What Can Be Computed in Algebraic Geometry?

Dave Bayer * David Mumford

May 7, 1992

This paper evolved from a long series of discussions between the two authors, going back to around 1980, on the problems of making effective computations in algebraic geometry, and it took more definite shape in a survey talk given by the second author at a conference on Computer Algebra in 1984. The goal at that time was to bring together the perspectives of theoretical computer scientists and of working algebraic geometers, while laying out what we considered to be the main computational problems and bounds on their complexity. Only part of the talk was written down and since that time there has been a good deal of progress. However, the material that was written up may still serve as a useful introduction to some of the ideas and estimates used in this field (at least the editors of this volume think so), even though most of the results included here are either published elsewhere, or exist as "folk-theorems" by now.

The article has four sections. The first two parts are concerned with the theory of Gröbner bases; their construction provides the foundation for most computations, and their complexity dominates the complexity of most techniques in this area. The first part introduces Gröbner bases from a geometric point of view, relating them to a number of ideas which we take up in more detail in subsequent sections. The second part develops the theory of Gröbner bases more carefully, from an algebraic point of view. It could be read independently, and requires less background. The third part is an investigation into bounds in algebraic geometry of relevance to

*Partially supported by NSF grant DMS-90-06116.

these computations. We focus on the *regularity* of an algebraic variety (see Definition 3.2), which, beyond its intrinsic interest to algebraic geometers, has emerged as a measure of the complexity of computing Gröbner bases (see [BS87a], [BS87b], [BS88]). A principal result in this part is a bound on the regularity of any smooth variety by the second author: Theorem 3.12(b). This bound has stimulated subsequent work, and has now been generalized by [BEL91]. Another result of interest is Proposition 3.13, which elucidates the scheme structure of the ideal membership problem. The fourth part is a short discussion of work on algorithms for performing some other key operations on varieties, some open problems about these operations and some general ideas about what works and what doesn't, reflecting the prejudices of the authors.

One of the difficulties in surveying this area of research is that mathematicians from so many specialties have gotten involved, and they tend both to publish in their own specialized journals and to have specific agendas corresponding to their area. Thus one group of researchers, the working algebraic geometers, are much more interested in actually computing examples than in worst-case complexity bounds. This group, including the first author, has put a great deal of work into building a functioning system, *Macaulay*, based on Gröbner bases, which has solved many problems and provided many examples to the algebraic geometry community [BS92a]. Another group comes from theoretical computer science and is much more interested in theoretical bounds than practical systems (cf. the provocative comments in Lenstra's survey [Len92]). It seems to us that more communication would be very helpful: On the one hand, the working algebraic geometer knows lots of facts about varieties that can be very relevant to finding fast algorithms. Conversely asymptotic and/or worst-case performance bounds are sometimes, at least, important indicators of real-time performance. These theoretical bounds may also reveal important distinctions between classes of procedures, and may pose new and deep problems in algebraic geometry. Thus we will see in Section 3 how regularity estimates flesh out a picture explaining why Gröbner basis computations can have such explosive worst case behavior, yet be so useful for the kinds of problems typically posed by mathematicians. Finally, to make this article more useful in bridging this gap, we have tried to include a substantial number of references in our discussions below.

1 A Geometric Introduction

Let X be a subvariety or a subscheme of projective n-space $\mathbf{P}^n$, over a field k. Let $\mathcal{F}$ be a vector bundle or a coherent sheaf supported on X. We would like to be able to manipulate such objects by computer. From algebra we get finite descriptions, amenable to such manipulations: Let $S = k[x_0, \ldots, x_n]$ be the homogeneous coordinate ring of $\mathbf{P}^n$. Then X can be taken to be the subscheme defined by a homogenous ideal $I \subset S$, and $\mathcal{F}$ can be taken to be the sheaf associated to a finitely generated S-module M. We can represent I by a list of generators $(f_1, \ldots, f_r)$, and M by a presentation matrix F, where

$$M_1 \xrightarrow{F} M_0 \longrightarrow M \longrightarrow 0$$

presents M as a quotient of finitely generated free S-modules M_0, M_1. We concentrate on the case of an ideal I; by working with the submodule $J = \mathrm{Im}(F) \subset M_0$, the module case follows similarly.

The heart of most computations in this setting is a deformation of the input data to simpler data, combinatorial in nature: We want to move through a family of linear transformations of $\mathbf{P}^n$ so that in the limit our objects are described by monomials. Via this family, we hope to pull back as much information as possible to the original objects of study.

Choose a one-parameter subgroup $\lambda(t) \subset GL(n+1)$ of the diagonal form

$$\lambda(t) = \begin{bmatrix} t^{w_0} & & & \\ & t^{w_1} & & \\ & & \cdots & \\ & & & t^{w_n} \end{bmatrix},$$

where $W = (w_0, \ldots, w_n)$ is a vector of integer weights. For each $t \neq 0$, $\lambda(t)$ acts on X via a linear change of coordinates of $\mathbf{P}^n$, to yield the subscheme $X_t = \lambda(t)X \cong X$. The limit

$$X_0 = \lim_{t \to 0} X_t$$

is usually a simpler object, preferable to X for many computational purposes.

Even if we start out by restricting X to be a subvariety rather than a subscheme of $\mathbf{P}^n$, it does not suffice to take the limit X_0 set-theoretically; often

all we will get pointwise in the limit is a linear subspace $L \subset \mathbf{P}^n$, reflecting little besides the dimension of the original variety X. By instead allowing this limit to acquire embedded components and a nonreduced structure, we can obtain an X_0 which reflects much more closely the character of X itself.

We compute explicitly with the generators $f_1, \ldots, f_r$ of I: Let λ act on S by mapping each x_i to $t^{w_i} x_i$; λ maps each monomial $\mathbf{x}^A = x_0^{a_0} \cdots x_n^{a_n}$ to $t^{W \cdot A} \mathbf{x}^A = t^{w_0 a_0 + \ldots + w_n a_n} x_0^{a_0} \cdots x_n^{a_n}$. If $f = a\mathbf{x}^A + b\mathbf{x}^B + \ldots$, then $\lambda f = a\, t^{W \cdot A} \mathbf{x}^A + b\, t^{W \cdot B} \mathbf{x}^B + \ldots$. We take the projective limit $\mathrm{in}(f) = \lim_{t \to 0} \lambda f$ by collecting the terms of λf involving the least power of t; $\mathrm{in}(f)$ is then the sum of the terms $a\mathbf{x}^A$ of f so $W \cdot A$ is minimal. For a given f and most choices of λ, $\mathrm{in}(f)$ consists of a single term.

The limit X_0 we want is defined with all its scheme structure by the ideal $\mathrm{in}(I) = \lim_{t \to 0} \lambda I$, generated by the set $\{\, \mathrm{in}(f) \,|\, f \in I \,\}$. For a given I and most choices of λ, $\mathrm{in}(I)$ is generated by monomials. Unfortunately, this definition is computationally unworkable because I is an infinite set, and $\mathrm{in}(I)$ need not equal $(\mathrm{in}(f_1), \ldots, \mathrm{in}(f_r))$ for a given set of generators $f_1, \ldots, f_r$ of I. To understand how to compute $\mathrm{in}(I)$, we need to look more closely at the family of schemes X_t defined by λ.

Let $S[t]$ be the polynomial ring $k[x_0, \ldots, x_n, t]$; we view $S[t]$ as the coordinate ring of a one-parameter family of projective spaces $\mathbf{P}^n_t$ over the affine line with parameter t. For each generator f_j of I, rescale λf_j so the lowest power of t has exponent zero: Let $g_j = t^{-\ell} \lambda f_j$, where $\ell = W \cdot A$ is the least exponent of t in λf_j. Then $f_j = g_j|_{t=1}$ and $\mathrm{in}(f_j) = g_j|_{t=0}$. Now, let $J \subset S[t]$ be the ideal generated by $(g_1, \ldots, g_r)$; J defines a family Y over $\mathbf{A}^1$ whose central fiber is cut out by $(\mathrm{in}(f_1), \ldots, \mathrm{in}(f_r))$.

What is wrong with the family Y? Y can have extra components over $t = 0$, which bear no relation to its limiting behavior as $t \to 0$. Just as the set-theoretic limit $\lim_{t \to 0} X_t$ can be too small (we need the nonreduced structure), this algebraically defined limit can be too big; the natural limit lies somewhere in between.

The notion of a *flat* family captures exactly what we are looking for here. For example, if Y is flat, then there are no extra components over $t = 0$. While the various technical definitions of flatness can look daunting to the newcomer, intuitively flatness captures exactly the idea that every fiber of a family is the natural scheme-theoretic continuation of its neighboring fibers.

In our setting, all the X_t are isomorphic for $t \neq 0$, so we only need to consider flatness in a neighborhood of $t = 0$. Artin [Art76] gives a criterion for flatness applicable here: The *syzygies* of $g_1, \ldots, g_r$ are the relations $h_1 g_1 + \ldots + h_r g_r = 0$ for $h_1, \ldots, h_r \in S[t]$. Syzygies correspond to elements $(h_1, \ldots, h_r)$ of the $S[t]$-module $S[t]^r$; the set of all syzygies is a submodule of $S[t]^r$. Y is a flat family at $t = 0$ if and only if the restrictions $(h_1|_{t=0}, \ldots, h_r|_{t=0})$ of these syzygies to the central fiber generate the S-module of syzygies of $g_1|_{t=0}, \ldots, g_r|_{t=0}$.

When $g_1|_{t=0}, \ldots, g_r|_{t=0}$ are single terms, their syzygies take on a very simple form: The module of syzygies of two terms $a\mathbf{x}^A$, $b\mathbf{x}^B$ is generated by the syzygy $b\mathbf{x}^C(a\mathbf{x}^A) - a\mathbf{x}^D(b\mathbf{x}^B) = 0$, where $\mathbf{x}^E = \mathbf{x}^C\mathbf{x}^A = \mathbf{x}^D\mathbf{x}^B$ is the least common multiple of $\mathbf{x}^A$ and $\mathbf{x}^B$. The module of syzygies of r such terms is generated (usually not minimally) by the syzygies on all such pairs.

We want to lift these syzygies to syzygies of $g_1, \ldots, g_r$, working modulo increasing powers of t until each syzygy lifts completely. Whenever we get stuck, we will find ourselves staring at a new polynomial g_{r+1} so $t^\ell g_{r+1} \in J$ for some $\ell > 0$. Including g_{r+1} in the definition of a new $J' \supset J$ has no effect on the family defined away from $t = 0$, but will cut away unwanted portions of the central fiber; what we are doing is removing t-torsion. By iterating this process until every syzygy lifts, we obtain explicit generators $g_1, \ldots, g_r, g_{r+1}, \ldots, g_s$ for a flat family describing the degeneration of $X = X_1$ to a good central fiber X_0. The corresponding generators $g_1|_{t=1}, \ldots, g_s|_{t=1}$ of I are known as a *Gröbner basis* for I.

This process is best illustrated by an example. Let $S = k[w, x, y, z]$ be the coordinate ring of $\mathbf{P}^3$, and let $I = (f_1, f_2, f_3) \subset S$ for

$$f_1 = w^2 - xy,\ f_2 = wy - xz,\ f_3 = wz - y^2.$$

I defines a twisted cubic curve $X \subset \mathbf{P}^3$; X is the image of the map $(r, s) \mapsto (r^2 s, r^3, rs^2, s^3)$. Let

$$\lambda(t) = \begin{bmatrix} t^{-16} & & & \\ & t^{-4} & & \\ & & t^{-1} & \\ & & & t^0 \end{bmatrix}.$$

If $w^a x^b y^c z^d$ is a monomial of degree < 4, then $\lambda \cdot w^a x^b y^c z^d = t^{-\ell} w^a x^b y^c z^d$ where $\ell = 16a + 4b + c$. Thus, sorting the monomials of S of each degree

< 4 by increasing powers of t with respect to the action of λ is equivalent to sorting the monomials of each degree in lexicographic order.

