SIAM J. COMPUT.
Vol. 19, No. 4, pp. 750-773, August 1990

014

THE STRUCTURE OF POLYNOMIAL IDEALS AND GRÖBNER BASES*

THOMAS W. DUBÉ†

Abstract. This paper introduces the cone decomposition of a polynomial ideal. It is shown that every ideal has a cone decomposition of a standard form. Using only this and combinatorial methods, the following sharpened bound for the degree of polynomials in a Gröbner basis can be produced. Let $K[x_1, \cdots, x_n]$ be a ring of multivariate polynomials with coefficients in a field K, and let F be a subset of this ring such that d is the maximum total degree of any polynomial in F. Then for any admissible ordering, the total degree of polynomials in a Gröbner basis for the ideal generated by F is bounded by $2((d^2/2) + d)^{2^{n-1}}$.

Key words. Gröbner bases, algebraic computation, Hilbert functions

AMS(MOS) subject classifications. 68Q40, 05A17

1. Introduction.

Many problems of symbolic computation can ultimately be reduced to determining if a given polynomial p is contained in the ideal generated by a set of polynomials F. Gröbner bases are special bases for polynomial ideals with several important computational properties including the ability to rapidly determine ideal membership. The term *Gröbner basis* was coined by Buchberger, who earlier had pioneered the idea in his thesis. Gröbner bases differ only slightly from the *standard bases* defined by Hironaka, and many of these concepts can be traced back to the H-bases of Macaulay.

The increasing interest in Gröbner bases as a computational tool is in large part due to the algorithm provided by Buchberger whereby for any set of polynomials F, it is possible to construct a Gröbner basis for the ideal generated by F.

Although modified versions of Buchberger's algorithm have shown success in practice (including some commercial systems), the complexity of the algorithm has not been well understood. Giusti [6] has shown a Gröbner basis construction that always produces a Gröbner basis containing only polynomials of the lowest possible degree. A first step in understanding the complexity of the algorithm then is to bound the degree of polynomials that occur in a minimal Gröbner basis.

It has been widely known (thanks to [8] and [10]) that in the worst case the degree of polynomials in a Gröbner basis is at least double exponential in the number of indeterminates in the polynomial ring. This lower bound precludes the existence of an upper bound that would show the Gröbner basis algorithm to be tractable, but it does not answer the following question: "How large can the polynomials in a Gröbner basis be?"

The direction for producing an upper bound was provided by Bayer [1]. Bayer's thesis, together with the results of [6] and [10], shows that the degree bound of elements in a Gröbner basis is bounded by $(2d)^{(2n+2)^{n+1}}$.

The steps in producing this former bound may be summarized as follows:

(1) Begin with a basis F for $I \subseteq K[x_1, \cdots, x_n]$, with d the maximum degree of polynomials in F.

(2) If the ideal I is affine, introduce a new variable x_{n+1} to homogenize the ideal I to ${}^hI$.

* Received by the editors June 27, 1988; accepted for publication October 5, 1989. This work was supported in part by National Science Foundation grants DCR-84-01898 and DCR-84-01633.

† Courant Institute, New York University, New York, New York 10012.

(3) Place ${}^hI$ into *generic coordinates* [1]. Here it must be assumed that K is of characteristic zero.

(4) In generic coordinates the degree of polynomials required in a Gröbner basis with respect to reverse lexicographic ordering is bounded by $(2d)^{2^{n-1}}$ ([6]).

(5) The degree bound in generic coordinates also serves as a bound on the *regularity* of ${}^hI$ ([6]). (An ideal has regularity m if for every degree m polynomial p, the ideal (I, p) has a different Hilbert polynomial than I.) Since ${}^hI$ has $n+1$ variables, the regularity m of ${}^hI$ is bounded by $(2d)^{2^n}$.

(6) A polynomial ideal over n variables with regularity m has its Macaulay constant (b_1 as used in this paper) bounded by $(m + 2n + 2)^{(2n+2)^n}$ ([10]). The Macaulay constant of ${}^hI$ is therefore bounded by $D = ((2d)^{2^n} + 2n + 2)^{(2n+2)^n} \approx (2d)^{(2n+2)^{n+1}}$.

(7) For any admissible ordering, the degree of polynomials in a Gröbner basis for ${}^hI$ is bounded by the maximum of m and b_1 ([1]). The degree of these polynomials is therefore bounded by the value D given above.

(8) Specializing ${}^hI$ back to I by setting $x_{n+1} = 1$ produces a Gröbner for I whose polynomials also satisfy this same degree bound.

A first remark concerning this procedure is that bounding the regularity of ${}^hI$ is an unnecessary detour. Reference [6] shows that in generic coordinates with respect to reverse lexicographic ordering a Gröbner basis G can contain a polynomial g with $\mathrm{Hterm}(g) \in \mathrm{PP}[x_1, \cdots, x_i]$ only if for every degree z such that $d \le z \le \deg(g)$, G contains a polynomial g_d with $\deg(g_d) = d$ and $\mathrm{Hterm}(g_d) \in \mathrm{PP}[x_1, \cdots, x_i]$. This condition is nearly equivalent to what is defined in this thesis as a *standard cone decomposition*, and in fact the standard cone decomposition was developed as a way to mimic this behavior. Directly from Giusti's decomposition, the Hilbert polynomials of $\tilde{I}$ and $K[X]/\tilde{I}$ can be written in the forms needed to produce the bound on the Macaulay constant given in Chapter 3.

Furthermore, the methods for obtaining the old bound use fairly specialized branches of commutative algebra and algebraic geometry. Expertise in these areas is not common among computer scientists. Since there are a growing number of computer scientists who will want to use Gröbner bases, there is a need for a self-contained treatment. The methods of algebraic geometry are concise and elegant, but there is often much more insight gained by using brute force.

A major result of this current study was to obtain a new upper bound for Gröbner basis degree. If F is a set of n variable polynomials of degree at most d, then we prove that a reduced Gröbner basis for the ideal generated by F has degree at most

$$2\left(\frac{d^2}{2} + d\right)^{2^{n-1}}.$$

I believe that the method of obtaining this bound is perhaps of greater importance than the bound itself. The method, which involves decomposing the ideal into disjoint *cones*, avoids the need to change to generic coordinates. This greatly simplifies the description of the proof and eliminates the requirement that the field K have characteristic zero. Moreover, it sheds much light onto the structure of an ideal I and the quotient ring $K[X]/I$, and it is expected that further applications for cone decompositions could be found.

2. Background. The material in this section is presented primarily for the purpose of establishing notations and terminology. For a more thorough introduction to

Gröbner basis, the reader is directed to [2], [3], [4], or [9]. The notations for homogeneity and Hilbert functions are borrowed primarily from [12] and [11], and a more detailed summary of all this information can be found in [5].

2.1. Admissible orderings and Gröbner bases.

DEFINITION. A total ordering $\underset{A}{\geq}$ on the power products $\mathrm{PP}[X] = \mathrm{PP}[x_1, \cdots, x_n]$ of the ring $\mathcal{A}$ is called an *admissible ordering* if the following axioms hold:

(1) For all power products $a, b, c \in \mathrm{PP}[X]$, $a \underset{A}{>} b \Longrightarrow ca \underset{A}{>} cb$.

(2) For all variables x_i, $x_i \underset{A}{>} 1$.

Closely related to the concept of admissible orderings is that of head terms. The $\underset{A}{>}$-greatest power product contained in a monomial of a polynomial h is called the *head term of h with respect to* $\underset{A}{>}$ and is denoted by $\mathrm{Hterm}_A(h)$. For an ideal I, $\mathrm{Head}_A(I)$ is used to denote the ideal generated by the set $\{\mathrm{Hterm}_A(h) : h \in I\}$.

DEFINITION. Let G be a basis for the ideal I and let $\underset{A}{>}$ be an admissible ordering. G is called a *Gröbner basis* of I (with respect to $\underset{A}{>}$) if $\mathrm{Head}_A(I)$ is generated by the set $\{\mathrm{Hterm}(g) : g \in G\}$.

Let F be a set of polynomials and $\underset{A}{>}$ a fixed admissible ordering. A polynomial h is said to be *F-reducible*, if there exists $f \in F$, and monomial $c \in \mathcal{A}$ such that $\mathrm{Hterm}_A(cf)$ is a monomial of h. The polynomial $g = h - cf$ is then called a reduct of h, and this relationship is denoted as $h \xrightarrow{F} g$. The transitive closure $h \xrightarrow[*]{F} g$ of the reduction operation is defined to mean that there exists a sequence of polynomials $p_1, \cdots, p_k$ such that $p_1 = h$, $p_k = g$, and for all $i < k$, $p_i \xrightarrow{F} p_{i+1}$. Finally, g is called an *F-normal form of h* if

(1) $h \xrightarrow[*]{F} g$, and

(2) g is not F-reducible.

The following conditions are all equivalent (e.g., [3], [9]):

(1) G is a Gröbner basis for I with respect to $\underset{A}{>}$.

(2) $G \subset I$ and for every $h \in I$ there exists a $g \in G$ such that $\mathrm{Hterm}_A(g)$ divides $\mathrm{Hterm}_A(h)$.

(3) For all $h \in \mathcal{A}$, 0 is a G-normal form of h if and only if $h \in I$.

(4) G is a basis for I and every $h \in \mathcal{A}$ has a unique G-normal form that may be denoted as $\mathrm{nf}_G(h)$.

One of the most important features of Gröbner bases is the existence of unique normal forms. The following lemma shows that these normal forms provide a system of representatives for the residue class ring $\mathcal{A}/I$.

LEMMA 2.1. *Let G be a Gröbner basis for I with respect to the admissible ordering $\underset{A}{>}$. Then the following properties hold for all $s, t \in \mathcal{A}$:*

(1) $s - \mathrm{nf}_G(s) \in I$.

(2) $s - t \in I \iff \mathrm{nf}_G(s) = \mathrm{nf}_G(t)$.

(3) $\mathrm{nf}_G(s + t) = \mathrm{nf}_G(s) + \mathrm{nf}_G(t)$.

In a slight abuse of notation, N_I will be used to denote the set of normal forms

$$N_I = \{\mathrm{nf}_G(a) : a \in \mathcal{A}\},$$

where G is an arbitrary fixed Gröbner basis for I.

