{0\}$.

(ii) $\mathbf{N}_U(I_1) \subset \mathbf{N}_U(I_2)$ *for any ideals* I_1, I_2 *in* $\Im$ *satisfying* $I_1 \supset I_2$.

(iii) For $m = (m_1, ..., m_n) \in N^n, m_i \neq 0$ *we have* $x_i : \mathbf{N}_U\langle x^m\rangle \to \mathbf{N}_U\langle x^{m-\epsilon_i}\rangle$.

Proof. (i) Suppose $v \in \mathbf{N}_U(R[x])$. Then since $1 \in R[x]$ we get $1(v) = 0$.

(ii) Let $f \in \mathbf{N}_U(I_1)$; then $P(x)f = 0$ for all $P(x) \in I_1$. In particular, if $P(x) \in I_2$, then $P(x)f = 0$, so that $f \in \mathbf{N}_U(I_2)$.

(iii) If $v \in \mathbf{N}_U\langle x^m\rangle$, then $x^{m-\epsilon_i}(x_i v) = 0$ so that $x_i v \in \mathbf{N}_U\langle x^{m-\epsilon_i}\rangle$. □

Remark 1. The simplest vector space satisfying the conditions of the theorem is $U = \{0\}$.

Now we consider a linear space which separates all zero ideals:

$$\Im_0 \equiv \{I \in \Im | Z(I) = \{0\}\}.$$

Let U be an R linear space on which the elements of $R[x]$ act as linear operators, and for $I \in \Im_0$ let $\mathbf{N}_U(I) \equiv \{f \in U | If = 0\}$ (where $If = 0$ means that $P(x)f = 0$ *for all* $P(x) \in I$). We say that U *separates the ideals* in $\Im_0$ if $I_1 \neq I_2$ implies $\mathbf{N}_U(I_1) \neq \mathbf{N}_U(I_2)$.

We will say that U is a *minimal separating* R linear space for the set of ideals $\Im_0$ if every separating vector space U^* contains an isomorphic copy of U. We now consider the consequences of U being a minimal linear space that separates the ideals of $\Im_0$.

Since

$$\langle x^{m+k}\rangle \subset_{\neq} \langle x^m\rangle \text{ for all } k \text{ such that } |k| > 0$$

we must also have that

$$\mathbf{N}_U\langle x^m\rangle \subset_{\neq} \mathbf{N}_U\langle x^{m+k}\rangle.$$

Hence for every k such that $|k| > 0$ there exists $T^{m+k} \in \mathbf{N}\langle x^{m+k}\rangle$ so that $T^{m+k} \notin \mathbf{N}\langle x^m\rangle$. Therefore the smallest possible R linear space which separates the ideals in $\Im_0$ has all T^{m+k} equal to one another for all $|k| > 0$. We now construct an R linear space satisfying this condition.

Let $\{T^m | m \in N^n\}$ be a set of distinct symbols related to one another by

$$x^k T^m = T^{m-k} \text{ if } (m-k) \in N^n \text{ and}$$

$$x^k T^m = 0 \text{ otherwise .}$$

Theorem 2. *The R linear space*

$$R[T] \equiv Span_R\{T^m | m \in N^n\}$$

separates the ideals in $\Im_0$.

Proof. For $m \in N^n$, $\mathbf{N}\langle x^m\rangle \equiv \mathbf{N}_{R[T]}\langle x^m\rangle = Span_R\{T^k | k \in N^n, (k-m) \notin N^n\}$. In particular, $T^m \notin \mathbf{N}\langle x^m\rangle$. However, for all k such that $|k| > 0$, we have $T^m \in \mathbf{N}\langle x^{m+k}\rangle$. By construction $a \cdot T^m$ $(a \in R, a \neq 0)$ are the only elements in all of the $\mathbf{N}\langle x^{m+k}\rangle$ which are not in $\mathbf{N}\langle x^m\rangle$. Hence $R[T]$ separates the ideals in $\Im_0$ and is minimal. □

Remark 2. In Section 6 we show that $R[[T]]$ (the completion of $R[T]$) separates *all* ideals of $R[x]$.