We have

$$\begin{aligned} g_1 &= t^{32}\lambda f_1 &&= w^2 - t^{27}xy, \\ g_2 &= t^{17}\lambda f_2 &&= wy - t^{13}xz, \\ g_3 &= t^{16}\lambda f_3 &&= wz - t^{14}y^2. \end{aligned}$$

The module of syzygies on w^2, wy, wz is generated by the three possible pairwise syzygies; we start with the syzygy $y(w^2) - w(wy) = 0$. Substituting g_1, g_2 for the lead terms w^2, wy we get

$$y(w^2 - t^{27}xy) - w(wy - t^{13}xz) = t^{13}wxz - t^{27}xy^2$$

which is a multiple $t^{13}x$ of g_3. Thus, the syzygy

$$yg_1 - wg_2 - t^{13}xg_3 = 0$$

of g_1, g_2, g_3 restricts to the monomial syzygy $y(w^2) - w(wy) = 0$ when we substitute $t = 0$, as desired.

Similarly, the syzygy

$$zg_1 - t^{14}yg_2 - wg_3 = 0$$

restricts to the monomial syzygy $z(w^2) - w(wz) = 0$. When we attempt to lift $z(wy) - y(wz) = 0$, however, we find that

$$z(wy - t^{13}xz) - y(wz - t^{14}y^2) = -t^{13}xz^2 + t^{14}y^3.$$

xz^2 is not a multiple of w^2, wy, or wz, so we cannot continue; $J = (g_1, g_2, g_3)$ does not define a flat family. Setting $t = 1$, the troublesome remainder is $-xz^2 + y^3$. Making this monic, let $f_4 = xz^2 - y^3$; $f_4 \in I$ and

$$g_4 = t^4\lambda f_4 = xz^2 - ty^3.$$

Adjoin g_4 to the ideal J, redefining the family Y. Now,

$$zg_2 - yg_3 + t^{13}g_4 = 0$$

restricts to $z(wy) - y(wz) = 0$ as desired.

The module of syzygies of w^2, wy, wz, and xz^2 is generated by the pairwise syzygies we have already considered, and by the syzygy $xz(wz) - w(xz^2) = 0$, which is the restriction of

$$-ty^2 g_2 + xzg_3 - wg_4 = 0.$$

Thus, $J = (g_1, g_2, g_3, g_4)$ defines a flat family Y, and

$$w^2 - xy,\ wy - xz,\ wz - y^2,\ xz^2 - y^3$$

is a Gröbner basis for I. The limit X_0 is cut out by the monomial ideal $\mathrm{in}(I) = (w^2, wy, wz, xz^2)$, which we shall see shares many properties with the original ideal I. Note that $xz^2 - y^3 = 0$ defines the projection of X to the plane $\mathbf{P}^2$ in x, y, and z.

The scheme structure of X_0 is closely related to the combinatorial structure of the monomial k-basis for $S/\mathrm{in}(I)$: For each degree d in our example, the monomials not belonging to $\mathrm{in}(I)$ consist of three sets $\{x^d, x^{d-1}y, \ldots, y^d\}$, $\{x^{d-1}z, x^{d-2}yz, \ldots, y^{d-1}z\}$, $\{y^d, y^{d-1}z, \ldots, z^d\}$, and a lone extra monomial $x^{d-1}w$. The first two sets correspond to a double line supported on $w = z = 0$, the third set to the line $w = x = 0$, and the extra monomial to an embedded point supported at $w = y = z = 0$. Together, this describes the scheme structure of X_0. The first two sets consist of $d+1$ and d monomials, respectively; the third set adds $d-1$ new monomials, and overlaps two monomials we have already seen. With the extra monomial, we count $3d+1$ monomials in each degree, which agrees with the dimensions of the graded pieces of S/I. The embedded point is crucial; it makes this count come out right, and it alone keeps X_0 nonplanar like X.

The new monomial generator xz^2 of $\mathrm{in}(I)$ excludes the line $w = y = 0$ from X_0; combinatorially, it excludes all but three monomials of the set $\{x^d, x^{d-1}z, \ldots, z^d\}$ from the monomial k-basis for each degree of the quotient $S/\mathrm{in}(I)$. We can see that this line is unwanted as follows: Away from $t = 0$, Y is parametrized by $(r, s, t) \mapsto (t^{16}r^2 s, t^4 r^3, trs^2, s^3, t)$. Thus, fixing r and s, the curve $(r, ts, t) \mapsto (t^{17}r^2 s, t^4 r^3, t^3 rs^2, t^3 s^3, t)$, with projective limit $(0, 0, r, s, 0)$ as $t \to 0$. Similarly, the curve $(r, t^3 s, t^2)$ has as its limit $(0, r^2, s^2, 0, 0)$. These calculations show that the lines $w = z = 0$ and $w = x = 0$ indeed belong set-theoretically to the limit X_0. We can find no

such curve whose limit is a general point on the line $w = y = 0$, for $(r, t^4 s, t^3)$ doesn't work. Thus, the line $w = y = 0$ sticks out of the good total space Y.

One usually computes Gröbner bases by working directly in the ring S, dispensing with the parameter t. The one-parameter subgroup λ is replaced by a total order on the monomials of each degree, satisfying the *multiplicative* property $\mathbf{x}^A > \mathbf{x}^B \Rightarrow \mathbf{x}^C \mathbf{x}^A > \mathbf{x}^C \mathbf{x}^B$ for all $\mathbf{x}^C$. In fact, for our purposes these are equivalent concepts: The weight vector W associated with λ induces the order $\mathbf{x}^A > \mathbf{x}^B \Longleftrightarrow W \cdot A < W \cdot B$, which is a total multiplicative order in low degrees as long as no two monomials have the same weight. Conversely, given any multiplicative order and a degree bound d, one can find many λ which induce this order on all monomials of degree $< d$. See [Bay82], [Rob85] for characterizations of such orders.

We shall be particularly interested in two multiplicative orders, the *lexicographic* order used in our example, and the *reverse lexicographic* order. The lexicographic order simply expands out the monomials of each degree into words, and sorts them alphabetically, i.e. $\mathbf{x}^A > \mathbf{x}^B$ iff the first nonzero entry in $A - B$ is positive. The reverse lexicographic order pushes highest powers of x_n in any expression back to the end, then within these groups pushes highest powers of x_{n-1} to the end, etc., i.e. $\mathbf{x}^A > \mathbf{x}^B$ iff the last nonzero entry of $A - B$ is negative.

What do these orders mean geometrically? The dominant effect of the lexicographic order is a projection from $\mathbf{P}^n$ to $\mathbf{P}^{n-1}$, eliminating x_0. A second order effect is a projection to $\mathbf{P}^{n-2}$, and so forth. We could compute the deformation from X to X_0 with respect to the lexicographic order in stages carrying out these projections, first applying a λ with $W = (-1, 0, \ldots, 0)$, then with $W = (-1, -1, 0, \ldots, 0)$, etc. Alternatively, for monomials of each degree $< d$, we can apply the single λ with $W = (-d^{n-1}, \ldots, -d, -1, 0)$, generalizing the λ used in our example. Use of the lexicographic order tends to muck up the family Y more than necessary in most applications, because projections tend to complicate varieties.

For the reverse lexicographic order, the dominant effect is a projection of $\mathbf{P}^n$ down to the last coordinate point $(0, \ldots, 0, 1)$. As a secondary effect, this order projects down to the last coordinate line, and so forth. In other words, this order first tries to make X into a cone over the last coordinate point, and only then tries to squash the result down to or cone it over the

last coordinate line, etc. For monomials of each degree $< d$, this can be realized by applying λ with $W = (0, 1, d, \ldots, d^{n-1})$. Like such cones, the reverse lexicographic order enjoys special properties with respect to taking linear sections of X or X_0 by intersection with the spaces defined by the last variable(s) (see [BS87a]). The preferred status of the reverse lexicographic order can be attributed to this relationship, because generic linear sections do not complicate varieties.

For example, if we take X to be three general points in $\mathbf{P}^2$, then using the lexicographic order X_0 becomes a triple point on a line, because the first order effect is the projection of the three points to a line, and the second order limiting process keeps the points within this line. By contrast, if we use the reverse lexicographic order then X_0 becomes the complete first order neighborhood of a point (a point doubled in all directions). This is because the first order limiting process brings the three points together from distinct directions, tracing out a cone over the three points. The first order neighborhood of the vertex in this cone has multiplicity 3, and is the same as the complete first order neighborhood in the plane of this vertex.

For those familiar with the theory of valuations in birational geometry [ZS76, Vol. II, Ch. VI], the lexicographic and reverse lexicographic orders have simple interpretations. Recall that if X is a variety of dimension n, and

$$F : X = Z_0 \supset Z_1 \supset Z_2 \supset \ldots \supset Z_n$$

is a flag of subvarieties, $\operatorname{codim}_X(Z_i) = i$, with Z_i smooth at the generic point of Z_{i+1}, then we can define a rank n valuation v_F on X as follows: For each $i = 1, \ldots, n-1$, fix f_i to be a function on Z_{i-1} with a 1^{st} order zero on Z_i. Then for any function f, we can define $e_1 = \operatorname{ord}_{Z_1}(f)$, $e_2 = \operatorname{ord}_{Z_2}((f/f_1^{e_1})\,|_{Z_1})$, etc., and $v_F(f) = (e_1, \ldots, e_n) \in \mathbf{Z}^n$, where the value group $\mathbf{Z}^n$ is ordered lexicographically. The arbitrarily chosen f_i are not needed to compare two functions f, g: We have $v_F(f) \succ v_F(g)$ if and only if $\operatorname{ord}_{Z_1}(f/g) > 0$, or if this order is zero and $\operatorname{ord}_{Z_2}((f/g)\,|_{Z_1}) > 0$, and so forth. Such a valuation also defines an order on each graded piece S_d of the homogeneous coordinate ring: take any $f_0 \in S_d$ and say $f > g$ if and only if $v_F(f/f_0) \succ v_F(g/f_0)$. More generally, one may take the Z_i to be subvarieties of a variety X' dominating X and pull back functions to X' before computing v_F.

The lexicographic order on monomials of each degree of $\mathbf{P}^n$ is now induced

by the flag

$$\mathbf{P}^n \supset V(x_0) \supset V(x_0, x_1) \supset \ldots \supset V(x_0, \ldots, x_{n-1}).$$

For example, the first step in the comparison defining $v_F(\mathbf{x}^A/f_0) \succ v_F(\mathbf{x}^B/f_0)$ has the effect of asking if $a_0 - b_0 > 0$.

The reverse lexicographic order is induced by a flag on a blowup X of $\mathbf{P}^n$: First blow up $V(x_0, \ldots, x_{n-1})$ and let E_1 be the exceptional divisor. Next blow up the proper transform of $V(x_0, \ldots, x_{n-2})$, and let E_2 be this exceptional divisor. Iterating, we can define a flag

$$X \supset E_1 \supset E_1 \cap E_2 \supset \ldots \supset E_1 \cap \ldots \cap E_n$$

which induces the reverse lexicographic order on monomials in each degree. For example, looking at the affine piece of the first blow up obtained by substituting $x_0 = x_0' x_{n-1}, \ \ldots, \ x_{n-2} = x_{n-2}' x_{n-1}$, the power of x_{n-1} in the transform of $\mathbf{x}^A$ is $a_0 + \ldots + a_{n-1}$, which is the order of vanishing of this monomial on E_1. Thus, the first step in the comparison defining $v_F(\mathbf{x}^A/f_0) \succ v_F(\mathbf{x}^B/f_0)$ has the effect of asking if $a_0 + \ldots + a_{n-1} - b_0 - \ldots - b_{n-1} > 0$, which is what we want.