2.2. Direct decompositions. Let T be a subset of the polynomial ring $\mathcal{A}$, and let $S_1, \cdots, S_m$ be a (possibly infinite) family of subsets of T. The sets S_i are said to be a *direct decomposition* of T if every $p \in T$ can be *uniquely* expressed in the form $p = \sum_{i=1}^{r} p_i$, where $p_i \in S_i$ and r is finite. The fact that the S_i form a direct decomposition of T is expressed using the notation

$$T = S_1 \oplus S_2 \oplus \cdots \oplus S_m .$$

The following two important properties of direct decompositions can be easily verified.

(1) Let $S_1, \cdots, S_k$ be a direct decomposition for T, and let $R_1, \cdots, R_m$ be a direct decomposition for S_1. Then $S_2, \cdots, S_k, R_1, \cdots, R_m$ is a direct decomposition for T.

(2) Let $P = \{hf : h \in T\}$ for some polynomial f, and let $S_1, \cdots, S_k$ be a direct decomposition of T. Then the sets $Q_i = \{hf : h \in S_i\}$ form a direct decomposition of P.

Example 1. For any ideal $I \subseteq \mathcal{A}$, I and N_I form a direct decomposition of $\mathcal{A}$.

Proof. Let G be the Gröbner basis of I used to form $N_I = \mathrm{nf}_G(\mathcal{A})$. Since G is a Gröbner basis, each polynomial h has a unique G-normal form, and the decomposition $h = \mathrm{nf}_G(h) + (h - \mathrm{nf}_G(h))$ is unique. □

DEFINITION. Let I be any ideal of $\mathcal{A}$, and $h \in \mathcal{A}$. The ideal quotient operation $I : h$ is defined by $I : h = \{f \in \mathcal{A} : fh \in I\}$. Note that it trivially follows that $(I : g) : h = I : (gh)$.

Example 2. For an ideal $J \subset \mathcal{A}$ and $f \in \mathcal{A}$, let

$$\begin{aligned} I &= (J, f), \quad L = J : f, \\ S &= \{af : a \in N_L\} ; \end{aligned}$$

then $I = J \oplus S$.

Proof. Let G be the Gröbner basis for L used to form N_L and $S = fN_L$. The sets J and S are clearly subsets of I, so it need only be shown that each $h \in I$ can be uniquely expressed as $h = h_J + h_S$. It will first be shown that such a decomposition exists, and then that the decomposition is unique.

Every polynomial $h \in I$ can be written as $h = a_J + a_f f$ with $a_J \in J$. It is now claimed that a decomposition of h exists with $h_J = h - \mathrm{nf}_G(a_f)f$ and $h_S = \mathrm{nf}_G(a_f)f$. Since the sum of these two polynomials is trivially h, it must only be shown that $h_J \in J$. This follows directly from the definitions of the sets involved:

$$\begin{aligned} &a_f - \mathrm{nf}_G(a_f) \in L, \\ &(a_f - \mathrm{nf}_G(a_f))f \in J, \\ &h_J = a_J + (a_f - \mathrm{nf}_G(a_f))f \in J . \end{aligned}$$

Now consider any two decompositions of h:

$$h = a_1 + \mathrm{nf}_G(b_1)f_r = a_2 + \mathrm{nf}_G(b_2)f_r ,$$

where $a_1, a_2 \in J$.

$$\begin{aligned} &(\mathrm{nf}_G(b_1) - \mathrm{nf}_G(b_2))f_r = a_2 - a_1 \in J, \\ &\mathrm{nf}_G(b_1) - \mathrm{nf}_G(b_2) \in L, \\ &\mathrm{nf}_G(b_1) - \mathrm{nf}_G(b_2) = 0 . \end{aligned}$$

Therefore the decomposition is unique. □

Applying this technique recursively, we obtain the following decomposition of an ideal.

Example 3. Let $F = \{f_1, \cdots, f_r\}$ be a basis for an ideal I. Let S_1 be the principal ideal (f_1), and for each $i = 2, \cdots, r$, let

$$\begin{aligned} L_i &= (f_1, \cdots, f_{i-1}) : f_i, \quad \text{and} \\ S_i &= \{hf_i : h \in N_{L_i}\} . \end{aligned}$$

Then, $I = S_1 \oplus \cdots \oplus S_r$.

In summary, for any ideal I, the ring $\mathcal{A}$ can be decomposed into I and N_I. Furthermore, I itself can be decomposed into sets of the form $S_i = \{hf_i : h \in N_{L_i}\}$, which in turn could be further decomposed if we could decompose N_{L_i}. Sets of the form N_I need to be studied more closely.

2.3. Homogeneity. Let f be a polynomial in $\mathcal{A}$; then f can be written as a finite sum $f = f_k + f_{k-1} + \cdots + f_0$, where each f_z is either zero, or a sum of monomials each of which has total degree z. In such a decomposition of f, each nonzero f_z is called the *homogeneous component of f of degree z*. The nonzero homogeneous component f_k of greatest total degree is called the initial form of f and is denoted by $\text{in}(f)$. A polynomial f is called a *homogeneous polynomial* if f consists of at most one nonzero homogeneous component.

DEFINITION. A set $S \subseteq \mathcal{A}$ is called *homogeneous* if it satisfies the following two properties:

(1) $f \in S$ implies that each homogeneous component of f is also in S.

(2) f is a K-module.

A homogeneous set S that is an ideal of $\mathcal{A}$ is called simply a *homogeneous ideal.* A direct decomposition $S_1, \cdots, S_r$ of a homogeneous set T is called a *homogeneous direct decomposition* if each S_i is homogeneous.

For a homogeneous set T, the subset of degree z homogeneous polynomials will be denoted by T_z, i.e.,

$$T_z = \{f \in T : f \text{ is homogeneous of degree } z\} .$$

If T is closed under addition, then the collection of sets $\{T_0, T_1, \cdots\}$ trivially form a homogeneous direct decomposition of T.

For p a polynomial in the affine ring $\mathcal{A} = K[x_1, \cdots, x_n]$, let p be written as a sum of monomials $p = p_1 + \cdots + p_m$. The homogenization function ${}^h p$ is a mapping from the affine ring $\mathcal{A}$ to the projective ring $K[x_1, \cdots, x_n; y]$ where y is a new variable and the mapping is defined as

$${}^h p = \sum_{i=1}^{m} p_i y^{\deg(p) - \deg(p_i)}.$$

Throughout this paper, y will be used to denote the extra variable, which is introduced by homogenization. ${}^h\mathcal{A}$ will denote the projective ring ${}^h\mathcal{A} = K[x_1, \cdots, x_n; y]$, which results from the introduction of y.

To return from ${}^h\mathcal{A}$ to the original ring, use the natural homomorphism ${}^a p$ defined by partially evaluating p at $y = 1$. For example, ${}^h(x_1^3 + x_2) = x_1^3 + x_2 y^2$, and ${}^a(x_2^2 y + 3x_1 x_3^2) = x_2^2 + 3x_1 x_3^2$.

2.4. Hilbert functions. The Hilbert function of a homogeneous set T is denoted by $\varphi_T(z)$ and is defined as follows:

$$\varphi_T(z) = \text{the dimension of } T_z \text{ as a vector space over } K.$$

Equivalently, let $\underset{A}{>}$ be any fixed admissible ordering. The Hilbert function may be defined to be the number of degree z power products that occur as the head monomial of a polynomial of T. That is,

$$\varphi_T(z) = |\{p \in \mathrm{PP}[X] : p \in \mathrm{Head}_A(T_z)\}| .$$

The definitions of homogeneous direct decompositions and Hilbert functions lead immediately to Lemma 2.2.

LEMMA 2.2. *Let $S_1, \cdots, S_r$ be a homogeneous direct decomposition of T; then* $\varphi_T(z) = \sum_{i=1}^{r} \varphi_{S_i}(z)$.

Let $I \subseteq {}^h\mathcal{A}$ be a homogeneous ideal. If N is any homogeneous set of representatives for the quotient ring ${}^h\mathcal{A}/I$, then $I \oplus N$ is a homogeneous direct decomposition for the entire ring ${}^h\mathcal{A}$. Therefore the Hilbert function of N (and hence ${}^h\mathcal{A}/I$) satisfies the relation

$$\varphi_{{}^h\mathcal{A}/I}(z) = \varphi_N(z) = \varphi_{{}^h\mathcal{A}}(z) - \varphi_I(z) .$$

In particular, since I is a homogeneous ideal, a homogeneous system of representatives for the ring ${}^h\mathcal{A}/I$ can be constructed as

$$N_I = \{\mathrm{nf}_G(a) : a \in {}^h\mathcal{A}\} ,$$

where G is any Gröbner basis for I.

It is a classic result that for any ideal I, at sufficiently large z, the Hilbert functions $\varphi_I(z)$ and $\varphi_{{}^h\mathcal{A}/I}(z)$ become polynomials in z. These polynomials will be denoted using the notation $\overline{\varphi}_I(z)$ and $\overline{\varphi}_{{}^h\mathcal{A}/I}(z)$.

3. Cone decompositions of the polynomial ring. The main goal in finding a direct decomposition for an ideal I is to partition I into subsets whose Hilbert function can easily be described. In particular, the types of elements desired are sets of the form $\{ah : a \in K[u]\}$, where h is a homogeneous polynomial and u is a subset of $X = \{x_1, \cdots, x_n\}$.

DEFINITION. For h a homogeneous polynomial and $u \subseteq X$, the set $\{ah : a \in K[u]\}$ is called a *cone* and is denoted by $C(h, u)$.

Some insight into the behavior of cones can be gained from considering monomial ideals with $n = 2$. This case can be well understood because the cones may be depicted graphically. However, since many interesting phenomena occur only at the higher dimensions this simple case can at times be misleading. For example, in two dimensions all Borel-fixed ideals are lexicographic. Furthermore, there are many features of general polynomial ideals that do not appear in the monomial case. This problem is not so important here though, because the cone decomposition will be applied mainly to monomial ideals.