Now

$$R[[T]] = \{ \sum_{k \in N^n} a_k T^k | a_k \in R\}$$

and for every $m \in N^n$ there exists a linear functional

$$\pi_m : R[[T]] \to R$$

so that

$$\pi_m(\sum_{k \in N^n} a_k T^k) \equiv a_m.$$

One major difference between $Z(I)$ and $\mathbf{N}(I)$ is already apparent: $\mathbf{N}(I)$ is always an R linear space whereas $Z(I)$ is a linear space only for very special cases of I. Generally speaking $\mathbf{N}(I)$ contains a multitude of points which are not of the form $\rho(r), r \in R^n$. As a simple example of this last statement we offer the following: Let $n = 1$ and let $P(x) = (x-1)^2$. Then $Z\langle P(x)\rangle = \{1\}$ whereas $\mathbf{N}\langle P(x)\rangle$ contains the R linearly independent points $\sum_{0 \leq k} T^k$ and $\sum_{0 \leq k} kT^k$. In fact these linearly independent points form a basis for $\mathbf{N}\langle P(x)\rangle$

5. A basis for $\mathbf{N}(I)$

In this section we give an algorithm which is used to find a countable basis for $\mathbf{N}(I)$ for any $I \in \Im$ which is then used to prove various versions of Nullstellensatz and related theorems.

5.1. **A total order.** We *totally order* N^n as follows: For $k \in N^n$ we set

$$|k| = k_1 + \cdots + k_n.$$

For $k, m \in N^n$ we write $k < m$ if either (i) $|k| < |m|$ or (ii) if $|k| = |m|$ and the leftmost nonzero element of $m - k$ is negative.

It follows that $<$ is a total order on N^n. It also follows that if $u, v, w \in N^n$ and

$$\text{if } u < v, \text{ then } u + w < v + w. \tag{5}$$

The ordering given by $<$ is called a *graded lexicographic ordering* on N^n with $T_1 < T_2 < \cdots < T_n$. It is described on page 17 of [4]. We also introduce the *partial order* $\leq\leq$ on N^n defined by $k \leq\leq m$ if $k_i \leq m_i$ for $1 \leq i \leq n$. Similarly we write $k << m$ if $k \leq\leq m$ and $k_i < m_i$ for some $1 \leq i \leq n$. It follows that if $k << m$, then $k \leq\leq m$ and that the orderings $\leq\leq, <$ are *compatible* in the sense that if $k \leq\leq m, k \neq m$, then $k < m$.

5.2. **Extreme terms.** We define the maps $lt : R[[T]] - \{0\} \to R[[T]]$ and $elt : R[[T]] - \{0\} \to N^n$ by

$$lt\left(\sum_{k\in N^n} a_k T^k\right) \equiv a_m T^m \text{ and } elt\left(\sum_{k\in N^n} a_k T^k\right) \equiv m$$

if $a_m \neq 0$ and $a_k = 0 \forall k < m$. The term *lt* stands for *least term* and *elt* stands for *exponent of least term*. The function *elt* is well defined for all elements of $R[[T]]-\{0\}$ since the exponents of these elements are bounded below by 0.

We also define the map $ht : R[x] - \{0\} \to R[x]$ by

$$ht\left(\sum b_k x^k\right) \equiv b_m x^m$$

if $b_m \neq 0$ and $b_k = 0$ $k > m$. The function *ht* stands for *highest term*. The function *ht* is well defined for all elements of $R[x] - \{0\}$ since a polynomial contains only a finite number of terms.

5.3. **An algebra and its ordered basis.** Given a finitely generated, commutative, and associative algebra $R[\alpha] = R[\alpha_1, \alpha_2, ..., \alpha_n]$, the α_i will solve a system of polynomial relations:

$$I(y) = \{Q(y) \in R[y] = R[y_1, ..., y_n] | Q(\alpha) = 0\},$$

and $I(y)$ is easily seen to be an ideal.

Conversely, if $I(y)$ is any ideal in $R[y]$, there exists a finitely generated, commutative, and associative algebra $R[\alpha]$ so that $I(y)$ is exactly the set of polynomial relations which $\alpha_1, ..., \alpha_n$ jointly solve. The following theorem describes the relationship between these two concepts.

Theorem 3. *Given an algebra $R[\alpha]$ there exists an ideal $I(y)$ so that $R[\alpha] \cong \frac{R[y]}{I(y)}$. Conversely, given an ideal $I(y)$ in $R[y]$ there exists an algebra $R[\alpha]$ so that $I(y) = \{Q(y) \in R[y] | Q(\alpha) = 0\}$ and $R[\alpha] \cong \frac{R[y]}{I(y)}$.*

Proof. This is a standard result in algebra. The isomorphism is $\alpha_i \leftrightarrow [y_i]$ where $[y_i]$ is the residue class corresponding to y_i in $\frac{R[y]}{I(y)}$. □

We now inductively define an *ordered basis* B for the vector space $R[\alpha]$. B will be a subset of $B^* \equiv \{\alpha^k = \alpha_1^{k_1} \cdots \alpha_n^{k_n} | k \in N^n\}$. We say that α^j *precedes* α^k in B^* if $j < k$ in the total order $<$.