Taking into account the equivalence between multiplicative orders and one-parameter subgroups, the process we have described in $S[t]$ is exactly the usual algorithm for computing Gröbner bases. It is computationally advantageous to set $t = 1$ and dismiss our extra structure as unnecessary scaffolding, but it is conceptually advantageous to treat our viewpoint as what is "really" going on; many techniques of algebraic geometry become applicable to the family Y, and assist in analyzing the complexity of Gröbner bases. Moreover, this picture may help guide improvements to the basic algorithm. For example, for very large problems, it could be computationally more efficient to degenerate to X_0 in several stages; this has not been tried in practice.

The coarsest measure of the complexity of a Gröbner basis is its maximum degree, which is the highest degree of a generator of the ideal $\mathrm{in}(I)$ defining X_0. This quantity is bounded by the better-behaved *regularity* of $\mathrm{in}(I)$: The regularity of an ideal I is the maximum over all i of the degree minus i of any minimal i^{th} syzygy of I, treating generators as 0^{th} syzygies. When I is the largest (the *saturated*) ideal defining a scheme X, we call this the regularity

of X. We take up regularity in detail in Section 3; here it suffices to know that regularity is *upper semi-continuous* on flat families, i.e. the regularity can only stay the same or go up at special fibers.

Let $\mathrm{reg}(I)$ denote the regularity of I, and $\mathrm{reg}_0(I)$ denote the highest degree of a generator of I. In our case, $t = 0$ is the only special fiber, and the above says that

$$\mathrm{reg}_0(I) \;\leq\; \mathrm{reg}(I) \;\leq\; \mathrm{reg}(\mathrm{in}(I)) \;\geq\; \mathrm{reg}_0(\mathrm{in}(I)),$$

where $\mathrm{reg}_0(I)$ can be immediately determined from the input data, and $\mathrm{reg}_0(\mathrm{in}(I))$ is the degree-complexity of the Gröbner basis computation. In practice, each of these inequalities are often strict.

However when k is infinite, then for any set of coordinates for $\mathbf{P}^n$ chosen from a dense open set $U \subset GL(n+1)$ of possibilities, Galligo ([Gal74]; see also [BS87b]) has shown that the limiting ideal $\mathrm{in}(I)$ takes on a very special form: $\mathrm{in}(I)$ is invariant under the action of the Borel subgroup of upper triangular matrices in $GL(n+1)$. This imposes strong geometric conditions on X_0. In particular, the associated primes of $\mathrm{in}(I)$ are also Borel-fixed, so they are all of the form $(x_0, \ldots, x_i)$ for various i. This means that the components of X_0 are supported on members of a flag.

In characteristic zero, it is shown in [BS87a] that the regularity of a Borel-fixed ideal is exactly the maximum of the degrees of its generators, or in our notation, that $\mathrm{reg}(\mathrm{in}(I)) = \mathrm{reg}_0(\mathrm{in}(I))$ when $\mathrm{in}(I)$ is Borel-fixed. Thus, for generic coordinates in characteristic zero, the degree-complexity of computing Gröbner bases breaks down into two effects: the gap $\mathrm{reg}_0(I) \leq \mathrm{reg}(I)$ between the input degrees and the regularity of X, and the gap $\mathrm{reg}(I) \leq \mathrm{reg}(\mathrm{in}(I))$ allowed by upper-semicontinuity.

A combination of theoretical results, hunches and experience guides the practitioner in assessing the first gap; what about the second? Does the regularity have to jump at all? One can easily find examples of ideals and total orders exhibiting such a jump, but in [BS87a], it is shown that for the reverse lexicographic order, in generic coordinates and any characteristic, there is no jump: $\mathrm{reg}(I) = \mathrm{reg}(\mathrm{in}(I))$, so in characteristic zero we have

$$\mathrm{reg}_0(\mathrm{in}(I)) = \mathrm{reg}(I).$$

In this sense, this order is an optimal choice: *For the reverse lexicographic order, the degree-complexity of a Gröbner basis computation is exactly the*

regularity of the input data. This agrees with experience; computations made on the same inputs using the lexicographic order can climb to much higher degrees than the reverse lexicographic order, in practice.

For many applications, one is free to choose any order, but some problems restrict us to using orders satisfying combinatorial properties which the reverse lexicographic order fails to satisfy. An example, developed further in Section 2, is that of eliminating variables, or equivalently, of computing projections. To compute the intersection of I with a subring $R = k[x_i, \ldots, x_n]$, it is necessary to use an order which in each degree sorts all monomials not in R ahead of any monomial in R. The lexicographic order is an example of such an order, for each i simultaneously. This strength comes at a cost; we are paying in regularity gaps for properties we may not need in a particular problem. An optimal order if you need one specific projection (in the same sense as above) is constructed by sorting monomials by total degree in the variables to be eliminated, and then breaking ties using the reverse lexicographic order. See [BS87b] for this result, and a generalization to the problem of optimally refining any nonstrict order.

Using this elimination order, one finds that the inherent degree-complexity of a computation is given not by the regularity of X itself, but rather by the regularity of the *flat projection* X' of X, which is the central fiber of a flat family which animates the desired projection of X as $t \to 0$. The jump in regularity between X and X' is unavoidable; by choosing an optimal order, we avoid the penalty of a further jump in regularity between X' and X_0.

The regularity of algebraic varieties or schemes X is far from being well understood, but there is considerable interest in its study; this computational interpretation of regularity as the inherent degree-complexity of an ideal is but one more log on the fire.

From a theoretical computer science perspective, the full complexity of computing Gröbner bases is determined not merely by the highest degree $\mathrm{reg}_0(I)$ in the basis, but by the total number of arithmetic operations in the field k required to compute this basis. This has not been analyzed in general, but for 0-dimensional ideals I, Lakshman and Lazard ([Lak91], [LL91]) have shown that the complexity of computing reduced Gröbner bases is bounded by a polynomial in d^n, where d is the maximum degree of the generators,

and n is the number of variables.

2 Gröbner Bases

Let $S = k[x_0, \ldots, x_n]$ be a graded polynomial ring over the field k, and let $I \subset S$ be a homogeneous ideal.

Let S_d denote the finite vector space of all homogeneous, degree d polynomials in S, so $S = S_0 \oplus S_1 \oplus \ldots \oplus S_d \oplus \ldots$. Writing I in the same manner as $I = I_0 \oplus I_1 \oplus \ldots \oplus I_d \oplus \ldots$, we have $I_d \subset S_d$ for each d. Recall that the Hilbert function of I is defined to be the function $p(d) = \dim(I_d)$, for $d \geq 0$.

A total order $>$ on the monomials of S is said to be *multiplicative* if whenever $\mathbf{x}^A > \mathbf{x}^B$ for two monomials $\mathbf{x}^A$, $\mathbf{x}^B$, then $\mathbf{x}^C\mathbf{x}^A > \mathbf{x}^C\mathbf{x}^B$ for all monomials $\mathbf{x}^C$. This condition insures that if the terms of a polynomial are in order with respect to $>$, then they remain in order after multiplication by a monomial.

Definition 2.1 *Let $>$ be a multiplicative order. For a homogeneous polynomial $f = c_1\mathbf{x}^{A_1} + \ldots + c_m\mathbf{x}^{A_m}$ with $\mathbf{x}^{A_1} > \ldots > \mathbf{x}^{A_m}$, define the initial term* $\mathrm{in}(f)$ *to be the lead (that is, the largest) term $c_1\mathbf{x}^{A_1}$ of f. For a homogeneous ideal $I \subset S$, define the initial ideal* $\mathrm{in}(I)$ *to be the monomial ideal generated by the lead terms of all elements of I.*

Note that the definitions of $\mathrm{in}(f)$ and $\mathrm{in}(I)$ depend on the choice of multiplicative order $>$. See [BM88] and [MR88] for characterizations of the finite set of $\mathrm{in}(I)$ realized as the order $>$ varies.

Fix a multiplicative order $>$ on S.

Proposition 2.2 (Macaulay) *I and* in*(I) have the same Hilbert function.*

Proof. ([Mac27]) The lead terms of I_d span $\mathrm{in}(I)_d$, because every monomial $\mathbf{x}^A \in \mathrm{in}(I)$ is itself the lead term $\mathrm{in}(f)$ of some polynomial $f \in I$: Since $\mathbf{x}^A = \mathbf{x}^C\mathbf{x}^B$ for some $\mathbf{x}^B = \mathrm{in}(g)$ with $g \in I$, we have $\mathbf{x}^A = \mathrm{in}(f)$ for $f = \mathbf{x}^C g$.

Choose a k-basis $B_d \subset I_d$ with distinct lead terms, and let $\mathrm{in}(B_d)$ be the set of lead terms of B_d; $\mathrm{in}(B_d)$ has cardinality $p(d) = \dim(I_d)$. Since any

element of I_d is a linear combination of elements of B_d, any lead term of I_d is a scalar multiple of an element of $\text{in}(B_d)$. Thus, $\text{in}(B_d)$ is a basis for $\text{in}(I)_d$, so $p(d) = \dim(\text{in}(I)_d)$. ∎

One can compute the Hilbert function of I by finding $\text{in}(I)$ and applying this result; see [MM83], [BCR91], and [BS92b].

Corollary 2.3 *The monomials of S which don't belong to* $\text{in}(I)$ *form a k-basis for S/I.*

Proof. These monomials are linearly independent in S/I, because any linear relation among them is a polynomial belonging to I, and all such polynomials have lead terms belonging to $\text{in}(I)$. These monomials can be seen to span S/I by a dimension count, applying Proposition 2.2. ∎

Two examples of multiplicative orders are the lexicographic order and the reverse lexicographic order. $\mathbf{x}^A > \mathbf{x}^B$ in the lexicographic order if the first nonzero coordinate of $A - B$ is positive. For example, if $S = k[w, x, y, z]$, then $w > x > y > z$ in S_1, and

$$w^2 > wx > wy > wz > x^2 > xy > xz > y^2 > yz > z^2$$

in S^2.

$\mathbf{x}^A > \mathbf{x}^B$ in the reverse lexicographic order if the last nonzero coordinate of $A - B$ is negative. For example, if $S = k[w, x, y, z]$, then $w > x > y > z$ in S_1, and

$$w^2 > wx > x^2 > wy > xy > y^2 > wz > xz > yz > z^2$$

in S^2. These two orders agree on S_1, but differ on the monomials of S of degree > 1 when $n \geq 2$.

The lexicographic order has the property that for each subring $k[x_i, \ldots, x_n] \subset S$ and each polynomial $f \in S$, $f \in k[x_i, \ldots, x_n]$ if and only if $\text{in}(f) \in k[x_i, \ldots, x_n]$. The reverse lexicographic order has the property that for each $f \in k[x_0, \ldots, x_i]$, x_i divides f if and only if x_i divides $\text{in}(f)$.

One can anticipate the applications of these properties by considering a k-basis $B_d \subset I_d$ with distinct lead terms, as in the proof of Proposition 2.2.

