J. Symbolic Computation (1988) **6**, 135–147

On the Complexity of Computing Syzygies

DAVID BAYER AND MICHAEL STILLMAN‡

Department of Mathematics,
Columbia University, New York, NY 10027, *U.S.A.*
Department of Mathematics,
Cornell University, Ithaca, NY 14853, U.S.A.

We give a self-contained exposition of Mayr & Meyer's example of a polynomial ideal exhibiting double exponential degrees for the ideal membership problem, and generalise this example to exhibit minimal syzygies of double exponential degree. This demonstrates the existence of subschemes of projective space of double exponential regularity.

Introduction

Let k be a field, and let $R = k[x_1, \ldots, x_n]$ be a polynomial ring. Let $I \subset R$ be an ideal generated by $h_1, \ldots, h_s$ of degree $\leqslant d$. Many important operations in computational ring theory rely on the basic operations of constructing a standard or Gröbner basis for the ideal I, and constructing a basis of syzygies of I. Corresponding to these two operations are the following two problems, which are closely related:

(*IM*) *Ideal membership.* If $h \in I$, what degrees can occur for $g_1, \ldots, g_s \in R$ of minimal degree so $h = \Sigma_{i=1,s} g_i h_i$?

(*SYZ*) *Syzygies.* If I is homogeneous, what degrees can occur for the minimal (first) syzygies of I?

Once a standard or Gröbner basis has been constructed for I, the membership of $h \in I$ can be easily determined, and the $g_1, \ldots, g_s$ found, for any h (Buchberger, 1976). Thus, an answer to the problem (IM) gives a lower bound to the complexity of computing standard or Gröbner bases. Similarly, the complexity of computing a basis of syzygies of I is controlled by the degrees in which these syzygies occur.

Hermann (1926) (see also Seidenberg (1974) and Masser & Wüstholz (1983)) gave an upper bound, double exponential in the number of variables n, which applies equally to both of the above problems (IM) and (SYZ). More recently, Mayr & Meyer (1982) proved using methods of complexity theory that for (IM), the double exponential form of Hermann's bound for (IM) cannot be improved. This came as some surprise, since the rates of growth familiar to algebraic geometers had all been single exponential, such as Bezout's theorem on the number of points in the intersection of n hypersurfaces in n variables.

A careful look at Mayr & Meyer's construction yields the corresponding statement for (SYZ): The double exponential form of Hermann's bound for (SYZ) cannot be improved.

---

‡ Partially supported by N.S.F. grant DMS 8403168.

0747-7171/88/020135+13 $03.00/0

This is not apparent from the result of Mayr & Meyer (1982) alone, and in fact it was conjectured for some time after their work that Hermann's bound for (SYZ) could be substantially improved.

On a different front, Mayr & Meyer's result created widespread pessimism in the field of computer algebra about the viability of computing with standard or Gröbner bases, even as others were successfully solving large, naturally occurring problems using these bases.

These different points of view can be reconciled with the aid of the concept of regularity. Recall that a homogeneous ideal $I \subset R$ is defined to be m-regular if for $j \geqslant 0$, the jth syzygies of I are of degree $\leqslant m+j$ (Mumford, 1966; Eisenbud & Goto, 1984); the regularity of I is defined to be the least m for which I is m-regular.

There is considerable interest in finding sharp bounds for the regularity of ideals I with good geometric properties; see Eisenbud & Goto (1984). Mumford has shown (unpublished) that if I is the ideal of a characteristic zero non-singular variety of dimension p, degree d, then I is m-regular for $m=(p+1)(d-2)+2$. In Eisenbud & Goto (1984), it is conjectured that if I is a prime ideal defining a variety of codimension r, degree d in $\mathbf{P}^{n-1}$, then I is $(d+1-r)$-regular. In Gruson, Lazarsfeld & Peskine (1983), this statement is proved for reduced, irreducible curves. In Pinkham (1988), non-singular surfaces in $\mathbf{P}^5$ are shown to be $(d-1)$-regular. We conjecture that any reduced subscheme of total degree d in $\mathbf{P}^{n-1}$ is d-regular.

Where does the work of Mayr & Meyer (1982) fit into this picture? Their construction can be used to demonstrate that some ideals have regularity double exponential in the number of variables n. Thus, problem instances in this domain can be highly intractable. On the other hand, the study of the regularity of ideals with good geometric properties shows that many syzygy problem instances which arise naturally in algebraic geometry are tractable; this agrees with practical experience in finding syzygies by computer. Thus, regularity provides a framework for grading problem instances according to their computational complexity, and for resolving the conflict between theory and practice observed above.

It is imperative that the boundary between "nice" ideals with low regularity, and "wild" ideals such as the example of Mayr & Meyer (1982), be better understood; this paper is concerned with the wild side of the above boundary.

In this paper, we give a self-contained exposition of the key construction of Mayr & Meyer (1982). In section 1, we give a sufficient condition for a syzygy of a homogeneous ideal generated by differences of monomials to be minimal. In section 2, we construct an ideal $J_n \subset A$ which exhibits double exponential degrees for the ideal membership problem, and an ideal $K_n \subset A[z]$ having a minimal syzygy of double exponential degree. In section 3, we examine the underlying geometry of the ideal membership problem, and consider the related question of the complexity of ideal membership of 1. An exciting recent result of D. Brownawell (1986) establishes that this restricted problem is computationally more tractable than ideal membership in general.