Now for every $k \in N^n$ one of the following two conditions holds for α^k: either

(i) there exists $c_{k,i} \in R (i < k)$ so that $\alpha^k = \sum_{i<k} c_{k,i}\alpha^i$ or
(ii) α^k *cannot* be written as a R linear combination of preceding elements.

We define an ordered basis B of $R[\alpha]$ by

$$B \equiv \{\alpha^k | \nexists c_{k,i} \in R \text{ so that } \alpha^k = \sum_{i<k} c_{k,i}\alpha^i\} . \tag{6}$$

(Note: $\alpha^0 = 1 \in B$ since no term precedes 1 in B^*.) Corresponding to B is the set of exponents of the basis elements:

$$J \equiv \{k \in N^n | \alpha^k \in B\} . \tag{7}$$

We will refer to B as the *ordered basis* defined in terms of the total order $<$. It follows that every $\alpha^k \notin B$ can be written *uniquely* in the form

$$\alpha^k = \sum_{i \in J, i<k} c_{k,i}\alpha^i \tag{8}$$

where $c_{k,i} \in R$.

5.4. **The unique reduced Grobner basis for an ideal.**

Lemma 1. *Let B, J be defined as in* (6) *and* (7)*. Then there exists a unique, minimal, and finite set of incomparable (with respect to $\leq\leq$) exponents $V = \{v_1, ..., v_t\} \in \tilde{J}$ (= the complement of J in N^n) so that $\alpha^j \notin B$ if and only if $v_i \leq\leq j$ for some $1 \leq i \leq t$.*

Proof. The proof uses Dickson's Lemma and is given in [2]. □

We will refer to the unique set $\{\alpha^{v_i} | 1 \leq i \leq t\}$ as the *reduced, generating* set of $\tilde{B} = \{\alpha^k | k \in \tilde{J}\}$. Recalling the definition of I, we see that $\alpha^{v_i} \notin B$ for $1 \leq i \leq t$. So there exists a unique $c_{v_i,k} \in R$ so that

$$\alpha^{v_i} = \sum_{k \in J, k<v_i} c_{v_i,k}\alpha^k.$$

Hence, letting

$$Q_i(y) = (y^{v_i} - \sum_{k \in J, k<v_i} c_{v_i,k}y^k), \tag{9}$$

we see that $Q_i(y) \in I(y)$ for $1 \leq i \leq t$ and that

$$ht(Q_i(y)) = y^{v_i}.$$

Although we won't use the fact we note that $\{Q_i(y) | 1 \leq i \leq t\}$ is the *unique, reduced Grobner basis* for $I(y)$.

5.5. **The Hilbert Characteristic Function.** For ideal I having index set J let

$$\chi_m(I) \equiv |\{k \in J | |k| \leq m\}|$$

where $|\cdot|$ of a set is its cardinality. $\chi_m(I)$ is the Hilbert Characteristic Function of I and is a polynomial in m for large m.

In section 9 we discuss the relationship between $\mathbf{N}(I)$ and $\chi_m(I)$.

5.6. **The solution space of an ideal I.** Let I be an ideal in $R[x]$. We define

$$\mathbf{N}(I) = \{f = f(T) \in R[[T]] | P(x)f = 0 \text{ for all } P(x) \in I\} .$$

$\mathbf{N}(I)$ is an R linear subspace of $R[[T]]$. One goal of this paper is to find an explicit basis for $\mathbf{N}(I)$ for any I in $\Im$. We solve this problem by finding an ordered basis for the algebra $R[\alpha] \cong \frac{R[y]}{I(y)}$ and then rewrite the expression $w(T) = \sum_{k \in N^n} T^k \otimes \alpha^k$ in terms of this basis (say $w(T) = \sum_{k \in J} w_k(T) \otimes \alpha^k$). We show that $\{w_k(T) | k \in J\}$ forms a basis for $\mathbf{N}(I)$ meaning that if $f \in \mathbf{N}(I)$, then there exists a *unique* $a_k \in R$ so that $f = \sum_{k \in J} a_k w_k(T)$.