With respect to the lexicographic order, $B_d \cap k[x_i, \ldots, x_n]$ is then a k-basis for $I_d \cap k[x_i, \ldots, x_n]$ for each i. With respect to the reverse lexicographic order, $B_d \cap (x_n)$ is then a k-basis for $I_d \cap (x_n)$. Thus, these orders enable us to find polynomials in an ideal which do not involve certain variables, or which are divisible by a certain variable. For a given degree d, one could construct such a basis B_d by applying Gaussian elimination to an arbitrary k-basis for I_d. However, this cannot be done for all d at once; such a computation would be infinite. We will finesse this difficulty by instead constructing a finite set of elements of I whose monomial multiples yield polynomials in I with every possible lead term.

Such sets can be described as follows:

Definition 2.4 *A list $F = [f_1, \ldots, f_r] \subset I$ is a (minimal) Gröbner basis for I if $\mathrm{in}(f_1), \ldots, \mathrm{in}(f_r)$ (minimally) generate $\mathrm{in}(I)$.*

$\mathrm{in}(I)$ is finitely generated because S is Noetherian, so Gröbner bases exist for any ideal I.

The order of the elements of F is immaterial to this definition, so F can be thought of as a set. We are using list notation for F because we are going to consider algorithms for which the order of the elements is significant. For convenience, we shall extend the notation of set intersections and containments to the lists F.

A minimal set of generators for an ideal I need not form a Gröbner basis for I. For example, if $S = k[x, y]$ and $I = (x^2 + y^2, xy)$, then with respect to the lexicographic order, $\mathrm{in}(x^2 + y^2) = x^2$ and $\mathrm{in}(xy) = xy$. Yet $y(x^2 + y^2) - x(xy) = y^3 \in I$, so $y^3 \in \mathrm{in}(I)$. Thus, any Gröbner basis for I must include y^3; it can be shown that $\mathrm{in}(I) = (x^2, xy, y^3)$ and $[x^2 + y^2, xy, y^3]$ is a Gröbner basis for I.

On the other hand,

Lemma 2.5 *If $F = [f_1, \ldots, f_r]$ is a Gröbner basis for I, then $f_1, \ldots, f_r$ generate I.*

Proof. For each degree d, we can construct a k-basis $B_d \subset I_d$ with distinct lead terms, whose elements are monomial multiples of $f_1, \ldots, f_r$: For each

$\mathbf{x}^A \in \text{in}(I)_d$, $\mathbf{x}^A$ is a scalar multiple of $\mathbf{x}^C \text{in}(f_i)$ for some $\mathbf{x}^C$ and some i; include $\mathbf{x}^C f_i$ in the set B_d. Thus, the monomial multiples of $f_1, \ldots, f_r$ span I. ∎

Proposition 2.6 (Spear, Trinks) *Let $R \subset S$ be the subring $R = k[x_i, \ldots, x_n]$. If $F = [f_1, \ldots, f_r]$ is a Gröbner basis for the ideal I with respect to the lexicographic order, then $F \cap R$ is a Gröbner basis for the ideal $I \cap R$. In particular, $F \cap R$ generates $I \cap R$.*

Proof. ([Spe77], [Zac78], [Tri78]) Let $f \in I \cap R$; $\text{in}(f)$ is a multiple of $\text{in}(f_i)$ for some i. Since $\text{in}(f) \in R$, $\text{in}(f_i) \in R$, so $f_i \in R$. Thus, $F \cap R$ is a Gröbner basis for $I \cap R$. By Lemma 2.5, $F \cap R$ generates $I \cap R$. ∎

Proposition 2.6 has the following geometric application: If I defines the subscheme $X \subset \mathbf{P}^n$, then $I \cap k[x_i, \ldots, x_n]$ defines the projection of X to $\mathbf{P}^{n-i} = \text{Proj}(k[x_i, \ldots, x_n])$.

Recall that the saturation I^{sat} of I is defined to be the largest ideal defining the same subscheme $X \subset \mathbf{P}^n$ as I. I^{sat} can be obtained by taking an irredundant primary decomposition for I, and removing the primary ideal whose associated prime is the irrelevant ideal $(x_0, \ldots, x_n)$. I is saturated if $I = I^{\text{sat}}$.

If the ideal I is saturated, and defines a finite set of points $X \subset \mathbf{P}^n$, then $I \cap k[x_{n-1}, x_n]$ is a principal ideal (f), where $\{f = 0\}$ is the image of the projection of X to $\mathbf{P}^1 = \text{Proj}(k[x_{n-1}, x_n])$. Given a linear factor of f of the form $(bx_{n-1} - ax_n)$, we can make the substitution $x_{n-1} = az$, $x_n = bz$ for a new variable z, to obtain from I an ideal $J \subset k[x_0, \ldots, x_{n-2}, z]$ defining a finite set of points in $\mathbf{P}^{n-1}$. For each point $(c_0, \ldots, c_{n-2}, d)$ in the zero locus of J, $(c_0, \ldots, c_{n-2}, ad, bd)$ is a point in the zero locus of I.

If $X \subset \mathbf{P}^{n-1}$ is of dimension 1 or greater, then in general $I \cap k[x_{n-1}, x_n] = (0)$, because a generic projection of X to $\mathbf{P}^1$ is surjective. In this case, an arbitrary substitution $x_{n-1} = az$, $x_n = bz$ can be made, and the process of projecting to $\mathbf{P}^1$ iterated. Thus, the lexicographic order can be used to find solutions to systems of polynomial equations.

Recall that the ideal quotient $(I : f)$ is defined to be the ideal $\{\, g \in S \mid fg \in I \,\}$. Since S is Noetherian, the ascending chain of ideals $(I : f) \subset$

$(I : f^2) \subset (I : f^3) \subset \ldots$ is stationary; call this stationary limit $(I : f^\infty) = \{\, g \in S \mid f^m g \in I \text{ for some } m \,\}$.

Proposition 2.7 *If $[x_n^{a_1} f_1, \ldots, x_n^{a_r} f_r]$ is a Gröbner basis for the ideal I with respect to the reverse lexicographic order, and if none of $f_1, \ldots, f_r$ are divisible by x_n, then $F = [f_1, \ldots, f_r]$ is a Gröbner basis for the ideal $(I : x_n^\infty)$. In particular, $f_1, \ldots, f_r$ generate $(I : x_n^\infty)$.*

Proof. ([Bay82], [BS87a]) We have $F \subset (I : x_n^\infty)$. Let $f \in (I : x_n^\infty)$; $x_n^m f \in I$ for some m, so $\text{in}(x_n^m f)$ is a multiple of $\text{in}(x_n^{a_i} f_i)$ for some i. Since f_i is not divisible by x_n, $\text{in}(f_i)$ is not divisible by x_n, so $\text{in}(f)$ is a multiple of $\text{in}(f_i)$. Thus, F is a Gröbner basis for $(I : x_n^\infty)$. By Lemma 2.5, $f_1, \ldots, f_r$ generate $(I : x_n^\infty)$. ∎

If $I = \mathbf{q}_0 \cap \mathbf{q}_1 \cap \ldots \cap \mathbf{q}_t$ is a primary decomposition of I, then $(I : x_n^\infty) = (\cap \mathbf{q}_i : x_n^\infty) = \cap (\mathbf{q}_i : x_n^\infty)$. We have $(\mathbf{q}_i : x_n^\infty) = (1)$ if the associated prime $\mathbf{p}_i$ of $\mathbf{q}_i$ contains x_n, and $(\mathbf{q}_i : x_n^\infty) = \mathbf{q}_i$ otherwise. Thus, if I defines the subscheme $X \subset \mathbf{P}^n$, then $(I : x_n^\infty)$ defines the subscheme consisting of those primary components of X not supported on the hyperplane $\{x_n = 0\}$.

$(I : x_n^\infty)$ is saturated, because it cannot have $(x_0, \ldots, x_n)$ as an associated prime. If x_n belongs to none of the associated primes of I except $(x_0, \ldots, x_n)$, or equivalently if $\{x_n = 0\}$ is a generic hyperplane section of $X \subset \mathbf{P}^n$, then $(I : x_n^\infty) = I^{\text{sat}}$. Thus, the reverse lexicographic order can be used to find the saturation of I.

One of the most important uses of Gröbner bases is that they lead to canonical representations of polynomials modulo an ideal I, i.e. a division algorithm in which every $f \in S$ is written canonically as $f = \sum g_i f_i + h$, where $[f_1, \ldots, f_r]$ is a Gröbner basis for I, and h is the remainder after division.

Recall the division algorithm for inhomogeneous, univariate polynomials $f(x)$, $g(x) \in k[x]$: Let $\text{in}(f)$ denote the highest degree term of f. The remainder of g under division by f can be recursively defined by

$$R_f(g) = R_f(g - cx^a f)$$

if $\text{in}(f)$ divides $\text{in}(g)$, where $cx^a = \text{in}(g)/\text{in}(f)$, and by

$$R_f(g) = g$$

otherwise.

Division can be generalized to homogeneous polynomials $f_1, \ldots, f_r, g \in S$, given a multiplicative order on S ([Hir64], [Bri73], [Gal74], [Sch80]): The remainder $R_F(g)$ of g under division by the list of polynomials $F = [f_1, \ldots, f_r]$ can be recursively defined by

$$R_F(g) = R_F(g - c\mathbf{x}^A f_i)$$

for the least i so $\text{in}(g)$ is a multiple $c\mathbf{x}^A$ of $\text{in}(f_i)$, and by

$$R_F(g) = \text{in}(g) + R_F(g - \text{in}(g))$$

if $\text{in}(g)$ is not a multiple of any $\text{in}(f_i)$. $R_F(g)$ is an element of S.

Thus, the fate of $\text{in}(g)$ depends on whether or not $\text{in}(g) \in (\text{in}(f_1), \ldots, \text{in}(f_r))$. Let I be the ideal generated by $f_1, \ldots, f_r$. If $F = [f_1, \ldots, f_r]$ fails to be a Gröbner basis for I, then the remainder is poorly behaved. For example, with respect to the lexicographic order on $k[x, y]$,

$$R_{[xy, x^2+y^2]}(x^2y) = x^2y - x(xy) = 0,$$

but

$$R_{[x^2+y^2, xy]}(x^2y) = x^2y - y(x^2 + y^2) = -y^3,$$

so the remainder $R_F(g)$ is dependent on the order of the list F. Note that $x^2y \in (x^2 + y^2, xy)$.

If on the other hand, F is a Gröbner basis for the ideal I, then $R_F(g)$ is a k-linear combination of monomials not belonging to $\text{in}(I)$. By Corollary 2.3, these monomials form a k-basis for S/I, so each polynomial in S has a unique representation in terms of this k-basis, modulo the ideal I. The remainder gives this unique representation, and is independent of the order of F (but dependent on the multiplicative order chosen for the monomials of S). In particular, $R_F(g) = 0$ if and only if $g \in I$.

An algorithm for computing a Gröbner basis for I from a set of generators for I was first given by Buchberger ([Buc65], [Buc76]). This algorithm was discovered independently by Spear ([Spe77], [Zac78]), Bergman [Ber78], and Schreyer [Sch80]. It was termed the division algorithm by Schreyer, after the division theorem of Hironaka ([Hir64], [Bri73], [Gal74]).

Define $S(f_i, f_j)$ for $i < j$ by

$$S(f_i, f_j) = b\mathbf{x}^B f_i - c\mathbf{x}^C f_j,$$

where $\mathbf{x}^A = b\mathbf{x}^B \text{in}(f_i) = c\mathbf{x}^C \text{in}(f_j)$ is the least common multiple of $\text{in}(f_i)$ and $\text{in}(f_j)$. $b\mathbf{x}^B f_i$ and $c\mathbf{x}^C f_j$ each have $\mathbf{x}^A$ as lead term, so $\mathbf{x}^A$ cancels out in $S(f_i, f_j)$, and $\mathbf{x}^A > \text{in}(S(f_i, f_j))$.