The graphical representation in two dimensions is illustrated in Fig. 1. The power products are represented in a triangular grid with 1 at the bottom. Powers of x run along the left side of the grid, and powers of y to the right. To reach the vertex associated with a given power product $x^a y^b$ count a places upward to the left and then

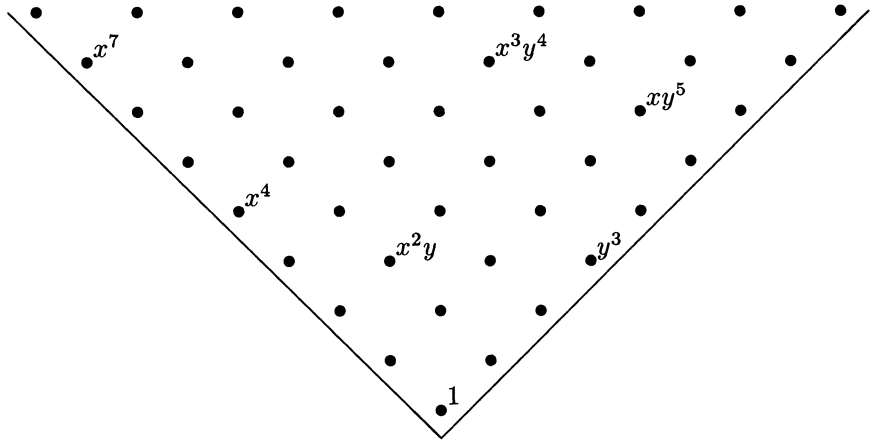

FIG. 1. *The power-products of* $K[x, y]$.

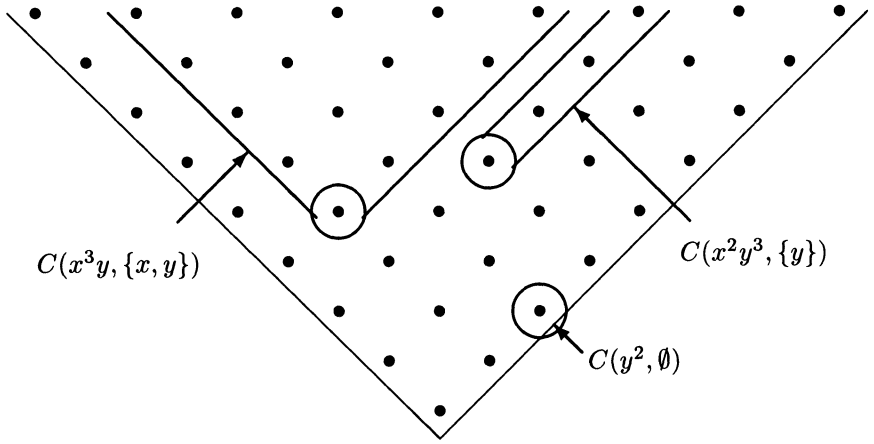

FIG. 2. *Examples of cones.*

b places upward to the right. Figure 2 illustrates that cones may then be represented in the diagram by encircling the power products which the cones contain.

For a cone $C(h,u)$, the Hilbert function of $C(h,u)$ is dependent only on $\deg(h)$ and $|u|$. Counting the number of power products in $\text{Head}(C(h,u))$, we find that if $u = \emptyset$, then

$$\varphi_{C(h,\emptyset)}(z) = \begin{cases} 0, & z \neq \deg(h), \\ 1, & z = \deg(h), \end{cases}$$

and for $|u| > 0$,

$$\varphi_{C(h,u)}(z) = \begin{cases} 0, & z < \deg(h), \\ \binom{z - \deg(h) + |u| - 1}{|u| - 1}, & z \geq \deg(h). \end{cases}$$

DEFINITION. Let $h_1, \cdots, h_r$ be homogeneous polynomials of $\mathcal{A}$, and let $u_1, \cdots, u_r$ be subsets of X. A finite set $P = \{\langle h_1, u_1\rangle, \cdots, \langle h_r, u_r\rangle\}$ is a *cone decomposition* of $T \subseteq \mathcal{A}$ if the cones $C(h_i, u_i)$ form a direct decomposition of T.

The cones $\langle h_i, u_i\rangle \in P$ that have $u_i = \emptyset$ form a finite part of T and do not contribute to the Hilbert polynomial of T. The remaining cones, for which $u_i \neq \emptyset$ form a direct decomposition of a set that is equivalent to T at large degrees. This portion of the cone decomposition will be denoted as

$$P^+ = \{\langle h, u\rangle \in P : u \neq \emptyset\} .$$

A cone decomposition P for T is said to be *k-standard* (k an integer) if the following two conditions hold:

(1) There is no pair $\langle h, u\rangle \in P^+$ with $\deg(h) < k$.

(2) For every $\langle g, v\rangle \in P^+$ and degree d such that $k \leq d \leq \deg(g)$, P contains a pair $\langle h, u\rangle$ with $\deg(h) = d$ and $|u| \geq |v|$.

Note that if P^+ is the empty set, then P is k-standard for all natural numbers k. On the other hand, if P^+ is nonempty, the only possible value for k is $\min\{\deg(h) : \langle h, u\rangle \in P^+\}$.

The following list contains an assortment of easily verifiable properties of cone decompositions and k-standard cone decompositions.

(1) $\emptyset$ is a 0-standard cone decomposition for $\emptyset$.

(2) $\{\langle h, u\rangle\}$ is a $\deg(h)$-standard cone decomposition of $C(h,u)$.

(3) $\{\langle 1, X\rangle\}$ is a 0-standard cone decomposition of $\mathcal{A}$.

(4) Let S_1 and S_2 be a direct decomposition of T, and let P_1 and P_2 be cone decompositions of S_1 and S_2, respectively. Then $P_1 \cup P_2$ is a cone decomposition of T.

(5) Let S_1 and S_2 be a direct decomposition of T, and let P_1 and P_2 be k-standard cone decompositions of S_1 and S_2, respectively. Then $P = P_1 \cup P_2$ is a k-standard cone decomposition of T.

(6) If $P = \{\langle h_1, u_1\rangle, \cdots, \langle h_s, u_s\rangle\}$ is a k-standard cone decomposition for T, then for any homogeneous polynomial c, the set $P' = \{\langle ch_1, u_1\rangle, \cdots, \langle ch_s, u_s\rangle\}$ is a $(k + \deg(c))$-standard cone decomposition for $\{ch : h \in T\}$.

There is one special cone decomposition that provides a useful function for manipulations.

DEFINITION. Let $u = \{x_{j_1}, \cdots, x_{j_m}\} \subseteq X$. Then define the set $E(h,u)$ as

$$E(h,u) = \{\langle h, \emptyset\rangle\} \cup \{\langle x_{j_i} h, \{x_{j_i}, \cdots, x_{j_m}\}\rangle : i = 1, \cdots, m\} .$$

It is easy to verify that $E(h,u)$ is a $(\deg(h)+1)$-standard cone decomposition of $C(h,u)$.

LEMMA 3.1. *Let P be a k-standard cone decomposition for T. Then, for any $d \geq k$, there exists a d-standard cone decomposition P_d for the set T.*

Proof. If $P^+ = \emptyset$, then the result holds trivially, so assume that P^+ is nonempty. It suffices to show that $(k+1)$-standard cone decomposition exists for T. Let $R = \{\langle h,u\rangle \in P : \deg(h) = k\}$, and $S = P - R$. The original set P was k-standard, so after removing the cones in R, the remaining set S is $(k+1)$ standard.

Since R contains only pairs $\langle h,u\rangle$ for which $\deg(h) = k$, R is trivially k-standard. The set spanned by the cones in R also has a $(k+1)$-standard cone decomposition, namely,

$$R' = \bigcup_{\langle h,u\rangle \in R} E(h,u) .$$

Finally, $P_{k+1} = R' \cup S$ is a $(k+1)$-standard cone decomposition for T. □

COROLLARY 3.2. *Let $S_1, \cdots, S_r$ be a direct decomposition of T, where for each S_i there exists a k_i-standard cone decomposition P_i. Then there exists a k-standard cone decomposition P of T with $k = \max\{k_1, \cdots, k_r\}$.*

4. Splitting a system of representatives. In this section it will be shown that for any homogeneous ideal I, it is possible to construct a 0-standard cone decomposition for N_I. Recall that once the ordering $>_A$ and a Gröbner basis G for I are fixed, then N_I and $N_{\mathrm{Head}_A(G)}$ have a termwise agreement as sets. Thus, only monomial ideals need be considered.

Let I be an ideal of $\mathcal{A}$ generated by the set of monomials $F = \{f_1, \cdots, f_r\}$. For a given variable x_j, there is a direct decomposition of I consisting of I_0 and I_1, where

$$I_0 = I \cap K[X - \{x_j\}]$$

and,

$$I_1 = I \cap (x_j) = \{x_j h : h \in \mathcal{A} \text{ and } x_j h \in I\} .$$

Clearly, I_0 is an ideal of $K[X - \{x_j\}]$ generated by $F \cap K[X - \{x_j\}]$. It is also easy to verify that I_1 is an ideal of $\mathcal{A}$ generated by the set $G = \{g_1, \cdots, g_r\}$, where

$$g_i = \begin{cases} x_j f_i, & f_i \in K[X - \{x_j\}], \\ f_i & \text{otherwise.} \end{cases}$$

Comparing the ideal I_1 defined above with the quotient $I : x_j$, shows that $I_1 = \{x_j h : h \in I : x_j\}$. Furthermore, this leads to the fact that $I : x_j$ is generated by $H = \{h_1, \cdots, h_r\}$ where

$$h_i = x_j^{-1} g_i = \begin{cases} f_i, & f_i \in K[X - \{x_j\}], \\ x_j^{-1} f_i & \text{otherwise.} \end{cases}$$

This method of forming a basis for $I : x_j$ is restated as an algorithm in Fig. 3.