5.7. **Construction of the nullspace basis.** We fix I and define α by $R[\alpha] \cong \frac{R[y]}{I(y)}$. We explore the expression

$$w(T) = \sum_{k \in N^n} T^k \otimes \alpha^k \tag{10}$$

where $P(\alpha) = 0$ for all $P(y) \in I(y)$ (and only $P(y) \in I(y)$). In (10), for each $\alpha^k \notin B$, we replace α^k by its unique R linear combination of preceding elements $\sum_{i \in J, i<k} c_{k,i}\alpha^i$ (see (8)). So

$$\begin{aligned} w(T) &\equiv \sum_{k \in \tilde{J}} T^k \otimes (\sum_{i \in J, i<k} c_{k,i}\alpha^i) + \sum_{k \in J} T^k \otimes \alpha^k \\ &= \sum_{k \in J} (\sum_{k<i, i \notin J} c_{i,k} T^i) \otimes \alpha^k + \sum_{k \in J} T^k \otimes \alpha^k \\ &= \sum_{k \in J} g_k(T) \otimes \alpha^k + \sum_{k \in J} T^k \otimes \alpha^k \qquad (11) \\ &= \sum_{k \in J} w_k(T) \otimes \alpha^k \qquad (12) \end{aligned}$$

where (11) and (12) also serve as the definitions of $g_k(T)$ and $w_k(T)$ for $k \in J$. Now $g_k(T) = \sum_{k<i, i \notin J} c_{i,k} T^i$ and

$$w_k(T) = T^k + g_k(T) \tag{13}$$

are defined for all $k \in J$ (and only $k \in J$).

Also by (11) and (12) and by the definition of $w_k(T)$, $g_k(T)$ we have

$$\pi_j g_k(T) = 0 \text{ for all } j, k \in J, \tag{14}$$

$$elt(g_k(T)) > k \text{ , and} \tag{15}$$

$$\pi_j w_k(T) = \delta_{j,k} \text{ for all } j, k \in J \text{ .} \tag{16}$$

5.8. **The Nullspace Basis Theorem.** We can now state and prove:

Theorem 4 (The Nullspace Basis Theorem)**.** *Let I be an ideal in $R[x]$. If $f = f(T) \in \mathbf{N}(I)$, then $f = \sum_{k \in J} (\pi_k f) w_k(T)$ is the unique representation for f as a real, linear combination of $w_k(T), k \in J$.*

5.8.1. *Proof that $w_k(T)$ are linearly independent and are in $\mathbf{N}(I)$.* Let $\phi : R[y] \to R[\alpha]$ be the homomorphism mapping $y_i \to \alpha_i$ (and so $P(\alpha) = 0$ for all $P(y) \in I(y)$). Hence

$$w(T) = (1 \otimes \phi)(\sum_{k \in N^n} T^k \otimes y^k) \text{ .} \tag{17}$$

We also note that (for $m \in N^n$)

$$(x^m \otimes 1) \left(\sum_{k \in N^n} T^k \otimes y^k \right) = \sum_{k \in N^n} T^k \otimes y^{k+m} \text{ .} \tag{18}$$

Let $P(x) = \sum_{i \in F} b_i x^i \in I$ (where F is the finite set of nonzero exponents occurring in $P(x)$).

We compute $(P(x)\otimes\phi)(\sum_{k\in N^n} T^k\otimes y^k)$ in two different ways. Now

$$(P(x)\otimes\phi)(\sum_{k\in N^n} T^k\otimes y^k) = (1\otimes\phi)(\sum_{k\in N^n} T^k\otimes y^k P(y)) \quad \text{(by (18))}$$
$$= \sum_{k\in N^n} T^k\otimes\alpha^k P(\alpha) = 0 \quad \text{(since } P(\alpha)=0\text{)}.$$

So

$$0 = (P(x)\otimes\phi)(\sum_{k\in N^n} T^k\otimes y^k) = (P(x)\otimes 1)w(T)$$
$$= (P(x)\otimes 1)(\sum_{k\in J} w_k(T)\otimes\alpha^k) \quad \text{(by (12))}$$
$$= \sum_{k\in J} P(x)w_k(T)\otimes\alpha^k.$$

Since $\{\alpha^k | k\in J\}$ is a linearly independent set over R we conclude that $P(x)w_k(T) = 0$ for all $k\in J$. This argument shows that $w_k(T)\in \mathbf{N}(I)$ for all $k\in J$ and that $\{w_k(T)|k\in J\}$ is a linearly independent set over R.

5.8.2. *Proof that $w_k(T)$ spans $\mathbf{N}(I)$.* We now show that every $f\in\mathbf{N}(I)$ can be written as a linear combination of the $w_k(T)$. We calculate $\pi_k f$ for all $k\in J$ and then we form

$$\hat{f} = f - \sum_{k\in J}(\pi_k f)w_k(T)$$

and so we know that $\hat{f}\in\mathbf{N}(I)$. We will show that $\hat{f}=0$ by showing that $\pi_k\hat{f}=0$ for all $k\in N^n$ from which we can conclude that $f=\sum_{k\in J}(\pi_k f)w_k(T)$.