If F is a Gröbner basis for the ideal I, then $R_F(S(f_i, f_j)) = 0$ for each $i < j$, since $S(f_i, f_j) \in I$. Conversely,

Proposition 2.8 (Buchberger) *If $R_F(S(f_i, f_j)) = 0$ for each $i < j$, then $F = [f_1, \ldots, f_r]$ is a Gröbner basis for the ideal $I = (f_1, \ldots, f_r)$.*

See [Buc65], [Buc76]. We postpone a proof until the theory has been extended to S-modules. This result can also be thought of as an explicit converse to the assertion that if F is a Gröbner basis, then division is independent of the order of F: Whenever we have a choice in division between subtracting off a multiple of f_i and a multiple of f_j, the difference is a multiple of $S(f_i, f_j)$. If division is independent of the order of F, then these differences must have remainder zero, so by Proposition 2.8, F is a Gröbner basis.

As sketched in Section 1, Proposition 2.8 can be used to compute a Gröbner basis from a set of generators $f_1, \ldots, f_r$ for the ideal I: For each $i < j$ so $f_{r+1} = R_F(S(f_i, f_j)) \neq 0$, adjoin f_{r+1} to the list $F = [f_1, \ldots, f_r]$. Note that $f_{r+1} \in I$. By iterating until no new polynomials are found, a Gröbner basis F is obtained for I. This process terminates because S is Noetherian, and each new basis element corresponds to a monomial not in the ideal generated by the preceding lead terms.

We now extend this theory to S-modules. Let M be a graded, finitely generated S-module, given by the exact sequence of graded S-modules

$$M_1 \xrightarrow{F} M_0 \longrightarrow M \longrightarrow \mathbf{0},$$

where $M_0 = Se_{01} \oplus \ldots \oplus Se_{0q}$ and $M_1 = Se_{11} \oplus \ldots \oplus Se_{1r}$ are free S-modules with $\deg(e_{ij}) = d_{ij}$ for each i, j. We now think of F both as a list $[f_1, \ldots, f_r]$ of module elements, and as a map between free modules: Let $f_i = F(e_{1i}) \neq 0$

for $i = 1, \ldots, r$, and let $I \subset M_0$ be the homogeneous submodule generated by $f_1, \ldots, f_r$. Thus, $M = M_0/I$.

A monomial of M_0 is an element of the form $\mathbf{x}^A e_{0i}$; such an element has degree $\deg(\mathbf{x}^A) + d_{0i}$. An order on the monomials of M_0 is multiplicative if whenever $\mathbf{x}^A e_{0i} > \mathbf{x}^B e_{0j}$, then $\mathbf{x}^C \mathbf{x}^A e_{0i} > \mathbf{x}^C \mathbf{x}^B e_{0j}$ for all $\mathbf{x}^C \in S$. For some applications, such as developing a theory of Gröbner bases over quotients of S, one wants this order to be compatible with an order on S: If $\mathbf{x}^A > \mathbf{x}^B$, then one wants $\mathbf{x}^A e_{0i} > \mathbf{x}^B e_{0i}$ for $i = 1, \ldots, r$. The orders encountered in practice invariably satisfy this second condition, but it does not follow from the first, and we do not require it here.

One way to extend a multiplicative order on S to a compatible multiplicative order on M_0 is to declare $\mathbf{x}^A e_{0i} > \mathbf{x}^B e_{0j}$ if $i < j$, or if $i = j$ and $\mathbf{x}^A > \mathbf{x}^B$. Another way is to assign monomials $\mathbf{x}^{C_1}, \ldots, \mathbf{x}^{C_q}$ in S to the basis elements $e_{01}, \ldots, e_{0q}$ of M_0, and to declare $\mathbf{x}^A e_{0i} > \mathbf{x}^B e_{0j}$ if $\mathbf{x}^{A+C_i} > \mathbf{x}^{B+C_j}$, or if $A + C_i = B + C_j$ and $i < j$.

Fix a choice of a multiplicative order $>$ on M_0. The constructions developed for S carry over intact to M_0, with the same proofs ([Gal79], [Sch80], [Bay82]): Given an element $f \in M_0$, define $\mathrm{in}(f)$ to be the lead term of f. Define $\mathrm{in}(I)$ to be the submodule generated by the lead terms of all elements of $I \subset M_0$; $\mathrm{in}(I)$ is a monomial submodule of M_0 with the same Hilbert function as I. Define $F = [f_1, \ldots, f_r] \subset I$ to be a Gröbner basis for I if $\mathrm{in}(f_1), \ldots, \mathrm{in}(f_r)$ generate $\mathrm{in}(I)$; a set of generators for I need not be a Gröbner basis for I, but a Gröbner basis for I generates I. Given an element $g \in M_0$, define $R_F(g) \in M_0$ exactly as was done for the free module S. If F is a Gröbner basis for I, then $R_F(g) = 0$ if and only if $g \in I$.

The quotient of g under division by $f_1, \ldots, f_r$ can be recursively defined by

$$Q_F(g) = c\mathbf{x}^A e_{1i} + Q_F(g - c\mathbf{x}^A f_i)$$

for the least i so $\mathrm{in}(g)$ is a multiple $c\mathbf{x}^A$ of $\mathrm{in}(f_i)$, and by

$$Q_F(g) = Q_F(g - \mathrm{in}(g))$$

if $\mathrm{in}(g)$ is not a multiple of any $\mathrm{in}(f_i)$. The quotient is an element of M_1.

Following the recursive definitions of the remainder and quotient, it can be inductively verified that

$$g = F(Q_F(g)) + R_F(g).$$

If F is a Gröbner basis for I, and $g \in I$, then $R_F(g) = 0$, so the quotient lifts g to M_1. In this case, the quotient can be thought of as expressing g in terms of $f_1, \ldots, f_r$.

Define $S(f_i, f_j)$ for $i < j$ by

$$S(f_i, f_j) = b\mathbf{x}^B f_i - c\mathbf{x}^C f_j,$$

if $\mathrm{in}(f_i)$ and $\mathrm{in}(f_j)$ have a least common multiple $\mathbf{x}^A e_{0k} = b\mathbf{x}^B \mathrm{in}(f_i) = c\mathbf{x}^C \mathrm{in}(f_j)$. Leave $S(f_i, f_j)$ undefined if $\mathrm{in}(f_i)$ and $\mathrm{in}(f_j)$ lie in different summands of M_0, and so don't have common multiples.

Recall that the module of syzygies of $f_1, \ldots, f_r$ is defined to be the kernel of the map F, which is the submodule of M_1 consisting of all $h \in M_1$ so $F(h) = 0$. Thus, if $h = h_1 e_{11} + \ldots + h_r e_{1r}$ is a syzygy, then $h_1 f_1 + \ldots + h_r f_r = 0$. Let $J \subset M_1$ denote the module of syzygies of $f_1, \ldots, f_r$, and let $K \subset M_1$ denote the module of syzygies of $\mathrm{in}(f_1), \ldots, \mathrm{in}(f_r)$.

Define the map $\mathrm{in}(F) : M_1 \to M_0$ by $\mathrm{in}(F)(e_{1i}) = \mathrm{in}(f_i)$; K is the kernel of $\mathrm{in}(F)$. For each $i < j$ so $S(f_i, f_j)$ is defined, define t_{ij} to be the element

$$t_{ij} = b\mathbf{x}^B e_{1i} - c\mathbf{x}^C e_{1j} \in M_1,$$

where $\mathbf{x}^A e_{0k} = b\mathbf{x}^B \mathrm{in}(f_i) = c\mathbf{x}^C \mathrm{in}(f_j)$ is the least common multiple of $\mathrm{in}(f_i)$ and $\mathrm{in}(f_j)$, as before. $\mathrm{in}(F)(t_{ij}) = 0$, so each t_{ij} belongs to the syzygy module K. Observe that $F(t_{ij}) = S(f_i, f_j)$.

Assign the following multiplicative order on M_1, starting from the order on M_0 ([Sch80]; see also [MM86]): Let $\mathbf{x}^A e_{1i} > \mathbf{x}^B e_{1j}$ if $\mathbf{x}^A \mathrm{in}(f_i) > \mathbf{x}^B \mathrm{in}(f_j)$, or if these terms are k-multiples of each other and $i < j$. If the order on M_0 is compatible with an order on S, then this order on M_1 is compatible with the same order on S.

With respect to this order on M_1, we have

Lemma 2.9 *The list* $[\, t_{ij} \,]$ *is a Gröbner basis for the module* K *of syzygies of* $\mathrm{in}(f_1), \ldots, \mathrm{in}(f_r)$.

Proof. Let $h \in M_1$, so $\mathrm{in}(F)(h) = 0$. Then $\mathrm{in}(F)(\mathrm{in}(h))$ is canceled by $\mathrm{in}(F)(h - \mathrm{in}(h))$ in M_0. Therefore, if $\mathrm{in}(h) = \mathbf{x}^A e_{1i}$, then h has another term $\mathbf{x}^B e_{1j}$ so $\mathbf{x}^A \mathrm{in}(f_i)$ and $\mathbf{x}^B \mathrm{in}(f_j)$ are k-multiples of each other and $i < j$.

Thus, t_{ij} is defined and $\mathrm{in}(t_{ij})$ divides $\mathrm{in}(h)$, so $[\, t_{ij}\,]$ is a Gröbner basis for K. ∎

Thus, the set $\{t_{ij}\}$ generates K. In general, the $[\, t_{ij}\,]$ are far from being a minimal Gröbner basis for K; we consider the effects of trimming this list in Proposition 2.10 below.

Define

$$s_{ij} = t_{ij} - Q_F(S(f_i, f_j))$$

whenever $R_F(S(f_i, f_j)) = 0$. Note that $\mathrm{in}(s_{ij}) = \mathrm{in}(t_{ij})$. Each s_{ij} is the difference of two distinct elements of M_1, each of which is mapped by F to $S(f_i, f_j)$, so $F(s_{ij}) = 0$. In other words, s_{ij} belongs to the syzygy module J. Conversely,

Proposition 2.10 (Richman, Spear, Schreyer) *Choose a set of pairs $T = \{(i,j)\}$ such that the set $\{t_{ij}\}_{(i,j)\in T}$ generates the module K of syzygies of $\mathrm{in}(f_1), \ldots, \mathrm{in}(f_r)$. If $R_F(S(f_i, f_j)) = 0$ for each $(i,j) \in T$, then*

(a) $F = [f_1, \ldots, f_r]$ is a Gröbner basis for I;

(b) the set $\{s_{ij}\}_{(i,j)\in T}$ generates the module J of syzygies of $f_1, \ldots, f_r$.

Moreover,

(c) if $[\, t_{ij}\,]_{(i,j)\in T}$ is a Gröbner basis for K, then $[\, s_{ij}\,]_{(i,j)\in T}$ is a Gröbner basis for J.

Proof. ([Ric74], [Spe77], [Zac78], [Sch80]) First, suppose that $[\, t_{ij}\,]_{(i,j)\in T}$ is a Gröbner basis for K. Let $h \in J$, so $F(h) = 0$. By the same reasoning as in the proof of Lemma 2.9, we can find $(i,j) \in T$ so $\mathrm{in}(t_{ij})$ divides $\mathrm{in}(h)$. Since $\mathrm{in}(s_{ij}) = \mathrm{in}(t_{ij})$, $\mathrm{in}(s_{ij})$ also divides $\mathrm{in}(h)$, so $[\, s_{ij}\,]_{(i,j)\in T}$ is a Gröbner basis for J, proving (c).