The bound given by Hermann (1926), Seidenberg (1974), Mayr & Meyer (1982) and Masser & Wüstholz (1983) is as follows: Assume that k is an infinite field, and let $b_i=\Sigma_{j=1,s} g_j h_{ij}$, $i=1,\ldots,t$, be a system of linear equations in $g_1,\ldots,g_s$, with each $b_i, h_{ij} \in R$. Let $d=\max_{i,j}\{\deg(h_{ij})\}$ and $B=\max_i\{\deg(b_i)\}$, taking $\deg(0)=0$. If these equations have a solution, then there exists a solution $(g_1,\ldots,g_s)$ where $\deg(g_i) \leqslant B+2(sd)2^{n-1}$ for each i. Furthermore, when each $b_i=0$ the R-module of solutions is generated by elements $(g_1,\ldots,g_s)$ where $\deg(g_i) \leqslant 2(sd)2^{n-1}$. It follows for (IM), the degree of each g_i is bounded by $\deg(h)+2(sd)2^{n-1}$; the corresponding degree bound for (SYZ) is $d+2(sd)2^{n-1}$. In Masser & Wüstholz (1983), a sharper bound is given for ideal membership of h: $\deg(g_i) \leqslant 2(2d)2^{n-1}$.

Let $d \geqslant 2$, and define $e_n = d^{2^n}$. Mayr & Meyer (1982) construct a polynomial ring A in $10n$ variables, and an ideal $I_n \subset A$ which can in effect count to e_n: Included among the variables of A are a "start" variable S, a "finish" variable F, four "counter" variables $B_1, \ldots, B_4$, and four "catalyst" variables $C_1, \ldots, C_4$. The ideal I_n is generated by differences of monomials, and contains the four differences $S\,C_i - F\,C_i\,B_i^{e_n}$. I_n is defined recursively in terms of I_{n-1}; its construction relies on the identity $e_n = (e_{n-1})^2$.

The relations $SC_i - FC_iB_i^{e_n} \in I_n$ are used in Mayr & Meyer (1982) as part of a construction which realises the halting problem for a bounded 3-counter machine as an instance of the decision problem for ideal membership: Given h, and $h_1, \ldots, h_s \in R$, does $h \in (h_1, \ldots, h_s)$? Thus, the decision problem for ideal membership is seen to be exponential space complete; the argument in Mayr & Meyer (1982) is valid over any field k. Since there are problems solvable in exponential space which are known to require exponential space, a degree bound for ideal membership which grows double exponentially in the maximum of the number of variables and the number of generators is seen to be inevitable.

By setting $B_1, \ldots, B_4, C_1, \ldots, C_4 = 1$ at the top level of recursion in the above construction, we obtain an ideal $J_n \subset A$ so $S - F \in J_n$; this instance of ideal membership directly exhibits double exponential degrees for any $g_1, \ldots, g_s$ so $S - F = \Sigma_{i=1,s} g_i h_i$, where $h_1, \ldots, h_s$ denote the generators of J_n. After homogenising with z, and adjoining $S - F$, we obtain an ideal $K_n \subset A[z]$ exhibiting a minimal syzygy of double exponential degree. K_n is thus of double exponential regularity.

We would like to thank David Mumford for many helpful conversations.

1. Syzygies Formed by Differences of Monomials

Let $I \subset R$ be an ideal generated by $h_1, \ldots, h_s$, where each $h_i = x^{A_i} - x^{B_i}$, and the differences $A_i - B_i \in \mathbf{Z}^n$, $i = 1, \ldots, s$ are all distinct.

1.1. Definition. In the above situation, define $G = G(h_1, \ldots, h_s)$ to be the directed graph having as vertex set the monomials of R, and having as edge set all directed edges (A, B) from x^A to x^B, so $A - B = A_i - B_i$ for some (unique) i.

1.2. Example. *Let* $R = k[x, y, z]$, and let

$$h_1 = xy^3z - y^2z^3, \qquad h_2 = xyz^3 - x^3z^2, \qquad h_3 = x^3yz - x^2y^3.$$

The ideal $I = (h_1, h_2, h_3)$ is homogeneous, and has $xh_1 + yh_2 + zh_3 = 0$ as a minimal syzygy.

A portion of the graph $G = G(h_1, h_2, h_3)$, with vertex set consisting of the monomials in R_5 and R_6, is shown in Fig. 1. The monomials are arranged as their exponents appear on antidiagonal slices of $\mathbf{N}^3 \subset \mathbf{R}^3$; the monomial nearest the label x is x^5, for example. The three edges shown on the monomials for R_5 correspond to the generators h_1, h_2, h_3 of I. The edges for R_6 all correspond to monomial multiples of generators; the closed triangle formed by three of the edges corresponds to the syzygy $xh_1 + yh_2 + zh_3 = 0$.

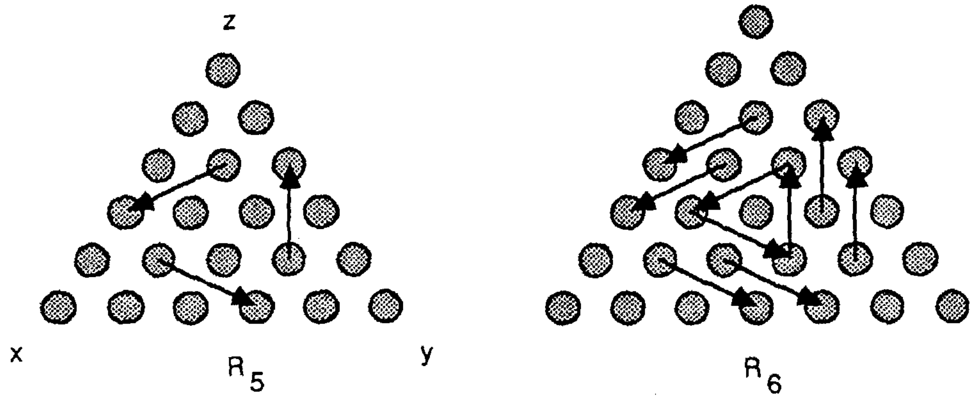

Fig. 1. A syzygy formed by differences of monomials.

1.3. Definition. A chain C in $G = G(h_1, \ldots, h_s)$ is a formal sum $\Sigma c_{A,B}(A, B)$ of edges (A, B) in G with coefficients $c_{A,B} \in k$; $C(G)$ will denote the set of chains in G. The value $|C| \in R$ of a chain C is the sum $\Sigma (c_{A,B}x^A - c_{A,B}x^B)$. C is a cycle if $|C| = 0$; $Z(G)$ will denote the set of cycles in G.