We suppose by way of contradiction that $\hat{f}\neq 0$. So it must be that either $\exists j\in J$ so that $\pi_j\hat{f}\neq 0$ or $\exists j\in\tilde{J}$ so that $\pi_j\hat{f}\neq 0$. Now for $j\in J$ we have $\pi_j\hat{f} = \pi_j f - \sum_{k\in J}(\pi_k f)(\pi_j w_k(T)) = \pi_j f - \sum_{k\in J}(\pi_k f)\delta_{j,k}$ (by (16)). So $\pi_j\hat{f} = \pi_j f - \pi_j f = 0$.

Hence it must be that $\pi_j\hat{f}\neq 0$ for some $j\in\tilde{J}$. So we assume that $\hat{f} = \sum_{k\in\tilde{J}} a_k T^k, a_k\in R$, where some $a_k\neq 0$. Let $lt(\sum_{k\in\tilde{J}} a_k T^k) = a_m T^m$ be the least term (in the total order $<$). So $a_m\neq 0$. By Lemma 1 there exists $s, 1\leq s\leq t$ so that $v_s\leq\leq m$. We now look at $Q_s(x)\hat{f}$ where $Q_s(x)$ is defined in (9).

Claim. $lt(Q_s(x)\hat{f}) = a_m T^{m-v_s}$ (so that $\hat{f}\notin\mathbf{N}(I)$ which gives us our contradiction). For simplicity we write $\hat{f} = a_m T^m + \sum_{i\in M} a_i T^i$ and $Q_s(x) = x^{v_s} + \sum_{j\in U} b_j x^j$ where these expressions define the subsets M and U of N^n and where

$$m < i\forall i\in M, j < v_s\forall j\in U, \text{ and } v_s\leq\leq m\ . \tag{19}$$

So

$$Q_s(x)\hat{f} = a_m T^{m-v_s} + \sum_{i\in M} a_i T^{i-v_s} + \sum_{j\in U} a_m b_j T^{m-j} + \sum_{i\in M}\sum_{j\in U} a_i b_j T^{i-j}\ .$$

But (5) and (19) imply the following:

$$m - v_s < i - v_s\forall i\in M, \text{ so we have } \pi_{m-v_s}(T^{i-v_s}) = 0\ \forall i\in M\ ,$$
$$m - v_s < m - j\forall j\in U, \text{ so we have } \pi_{m-v_s}(T^{m-j}) = 0\forall j\in U\ ,$$

and

$m - v_s < i - j \forall i \in M, \forall j \in U$, so we have $\pi_{m-v_s}(T^{i-j}) = 0 \forall i \in M, \forall j \in U$.

Hence $lt(Q_s(T)\hat{f}) = a_m T^{m-v_s} \neq 0$. This contradiction proves the theorem. □

6. $R[[T]]$ separates the ideals of $R[x]$

In section 4 we showed that if $I_1 \subset I_2$, then $\mathbf{N}(I_1) \supset \mathbf{N}(I_2)$. In this section we show that $\mathbf{N}(I_1) \supset \mathbf{N}(I_2)$ implies $I_1 \subset I_2$.

Let $R[\beta] \cong \frac{R[y]}{I_1(y)}$.

Lemma 2. *Let $g_k(\beta) \in R[\beta]$ for $k \in N^n$. If $\sum_{k\in N^n} T^k g_k(\beta) = 0$, then $g_k(\beta) = 0 \forall k \in N^n$.*

Proof. We extend the definition of $R[[T]]$ to $R[\beta][[T]]$ which is to say that $R[\beta][[T]]$ is the set of all functions from N^n to $R[\beta]$. So $g \in R[\beta][[T]]$ has the unique representation $g(T) = \sum_{k\in N^n} g_k(\beta)T^k$. Consequently $g(T) = 0$ if and only if $g_k(\beta) = 0$ for all $k \in N^n$. □

Theorem 5. *$I_1 \supset I_2$ if and only if $\mathbf{N}(I_1) \subset \mathbf{N}(I_2)$.*

Proof. We need only show that if $\mathbf{N}(I_1) \subset \mathbf{N}(I_2)$ and if $q(y) \in I_2(y)$, then $q(y) \in I_1(y)$.

As before we form $\sum_{k\in N^n} T^k \otimes \beta^k = \sum_{k\in J_1} w_k(T) \otimes \beta^k$. So $w_k(T) \in \mathbf{N}(I_1)$ for all $k \in J_1$ (where J_1 is the set of exponents which correspond to the basis elements in $R[\beta]$).