Now, suppose that $\{t_{ij}\}_{(i,j)\in T}$ merely generates K. Let T' be a set of pairs so $[\, t_{\ell m}\,]_{(\ell,m)\in T'}$ is a Gröbner basis for K. It is enough to construct a list $[\, u_{\ell m}\,]_{(\ell,m)\in T'}$ of elements of J, generated by $\{s_{ij}\}_{(i,j)\in T}$, so $\mathrm{in}(u_{\ell m}) = \mathrm{in}(t_{\ell m})$ for all $(\ell, m) \in T'$. Then by the preceding argument, $[\, u_{\ell m}\,]_{(\ell,m)\in T'}$ is a Gröbner basis for J, so $\{s_{ij}\}_{(i,j)\in T}$ generates J.

Write each $t_{\ell m} = \sum g_{\ell m ij} t_{ij}$, for $(\ell, m) \in T'$ and $(i,j) \in T$, in such a way that the terms of $t_{\ell m}$ and each term of each product $g_{\ell m ij} t_{ij}$ map via $\mathrm{in}(F)$

to multiples of the same monomial in M_0. In other words, find a minimal expression for each $t_{\ell m}$, which avoids unnecessary cancellation. Then define

$$u_{\ell m} = \sum g_{\ell m i j} s_{ij}.$$

We have $\mathrm{in}(u_{\ell m}) = \mathrm{in}(t_{\ell m})$, proving (b).

Let $f \in I$, and choose $g \in M_1$ so $f = F(g)$. Let $h \in M_1$ be the remainder of g under division by $[\, u_{\ell m}\,]_{(\ell,m)\in T'}$; $f = F(h)$. Since $\mathrm{in}(h)$ is not a multiple of any $\mathrm{in}(u_{\ell m}) = \mathrm{in}(t_{\ell m})$, the lead term of $F(\mathrm{in}(h))$ is not canceled by any term of $F(h - \mathrm{in}(h))$. Therefore, if $\mathrm{in}(h) = a\mathbf{x}^A e_{1i}$, then $\mathrm{in}(f_i)$ divides $\mathrm{in}(F)$. Thus, $F = [f_1, \ldots, f_r]$ is a Gröbner basis for I, proving (a). ∎

Proposition 2.8 follows as a special case of this result.

The above proof can be understood in terms of an intermediate initial form $\mathrm{in}_0(h)$ for $h \in M_1$: Apply the map $\mathrm{in}(F)$ separately to each term of h, and let $\mathbf{x}^A \in M_0$ be the greatest monomial that occurs in the set of image terms. Define $\mathrm{in}_0(h)$ to be the sum of all terms of h which map via $\mathrm{in}(F)$ to multiples of $\mathbf{x}^A$. Then in refines in_0, for according to the order we have defined on M_1, $\mathrm{in}(h)$ is the term of $\mathrm{in}_0(h)$ lying in the summand of M_1 whose basis element e_i has the smallest index i.

In this language, $t_{ij} = \mathrm{in}_0(t_{ij}) = \mathrm{in}_0(s_{ij})$. Our expressions for the $t_{\ell m}$ have the property that each $g_{\ell m i j} t_{ij} = \mathrm{in}_0(g_{\ell m i j} t_{ij})$, with each term of each product for a given $t_{\ell m}$ mapping via $\mathrm{in}(F)$ to multiples of the same monomial $\mathbf{x}^A$. Thus, each $\mathrm{in}_0(g_{\ell m i j} s_{ij}) = g_{\ell m i j} t_{ij}$; the tails $g_{\ell m i j}(s_{ij} - \mathrm{in}_0(s_{ij}))$ stay out of our way, mapping termwise via $\mathrm{in}(F)$ to monomials which are less than $\mathbf{x}^A$ with respect to the order on M_0.

Observe that $Q_F(g)$ is a linear combination of monomials not belonging to $\mathrm{in}(J)$, for any $g \in M_0$.

In [Buc79], Buchberger gives a criterion for selecting a set T of pairs (i, j) in the case where I is an ideal: If $(i_0, i_1), (i_1, i_2), \ldots, (i_{s-1}, i_s) \in T$, and the least common multiple of $\mathrm{in}(f_{i_0}), \mathrm{in}(f_{i_1}), \ldots, \mathrm{in}(f_{i_s})$ is equal to the least common multiple of $\mathrm{in}(f_{i_0})$ and $\mathrm{in}(f_{i_s})$, then (i_0, i_s) need not belong to T. In other words, if $t_{i_0 i_s} \in (t_{i_0 i_1}, \ldots, t_{i_{s-1} i_s})$, then the pair (i_0, i_s) is unnecessary; this condition is equivalent to the condition of Proposition 2.10, for the case of an ideal.

Suppose that we wish to compute the syzygies of a given set of elements

$g_1, \ldots, g_s$ of M_0. To do this, compute a Gröbner basis $f_1, \ldots, f_r$ for the submodule $I \subset M_0$ generated by $g_1, \ldots, g_s$. Keep track of how to write each f_i in terms of $g_1, \ldots, g_s$. Using these expressions, each syzygy of $f_1, \ldots, f_r$ can be mapped to a syzygy of $g_1, \ldots, g_s$. These images generate the module of syzygies of $g_1, \ldots, g_s$; the set of syzygies obtained in this way is not in general minimal.

Syzygies can be used to find a minimal set of generators for a submodule $I \subset M_0$ from a given set of generators $g_1, \ldots, g_s$: If $h_1 g_1 + \ldots + h_r g_r = 0$ is a syzygy of $g_1, \ldots, g_s$ with $h_1 \in k$, then $g_1 = (h_2 g_2 + \ldots + h_r g_r)/h_1$, so g_1 is not needed to generate I. All unnecessary generators can be removed in this way.

Alternatively, a careful implementation of Gröbner bases can directly find minimal sets of generators for submodules: Starting from an arbitrary set of generators, we can eliminate unnecessary generators degree by degree, by removing those which reduce to zero under division by a Gröbner basis for the ideal generated by the preceding generators.

Either way, we can trim the set of syzygies computed via Gröbner bases for a given set of generators $g_1, \ldots, g_s$ of I, to obtain a minimal set of generators for the syzygy module J. By starting with a minimal generating set for I, and iterating this method, a minimal free resolution can be found for I.

A beautiful application of these ideas yields a proof of the Hilbert syzygy theorem, that minimal free resolutions terminate (Schreyer [Sch80], [Sch91], for an exposition see also Eisenbud [Eis92]). At each stage of a resolution, order the Gröbner basis F for I in such a way that for each $i < j$, letting $\mathrm{in}(f_i) = a\mathbf{x}^A e_{0k}$ and $\mathrm{in}(f_j) = b\mathbf{x}^B e_{0\ell}$, we have $\mathbf{x}^A > \mathbf{x}^B$ in the lexicographic order. If the variables $x_1, \ldots, x_m$ are missing from the initial terms of the f_i, then the variables $x_1, \ldots, x_{m+1}$ will be missing from the initial terms of the syzygies s_{ij}. Iterating, we run out of variables, so the resolution terminates.

3 Bounds

How hard are the algorithms in algebraic geometry? We describe some key bounds. The best known example is the bound established by G. Hermann [Her26] for ideal membership:

Theorem 3.1 (G. Hermann) *Let k be any field, let $(f_1, ..., f_k) \subset k[x_1, ..., x_n]$ and let $d = \max(\deg(f_i))$. If $g \in (f_1, ..., f_k)$, then there is an expression*

$$g = \sum_{i=1}^{k} a_i f_i$$

where $\deg(a_i) \leq \deg(g) + 2(kd)^{2^{n-1}}$.

This type of bound is called "doubly exponential". However, with the advent of the concept of coherent sheaf cohomology [Ser55] and the systematic study of vanishing theorems, it has become apparent that the vanishing of these groups in high degrees is almost always the most fundamental bound. The concept of an ideal being "m-regular" or "regular in degrees $\geq m$" was introduced by one of us [Mum66] by generalizing ideas of Castelnuovo:

Definition 3.2 [1] *Let k be any field, let $I \subset k[x_0, \ldots, x_n]$ be an ideal generated by homogeneous polyomials, let I_d be the homogeneous elements in I of degree d, let I be the corresponding sheaf of ideals in $\mathcal{O}_{\mathbf{P}^n}$, and let $I(d)$ be the d^{th} twist of I. Then the following properties are equivalent and define the term "m-regular":*

(a) the natural map $I_m \to H^0(\mathcal{I}(m))$ is an isomorphism and $H^i(\mathcal{I}(m-i)) = (0)$, $1 \leq i \leq n$

(b) the natural maps $I_d \to H^0(\mathcal{I}(d))$ are isomorphisms for all $d \geq m$ and $H^i(\mathcal{I}(d)) = (0)$ if $d + i \geq m$, $i \geq 1$.

(c) Take a minimal resolution of I by free graded $k[X]$-modules:

$$0 \to \bigoplus_{\alpha=1}^{r_n} k[\mathbf{x}] \cdot e_{\alpha,n} \xrightarrow{\phi_n} \ldots \xrightarrow{\phi_1} \bigoplus_{\alpha=1}^{r_0} k[\mathbf{x}] \cdot e_{\alpha,0} \xrightarrow{\phi_0} k[\mathbf{x}] \longrightarrow k[\mathbf{x}]/I \to 0.$$

Then $\deg(e_{\alpha,i}) \leq m + i$ for all α, i. (In particular, if $f_\alpha = \phi_0(e_{\alpha,0})$, then $f_1, \ldots, f_{r_0}$ are minimal generators of I, and $\deg(e_{\alpha,0}) = \deg(f_\alpha) \leq m$.)

The intuitive idea is that past degree m, nothing tricky happens in the ideal I. Unfortunately, neither (a), (b) nor (c) can be verified by any obvious finite algorithm. This lack of a finitely verifiable criterion for m-regularity has been remedied by a joint result of the first author and M. Stillman [BS87a]:

[1] The definition has been slightly modified so as to apply to ideals I instead of the corresponding sheaf of ideals $\mathcal{I}$.

Theorem 3.3 (Bayer-Stillman) *I is m-regular if and only if the degrees of the minimal set of generators of I are at most m, and there exists a set $y_0, \ldots, y_\ell$ of linear combinations of $x_0, ..., x_n$ such that for all homogeneous f of degree m,*

$$\begin{aligned} y_0 f \in I &\Rightarrow f \in I \\ y_1 f \in I &\Rightarrow f \in I + k[\mathbf{x}] \cdot y_0 \\ &\cdots \\ y_\ell f \in I &\Rightarrow f \in I + \sum_{i=0}^{\ell-1} k[\mathbf{x}] \cdot y_i \end{aligned}$$

and

$$f \in I + \sum_{i=0}^{\ell} k[\mathbf{x}] \cdot y_i.$$

Moreover, if this holds at all, it holds for $y_0, \ldots, y_\ell$ taken arbitrarily from a Zariski-open set in the space of $\ell + 1$ linear forms.