1.4. Example. One can see from the diagram of example (1.2) that

$$((2, 3, 1), (1, 2, 3)) + ((1, 2, 3), (3, 1, 2)) + ((3, 1, 2), (2, 3, 1))$$

is a cycle in G, corresponding to the closed triangle of edges on R_6.

For a chain $C = \Sigma c_{A,B}(A, B)$, and a monomial $x^D \in R$, define

$$x^D C = \Sigma c_{A,B}(A + D, B + D).$$

This gives the set of chains $C(G)$ an R-module structure. The value map $C \to |C|$ is seen to be an R-module homomorphism. Thus its kernel, the set of cycles $Z(G)$, is an R-module.

1.5. Lemma. *The module of syzygies, $SYZ(I)$, of $I = (h_1, \ldots, h_s)$ is isomorphic to the module of cycles $Z(G)$.*

Proof. The map from the module $M = Rh_1 \oplus \ldots \oplus Rh_s$ to the module of chains $C(G)$ obtained by sending $(g_1, \ldots, g_s) \in M$ to the chain $\Sigma g_i \cdot (A_i, B_i)$ is an R-module isomorphism. Since $\Sigma g_i h_i = 0$ iff the chain $\Sigma g_i \cdot (A_i, B_i)$ is a cycle, the submodule $\mathrm{SYZ}(I) \subset M$ is mapped isomorphically onto $Z(G) \subset C(G)$. ■

In fact, we have shown that the two exact sequences

$$0 \to \mathrm{SYZ}(I) \to Rh_1 \oplus \ldots \oplus Rh_s \to R$$

and

$$0 \to Z(G) \to C(G) \to R$$

are isomorphic, where the right-hand map of the second sequence is the value map for chains. The formalism of chains and cycles on G provides a graph-theoretic language for discussing the syzygy exact sequence; we shall make particular use of the ability to consider connected components of G.

Define the monomials of a chain $C=\Sigma c_{A,B}\cdot(A,B)$ to be those monomials x^A so $c_{A,B}\neq 0$ or $c_{B,A}\neq 0$ for some monomial x^B; this set of monomials will be called the support of C. Define

$$gcd(C)=gcd\{x^A|x^A \text{ is a monomial of } C\}.$$

The following is a sufficient criterion for a syzygy of I to be minimal, when I is homogeneous:

1.6. LEMMA. *Let I be a homogeneous ideal, and let $G=G(h_1,\ldots,h_s)$. Let G_A be the (connected) component of G containing the monomial x^A. Suppose that*

- (*i*) *the set of cycles with support in G_A is a one-dimensional vector space $\{aC|a\in k\}$ for some cycle C, and*
- (*ii*) $gcd(C)=1$.

Then the syzygy corresponding to C is a minimal syzygy of I.

PROOF. We show that in the module of cycles $Z(G)$, C cannot be expressed as $C=\Sigma\alpha_iC_i$ for cycles C_i, where each $\alpha_i=ax^D$ for some $a\in k$, $D\in\mathbf{N}^n$, and each $\deg(\alpha_i)>0$. The result follows by the isomorphism of lemma 1.5.

Suppose on the contrary that C can be so expressed. Write each C_i as $C_i'+C_i''$, where C_i' is a cycle so α_iC_i' has support in G_A, and C_i'' is a cycle so α_iC_i'' has support in $G-G_A$. Then $C=\Sigma\alpha_iC_i'$. However, each $\alpha_iC_i'=a_iC$ for some $a_i\in k$, by (i). At least one $a_i\neq 0$; for this i, we have $\alpha_i|gcd(C)$. Since $\deg(\alpha_i)>0$, this contradicts (ii). ■

2. The Example of Mayr and Meyer

In this section, we describe the construction of Mayr & Meyer (1982), and modify it to give examples for ideal membership and syzygies.

Let $n\geqslant 0$ and $d\geqslant 2$ be integers, and define $e_n=d^{2^n}$. Let V_j denote the set of variables $\{s_j,f_j,b_{j1},\ldots,b_{j4},c_{j1},\ldots,c_{j4}\}$, for $j=0,\ldots,n$; the variables in V_j will be said to be of level j. Let $A=k[V_0,\ldots,V_n]$.

As a notational convenience, when an integer r, $0\leqslant r\leqslant n$, is fixed, let upper-case letters denote level r variables, and let lower-case letters denote level $r-1$ variables: Let $S=s_r$, $F=f_r$, $B_i=b_{ri}$, $C_i=c_{ri}$, $s=s_{r-1}$, $f=f_{r-1}$, $b_i=b_{r-1,i}$, and $c_i=c_{r-1,i}$.

Define $I_0\subset A$ to be the ideal generated by

$$s_0c_{0i}-f_0c_{0i}b_{0i}^d \quad \text{for } i=1,\ldots,4;$$

these generators will be said to be of level 0. Given the ideal I_{r-1}, for $1\leqslant r\leqslant n$, define I_r to be the ideal generated by I_{r-1} and the new generators of level r,

$$\begin{array}{ll}
S-sc_1, & sc_4-F,\\
fc_1-sc_2, & sc_3-fc_4,\\
fc_2b_1-fc_3b_4, & sc_3-sc_2,\\
fc_2C_ib_2-fc_2C_iB_ib_3 & \text{for } i=1,\ldots,4.
\end{array}$$

A monomial $x^D\in A$ will be said to be of level j if

- (i) x^D involves only variables of level $\geqslant j$, and
- (ii) x^D is linear in $\{s_j,f_j\}$, and in $\{c_{j1},\ldots,c_{j4}\}$, and is not divisible by any of $s_{j+1},\ldots,s_n$ or $f_{j+1},\ldots,f_n$.