Since $\mathbf{N}(I_1) \subset \mathbf{N}(I_2)$ we conclude that $w_k(T) \in \mathbf{N}(I_2)$ for all $k \in J_1$. Hence $q(x)w_k(T) = 0 \forall k \in J_1$ which implies that $q(x)\sum_{k\in J_1} w_k(T) \otimes \beta^k = 0$ so that $q(x)\sum_{k\in N^n} T^k \otimes \beta^k = 0$. Hence $\sum_{k\in N^n} T^k \otimes q(\beta)\beta^k = 0$. In particular, by using Lemma 2, this implies that $q(\beta)\beta^0 = 0$. Whence $q(y) \in I_1(y)$ by Theorem 3. □

7. The Nullspace Theorem

For this discussion of Hilbert's Nullstellensatz, let R be an algebraically closed field, let I be an ideal in $R[x]$, and as before we let $Z(I) = \{r \in R^n | P(r) = 0 \forall P(x) \in I\}$. Additionally the *radical* of I is defined by

$$rad(I) = \{g = g(x) \in R[x] | \exists m \in N \text{ so that } g^m \in I\}.$$

One version of Hilbert Nullstellensatz says that

$$rad(I) = \mathbf{I}(Z(I))$$

where the $\mathbf{I}$ of a subset of points of R^n is the smallest ideal in $\Im$ containing those points as zeros. (Similarly, we will let $\mathbf{I}$ of a subset of points of $R[[T]]$ be the smallest ideal in $\Im$ containing those points in its nullspace.)

Now, once again, we let R be a field. Then for $R[[T]], R[x]$ we have

Theorem 6 (Nullspace Theorem)**.** *$\mathbf{I}(\mathbf{N}(I_1)) = I_1$ for all $I_1 \in \Im$.*

Proof. By definition we have

$$\mathbf{I}(\mathbf{N}(I_1)) = \{Q(x) \in R[x] | Q(x)f = 0 \text{ for all } f \in \mathbf{N}(I_1)\}.$$

Set $I_2 = \mathbf{I}(\mathbf{N}(I_1))$. It follows that $I_1 \subset I_2$. By way of contradiction we suppose that $I_1 \subset_{\neq} I_2$ (so that $\mathbf{N}(I_1) \supset_{\neq} \mathbf{N}(I_2)$). Then there exists $f \in R[[T]]$ so that $I_1(f) =$

0 but $I_2(f) \neq \{0\}$. But this implies that there exists $P(x) \in I_2$ so that $P(x)f \neq 0$ even though $f \in \mathbf{N}(I_1)$. This contradicts the definition of $\mathbf{I}(\mathbf{N}(I_1)) = I_2$. □

For any set of polynomials $\mathcal{Q} \subset R[x]$ we define

$$\mathbf{N}(\mathcal{Q}) \equiv \{f \in R[[T]] \mid Q(x)(f) = 0 \forall Q(x) \in \mathcal{Q}\}.$$

We have the obvious

Corollary 1. $\mathbf{N}(\mathbf{I}(\eta)) = \eta$ *for any algebraic nullspace* η.

Remark 3. Although Hilbert's Nullstellensatz requires that R be algebraically closed, algebraic closure is not required for the Nullspace Theorem.

8. The ideal-algebraic solution space correspondence

Theorem 7. *Let* I_1, I_2 *be any ideals in* $\Im$ *and let* $\mathbf{N}(I_1), \mathbf{N}(I_2)$ *be the corresponding algebraic nullspaces. Then*

(i) $\mathbf{N}(I_1 \cap I_2) = \mathbf{N}(I_1) \ + \ \mathbf{N}(I_2)$,
(ii) $\mathbf{N}(I_1 \ + \ I_2) = \mathbf{N}(I_1) \ \cap \ \mathbf{N}(I_2)$,
(iii) $\mathbf{I}(\mathbf{N}(I_1) \ + \ \mathbf{N}(I_2)) = \mathbf{I}(\mathbf{N}(I_1)) \cap \mathbf{I}(\mathbf{N}(I_2)) = I_1 \cap I_2$, *and*
(iv) $\mathbf{I}(\mathbf{N}(I_1) \ \cap \ \mathbf{N}(I_2)) = \mathbf{I}(\mathbf{N}(I_1)) \ + \ \mathbf{I}(\mathbf{N}(I_2)) = I_1 \ + \ I_2$.