To see why m-regularity is a key bound, we want to show that it controls some of the geometric features of the ideal I. Let's introduce several refined notions of the "degree" of I:

Definition 3.4 *If $I = \mathbf{q}_0 \cap \mathbf{q}_1 \cap \ldots \cap \mathbf{q}_t$ is a primary decomposition of I, $\sqrt{\mathbf{q}_i} = \mathbf{p}_i$ is prime and $V(\mathbf{p}_i)$ is the subvariety Z_i of $\mathbf{P}^n$ for $i \geq 1$, while $\mathbf{p}_0 = (x_0, \ldots, x_n)$ (so that $V(\mathbf{p}_0) = \emptyset$), then first let $\mathbf{q}_1, \ldots, \mathbf{q}_s$ be the isolated components, (i.e., $Z_i \not\subset Z_j$ if $1 \leq i \leq s$, $1 \leq j \leq t$, $i \neq j$, or equivalently, $V(I) = Z_1 \cup \ldots \cup Z_s$ is set-theoretically the minimal decomposition of $V(I)$ into varieties). Then let*

$$\begin{aligned} \text{mult}(\mathbf{q}_i) &= \textit{length } \ell \textit{ of a maximal chain of } \mathbf{p}_i\textit{-primary ideals:} \\ &\quad \mathbf{q}_i = J_\ell \subsetneq J_{\ell-1} \subsetneq \ldots \subsetneq J_1 = \mathbf{p}_i \end{aligned}$$

(Equivalently, this is the length of the local ring $k[\mathbf{x}]_{\mathbf{p}_i}/Ik[\mathbf{x}]_{\mathbf{p}_i}$, or, in the language of schemes, if η is the generic point of Z_i, then this is the length of $\mathcal{O}_{\eta,\mathbf{P}^n}$.)

$$\deg(Z_i) = \textit{usual geometric degree of } Z_i\textit{:}$$

$$\begin{aligned}
& \textit{the cardinality of } Z_i \cap L \textit{ for almost all} \\
& \textit{linear spaces } L \textit{ of complementary dimension.} \\
\text{geom-deg}_r(I) \;\; &= \sum_{\substack{i \text{ such that } \dim Z_i = r \\ 1 \le i \le s}} \text{mult}(\mathbf{q}_i) \deg(Z_i)
\end{aligned}$$

If $\mathbf{q}_i$ is one of the non-isolated, or embedded components, then we extend the concept of multiplicity more carefully: Let

$$I_i = \left\{ \cap \mathbf{q}_j \Big| j \text{ such that } \mathbf{p}_j \subsetneqq \mathbf{p}_i \text{ or equivalently } Z_j \supsetneqq Z_i \right\} \cap \mathbf{p}_i$$

and

$$\begin{aligned}
\text{mult}_I(\mathbf{q}_i) \;\; &= \; \text{length } \ell \text{ of a maximal chain of ideals:} \\
& \quad \mathbf{q}_i \cap I_i = J_\ell \subsetneqq J_{\ell-1} \subsetneqq \ldots \subsetneqq J_0 = I_i \\
& \quad \text{where each } J_k \text{ satisfies: } ab \in J_k, a \notin \mathbf{p}_i \Rightarrow b \in J_k.
\end{aligned}$$

(Equivalently, J_k equals $\mathbf{q}_k \cap I_i$ for some $\mathbf{p}_i$-primary ideal $\mathbf{q}_k$.) In particular:

$$\begin{aligned}
I_0 \;\; &= \; \bigcap_{j=1}^{t} \mathbf{q}_j \text{ is known as } I^{\text{sat}}, \text{ and} \\
\text{mult}_I(\mathbf{q}_0) \;\; &= \; \text{length } \ell \text{ of a maximal chain of ideals} \\
& \quad I = J_\ell \subsetneqq J_{\ell-1} \subsetneqq \ldots \subsetneqq J_0 = I^{\text{sat}} \\
& = \; \dim_k(I^{\text{sat}}/I).
\end{aligned}$$

For $s+1 \le i \le t$, an equivalent way to define $\text{mult}_I(\mathbf{q}_i)$ is as the length of the module

$$I_i k[\mathbf{x}]_{\mathbf{p}_i} / I k[\mathbf{x}]_{\mathbf{p}_i}$$

or, in the language of schemes, the length of

$$I_i \mathcal{O}_{\eta, \mathbf{P}^n} / I \mathcal{O}_{\eta, \mathbf{P}^n}$$

where η is the generic point of Z_i.

Then write

$$\text{arith-deg}_r(I) = \sum_{\substack{i \text{ such that } \dim Z_i = r \\ 1 \le i \le s}} \text{mult}_I(\mathbf{q}_i) \deg(Z_i)$$

and

$$\text{arith-deg}_{-1}(I) = \text{mult}_I(\mathbf{q}_0).$$

The idea here is best illustrated by an example: let

$$I = (x_1^2, x_1x_2) \subset k[x_0, x_1, x_2].$$

Then

$$\begin{aligned} I &= \mathbf{q}_1 \cap \mathbf{q}_2 \\ q_1 &= (x_1),\ \mathbf{p}_1 = (x_1),\ Z_1 = \{\text{line } x_1 = 0\} \\ q_2 &= (x_1^2, x_1x_2, x_2^2),\ \mathbf{p}_2 = (x_1, x_2),\ Z_2 = \{\text{point } (1,0,0)\}. \end{aligned}$$

Then

$$\deg(Z_1) = 1,\ \text{mult}(q_1) = 1$$

so

$$\text{geom-deg}_1(I) = \text{arith-deg}_1(I) = 1.$$

One might be tempted to simply define

$$\text{mult}_I(\mathbf{q}_2) = \text{ length of chain of } \mathbf{p}_2\text{-primary ideals between } \mathbf{q}_2, \mathbf{p}_2$$

and since

$$k[\mathbf{x}]_{\mathbf{q}_2}/\mathbf{q}_2k[\mathbf{x}]_{\mathbf{p}_2} \cong K \cdot 1 + K \cdot x_1 + K \cdot x_2, K = k(x_0)$$

this is 3. But embedded components are not unique! In fact,

$$\begin{aligned} I &= \mathbf{q}_1 \cap \mathbf{q}_2' \\ \mathbf{q}_2' &= (x_1^2x_2) \text{ also,} \end{aligned}$$

which leads to

$$k[\mathbf{x}]_{\mathbf{p}_2}/\mathbf{q}_2'k[\mathbf{x}]_{\mathbf{p}_2} \cong K \cdot 1 + K \cdot x_2$$

which has length 2. The canonical object is not the local ring $k[\mathbf{x}]_{\mathbf{p}_2}/\mathbf{q}_2k[\mathbf{x}]\mathbf{p}_2$ but the ideal

$$\text{Ker}\,(k[\mathbf{x}]_{\mathbf{p}_2}/Ik[\mathbf{x}]_{\mathbf{p}_2} \to k[\mathbf{x}]_{\mathbf{p}_2}/\mathbf{p}_2k[\mathbf{x}]_{\mathbf{p}_2}) \cong k \cdot x_1$$

which has length 1. Thus, the correct numbers are

$$\text{mult}_I(\mathbf{q}_2) = 1$$

and

$$\begin{aligned} \text{geom-deg}_0(I) &= 0 \\ \text{arith-deg}_0(I) &= 1. \end{aligned}$$

Now the question arises: find bounds on these degrees in terms of generators of I. For geometric degrees, a straightforward extension of Bezout's theorem gives:

Proposition 3.5 *Let $d(I)$ be the maximum of the degrees of a minimal set of generators of I. Then*

$$\text{geom-deg}_r(I) \leq d(I)^{n-r}.$$

A proof can be found in [MW83]. The idea is clear from a simple case: Suppose $f, g, h \in K[x, y, z]$ and $f = g = h = 0$ consists of a curve C and ℓ points P_i off C. We can bound ℓ like this: Choose 2 generic combinations f', g' of f, g, h so that $f' = g' = 0$ does not contain a surface. It must be of the form $C \cup C'$, C' one-dimensional, containing all the P_i but not the generic point of C. Then by the usual Bezout theorem

$$\deg C' \leq \deg f' \ \deg g' = d(I)^2.$$

Let h' be a 3^{rd} generic combination of f, g, h. Then $C' \cap \{h' = 0\}$ consists of a finite set of points including the P_i's. Thus

$$\begin{aligned} \ell &= \#P_i \leq \#(C' \cap \{h' = 0\}) \\ &\leq \deg C' \cdot d(I) \text{ by Bezout's theorem} \\ &\leq d(I)^3. \end{aligned}$$

Can arith-deg(I) be bounded in the same way? In fact, it cannot, as we will show below. Instead, we have

Proposition 3.6 *If $m(I)$ is the regularity of I, then for $-1 \leq r \leq n$,*

$$\text{arith-deg}_r(I) \leq \binom{m(I)+n-r-1}{n-r} \leq m(I)^{n-r}$$

which replaces $d(I)$ by the regularity of I. A proof is given in the technical appendix.

We have introduced two measures of the complexity of a homogeneous ideal I. The first is $d(I)$, the maximum degree of a polynomial in a minimum set of generators of I. The second is $m(I)$, which bounds the degrees of generators and of all higher order syzygies in the resolution of I (Definition 3.2 (c)). Obviously,

$$d(I) \leq m(I).$$

A very important question is how much bigger can $m(I)$ be than $d(I)$? The nature of the answer was conjectured by one of us in his thesis [Bay82] and this conjecture is being borne out by subsequent investigations. This conjecture is that in the worst case $m(I)$ is roughly the $(2^n)^{\text{th}}$ power of $d(I)$ – a bound like G. Hermann's. But that if $I = I(Z)$ where Z is geometrically nice, e.g. is a smooth irreducible variety, then $m(I)$ is much smaller, like the n^{th} power of $d(I)$ or better. This conjecture then has three aspects:

(1) a doubly exponential bound for $m(I)$ in terms of $d(I)$,
which is always valid,
(2) examples of I where the bound in (1) is best possible, or nearly so,
(3) much better bounds for $m(I)$
valid if $V(I)$ satisfies various conditions.

All three aspects are partially proven, but none are completely clarified yet. We will take them up one at a time.

A doubly exponential bound for $m(I)$ in terms of $d(I)$ may be deduced easily *in characteristic zero* from the work of M. Giusti [Giu84] and A. Galligo [Gal79]:

Theorem 3.7 *If $\text{char}(k) = 0$ and $I \subset k[x_0, \ldots x_n]$ is any homogeneous ideal, then*

$$m(I) \leq (2d(I))^{2^{n-1}}.$$

It seems likely that Theorem 3.7 holds in characteristic p, too. A weaker result can be derived quickly in any characteristic by straightforward cohomological methods:

Proposition 3.8 *If $I \subset k[x_0, \ldots x_n]$ is any homogeneous ideal, then*

$$m(I) \leq (2d(I))^{n!}.$$

The proof is given in the technical appendix.

Next, we ask whether Theorem 3.7 is the best possible, or nearly so. The answer is yes, because of a very remarkable example due to E. Mayr and A. Meyer [MM82].

Example 3.9 Let I_n^A be the ideal in $10n$ variables $S^{(m)}, F^{(m)}, C_i^{(m)}, B_i^{(m)}, 1 \leq i \leq 4, 1 \leq m \leq n$ defined by the $10n - 6$ generators

$$2 \leq m \leq n \quad \begin{cases} S^{(m)} - S^{(m-1)} C_1^{(m-1)} \\ F^{(m)} - S^{(m-1)} C_4^{(m-1)} \\ C_i^{(m)} F^{(m-1)} B_2^{(m-1)} - C_i^{(m)} B_i^{(m)} F^{(m-1)} B_3^{(m-1)}, \quad 1 \leq i \leq 4 \end{cases}$$

$$1 \leq m \leq n-1 \quad \begin{cases} F^{(m)} C_1^{(m)} B_1^{(m)} - S^{(m)} C_2^{(m)} \\ F^{(m)} C_2^{(m)} - F^{(m)} C_3^{(m)} \\ S^{(m)} C_3^{(m)} B_1^{(m)} - S^{(m)} C_2^{(m)} B_4^{(m)} \\ S^{(m)} C_3^{(m)} - F^{(m)} C_4^{(m)} B_4^{(m)} \end{cases}$$

$$C_i^{(1)} S^{(1)} - C_i^{(1)} F^{(1)} (B_i^{(1)})^2, \quad 1 \leq i \leq 4$$

Let I_n^H be the ideal gotten from I_n^A by homogenizing with an extra variable u. Then Mayr and Meyer [MM82, lemma 8, p. 318] prove:

Lemma 3.10 *Let $e_n = 2^{2^n}$. If M is any monomial in these variables, $S^{(n)} C_i^{(n)} - F^{(n)} M \in I_n^A$ if and only if*

$$M = C_i^{(n)} (B_i^{(n)})^{e_n},$$

and $S^{(n)} C_i^{(n)} - S^{(n)} M \in I_n^A$ if and only if

$$M = C_i^{(n)}.$$

Now note that the generators of I_n^A and I_n^H are all of the very simple type given by a difference of two monomials. Quite generally, if

$$\begin{aligned} J &\subset k[x_1, \ldots, x_n] \\ J &= (\ldots, \mathbf{x}^{\alpha_i} - \mathbf{x}^{\beta_i}, \ldots)_{1 \le i \le k} \end{aligned}$$

then the quotient ring $k[\mathbf{x}]/J$ has a very simple form. In fact, we get an equivalence relation between monomials generated by

$$\mathbf{x}^{\alpha_i + \gamma} \sim \mathbf{x}^{\beta_i + \gamma}, \text{ any } i, \gamma$$

and

$$k[\mathbf{x}]/J \cong \oplus_\delta \, k \cdot \mathbf{x}^\delta$$

where δ runs over a set of representatives of each equivalence class.