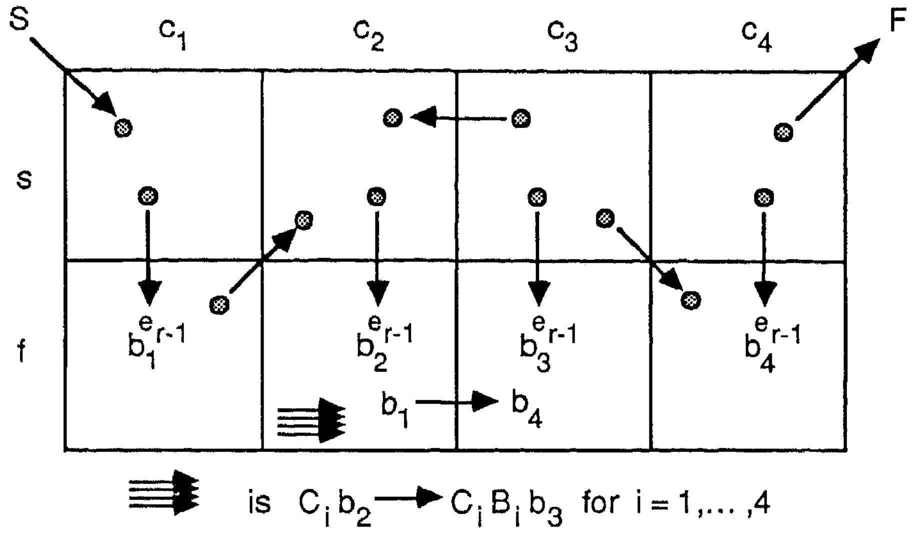

Fig. 2. The example of Mayr and Meyer.

Note that some monomials of A have no level; we shall be studying chains whose support consists of monomials which are of some level.

In abuse of notation, write $G(I_r)$ for the graph of definition 1.1 corresponding to the given generators of I_r. Figure 2 can be used as an aid in visualising $G(I_r)$. The eight boxes shown have row labels s, f, and column labels $c_1, \ldots, c_4$. Each monomial xD of level $r-1$ in A is divisible by unique row and column labels; assign x^D to the box with these labels. Assign the monomials of level r in A to the region outside of all eight boxes. The other monomials of A are not represented in Fig. 2.

The level r generators for I_r induce a multitude of directed edges on these monomials; each solid arrow is meant to denote all edges induced on diagram monomials by a given level r generator $x^D - x^E$. Each such x^D (or x^E) divides the monomials in exactly one diagram region; the corresponding arrow end is located in this region, and labelled with x^D (or x^E). If the region is a box, the variables s, f, and $c_1, \ldots, c_4$ have been suppressed from this label.

We shall see in lemma 2.2 that $sc_i - fc_i b_i^{e_{r-1}} \in I_{r-1}$ for $i = 1, \ldots, 4$; the vertical arrows represent these relations. In the graph $G(I_r)$, these relations can be obtained when $r \geqslant 2$ via multistep chains through vertices corresponding to monomials of level $< r-1$.

2.1. Example. Let $r = 1$, and $d = 3$. Then the ideal I_r contains the relations $S C_i - F C_i B_i^9$ for $i = 1, \ldots, 4$.

Figure 3 illustrates a chain whose value is this relation for a fixed i; each arrow which lies entirely inside box fc_2 is meant to represent a sequence of three edges. To interpret the path shown as a chain, follow it from left to right. Assign coefficient $+1$ to those edges that point in the direction of the path, and coefficient -1 to the remaining edges.

The relation $S C_i - F C_i B_i^9$ can be best understood by keeping track of the variables $b_1, \ldots, b_4$, which act as counters, as one moves along the chain. The chain obtains e_{r-1} copies of the counter b_1 on entering the box fc_1. It crosses from the box fc_2 to the box fc_3 a total of e_{r-1} times, as it laps around in the middle four boxes; a copy of b_1 is

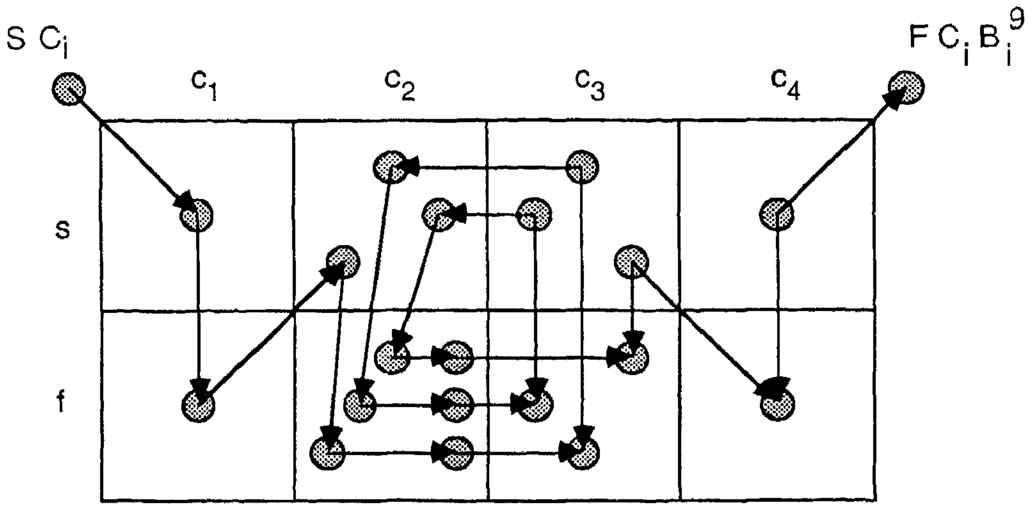

Fig. 3. A chain of equivalences.

converted to b_4 to count each crossing. The e_{r-1} copies of b_2 obtained on entering box fc_2 are each converted there into B_ib_3 in the presence of the catalyst C_i. The copies of b_3 are all consumed on exiting box fc_3. In this way, a total of $e_r=(e_{r-1})^2$ copies of B_i are obtained. The e_{r-1} copies of b_4 accumulated by the chain are consumed on exiting box fc_4.