Proof of (i). By the Nullspace Theorem we need only show that

$$I_1 \cap I_2 = \mathbf{I}(\mathbf{N}(I_1) \ + \ \mathbf{N}(I_2)).$$

Now if $Q(x) \in \mathbf{I}(\mathbf{N}(I_1) \ + \ \mathbf{N}(I_2))$, then $Q(x)\mathbf{N}(I_1) = 0$ and $Q(x)\mathbf{N}(I_2) = 0$. These imply that $\mathbf{N}(I_1) \subset \mathbf{N}\langle Q(x)\rangle$ and $\mathbf{N}(I_2) \subset \mathbf{N}\langle Q(x)\rangle$ which imply that $Q(x) \in I_1$ and $Q(x) \in I_2$ by Theorem 4. Hence $Q(x) \in I_1 \cap I_2$. The proofs of the remaining parts of the theorem are similar and are left to the reader. □

9. $w_k(T)$ is in reduced echelon form

In this section we generalize the matrix concepts of echelon and reduced echelon form to certain subsets of $R[[T]]$. The fact that the rows of a reduced echelon form matrix form a basis for its row space and the fact that it is easy to write any element in the row space as a linear combination of these basis elements will generalize to our new setting. In particular we show that the basis $\{w_k(T)|k \in J\}$ we developed in the Basis Theorem is in *reduced echelon* form.

As before we have $R[[T]] = Span\{T^k|k \in N^n\}$ where N^n is totally ordered by lex order $<$ and any $f \in R[[T]]$ can be written uniquely in the form $f = \sum_{k \in N^n} \pi_k(f)T^k$. For $f \in R[[T]]$ we define $Support(f) = \{k \in N^n|\pi_k(f) \neq 0\}$ and $elt(f) = \{l \in Support(f)|l \leq k \forall k \in Support(f)\}$. $Support(f)$ and $elt(f)$ are well defined for all $f \in R[[T]]$. We also say that f is *monic* if $\pi_l(f) = 1$ for $l = elt(f)$. Let L be an R linear subspace of $R[[T]]$. We call $B \subset L$ a *basis* for L if every element of L is a unique countable linear combination of elements of B. We say $F \subset L$ is in *echelon form* if the elements of F are monic and if $elt : F \to N^n$ is injective. Let $elt(F) = \{elt(f)|f \in F\}$ and for $f, g \in F$ we will write $f < g$ if $elt(f) < elt(g)$. If F is in echelon form, then we totally order it using the total order from $elt(F)$. For F in echelon form we define $Span_R(F) = \{\sum_{j \in elt(F)} a_j F_j|a_j \in R\}$. It follows easily that if F is in echelon form, then F is a basis for $Span_R(F)$. We also define $Support(F) \equiv \{Support(f)|f \in F\}$.

We say that $F \subset R[[T]]$ is in *reduced echelon form* if F is in echelon form and if for all $k \in elt(F)$ there is only one $f \in F$ so that $k \in Support(F)$. The following theorems are straightforward applications of the definitions.

Theorem 8. *If $F = \{f_k | k \in elt(F) \subset N^n\} \subset R[[T]]$ is in reduced echelon form, then $g = \sum_{k \in elt(F)} \pi_k(g) f_k$ for all $g \in Span(F)$, and this representation is unique.*

Theorem 9. *If $Span(F) = Span(G)$ where F and G are in reduced echelon form (for the same lex order), then $F = G$.*

The point of this discussion is that $\{w_k(T) | k \in J\}$ defined in the proof of the Basis Theorem is in reduced echelon form, it is unique for a given lex order, and it is a basis for the space that it spans. The proof that $\{w_k(T) | k \in J\}$ is in reduced echelon form follows from the construction given in the Basis Theorem. (To get the Basis Theorem we would still need to prove that $Span\{w_k(T) | k \in J\} = \mathbf{N}(I)$.)

Another consequence of our construction (see (12),(7)) is that Hilbert's Characteristic function describes the number of generalized solutions.

This is also a good place to note that for $r \in Z(I)$ we have (by (13), (14), and (16))

$$\rho(r) = \sum_{k \in N^n} r^k T^k = \sum_{k \in J} r^k w_k(T).$$

Remark 4. We will say that I has *finitely many generalized solutions* if $\mathbf{N}(I)$ is finite dimensional.

10. Examples

10.1. **Multiplicities in R^n and $R[[T]]$.** When $n = 1$ every ideal is of the form $\langle P(x) \rangle$ and $Z(I)$ is finite for any $I \in \Im$. To determine an ideal from its solutions we need to include multiplicity information. Let $Z^*(I)$ be the multiset consisting of $Z(I)$ with every solution repeated with its proper multiplicity (so $|Z^* \langle P(x) \rangle| = degree(P(x))$ for any $P(x) \in \Im$). There is then a one to one order reversing correspondence between $\Im$ and $Z^*(\Im) \equiv \{Z^*(I) | I \in \Im\}$.