DEFINITION. Let $P \cup Q$ be a cone decomposition of $T \subseteq \mathcal{A}$, and let I be an ideal of $\mathcal{A}$. Then P and Q are said to *split T relative to I* if $\langle h,u\rangle \in P$ implies $C(h,u) \subseteq I$ (i.e., $h \in I$), and $\langle h,u\rangle \in Q$ implies $C(h,u) \cap I = \{0\}$. We may easily verify that P is a cone decomposition of $T \cap I$. Furthermore, the following lemma shows that under proper restrictions Q is a cone decomposition for $T \cap N_I$.

```
QUOTIENT_BASIS(F, x_j)
    Input :   F a monomial basis for I ⊆ A
              x_j ∈ X a variable
    Output :  F' a monomial basis for I : x_j.

    F' := ∅
    For f_i ∈ F
        if f_i ∈ K[X − {x_j}] then F' := F' ∪ {f_i}
                             else F' := F' ∪ {x_j^{-1} f_i}
    return (F')
End.
```

FIG. 3. *The algorithm for forming a basis for* $I : x_j$.

LEMMA 4.1. *Let* $P = \{\langle g_1, u_1\rangle, \cdots, \langle g_r, u_r\rangle\}$ *and* $Q = \{\langle h_1, v_1\rangle, \cdots, \langle h_s, v_s\rangle\}$ *split* T *relative to a monomial ideal* I, *where for each* $\langle h_i, v_i\rangle \in Q$, h_i *is a monomial. Then* Q *is a cone decomposition for* $T \cap N_I$.

Proof. If I is a monomial ideal, then regardless of admissible ordering $\underset{A}{>}$ and Gröbner basis G,

$$f \in N_I \iff \text{each monomial of } f \text{ is not in } I.$$

Furthermore, if h_i is a monomial then

$$f \in C(h_i, v_i) \iff \text{each monomial of } f \text{ is in } C(h_i, v_i).$$

By the definition of a *splitting* set of cones, $C(h_i, v_i) \cap I = \{0\}$, so

$$f \in C(h_i, v_i) \iff \text{no monomial of } f \text{ is in } I.$$

Let $f \in T \cap N_I$. $f \in T$ implies that f can be written uniquely as

$$f = f_{P_1} + f_{P_2} + \cdots + f_{P_r} + f_{Q_1} + \cdots + f_{Q_s}.$$

The partial sum $f_{\overline{P}} = f_{P_1} + f_{P_2} + \cdots + f_{P_r}$ is in I and therefore is a sum of monomials in I. Since no such monomials appear in the cones $C(h_i, v_i)$, all monomials of $f_{\overline{P}}$ must also appear in f. But $f \in N_I$ and can include no monomial of I. Hence, $f_{\overline{P}} = 0$ and f can be written uniquely as $f = f_{Q_1} + \cdots + f_{Q_s}$. □

As an example, consider the monomial ideal $I = (x^4y, xy^3, y^5)$. This ideal has a cone decomposition given by the set

$$\{C(x^4y, \{x\}),\ C(x^4y^2, \{x\}),\ C(xy^3, \{x, y\}),\ C(y^5, \{y\})\}.$$

This particular cone decomposition is illustrated in Fig. 4. Viewing this figure should make it clear that this cone decomposition is not unique.

For example, the cones $C(xy^3, \{x, y\})$ and $C(y^5, \{y\})$ can be equivalently replaced by $C(xy^3, \{x\})$, $C(xy^4, \{x\})$, and $C(y^5, \{x, y\})$.

Now, let T be the ideal generated by x^2. Then $T \cap I$ has a cone decomposition described by the set

$$P = \{\langle x^4y, \{x\}\rangle, \langle x^4y^2, \{x\}\rangle, \langle x^2y^3, \{x, y\}\rangle\}.$$

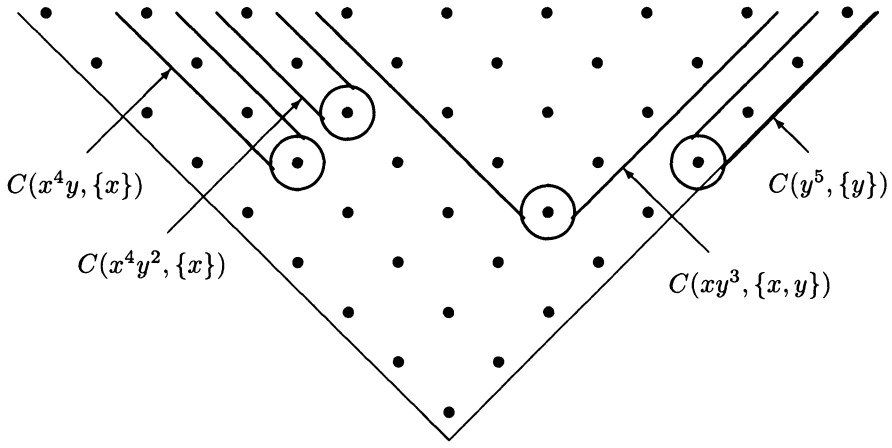

FIG. 4. *The cone decomposition of* $I = (x^4y, xy^3, y^5)$.

This cone decomposition is illustrated in Fig. 5.

But *splitting* I requires not only a cone decomposition P for the ideal, but also a cone decomposition Q for $T \cap N_I$. In this example, Fig. 6 shows that Q may be chosen as the set of cones described by

$$Q = \{\langle x^2, \{x\}\rangle, \langle x^2y, \emptyset\rangle, \langle x^2y^2, \emptyset\rangle, \langle x^3y, \emptyset\rangle, \langle x^3y^2, \emptyset\rangle\} .$$

From the definition of what is meant by a cone decomposition $P \cup Q$ *splitting* a set relative to an ideal I, it is immediate that a cone $C(h, u)$ can belong to such a decomposition only if either $C(h, u) \subseteq I$ or $C(h, u) \cap I = \emptyset$. The following lemma shows that if the ideal I is a monomial ideal and h is also a monomial, then this condition can be effectively determined. This will provide an algorithm to split the ring $\mathcal{A}$ relative to a monomial ideal I.

LEMMA 4.2. *Let I be a monomial ideal, $h \in \mathrm{PP}[X]$, $u \subset X$, and let F be a power product basis for $I : h$. Then,*

(1) $C(h, u) \subseteq I$ *if and only if* $1 \in F$.

(2) $C(h, u) \cap I = \emptyset$ *if and only if* $F \cap \mathrm{PP}[u] = \emptyset$.

Proof. (1) $1 \in F \iff 1 \in I : h \iff h \in I \iff C(h, X) \subseteq I$.

(2) ($\Longrightarrow$) Assume $C(h, u) \cap I = \emptyset$. Then for $g \in \mathrm{PP}[u]$,

$$\begin{aligned} hg \in C(h, u) &\implies hg \notin I \\ &\implies g \notin I : h \\ &\implies g \notin F . \end{aligned}$$

(2) ($\Longleftarrow$) Assume $F \cap \mathrm{PP}[u] = \emptyset$. Then for $g \in \mathrm{PP}[u]$, g cannot be in $I : h$ since otherwise F would have to contain a divisor of g and this divisor would also be in $\mathrm{PP}[u]$. So

$$g \notin I : h \implies hg \notin I .$$

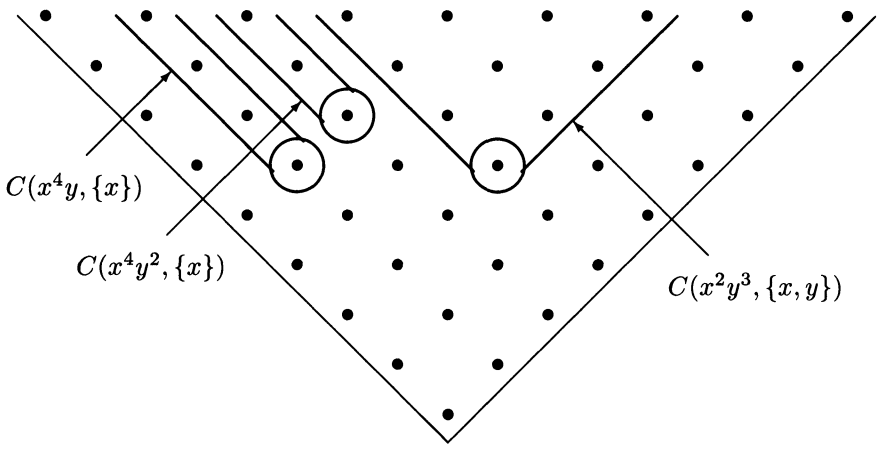

FIG. 5. *The cone decomposition of* $(x^2) \cap I$.

By the definition of $C(h, u)$, every polynomial contained in this cone is of the form hg with $g \in \mathrm{PP}[u]$ and hence not in the ideal I. □

Figure 7 provides an algorithm SPLIT for splitting a cone $C(h, u)$ with respect to a monomial ideal I.

LEMMA 4.3. *The algorithm* SPLIT *terminates.*

Proof. For a set of arguments h, u, and F, define the rank of the arguments as $|u| + \sum_{f \in F} \deg(f)$. It is now claimed that if SPLIT is invoked with arguments of rank r, then the two recursive calls (if reached) have arguments of rank $\leq r - 1$. For the first call, this is trivial. For the second call, it must be shown that there is some $f_i \in F$ such that $f_i \notin \mathrm{PP}[X - \{x_j\}]$. But, this must be true since otherwise $F \cap K[s \cup \{x_j\}] = \emptyset$, contradicting the choice of s, and hence x_j.

If $r = 0$, then F must either be $\{1\}$, or $\emptyset$. In either case, the recursion stops. Therefore the depth of recursion is at most r, and hence the algorithm terminates. □

LEMMA 4.4. *The algorithm* SPLIT *is correct.*

Proof. The previous lemma assures the termination of the algorithm, so the correctness of the algorithm can be proven using induction on the depth of recursion.

The basis case in which no recursive calls are made occurs if $1 \in F$ or $F \cap \mathrm{PP}[u] = \emptyset$. In both of these cases, Lemma 4.2 shows that the trivial decomposition (P, Q) satisfies the definition for splitting $C(h, u)$ relative to I.

Otherwise, the cone $C(h, u)$ is decomposed into

$$C(h, u) \;=\; C(h, u - \{x_j\}) \oplus C(x_j h, u) \; .$$

Since F is a power product basis for $I : h$, the function QUOTIENT_BASIS produces a power product basis F' for the ideal $I : x_j h$. Inductively, the algorithm SPLIT returns

(1) (P_0, Q_0), which splits $C(h, u - \{x_j\})$ relative to I, and

(2) (P_1, Q_1), which splits $C(x_j h, u)$ relative to I.