This chain corresponds to the following step by step computation in A/I_r, using generators of I_r:

$$\begin{aligned}
S\,C_i &= s\,c_1C_i = fc_1\,C_ib_1^3 \\
&= s\,c_2\,C_ib_1^3 = fc_2\,C_ib_1^3b_2^3 \\
&= \ldots = fc_2\,C_iB_i^3b_1^3b_3^3 \\
&= fc_3\,C_iB_i^3b_1^2b_3^3b_4 = s\,c_3\,C_iB_i^3b_1^2b_4 \\
&= s\,c_2\,C_iB_i^3b_1^2b_4 = fc_2\,C_iB_i^3b_1^2b_2^3b_4 \\
&= \ldots = fc_2\,C_iB_i^6b_1^2b_3^3b_4 \\
&= fc_3\,C_iB_i^6b_1b_3^3b_4^2 = s\,c_3\,C_iB_i^6b_1b_4^2 \\
&= s\,c_2\,C_iB_i^6b_1b_4^2 = fc_2\,C_iB_i^6b_1b_2^3b_4^2 \\
&= \ldots = fc_2\,C_iB_i^9b_1b_3^3b_4^2 \\
&= fc_3\,C_iB_i^9b_3^3b_4^3 = s\,c_3\,C_iB_i^9b_4^3 \\
&= fc_4\,C_iB_i^9b_4^3 = s\,c_4\,C_iB_i^9 = F\,C_iB_i^9.
\end{aligned}$$

Define $p_r\colon A\to A$ by $p_r(v)=v$ for all variables v of level $<r$ and for $v=s_r$ or f_r, and $p_r(v)=1$ for all other variables v. For $r\geqslant 1$, define J_r to be the ideal generated by $p_r(I_r)$. J_r differs from I_r only in that the four level r generators $fc_2\,C_ib_2-fc_2\,C_iB_ib_3$, $i=1,\ldots,4$, are replaced by the single generator $fc_2b_2-fc_2b_3$.

2.2. Lemma. *The following statements hold for I_r, $r\geqslant 0$, and for J_r, $r\geqslant 1$.*

(*i*) *I_r contains the four relations*

$$S\,C_i-F\,C_iB_i^{e_r},\ i=1,\ldots,4,$$

and J_r contains the relation

$$S-F.$$

(*ii*) *The monomials in the component of a level r monomial in the graph $G(I_r)$ are all of level $\leqslant r$; the monomials in the component of S and F in the graph $G(J_r)$ are S, F, or monomials of level $<r$.*

(*iii*) *The component of a level r monomial in the graph $G(I_r)$ contains no cycles, and the component of S in the graph $G(J_r)$ contains no cycles.*

(*iv*) *In the graph $G(J_r)$, there is a unique chain supported in the component of S with value $S-F$. In the graph $G(I_r)$, if x^D, x^E are distinct monomials of level $\geqslant r$, and $x^D-x^E \in I_r$, then there is a unique chain supported in the component of x^D in $G(I_r)$ with value x^D-x^E. In this case x^D-x^E is a multiple of one of the relations given in* (*i*) *for I_r.*

PROOF. Statements (i)–(iv) are evident for $r=0$; inductively assuming these statements for I_j when $j<r$, we prove them for I_r and J_r when $r \geqslant 1$.

(i) For I_r, these relations are constructed exactly as in example 2.1. Applying p_r to one of these relations, we obtain $S-F \in J_r$.

(ii) We are considering the component of a monomial x^D in $G=G(I_r)$, or $G=G(J_r)$, where x^D is of level r, or $x^D=S$, respectively. One edge of G is incident on x^D; it connects x^D to a monomial of level $r-1$. From statement (ii) for I_{r-1}, all paths in G from monomials of level $r-1$, consisting of edges arising from I_{r-1}, lead to monomials of level $\leqslant r-1$. Edges in G arising from level r generators are not incident on any monomials of level $<r-1$; starting from a level $r-1$ monomial, such an edge leads either to

(a) another level $r-1$ monomial, or to
(b) a monomial which is divisible by S or F, and not divisible by any variables of level $<r$, with the possible exception of $b_1, \ldots, b_4$. In this case, this edge is the only edge incident on the monomial described.

We need to show that a monomial in case (b) is not divisible by $b_1, \ldots, b_4$. For $G(I_r)$, this monomial is then of level r; for $G(J_r)$, this monomial is then either S or F.

The degrees in $\{b_1, b_4\}$ and in $\{b_2, b_3\}$ of monomials of level $r-1$ on these connected components can be given as a function of which box they occupy in Fig. 2: Monomials in boxes sc_1 and sc_4 are of degree 0 in $\{b_1, b_4\}$, and monomials in boxes sc_2, sc_3, fc_1, fc_2, fc_3, and fc_4 are of degree e_{r-1} in $\{b_1, b_4\}$. Monomials in boxes fc_2 and fc_3 are of degree e_{r-1} in $\{b_2, b_3\}$, and monomials in boxes sc_1, sc_2, sc_3, sc_4, fc_1, and fc_4 are of degree 0 in $\{b_2, b_3\}$. One verifies this by observing that the edges arising from level r generators of I_r or J_r, and the chains of statement (iv) for I_{r-1}, preserve this grading. Thus the monomials of case (b) are not divisible by $b_1, \ldots, b_4$.

(iii) Suppose that C is a cycle in the connected component of S in the graph $G(J_r)$. Assume without loss of generality that exactly two edges of C are incident on each monomial of C; any cycle C can be decomposed into simple cycles of this form. S or F cannot be monomials of C, since only one generator of J_r applies to each of S or F. Thus by statement (ii), the remaining monomials in this component are all of level $\leqslant r-1$.

Suppose that the highest level of a monomial of C is $j-1$, for some $j<r$. Then no edges corresponding to generators of level $>j$ are involved in C, so C is already a cycle in the connected component of a level j monomial in the graph $G(I_j)$, which contradicts statement (iii) for j.