When $n > 1$ the simple idea of including multiplicity information with each element of $Z(I)$ is not sufficient for separating ideals. So suppose that $|Z(I)|$ is finite and suppose that we know the appropriate arithmetic multiplicity ($\in N_1 = \{1, 2, 3, ...\}$) of each zero. It is easy to find two ideals having $Z(I_1) = Z(I_2)$, having identical arithmetic multiplicity information and yet so that $I_1 \neq I_2$. In particular let $I_1 = \langle {x_1}^2, {x_2}^2 - x_1 \rangle$ and $I_2 = \langle {x_1}^2, {x_2}^2 \rangle$. It is clear that $\{0\} = Z(I_1) = Z(I_2)$ and that (by Bezout's Theorem) this zero should be repeated 4 times. It is also easy to check that $I_1 \neq I_2$. So we cannot determine a finite dimensional ideal if we are given its solutions along with their multiplicities. Indeed these ideals also have the same Hilbert characteristic function, so the finer detail given by that function does not help us distinguish them.

On the other hand $\mathbf{N}(I_1) = Span\{1, T_1 + {T_2}^2, T_2, T_1 T_2 + {T_2}^3\}$ and $\mathbf{N}(I_2) = Span\{1, T_1, T_2, T_1 T_2\}$, and both of these sets are subsets of $R[T]$ (which we know separates the ideals of $\Im_0$). These bases can be found either directly or by resorting to the construction in the basis theorem. In either case, we clearly have $\mathbf{N}(I_1) \neq \mathbf{N}(I_2)$. We can also easily read both the arithmetic multiplicity and algebraic multiplicity from these bases. Arithmetic and algebraic multiplicity are discussed in the introduction of the paper *On the Multiplicities in Polynomial System Solving* [5], pg 3283.

10.2. **Resolving an inconsistency.** $Z\langle {x_2}^2 - x_1\rangle$ contains $(0,0)$ with multiplicity 1. When we add a *constraining* equation ${x_1}^2 = 0$ we get the paradoxical result that (by Bezout's theorem) $Z\langle {x_2}^2 - x_1, {x_1}^2\rangle$ has 4 solutions at $(0,0)$. Clearly there are two different meanings of *solution* being used here!

When this problem is looked at in $R[[T]]$ this paradox disappears. So by the Basis Theorem $\mathbf{N}\langle {x_2}^2 - x_1\rangle$ has an infinitely countable basis of the form $\{w_k(T_1,T_2)|k \in J\}$ where $J = \{(k_1,k_2)|k_1 \in N, k_2 \in \{0,1\}\}$. The first few (by lex order) $w_k(T)$ are $1, T_1 + {T_2}^2, T_2, {T_1}^2 + T_1{T_2}^2 + {T_2}^4, T_1T_2 + {T_2}^3$, and it is easy to check that these polynomials are indeed annihilated by ${x_2}^2 - x_1$. We also have that $\mathbf{N}\langle {x_1}^2\rangle = Span\{{T_1}^{k_1}{T_2}^{k_2}|k_2 \in N, k_1 \in \{0,1\}\}$. By Theorem 7 we have that $\mathbf{N}\langle {x_2}^2 - x_1, {x_1}^2\rangle = \mathbf{N}\langle {x_2}^2 - x_1\rangle \cap \mathbf{N}\langle {x_1}^2\rangle$ which is easily seen to equal $Span\{1, T_1 + {T_2}^2, T_2, T_1T_2 + {T_2}^3\}$, and so we have 4 distinct solutions at the origin.

References

[1] F.Macaulay, "The Algebraic Theory of Modular Systems", Cambridge U. Press, 1916. Reprint with new introduction, Cambridge U. Press, Cambridge, 1994. MR1281612 (95i:13001)

[2] P. S. Pedersen, "A Basis for Power Series Solutions to Systems of Linear Constant Coefficient Partial Differential Equations" *Adv. Math.*, **141** , 155-166 (1999). MR1667149 (99k:35022)

[3] D. Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry* (Springer-Verlag, New York, 1995). MR1322960 (97a:13001)

[4] D. Cox, J. Little, D. O'Shea, *Ideals, Varieties, and Algorithms* (Springer-Verlag, New York, 1992). MR1189133 (93j:13031)

[5] M.G.Marinari, H.M.Moller, T.Mora, "On Mulitplicities in Polynomial System Solving" *Trans. Amer. Math. Soc.* **348** (1996), no. 8, 3283-3321. MR1360228 (96k:13039)

Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, New Mexico 87545