These two decompositions are then joined to produce the desired decomposition of $C(h, u)$. □

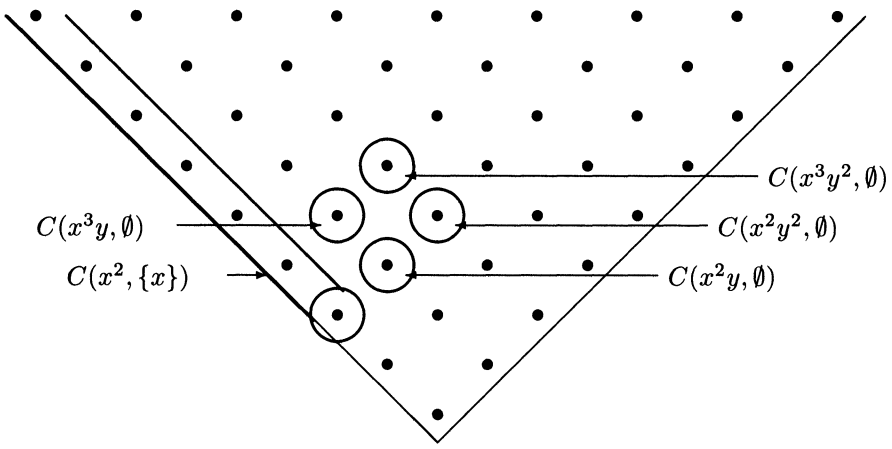

FIG. 6. *The cone decomposition of* $(x^2) \cap N_I$.

In the SPLIT algorithm, the choice of $s \subset u$ such that $F \cap \mathrm{PP}[s] = \emptyset$ as a *maximal* subset is not a necessary condition for the correctness of the algorithm in producing a splitting decomposition. However, it will soon be shown that the set Q returned by this algorithm has the additional property of being $\deg(h)$-standard. To prove that this is indeed true, we begin with a simple lemma regarding the condition $F \cap \mathrm{PP}[s] = \emptyset$.

LEMMA 4.5. *Let* h, u, I, *and* F *be as in algorithm* SPLIT. *Then for any* $v \subseteq X$,

$$C(h,v) \subseteq C(h,u) \cap N_I \iff v \subseteq u \text{ and } F \cap \mathrm{PP}[v] = \emptyset \ .$$

Proof. ($\Longrightarrow$) $C(h,v) \subseteq C(h,u)$ clearly implies $v \subseteq u$. To see that $F \cap \mathrm{PP}[v] = \emptyset$, let f be any nonzero element of $K[v]$. Then $hf \in C(h,v) \subseteq N_I$. But $I \cap N_I = \{0\}$ and neither h nor f is zero, so $hf \notin I$. Then,

$$\begin{aligned} hf \notin I &\implies f \notin I : h \\ &\implies f \notin F \ . \end{aligned}$$

($\Longleftarrow$) $v \subseteq u$ implies $C(h,v) \subseteq C(h,u)$, so it only remains to be shown that $C(h,v) \subseteq N_I$. To prove this, it is sufficient to show that no monomial of $C(h,v)$ belongs to I. Each monomial of $C(h,v)$ is of the form hf with f a monomial of $K[v]$. Then,

$$\begin{aligned} F \cap \mathrm{PP}[v] = \emptyset &\implies f \notin I : h \\ &\implies hf \notin I \ . \end{aligned}$$

□

LEMMA 4.6. *Let* h, u, I, *and* F *be valid input for algorithm* SPLIT, *and let* (P,Q) *denote the sets returned by* SPLIT(h,u,F). *Then for any power product* g, $C(g,v) \subseteq C(h,u) \cap N_I$ *implies that* Q *contains a pair* $\langle h,s \rangle$ *with* $|s| \geq |v|$.

Proof. Using the previous lemma,

$$\begin{aligned} C(g,v) \subseteq C(h,u) \cap N_I &\implies C(h,v) \subseteq C(h,u) \cap N_I \\ &\implies v \subseteq u \text{ and } F \cap \mathrm{PP}[v] = \emptyset. \end{aligned}$$

```
SPLIT(h,u,F)
    Input :   h ∈ PP[X]
              u ⊆ X is a set of variables
              F a power product basis for I : h
    Output :  (P,Q) which split C(h,u) relative to I.

    If 1 ∈ F then return (P = {⟨h,u⟩}, Q = ∅)
    If F ∩ PP[u] = ∅ then return (P = ∅, Q = {⟨h,u⟩})
    Otherwise
      Choose s ⊂ u a maximal subset such that F ∩ PP[s] = ∅
      Choose x_j ∈ u − s    [If s=u this point would not be reached.]

      (P_0,Q_0) := SPLIT(h, u − {x_j}, F)

      F' := QUOTIENT_BASIS(F, x_j)
      (P_1,Q_1) := SPLIT(x_j h, u, F')
      return (P = P_0 ∪ P_1, Q = Q_0 ∪ Q_1)
End.
```

FIG. 7. *The algorithm for splitting $C(h,u)$ relative to I.*

We proceed inductively on $|u|-|v|$. If $v = u$, then the algorithm returns $Q = \{\langle h,u\rangle\}$, satisfying the lemma. Otherwise, the choice of s as a *maximal* subset such that $F \cap \mathrm{PP}[s] = \emptyset$ implies that $|s| \geq |v|$. The previous lemma can now be applied in the opposite direction to get

$$C(h,s) \subseteq C(h,u-\{x_j\}) \cap N_I \ .$$

Using the induction hypothesis, the set Q_0 formed by the recursive call $\mathtt{SPLIT}(h,u-\{x_j\},F)$ contains a pair $\langle h,w\rangle$ with $|w| \geq |s| \geq |v|$. The lemma then follows from the fact that Q_0 is a subset of Q. □

A basis $R = \{f_1,\cdots,f_k\}$ for an ideal I is called a *reduced basis* if each f_i satisfies $f_i \notin (R-\{f_i\})$. On the other hand, suppose that F is not a reduced basis, and that $f_i \in (F-\{f_i\})$. Then, $F-\{f_i\}$ is also a basis for I. Successively removing redundant generators produces a subset of F that is a reduced basis for I.

LEMMA 4.7. *Let R be a reduced power product basis for a monomial ideal I, and let $P = \{\langle h_1,u_1\rangle,\cdots,\langle h_r,u_r\rangle\}$ be any cone decomposition of I where the h_i's are power products. Then, for each $f \in R$, there is a pair $\langle f,u\rangle \in P$.*

Proof. Let f be any element of R. Since $f \in I$, there is some $\langle h,u\rangle \in P$ such that $f \in C(h,u)$. But, now h is also in I, so $h = bg$ for some $g \in R$. But $f \in C(h,u)$, so f can be written as $f = ah = abg$. Since R is reduced, we have $ab = 1$, and $f = h$. □

LEMMA 4.8. *Let F be a power product basis for $I \neq \mathcal{A}$, $(P,Q) = \mathtt{SPLIT}(1,X,F)$, and let $R \subseteq F$ be a reduced basis for I. Then for every $f \in R$, Q contains a pair $\langle h,u\rangle$ with* $\deg(h) = \deg(f) - 1$.

Proof. Let f be any element of R. By the preceding lemma there is a pair $\langle f,v\rangle \in P$. Consider how this pair got into P. Since $\deg(f) > 0$, there must have been a recursive call $\mathtt{SPLIT}(f,v,F')$, where F' is a basis for $I : f$. This invocation of SPLIT must have been the child of either

(1) $\mathtt{SPLIT}(x_j^{-1}f,v,F'')$, or

(2) $\texttt{SPLIT}(f, v \cup \{x_j\}, F')$.

Using the first possibility as a basis case, inductively, we may step backward through the computation of $\texttt{SPLIT}(1, X, F)$ to find an invocation $\texttt{SPLIT}(x_j^{-1}f, v', F'')$ with $v' \supseteq v$.

The cone $C(x_j^{-1}f, v')$ could not have been a subset of I, since then a recursive call would not have been generated. Therefore, if

$$(P', Q') = \texttt{SPLIT}(x_j^{-1}f, v', F''),$$

then Q' is nonempty. But then Lemma 4.6 assures that Q' contains a pair of the form $\langle x_j^{-1}f, s\rangle$. Since $\deg(x_j^{-1}f) = \deg(f) - 1$, the existence of this pair in $Q' \subseteq Q$ satisfies the lemma. □

COROLLARY 4.9. *Let F be a power product basis for I, and let $(P, Q) = \texttt{SPLIT}(1, X, F)$. Then if $d = 1 + \max\{\deg(h) : \langle h, u\rangle \in Q\}$, I can be generated by the set $\{f \in F : \deg(f) \leq d\}$.*

LEMMA 4.10. *Let $(P, Q) = \texttt{SPLIT}(h, u, F)$. Q is a $\deg(h)$-standard cone decomposition.*

Proof. If Q is either $\emptyset$ or $\{\langle h, u\rangle\}$, then the lemma follows trivially. Otherwise, assume inductively (on the number of recursions) that Q_0 and Q_1 satisfy the lemma. That is Q_0 and Q_1 are, respectively, $\deg(h)$-standard and $(\deg(h) + 1)$-standard.

To show that Q is a $\deg(h)$-standard cone decomposition, it must be shown that for any $\langle g, v\rangle \in Q$ and degree d such that $\deg(h) \leq d \leq \deg(g)$, there is a pair $\langle p, t\rangle \in Q$ with $\deg(p) = d$ and $|t| \geq |v|$. Since $Q = Q_0 \cup Q_1$, there are two cases to consider.

(1) $\langle g, v\rangle \in Q_0$. Since Q_0 is itself a $\deg(h)$-standard cone decomposition, Q_0 contains all the pairs needed to satisfy the condition for $\langle g, v\rangle$, and Q_0 is a subset of Q.