Thus the support of C must include monomials of level $r-1$. Choose a traversal of C, and call two level $r-1$ monomials of C adjacent if no other level $r-1$ monomials occur between them in this traversal. Such adjacent monomials are either joined in C by an edge corresponding to a level r generator of J_r, or by a chain corresponding to one of the relations $sc_i - fc_i b_i^{e_{r-1}}$; this follows from statement (iv) for I_{r-1}. From this, we see that no level $r-1$ monomials of C can lie in the outer boxes sc_1, fc_1, sc_4, or fc_4 of Fig. 2: A monomial in boxes $s\,c_1$ or $s\,c_4$ cannot have two adjacent monomials, and once these boxes are excluded, the same claim can be made for boxes fc_1 and fc_4. Furthermore, a monomial in boxes sc_2, sc_3, or fc_3, and its two adjacent monomials, lie in three distinct boxes. A monomial in box fc_2 is of degree e_{r-1} in $\{b_2, b_3\}$, by the above proof of statement (ii). Thus it is part of a sequence of e_{r-1} adjacent monomials in box fc_2, joined by edges corresponding to the generator $b_2 - b_3$, with the monomials on each end adjacent to monomials in boxes sc_1 and fc_3, respectively.

From these local considerations along the cycle C, we see that a traversal of C must loop some non-zero number of times through the four boxes sc_2, fc_2, fc_3, and sc_3, always in the same direction. In particular, edges arising from the generator $fc_2 b_1 - fc_3 b_4$ are used in the cycle C, and are always traversed in the same direction. Since no other edges in C affect the degree of b_1 in monomials of C, the degree of b_1 must decrease or increase along a traversal of C. This is impossible, so statement (iii) holds for $G(J_r)$.

Applying the map p_r to a cycle in $G(I_r)$ produces a cycle in $G(J_r)$. Since p_r maps level r monomials of A to either S or F, statement (iii) follows for $G(I_r)$.

(iv) The existence of a chain in the component of S in $G(J_r)$ with value $S-F$ is guaranteed by statement (i). This chain is unique, since the existence of more than one such chain would permit the construction of a cycle contradicting statement (iii) for $G(J_r)$.

Let C be a chain in the component of x^D in $G(I_r)$ with value $x^D - x^E$. Let C' be the chain in the compoment of $p_r(x^D)$ in $G(J_r)$ which is the image under p_r of C; C' has value $p_r(x^D - x^E) \in J_r$.

x^D or x^E must both be of level r, for otherwise we would have $p_r(x^D) = 1$ or $p_r(x^E) = 1$, contradicting the fact that no relations in J_r involve the monomial 1. Thus p_r maps x^D to either S or F, and x^E to either S or F.

p_r cannot map x^D and x^E both to S, or both to F, for in these cases $p_r(x^D - x^E) = 0$, so C' is a cycle, contradicting statement (iii) for $G(J_r)$. Thus, we can assume without loss of generality that $p_r(x^D) = S$ and $p_r(x^E) = F$. C' is then the unique chain in the component of S in $G(J_r)$ with value $S-F$.

Since x^D and x^E are of level r, and no relation of I_r changes the variables $C_1, \ldots, C_4$, every monomial of C is divisible by C_i for some fixed i. C is determined by the chain C' and i, and is thus unique: C differs from C' only in that edges corresponding to the generator $fc_2 b_2 - fc_2 b_3 \in J_r$ are replaced by edges corresponding to the generator $fc_2 C_i b_2 - fc_2 C_i B_i b_3 \in I_r$. There are a total of e_r such edges, so for some monomial α, $x^D = \alpha\, S\, C_i$, and $x^E = \alpha\, F\, C_i B_i^{e_r}$. Thus, $x^D - x^E$ is a multiple of one of the relations given in (i) for I_r. ■

2.3. LEMMA. *Let $h = S - F$, and let $h_1, \ldots, h_s$ be the generators of J_r. For any expression $h = \Sigma_{i=1,s} g_i h_i$, some g_i will have degree $\geqslant r - 1 + 2e_0 + \ldots + 2e_{r-1}$.*

PROOF. One of the monomials occurring in the unique chain of lemma 2.2(iv) from S to F in the compoment of S in $G(J_r)$ is

$$f_0\, c_{03}\, b_{03}^{e_0}\, b_{04}^{e_0} \ldots c_{r-1,3}\, b_{r-1,3}^{e_{r-1}}\, b_{r-1,4}^{e_{r-1}}.$$

One of the two edges of the chain which are incident on this monomial is a degree $r - 1 + 2e_0 + \ldots + 2e_{r-1}$ multiple of the generator $s_0 c_{03} - f_0 c_{04}$ of J_r. Any expression

$\Sigma_{i=1,s} g_i h_i$ for $S-F$ corresponds to a chain which differs from this chain only by the addition of cycles supported in other components of $G(J_r)$, so some g_i has at least the stated degree. ■

2.4. THEOREM (*Mayr & Meyer*, 1982). *Any degree bound for the ideal membership problem must grow double exponentially in the maximum of the number of variables and the number of generators.*

PROOF. Taking $r=n$ in lemma 2.3, the polynomial ring A has $10n$ variables, and the ideal J_n has $10n+1$ generators of degree $\max\{5, d+2\}$. Any set of g_i's for the instance $S-F \in J_n$ of ideal membership have maximum degree exceeding $e_{n-1} = d2^{n-1}$, which grows double exponentially in $10n+1$. ■

Let $A' = A[z]$, where z is a homogenising variable. Define the projection $q: A' \rightarrow A$ to be the identity on $A \subset A'$, and to map z to 1. Let $r=n$, and let $J'_n \subset A'$ be the ideal generated by the homogenisations, using z, of the generators of J_n. Let $K_n \subset A'$ be the ideal generated by J'_n and $S-F$. J'_n and K_n are homogeneous ideals of A'.