(2) $\langle g, v\rangle \in Q_1$. Since Q_1 is a $(\deg(h) + 1)$-standard cone decomposition, Q_1 contains the pairs needed to satisfy the condition for $\langle g, v\rangle$ for $\deg(h)+1 \leq d \leq \deg(g)$. For $d = \deg(h)$, Lemma 4.6 assures that Q contains the needed pair. □

The remarks at the beginning of this section allow these results to be extended beyond monomial ideals.

THEOREM 4.11. *Let G be a Gröbner basis for I with respect to $>_A$. Let $(P, Q) = \texttt{SPLIT}(1, X, \text{Hterm}_A(G))$. Then Q is a 0-standard cone decomposition of $N_I = \text{nf}_G(\mathcal{A})$. Furthermore, if $d = 1 + \max\{\deg h : \langle h, u\rangle \in Q\}$, then $G' = \{g \in G : \deg(g) \leq d\}$ is also a Gröbner basis for I with respect to $>_A$.*

Proof. $\text{Hterm}_A(G)$ is a basis for $\text{in}_A(I)$. Therefore the SPLIT algorithm returns a 0-standard cone decomposition for $N_{\text{in}_A(I)} = N_I$.

By Corollary 4.9, the set

$$\{h \in \text{Hterm}_A(G) : \deg(h) \leq d\} \subseteq \text{Hterm}_A(G')$$

is a basis for $\text{in}_A(I)$, and hence G' is a Gröbner basis for I. □

5. Splitting a homogeneous ideal. So far we have seen that for any ideal I, there exists a 0-standard cone decomposition of N_I. But what about I itself? The construction SPLIT provides a cone decomposition of I that is only valid for monomial ideals, and even this does not produce a standard cone decomposition. The answer is found in the following lemma.

LEMMA 5.1. *Let $F = \{f_1, \cdots, f_r\}$ be a homogeneous basis for an ideal I; then there exists a k-standard cone decomposition P for I with*

$$k = \max\{\deg(f_i) \ : \ i = 1 \cdots r\}.$$

Proof. Let $S_1 = (f_1)$, and for $i = 2 \cdots r$ let $J_i = (f_1, \cdots, f_{i-1})$, $L_i = J_i : f_i$ and $S_i = \{cf_i \ : \ c \in N_{L_i}\}$. The sets $S_1, \cdots, S_r$ form a direct decomposition of I. S_1 is a principal ideal that has the $\deg(f_1)$-standard cone decomposition $P_1 = \{\langle f_1, X\rangle\}$. Using the construction provided by SPLIT, we form a 0-standard cone decomposition Q_i for each N_{L_i}. If $Q_i = \{\langle h_1, u_1\rangle, \cdots, \langle h_s, u_s\rangle\}$, then it follows that $P_i = \{\langle f_r h_1, u_1\rangle, \cdots, \langle f_r h_s, u_s\rangle\}$ is a $\deg(f_i)$-standard cone decomposition for S_i.

It then follows from Corollary 3.2 that there exists a k-standard cone decomposition P for I. □

For $I \neq \{0\}$, it will be preferable to use a slightly modified version of this result. When the sets P_i are united to form P, do not include P_1 in the union. This produces the following modified result.

COROLLARY 5.2. *Let $F = \{f_1, \cdots, f_r\}$ be a homogeneous basis for an ideal I with $r > 0$, and let $S_1, \cdots, S_r$ be as above. Then there exists a direct decomposition of I consisting of the primary ideal $S_1 = (f_1)$, and a k-standard cone decomposition P for $S_2 \oplus S_3 \oplus \cdots \oplus S_r$ with*

$$k = \max\{\deg(f_i) \ : \ i = 1 \cdots r\}.$$

6. The exact cone decomposition.

DEFINITION. For $T \subseteq K[X]$, Q is called an *exact cone decomposition* of T if Q is a k-standard cone decomposition of T for some k, and additionally for every degree d, Q^+ contains at most one pair $\langle h, u\rangle$ with $\deg(h) = d$.

If Q^+ is nonempty, then there is a unique value of k for which Q is k-standard. Let $\overline{a}_Q$ denote this value of k. In the case that Q^+ is empty, let $\overline{a}_Q = 0$. Both of these cases can be captured with the single definition: $\overline{a}_q$ is the least value of k such that Q is k-standard. However, this unified definition fails to emphasize the fact that in the more important case ($Q^+ \neq \emptyset$) the value of k is unique.

For $i = 0, \cdots, n+1$, let

$$b_i = \min\{d \geq \overline{a}_Q \ : \ \langle h, u\rangle \in Q \text{ and } |u| \geq i \implies \deg(h) < d\} \ .$$

It is a simple consequence of this definition that the b_i's satisfy $b_0 \geq b_1 \geq \cdots \geq b_{n+1} = \overline{a}_Q$. Furthermore,

$$b_1 \ = \ \begin{cases} 1 + \max\{\deg(h) \ : \ \langle h, u\rangle \in Q^+\}, & Q^+ \neq \emptyset, \\ 0, & Q^+ = \emptyset \ . \end{cases}$$

LEMMA 6.1. *Let Q be an exact cone decomposition, and let $b_0, \cdots, b_{n+1}$ be defined as above. Then for each $i = 1, \cdots, n$ and degree d such that $b_{i+1} \leq d < b_i$, there is exactly one pair $\langle h, u\rangle \in Q^+$ such that $\deg(h) = d$ and in that pair $|u| = i$.*

Proof. If Q^+ is empty, then $b_1 = b_2 = \cdots = b_{n+1} = 0$ and the lemma follows vacuously. Otherwise, for each $i = 1, \cdots, n$, the definition of b_i requires that $b_1 - 1$ be the largest degree such that Q contains a pair $\langle g, v\rangle$ with $|v| \geq i$. Since Q is b_{n+1}-standard, each degree $d = b_{n+1}, \cdots, b_i - 1$ must have a pair $\langle h_d, u_d\rangle \in Q$ with

$\deg(h_d) = d$ and $|u_d| \geq |v| \geq i$. Since Q is an exact cone decomposition $\langle h_d, u_d \rangle$ is the only pair $\langle h, u \rangle \in Q^+$ with $\deg(h) = d$.

Now if $b_i = b_{i+1}$ the range $d = b_{i+1}, \cdots, b_i - 1$ is vacuous. Otherwise, for each d in this range $|u_d| = i$ since $|u_d| > i$ would contradict the definition of b_{i+1}. □

The following trivial lemma provides a tool by which any standard cone decomposition may be transformed into an exact cone decomposition.

LEMMA 6.2. *Let Q be a k-standard cone decomposition of T, and let $\langle f, s \rangle$, $\langle g, v \rangle \in Q$ such that $\deg(f) = \deg(g)$, and $|v| \geq |s| > 0$. Then for any $x_j \in s$,*

$$Q' = (Q - \{\langle f, s \rangle\}) \cup \{\langle f, s - \{x_j\} \rangle, \langle x_j f, s \rangle\}$$

is also a k-standard cone decomposition of T.

Proof. It must be shown that for every pair $\langle \ell, w \rangle \in Q'$ and degree $d = k, \cdots, \deg(\ell)$ there is a pair $\langle h, u \rangle \in Q'$ with $\deg(h) = d$ and $|u| \geq |w|$. For $\langle \ell, w \rangle \in Q \cap Q'$, Q' *inherits* all the required pairs from Q. For the two new pairs, the presence of $\langle g, v \rangle \in Q'$ is sufficient to show that Q' must again contain the required pairs. □

This lemma provides a tool to *shift* pairs away from degrees occupied by other pairs.

One new term will be introduced only for the purposes of proving the correctness of the following algorithm. A k-standard cone decomposition P is called *m-exact* if for each degree d there is at most one pair $\langle h, u \rangle \in P$ such that $\deg(h) = d$ and $|u| > m$. With this definition, a cone decomposition is exact if and only if it is 0-exact. It also follows vacuously that any cone decomposition is n-exact. Consider the algorithm of Fig. 8.

```
SHIFT(Q, k, m)
    Input :   Q a k-standard m-exact cone decomposition for T
    Output :  Q' a k-standard (m − 1)-exact cone decomposition for T.
    Q' := Q
    If {⟨h, u⟩ ∈ Q : |u| ≥ m} = ∅ then return(Q').
    c := |{⟨h, u⟩ ∈ Q : |u| ≥ m}|.
    For d := k to k + c − 1 do
      B := {⟨h, u⟩ ∈ Q' : deg(h) = d and |u| ≥ m}
      While |B| > 1 loop
            Choose ⟨h, u⟩ ∈ B with |u| = m
            Choose x_j ∈ u
            B := B − {⟨h, u⟩}
            Q' := (Q' − {⟨h, u⟩}) ∪ {⟨h, u − {x_j}⟩, ⟨x_j h, u⟩}
      End While loop
    End For d loop
    return(Q')
End.
```

FIG. 8. *The algorithm for shifting pairs in a standard cone decomposition.*

LEMMA 6.3. *The algorithm* SHIFT *is correct.*

Proof. If Q is a k-standard partition, then it follows from the previous lemma that the set Q' will also be k-standard. Furthermore, the action of the algorithm assures that for each degree $d < k + c$, Q' will contain at most one pair $\langle h, u \rangle$ with $\deg(h) = d$ and $|u| \geq m$. But what about degrees $\geq k + c$? Checking the line at which Q' is modified will show that throughout the execution of this algorithm the

size of the set $\{\langle h,u\rangle \in Q' : |u| \geq m\}$ remains invariantly c. Now since Q' is k-standard, a pair $\langle g,v\rangle \in Q'$ with $|v| \geq m$ requires that Q' contain a pair $\langle h_d,u_d\rangle$ with $|u_d| \geq m$ and $\deg(h_d) = d$, for every degree $d = k,\cdots,\deg(g)$. The c pairs in the set $\{\langle h,u\rangle \in Q' : |u| \geq m\}$ must then include the $\deg(g) - k + 1$ pairs of the form $\langle h_d,u_d\rangle$. Therefore, $\deg(g) \leq c + k - 1$. □

Now, the SHIFT algorithm can be used to produce an exact cone decomposition using the algorithm in Fig. 9. Note that the action of the EXACT and SHIFT algorithms

```
EXACT(Q, k)
    Input :   Q a k-standard cone decomposition for T
    Output :  Q' an exact cone decomposition for T.
    Q_n := Q
    For m := n down to 1 do
        Q_{m-1} := SHIFT(Q_m, k, m)
    End For m loop
    return(Q_0)
End.
```

FIG. 9. *The algorithm for producing an exact partition.*

assures that if Q' is the exact cone decomposition produced by EXACT(Q,k), then the Macaulay constant b_0 for Q' satisfies

$$b_0 \geq 1 + \max\{\deg(h) : \langle h,u\rangle \in Q\}.$$

7. Exact cone decomposition and Hilbert function. For any cone decomposition P of a set T, the Hilbert function of T can be described by summing the Hilbert functions of the cones in P:

$$\varphi_T(z) = \sum_{\langle h,u\rangle \in P} \varphi_{C(h,u)}(z) .$$

For degrees z greater than or equal to

$$z' = \max\{\deg(h) : \langle h,u\rangle \in P\}$$

each of the cones has a Hilbert function described by the binomial coefficient

$$\overline{\varphi}_{C(h,u)} = \binom{z - \deg(h) + |u| - 1}{|u| - 1},$$

and so

$$\overline{\varphi}_T(z) = \sum_{\langle h,u\rangle \in P^+} \binom{z - \deg(h) + |u| - 1}{|u| - 1} .$$

But, if P is exact, then the constants $b_1,\cdots,b_{n+1}$ describe all of the cones in P^+, so

$$\overline{\varphi}_T(z) = \sum_{j=1}^{n} \sum_{d=b_{j+1}}^{b_j - 1} \binom{z - d + j - 1}{j - 1} .$$

Furthermore, the constant b_0 is defined to be the same as the constant z' given above, so the Hilbert function attains this polynomial form for degrees $z \geq b_0$.

Using the combinatorial identity

$$\sum_{d=b_{j+1}}^{b_j-1} \binom{z-d+j-1}{j-1} = \binom{z-b_{j+1}+j}{j} - \binom{z-b_j+j}{j},$$

the Hilbert function of T can be written in the form:

$$\begin{aligned}
\varphi_T(z) &= \sum_{j=1}^{n} \left[\binom{z-b_{j+1}+j}{j} - \binom{z-b_j+j}{j} \right] \\
&= \binom{z-b_{n+1}+n}{n} - \binom{z-b_1+1}{1} \\
&\quad + \sum_{j=1}^{n-1} \left[\binom{z-b_{j+1}+j}{j} - \binom{z-b_{j+1}+j+1}{j+1} \right] \\
&= \binom{z-b_{n+1}+n}{n} - 1 - \binom{z-b_1}{1} - \sum_{j=1}^{n-1} \binom{z-b_{j+1}+j}{j+1} \\
&= \binom{z-b_{n+1}+n}{n} - 1 - \sum_{j=0}^{n-1} \binom{z-b_{j+1}+j}{j+1}.
\end{aligned}$$

Replacing the summation variable with $i = j + 1$, this can be restated as

$$(*) \qquad \varphi_T(z) = \binom{z-b_{n+1}+n}{n} - 1 - \sum_{i=1}^{n} \binom{z-b_i+i-1}{i}.$$

In the classic paper [7], Macaulay first proved that for sufficiently high degree z, the Hilbert function of a polynomial quotient ring always attains the form of a polynomial such as the one given in (*). For this reason the constants $b_0, \cdots, b_{n+1}$ will be referred to as the Macaulay constants of T. The formulation given above has the added benefit of the additional constant b_0, which provides a bound on the point at which the Hilbert function $\varphi_T(z)$ attains its polynomial form $\overline{\varphi}_T(z)$ as given in (*).

For z in the range $b_1 \leq z < b_0$, the Hilbert functions of the cones in P^+ attain the polynomial forms used in calculating $\overline{\varphi}_T(z)$. For z in this range however, there are also some cones $C(h, \emptyset) - \in P - P^+$, which contribute to the Hilbert function of T. Therefore, for $z \geq b_1$ the following form of the Hilbert function is valid:

$$\begin{aligned}
\varphi_T(z) &= \overline{\varphi}_T(z) + \sum_{\langle h, \emptyset \rangle \in P} \varphi_{C(h,\emptyset)}(z) \\
&= \overline{\varphi}_T(z) + |\{\langle h, \emptyset \rangle \in P : \deg(h) = z\}|.
\end{aligned}$$

LEMMA 7.1. *Let P be any exact cone decomposition for a set T. Once the constant $b_{n+1} = \overline{a}_Q$ is fixed, the constants $b_0, b_1, \cdots, b_n$ are uniquely determined.*

Proof. The Hilbert polynomial of T can be written in the form

$$\overline{\varphi}_T(z) = a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \cdots + a_1 z + a_0 .$$

Assume inductively that the constants $b_{j+1}, \cdots, b_{n+1}$ have been uniquely determined such that Hilbert function given by (*) agrees with the coefficients $a_{n-1}, \cdots, a_j$. The binomial coefficient

$$\binom{z - b_i + i - 1}{i}$$

is a degree i monic polynomial in z. So, the coefficients $b_{j-1}, \cdots, b_1$ do not effect the coefficient of z^{j-1} in the Hilbert polynomial (*). Therefore, matching the coefficient a_{j-1} requires a unique choice for b_j.

We may then also uniquely determine b_0 as

$$b_0 = \min\{d \geq b_1 : \forall_{z>d} \overline{\varphi}_T(z) = \varphi_T(z)\} . \qquad \square$$

LEMMA 7.2. *Let I be a homogeneous ideal, then the Hilbert function of N_I is described by a unique set of Macaulay constants $b_0 \geq b_1 \geq \cdots \geq b_{n+1} = 0$. Furthermore, for any admissible ordering $\underset{A}{>}$ the degree of polynomials in a reduced Gröbner basis for I with respect to $\underset{A}{>}$ is bounded by b_0.*

Proof. Since N_I has a 0-standard cone decomposition, it is possible to find an exact cone decomposition for N_I with $b_{n+1} = 0$. Once b_{n+1} is fixed as zero, the other Macaulay constants are uniquely determined.

Let G be a Gröbner basis for I w.r.t. $\underset{A}{>}$. The set N_I admits a 0-standard cone decomposition Q, which may be found using the algorithm SPLIT$(1, X, \text{Hterm}_A(G))$. Let $d = 1 + \max\{\deg(h) : \langle h, u\rangle \in Q\}$. Theorem 4.11 assures that $\{g \in G : \deg(g) \leq d\}$ is also a Gröbner basis for I. The construction using algorithm EXACT then shows that the unique Macaulay constant b_0 is $\geq d$, and hence is also a bound on the degree of polynomials required in the Gröbner basis. $\square$

8. A bound for Gröbner basis degree. Let $F = \{f_1, \cdots, f_r\}$ be a homogeneous basis for an ideal I. Assume without loss of generality that f_1 has the largest degree $\deg(f_1) = d$. In the previous section, it has been shown that for any ideal I, there exists an exact partition Q for N_I in which the constant $\overline{a}_Q$ is zero. Furthermore, if the Macaulay constants associated with Q are $b_0 \geq b_1 \geq \cdots \geq b_{n+1} = 0$, then for degrees $z \geq b_0$, the Hilbert function of N_I attains the polynomial form

$$\varphi_{N_I}(z) = \binom{z+n}{n} - 1 - \sum_{i=1}^{n} \binom{z - b_i + i - 1}{i} .$$

It also has been shown that I itself has a direct decomposition consisting of the principal ideal (f_1) and an exact partition P with $\overline{a}_P = d$. Let $a_0 \geq a_1 \geq \cdots \geq a_{n+1} = d$ be the Macaulay constants for the portion of I partitioned by P. Then, for degrees $z \geq a_0$ the Hilbert function of I is equal to the polynomial

$$\varphi_I(z) = \binom{z-d+n-1}{n-1} + \binom{z-d+n}{n} - 1 - \sum_{i=1}^{n} \binom{z - a_i + i - 1}{i} .$$

Now since I and N_I form a direct decomposition of $K[X]$, the sum of their Hilbert functions must be equal to the Hilbert function of $K[X]$, which is

$$\varphi_{K[X]}(z) = \binom{z+n-1}{n-1} .$$

Therefore, for $z \geq \max\{a_0, b_0\}$,

$$\begin{aligned}\binom{z+n-1}{n-1} &= \binom{z-d+n-1}{n-1} + \binom{z-d+n}{n} \\ &\quad + \binom{z+n}{n} - 2 \\ &\quad - \sum_{i=1}^{n} \left[\binom{z-a_i+i-1}{i} + \binom{z-b_i+i-1}{i} \right] . \end{aligned} \tag{1}$$

The backwards difference operator ∇ is defined for any function $F(z)$ by $\nabla F(z) = F(z) - F(z-1)$, and $\nabla^j F(z) = \nabla(\nabla^{j-1} F(z))$. Using the identity

$$\binom{z+k}{n} - \binom{(z-1)+k}{n} = \binom{z+k-1}{n-1}$$

we have

$$\nabla \binom{z+k}{n} = \binom{z+k-1}{n-1} .$$

It then follows inductively that

$$\nabla^j \binom{z+k}{n} = \binom{z+k-j}{n-j} .$$

If $F_1(z) = F_2(z)$ for $z > k$, then clearly $\nabla F_1(z) = \nabla F_2(z)$ for $z > k+1$. For each j in the range $j = 0, \cdots, n-1$, apply the operator ∇^j to (1). [1] This yields the following set of equations for $j = 1, \cdots, n-1$, which are valid for large enough z:

$$\begin{aligned}\binom{z+n-j-1}{n-j-1} &= \binom{z-d+n-j-1}{n-j-1} + \binom{z-d+n-j}{n-j} \\ &\quad + \binom{z+n}{n} - 2 \\ &\quad - \sum_{i=j+1}^{n} \left[\binom{z-a_i+i-j-1}{i-j} + \binom{z-b_i+i-j-1}{i-j} \right] . \end{aligned}$$

Each side of these equations is a polynomial in z, so they must agree for each power of z. In particular, they must have the same constant term. Note that the constant term of $\binom{z+k}{n}$ is given by

$$\binom{0+k}{n} = \begin{cases} \binom{k}{n}, & k \geq 0, \\ (-1)^n \binom{n-1-k}{n}, & k < 0 . \end{cases}$$

Taking the constant terms of the previous set of equations, we obtain

$$\begin{aligned} 1 &= (-1)^{n-j-1} \binom{d-1}{n-j-1} + (-1)^{n-j} \binom{d-1}{n-j} \\ &\quad -1 - \sum_{i=j+1}^{n} (-1)^{i-j} \left[\binom{a_i}{i-j} + \binom{b_i}{i-j} \right] . \end{aligned}$$

[1] The technique of using the backwards difference operator has been used in a slightly different manner in [10].