2.5. LEMMA. *K_n has a minimal syzygy of degree $m+1$, where*

$$m = n + 2e_0 + \ldots + 2e_{n-1}.$$

PROOF. Let H denote the component of $z^m S$ in $G(J'_n)$. All of the monomials of H have the same degree, so q maps this set of monomials injectively into A. This induces an embedding of the graph H into the component of S in $G(J_n)$. Thus by lemma 2.2(iii) for J_n, H has no cycles.

The component of $z^m S$ in $G(K_n)$ has the same monomials as H, and in addition to the edges of H, one new edge e corresponding to $z^m(S-F)$. There are no other monomials or edges, since $z^m S$ and $z^m F$ are the only monomials of H which are divisible by S or F, so $z^m(S-F)$ is the only multiple of $S-F$ giving an edge incident on monomials of H. Denote this component of $G(K_n)$ by $H \cup \{e\}$.

Let C be the unique chain with value $S-F$ supported in the component of S in $G(J_n)$, given by lemma 2.2(iv). One verifies that every monomial of C has degree $\leqslant m+1$. The chain C can be homogenised to a chain C' in $G(J'_n)$ all of whose monomials are of degree $m+1$, by multiplying the monomials of C by appropriate powers of z. Each edge of C' is then a multiple, by some power of z, of the corresponding edge in C. C' is supported on $H \cup \{e\}$, and has value $z^m(S-F)$.

$C' - \{e\}$ is then a cycle supported on $H \cup \{e\}$. $gcd(C' - \{e\}) = 1$, since

$$f_0 c_{03} b_{03}^{e_0} b_{04}^{e_0} \ldots c_{r-1,3} b_{r-1,3}^{e_{r-1}} b_{r-1,4}^{e_{r-1}},$$

$z^m S$ and $z^m F$ are all monomials of this cycle. Since H has no cycles, every cycle supported on $H \cup \{e\}$ is a multiple $a(C' - \{e\})$, $a \in k$, of $C' - \{e\}$. Thus, by lemma 1.6, this cycle corresponds to a minimal syzygy of K_n, of degree $m+1$. ■

2.6. THEOREM. *Any bound for the regularity of homogeneous ideals must grow double exponentially in the maximum of the number of variables and the number of generators.*

PROOF. The polynomial ring A' has $10n+1$ variables, and the ideal K_n has $10n+1$ generators of degree $\max\{5, d+2\}$. By lemma 2.5, K_n has a minimal (first) syzygy of degree $m+1$, where $m \geqslant e_{n-1} = d2^{n-1}$. Thus, the regularity of K_n grows double exponentially in $10n+1$. ■

In theorems 2.4 and 2.6, we would like to be able to assert simply that double exponential growth in the number of variables is inevitable, to agree with the form of Hermann's bound (1926). However, the number of variables and the number of generators grow together in the construction of Mayr & Mayer (1982). It remains an open question whether or not these bounds must grow double exponentially in the number of variables, when both the degree of the generators and the number of generators are held fixed.

3. The Geometry of the Ideal Membership Problem

In this section, we discuss the ideal membership problem from a geometric point of view, and consider the related question of the complexity of ideal membership of 1.

Given an ideal I generated by $h_1, \ldots, h_s \in R$, homogenise each h_i with z, to produce $h'_1, \ldots, h'_s$ generating a homogeneous ideal $I' \subset R[z]$. I defines a subscheme X of the affine space $\mathbf{A}^n$, and I' defines a subscheme X' of the projective space $\mathbf{P}^n$. In general, I' is not the homogenisation of I; there will exist $h \in I$ whose homogenisations h' will fail to belong to I'. Corresponding to this, X' is not in general the projective closure of X. Let $H \subset \mathbf{P}^n$ be the hyperplane at infinity defined by $z=0$, so $\mathbf{P}^n = \mathbf{A}^n \cup H$; X' in general has primary components supported on H.

Let $h \in I$, so $h - \Sigma_{i=1,s} g_i h_i = 0$ for some $g_1, \ldots, g_s \in R$. Homogenising with z, we obtain $z^m h' - \Sigma_{i=1,s} z^{m_i} g'_i h'_i = 0$ for integers $m, m_1, \ldots, m_s$, where $g'_1, \ldots, g'_s$ are the homogenisations of $g_1, \ldots, g_s$, and at least one $m_i = 0$. In particular, $z^m h' \in I'$. Assume for simplicity that $\deg(h_i) = d$ for each h_i; we have $\max\{deg(g_i)\} = m + \deg(h) - d$. Conversely, if $z^j h' \in I'$ for some $j \geqslant 0$, then we can find $g_1, \ldots, g_s \in R$ so $h - \Sigma_{i=1,s} g_i h_i = 0$, with $\max\{\deg(g_i)\} = j + \deg(h) - d$. Thus, the degrees that can occur for $g_1, \ldots, g_s$ of minimal degree in the ideal membership problem are determined by the minimum m so $z^m h' \in I'$.

There is a relationship between the problem of ideal membership of 1, the ideal membership problem, and the syzygy problem: Consider the homogeneous versions of these problems. If $z^m h' \in I'$, then $z^m \in (I' : h')$, where $(I' : h') = \{f \in R[z] \mid f h' \in I'\}$. The generators of $(I' : h')$ can be computed from the generators of I', and the minimal syzygies of $I + (h)$. Thus, the ideal membership problem can be reduced to a combination of the syzygy problem and ideal membership of 1.

Since $h \in I$, h vanishes on X, so h' vanishes on the projective closure of X in $\mathbf{P}^n$. Thus, $z^m h' \in I'$ if z^m vanishes on the primary components of X' supported on H. This motivates the following definition.

3.1. Definition. Let $Y \subset \mathbf{A}^{n+1}$ be a scheme defined by the ideal $J \subset R[z]$, let $Z \subset \mathbf{A}^{n+1}$ be a reduced scheme defined by the ideal $K \subset R[z]$, and suppose that $\mathrm{Supp}(Y) \subset Z$. Define the thickness of Y relative to Z to be the least integer m so $K^m \subset J$.