At $j = n-1$, this is simply

$$1 - \binom{d-1}{1} - 1 + a_n + b_n = 1 .$$

So $a_n + b_n = d$. Together with the conditions $a_n \geq d$ and $b_n \geq 0$, this implies that $a_n = d$. When we substitute these values, the series of equations becomes

$$2(-1)^{n-j-1}\binom{d-1}{n-j-1} - 1 - \sum_{i=j+1}^{n-1} (-1)^{i-j}\left[\binom{a_i}{i-j} + \binom{b_i}{i-j}\right] = 1 .$$

Let c_{j+1} denote the sum $a_{j+1} + b_{j+1}$. Solving for this expression yields

$$c_{j+1} = 2 + 2(-1)^{n-j}\binom{d-1}{n-j-1} + \sum_{i=j+2}^{n-1} (-1)^{i-j}\left[\binom{a_i}{i-j} + \binom{b_i}{i-j}\right] .$$

At this point, we may note that the sum on the right is vacuous for $j = n-2$ and conclude that $c_{n-1} = 2 + 2(d-1) = 2d$. And since

$$\binom{a_{i+1}}{k} + \binom{b_{i+1}}{k} \leq \binom{c_{i+1}}{k} \tag{2}$$

is true for all i, for $j = n-3$ we have

$$c_{n-2} \leq 2 - 2\binom{d-1}{2} + \binom{2d}{2} = d^2 + 2d .$$

The remaining equations ($j < n-3$), all contain the expression

$$2 + (-1)^{n-j}\left[2\binom{d-1}{n-j-1} - \binom{a_{n-1}}{n-j-1} - \binom{b_{n-1}}{n-j-1}\right] .$$

The magnitude of this combination is bounded by

$$\binom{c_{n-1}}{n-j-1} ,$$

so the inequalities above may be replaced with the weaker inequalities:

$$c_{j+1} \leq \binom{c_{n-1}}{n-j-1} + \sum_{i=j+2}^{n-2} (-1)^{i-j}\left[\binom{a_i}{i-j} + \binom{b_i}{i-j}\right] .$$

The term in the sum for $i = j+3$ has a negative sign, and hence this term may be discarded. Giving all the remaining terms a positive sign produces the following still weaker inequalities :

$$\begin{aligned} c_{j+1} &\leq \binom{c_{n-1}}{n-j-1} + \left[\binom{a_{j+2}}{2} + \binom{b_{j+2}}{2}\right] + \sum_{i=j+4}^{n-2}\left[\binom{a_i}{i-j} + \binom{b_i}{i-j}\right] \\ &\leq \binom{c_{j+2}}{2} + \sum_{i=j+4}^{n-1}\binom{c_i}{i-j} . \end{aligned}$$

Or, upon repairing the subscripts by the change $j \to j-1$:

$$c_j \leq \binom{c_{j+1}}{2} + \sum_{i=j+3}^{n-1} \binom{c_i}{i-j+1}.$$

These inequalities may now be solved inductively to provide a bound on the magnitude of c_j.

LEMMA 8.1. *For $j \leq n-2$, the value of c_j satisfies the inequality $c_j \leq D_j$, where*

$$D_j = 2\left(\frac{d^2}{2}+d\right)^{2^{n-j-1}}.$$

Proof. It was already determined that $c_{n-2} \leq d^2+2d$, satisfying this claim. Now, assume inductively that c_i has the indicated bound for $j < i \leq n-2$.

For $i \geq j+3$ the inequality $2^{i-j-1} \geq i-j+1$ can be used to see that $(2^{n-i-1})(i-j+1) \leq 2^{n-j-2}$. Therefore,

$$\binom{D_i}{i-j+1} \leq \frac{D_i^{i-j+1}}{(i-j+1)!} \leq \frac{D_i^{2^{i-j-1}}}{(i-j+1)!} = D_{j+1}\frac{2^{i-j}}{(i-j+1)!}.$$

And so,

$$\begin{aligned}
c_j &\leq \binom{c_{j+1}}{2} + \sum_{i=j+3}^{n-1} \binom{c_i}{i-j+1} \\
&\leq \binom{D_{j+1}}{2} + \sum_{i=j+3}^{n-1} \binom{D_i}{i-j+1} \\
&\leq \frac{D_{j+1}^2 - D_{j+1}}{2} + \sum_{i=j+3}^{n-1} D_{j+1}\frac{2^{i-j}}{(i-j+1)!} \\
&\leq \frac{D_{j+1}^2}{2} - D_{j+1}\left[\frac{1}{2} - \sum_{i=j+3}^{n-1}\frac{2^{i-j}}{(i-j+1)!}\right] \\
&\leq \frac{D_{j+1}^2}{2} = D_j.
\end{aligned}$$

□

From this, we may conclude that the Macaulay constants a_1 and b_1 are each less than $D_1 = 2((d^2/2)+d)^{2^{n-2}}$. But what about the constants a_0 and b_0 that did not appear explicitly in the Hilbert function? For z in the range $\max\{a_1,b_1\} < z \leq \max\{a_0,b_0\}$, use the equality

$$\varphi_I(z) + \varphi_{N_I}(z) = \varphi_{K[X]}(z)$$

to obtain the relation

$$\begin{aligned}
&(\overline{\varphi}_I(z) + |\{\langle h,\emptyset\rangle \in P \,:\, \deg(h) = z\}|) \\
&\quad +(\overline{\varphi}_{L_I}(z) + |\{\langle h,\emptyset\rangle \in Q \,:\, \deg(h) = z\}|) \;=\; \varphi_{K[X]}(z).
\end{aligned}$$

At this point, we may note that the relationship $\overline{\varphi}_I(z) + \overline{\varphi}_{L_I}(z) = \varphi_{K[X]}(z)$, which was claimed valid for $z > \max\{a_0, b_0\}$ actually holds for $z \geq \max\{a_1, b_1\}$. Therefore, for z in the range $\max\{a_1, b_1\} < z \leq \max\{a_0, b_0\}$, it must be the case that

$$(|\{\langle h, \emptyset\rangle \in P \,:\, \deg(h) = z\}|) + (|\{\langle h, \emptyset\rangle \in Q \,:\, \deg(h) = z\}|) \;=\; 0.$$

This implies that $P \cup Q$ contains no pair $\langle h, \emptyset\rangle$ with $\deg(h) > \max\{a_1, b_1\}$. Therefore, the constant D_1 is also a bound on the value of b_0. Using this bound within Lemma 7.2 provides the proof of the following theorem.

THEOREM 8.2. *Let I be an ideal of $K[X] = K[x_1, \cdots, x_n]$ generated by a set of homogeneous polynomials F. Let $d = \max\{\deg(f) : f \in F\}$. Then for any admissible ordering $>_A$, the degree of polynomials required in a Gröbner basis for I with respect to $>_A$ is bounded by $2((d^2/2) + d)^{2^{n-2}}$.*

For I an affine ideal, we can homogenize a basis F for I using one additional variable x_{n+1}. Therefore, for any set of polynomials F we have Corollary 8.3.

COROLLARY 8.3. *Let $F \subset K[X]$, I the ideal generated by F, and let d be the maximum degree of any $f \in F$. Then for any admissible ordering $>_A$, the degree of polynomials required in a Gröbner basis for I with respect to $>_A$ is bounded by $2((d^2/2) + d)^{2^{n-1}}$.*

REFERENCES

[1] D. BAYER, *The division algorithm and the Hilbert scheme*, Ph.D. thesis, Harvard University, Cambridge, MA, 1982.

[2] B. BUCHBERGER, *A criterion for detecting unnecessary reductions in the construction of Gröbner-basis*, in Lecture Notes in Computer Science, Vol. 72, Springer-Verlag, Berlin, New York, 1979, pp. 3–21.

[3] ———, *Gröbner basis: An algorithmic method in polynomial ideal theory*, in Multidimensional Systems Theory, N. K. Bose, ed., D. Reidel, Boston, MA, 1985, pp. 184–229.

[4] ———, *History and basic features of the critical-pair/completion procedure*, J. Symb. Comput., 3 (1987), pp. 3–38.

[5] T. DUBÉ, *Quantitative analysis of problems in computer algebra: Gröbner bases and the Nullstellensatz*, Ph.D. thesis, New York University, New York, 1989.

[6] M. GIUSTI, *Some effectivity problems in polynomial ideal theory*, in Lecture Notes in Computer Science, Vol. 174, Springer-Verlag, Berlin, New York, 1984, pp. 159–171.

[7] F. S. MACAULAY, *Some properties of enumeration in the theory of modular systems*, Proc. London Math. Soc., 26 (1927), pp. 531–555.

[8] E. W. MAYR AND A. R. MEYER, *The complexity of the word problems for commutative semigroups and polynomial ideals*, Adv. in Math., 46 (1982), pp. 305–329.

[9] B. MISHRA AND C. K. YAP, *Notes on Gröbner basis*, in Information Sciences, An International Journal, Vol. 48, Elsevier Science, New York, 1989, pp. 219–252.

[10] M. MÖLLER AND F. MORA, *Upper and lower bounds for the degree of Groebner bases*, in Lecture Notes in Computer Science, Vol. 174, Springer-Verlag, Berlin, New York, 1984, pp. 172–183.

[11] R. P. STANLEY, *Hilbert functions of graded algebras*, Adv. in Math., 18 (1978), pp. 57–83.

[12] O. ZARISKI AND P. SAMUEL, *Commutative Algebra*, Vol. 2, Springer-Verlag, Berlin, New York, 1960.