Intuitively, the thickness m measures how much the structure of Y extends in directions normal to Z.

Let Y denote the union of those primary components of X' which are supported on the hyperplane H. It follows from the preceding discussion that if the thickness of Y relative to H is m, then $z^m h' \in I'$.

In this setting, theorem 2.4 asserts the existence of subschemes $X' \subset \mathbf{P}^n$ having primary components whose thickness relative to H is double exponential in the maximum of the number of variables and the number of ideal generators. This is startling when compared to other known bounds in algebraic geometry. For example, Bezout's theorem asserts that a complete intersection of n hypersurfaces of degree d in $\mathbf{P}^n$ has degree d^n, which is single exponential in n.

In considering the above geometric picture, one is lead to believe that the primary components of X' which cause this double exponential thickness relative to H are embedded components of X', and that these components have at most single exponential thickness relative to the support of X'. This, in fact, has been confirmed in characteristic zero by the following recent results of Brownawall (1986).

3.2. THEOREM (IDEAL MEMBERSHIP OF 1). (*Brownawell*, 1986). *Let*

$$I = (h_1, \ldots, h_s) \subset R = k[x_1, \ldots, x_n],$$

where k is of characteristic zero, and $\deg(h_i) \leqslant d$. *If* $1 \in I$, *then there exist* $g_1, \ldots, g_s \in R$ *so* $1 = \Sigma_{i=1,s} g_i h_i$, *where* $\deg(g_i h_i) \leqslant 3\mu n d^\mu$, *and* $\mu = \min(n, s)$.

3.3. COROLLARY (RADICAL IDEAL MEMBERSHIP). (*Brownawell*, 1986). *Following the same notation as in theorem* 3.2, *if f belongs to the radical rad(I) of I, then*

$$f^e \in I \text{ for some } e \leqslant e' = 3(\mu+1)(n+1)(d+1)^{\mu+1}.$$

The following corollary follows easily from these results:

3.4. COROLLARY. *Following the same notation as in theorem* 3.2, *let* $X \subset \mathbf{A}^n$ *be the scheme defined by the ideal* $I \subset R$, *let* $Y \subset \mathbf{A}^n$ *be a primary component of X defined by the ideal* $J \subset R$, *and let* $Z \subset \mathbf{A}^n$ *be the scheme defined by the ideal* $\mathrm{rad}(I) \subset R$, *so* $\mathrm{Supp}(Y) \subset Z$. *Then the thickness of Y relative to Z is given by some integer* $e \leqslant e' = 3(\mu+1)(n+1)(d+1)^{\mu+1}$.

PROOF. We need to show that $\mathrm{Rad}(I)^{e'} \subset J$. From corollary 3.3, $\mathrm{Rad}(I)^{e'} \subset I$. Since J is a primary component of I, $I \subset J$, and the corollary follows. ■

Conjecturally, theorem 3.2 and corollaries 3.3, 3.4 hold also in positive characteristic, but these statements are not yet known; they await an algebraic proof. The construction of Mayr & Meyer (1982) does not generalise to produce a counterexample, for the following reason: The variables s_j, f_j, and $c_{j1}, \ldots, c_{j4}$ limit the number of generators of I_n that apply at once to a monomial of level $\leqslant n$. If one introduces a new generator of the form $1 - x^D$, where any of these variables divide x^D, then one loses control of the degree of these variables. For example, the linearity in $\{s_0, \ldots, s_n, f_0, \ldots, f_n\}$ for monomials of level $\leqslant n$ cannot be preserved. Intuitively, one is no longer restricted to visiting a single box at a time in Fig. 2.

We have seen that the ideals exhibiting double exponential behaviour for the ideal membership problem are characterised by pathological primary components on the hyperplane at infinity; it can be shown that these ideals are of double exponential regularity. We have also seen that these primary components have double exponential thickness relative to the hyperplane at infinity, but that in characteristic zero they have only single exponential thickness relative to the support of the ideal I itself. In contrast, as discussed in the introduction, when geometric conditions are imposed that preclude such primary components, investigators have proven or conjectured regularity bounds exhibiting polynomial growth. For computer algebra systems to be able to safely navigate this problem domain, it is necessary both that the cases which exhibit polynomial behaviour be better understood, and that algorithms be able to recognise these cases.

Note added in proof

J. Kollàr (sharp effective Nullstellensatz, preprint 1988) has improved theorem 3.2 to include all characteristics. He obtains sharp bounds for $\deg(g_i h_i)$, if each $\deg h_i > 2$.

References

Brownawell, W. D. (1987). Bounds for the degrees in the Nullstellensatz. *Annals of Mathematics*, **126**, 577–591.

Buchberger, B. (1976). A theoretical basis for the reduction of polynomials to canonical forms. *ACM SIGSAM Bull.* 39, 19–29.

Eisenbud, D., Goto, S. (1984). Linear free resolutions and minimal multiplicity. *J. Algebra* **88**, 89–133.

Gruson, L., Lazarsfeld, R., Peskine, C. (1983). On a theorem of Castelnuovo, and the equations defining space curves. *Inventiones Math.* **72**, 491–506.

Hermann, G. (1926). Die Frage der endlich vielen Schritte in der Theorie der Polynomideale. *Math. Ann.* **95**, 736–788.

Masser, D. W., Wüstholz, G. (1983). Fields of large transcendence degree generated by values of elliptic functions. *Inventiones Math.* **72**, 407–464.

Mayr, E., Meyer, A. (1982). The complexity of the word problem for commutative semigroups and polynomial ideals. *Adv. Math.* **46**, 305–329.

Mumford, D. (1966). *Lectures on Curves on an Algebraic Surface*. Princeton University Press, Princeton, New Jersey.

Pinkham, H. A. (1986). Castelnuovo bound for smooth surfaces. *Invent. math.* **83**, 321–332.

Seidenberg, A. (1974). Constructions in algebra. *Trans. Amer. Math. Soc.* **197**, 273–313.