Bearing this in mind, let's look at the 1^{st} order syzygies for the homogeneous ideal:

$$J_n^H = (S^{(n)}, F^{(n)}, I_n^H).$$

$S^{(n)}$ and $F^{(n)}$ are part of a minimal set of generators, and let $f_\alpha \in I_n^H$ complete them. Then syzygies are equations:

$$p\, S^{(n)} + q\, F^{(n)} + \sum r_\alpha \, f_\alpha = 0.$$

One such is given by:

$$\left[u^{e_n + e}\, C_i^{(n)}\right] S^{(n)} + \left[-u^e \, (B_i^{(n)})^{e_n}\, C_i^{(n)}\right] F^{(n)} + \sum R_\alpha \, f_\alpha = 0$$

for some R_α, and some $e \ge 0$ (the extra power u^e is necessary because some terms $R_\alpha f_\alpha$ have degree greater than $e_n + 2$) whose degree is $2 + e_n + e$. Now express this syzygy as a combination of a minimal set of syzygies. This gives us in particular:

$$\begin{aligned} u^{e_n + e}\, C_i^{(n)} &= \sum a_\lambda \, p_\lambda \\ -u^e \, (B_i^{(n)})^{e_n}\, C_i^{(n)} &= \sum a_\lambda \, q_\lambda \\ p_\lambda \, S^{(n)} + q_\lambda \, F^{(n)} + \sum R_{\alpha\lambda} \, f_\alpha &= 0\,. \end{aligned}$$

Then for some λ, p_λ must have a term of the form u^ℓ or $u^\ell \, C_i^{(n)}$, hence the monomial $u_\ell \, S^{(n)}$ or $u_\ell \, C_i^{(n)} \, S^{(n)}$ occurs in $p_\lambda \, S^{(n)}$. But by the general

remark on quotient rings by such simple ideals, this means that this term must equal some second term $M\,S^{(n)}$ (M a monomial in p_λ) or $M\,F^{(n)}$ (M a monomial in q_λ) mod I_n^H. By the lemma, the first doesn't happen and the second only happens if the term $u^\ell\,C_i^{(n)}\,(B_i^{(n)})^{e_n}$ occurs in q_λ, in which case $e_n + 1 \leq \deg\, q_\lambda \,=\, \deg(\text{syzygy}(p_\lambda, q_\lambda, R_{\alpha\lambda})) - 1$. This proves:

Proposition 3.11 *J_n^H has for its bounds:*

$$\begin{aligned} d(J) &= 4 \\ m(J) &\geq 2^{2^n} + 1. \end{aligned}$$

Going on to the 3rd aspect of the conjecture, consider results giving better bounds for $m(I)$ under restrictive hypotheses on $V(I)$.

Theorem 3.12 *If $Z \subset \mathbf{P}^n$ is a reduced subscheme purely of dimension r, and $I = I(Z)$ is the full ideal of functions vanishing on Z, then*

(a) if $r \leq 1$, or Z is smooth, char$(k) = 0$ and $r \leq 3$, then:

$$m(I) \leq \deg Z - n + r + 1$$

(b) if char$(k) = 0$ and Z is smooth,

$$m(I) \leq (r+1)(\deg(Z) - 2) + 2.$$

Since $\deg(Z) \leq d(I)^{n-r}$ (Proposition 3.5), these bound $m(I)$ in terms of $d(I)$.

Part (a) of this are due to Gruson-Lazarsfeld-Peskine [GLP83] for $r \leq 1$, and to Pinkham [Pin86], Lazarsfeld [Laz87], and Ran [Ran90] for $r \leq 3$. It is *conjectured* by Eisenbud and Goto [EG84], and others, that the bound in (a) holds for all reduced irreducible Z, and it might well hold even for reduced equidimensional Z which are connected in codimension 1. As this problem is now understood, the needed cohomological arguments follow formally, once one can control the singularities of a projection of the variety. These singularities become progressively harder to subdue as the dimension of the variety increases, and are what impedes definitive progress beyond dimension 3.

Part (b) is due to the second author and is proven in the technical appendix. It has been generalized by Bertram, Ein, and Lazarsfeld [BEL91] to show that any smooth characteristic 0 variety of codimension e defined as a subscheme of $\mathbf{P}^n$ by hypersurfaces of degrees $d_1 \geq \ldots \geq d_m$ is $(d_1 + \ldots d_e - e + 1)$-regular. Since we cannot decide the previous conjecture, this is a result of considerable practical importance, for it strongly bounds the complexity of computing Gröbner bases of smooth characterisitic 0 varieties in terms of the degrees of the input equations.

The biggest missing link in this story is a decent bound on $m(I)$ for any reduced equidimensional ideal I. We would conjecture that if a linear bound as in part (a) doesn't hold, at the least a so-called "single exponential" bound, i.e. $m(I) \leq d^{0(n)}$ ought to hold. This is an essential ingredient in analyzing the worst-case behavior of all algorithms based on Gröbner bases, and would complete the story about what causes the bad examples discussed above. At least in some cases Ravi [Rav90] has proven that the regularity of the radical of a scheme is no greater than the regularity of the scheme itself.

There is a direct link between the bounds that we have given so far and the G. Hermann bound with which we started the section. This results from the following:

Proposition 3.13 *Let $I^A \subset k[x_1, \ldots, x_n]$ have generators $f_1, \ldots, f_k$ and let $I^H \subset k[x_0, x_1, \ldots, x_n]$ be the ideal generated by homogenizations $f_1^h, \ldots, f_k^h$ of the f_i. Let $I^H = \mathbf{q}_0 \cap \ldots \cap \mathbf{q}_t$ be the primary decomposition of I^H, let $Z_i = V(\mathbf{q}_i)$ and let*

$$\mathrm{mult}_\infty(I^H) = \max\left[\mathrm{mult}_I(\mathbf{q}_{i_1}) + \ldots + \mathrm{mult}_I(\mathbf{q}_{i_k}) + \mathrm{mult}_I(\mathbf{q}_0)\right]$$

where the max *is taken over chains $V((x_0)) \supset Z_{i_1} \supsetneq \ldots \supsetneq Z_{i_k}$. If $g \in I^A$, then we can write:*

$$g = \sum_{i=1}^{k} a_i f_i$$

where

$$\deg a_i \leq \deg g + \mathrm{mult}_\infty(I^H).$$

The proof goes like this: Let g^h be the homogenization of g. Consider the least integer m such that $x_0^m g^h \in I^H$. Since $g \in I$, this m is finite. Moreover,

if

$$x_0^m g^h = \sum x_0^{m_i} a_i^h f_i^h$$

then

$$g = \sum a_i f_i$$

and

$$\deg a_i = \deg(a^h) \leq \deg(x_0^m g^h) - \deg f_j \leq m + \deg(g).$$

Now in the primary decomposition of I^H, suppose that for some k,

$$x_0^k g^h \in \bigcap_{i \in S} q_i, \text{ and } x_0^k g^h \notin \mathbf{q}_j \text{ if } j \notin S.$$

Choose $\ell \notin S$ such that $V(q_\ell)$ is maximal. Since $g \in I^A$, we know $V(\mathbf{q}_\ell) \subset V((x_0))$, hence $x_0 \in \mathbf{p}_\ell$. Let

$$I_S = \bigcap_{i \in S} \mathbf{q}_i.$$

Then $\text{mult}_I(\mathbf{q}_\ell)$ is easily seen to be the length of a maximal chain of ideals between:

$$I \cdot k[\mathbf{x}]_{p_\ell} \text{ and } I_S \cdot k[\mathbf{x}]_{p_\ell}.$$

But look at the ideals J_p, for $p \geq 0$, defined by

$$I\, k[\mathbf{x}]_{\mathbf{p}_\ell} \subset \underbrace{(I, x_0^{k+p} g^h)\, k[\mathbf{x}]_{\mathbf{p}_\ell}}_{J_p} \subset I_S\, k[\mathbf{x}]_{\mathbf{p}_\ell}.$$

If $J_p = J_{p+1}$, then

$$\begin{array}{rcl} x_0^{k+p} g^h & \in & (I, x_0^{k+p+1} g^h) \\ \text{i.e., } x_0^{k+p} g^h & = & a\, x_0^{k+p+1} g^h + b, \quad b \in I. \end{array}$$

But $1 - ax_0$ is a unit in $k[\mathbf{x}]_{\mathbf{p}_\ell}$, so

$$J_p = x_0^{k+p} g^h = (1 - ax_0)^{-1} b \in I\, k[\mathbf{x}]_{\mathbf{p}_\ell}.$$

This means that in any case

$$x_0^{k+\text{mult}_I(\mathbf{q}_\ell)} g^h \in I \cdot k[\mathbf{x}]_{\mathbf{p}_\ell}$$

hence, because q_ℓ is p_ℓ-primary:

$$x_0^{k+\mathrm{mult}_I(\mathbf{q}_\ell)} g^h \in \mathbf{q}_\ell$$

Induction now shows that

$$x_0^{\mathrm{mult}_\infty(I^H)} g^h \in I^H$$

Corollary 3.14 *Let I^A, I^H be as above. If $g \in I^A$, then*

$$g = \sum a_i f_i$$

where $\deg(a_i) \leq \deg(g) + \binom{m(I)+n+1}{n+1}$.

Proof. Combine Propositions 3.6 and 3.13.

If we further estimate $m(I)$ by Theorem 3.7 in characteristic 0 or by Proposition 3.8, we get somewhat weaker versions of Hermann's Theorem 3.1. But if $I = V(Z)$, Z a good variety, we may expect the Corollary to give much better bounds than Theorem 3.1.

Corollary 3.14 shows that any example which demonstrates the necessity of double exponential growth in Hermann's ideal membership bound (Theorem 3.1) also demonstrates the necessity of double exponential growth in the bounds on $m(I)$ given in Theorem 3.7 and Proposition 3.8. Thus we can make use of the general arguments for the existence of such examples given in [MM82], rather than depending on the single example of Proposition 3.11, to show that the bounds on $m(I)$ inevitably grow double exponentially: Since in Corollary 3.14, the degrees of the a_i are bounded by a single exponential function of $m(I)$, in all examples where the degrees of the a_i grow double exponentially, $m(I)$ also grows double exponentially.

This line of argument gives a geometric link between the ideal membership problem and $m(I)$: In Corollary 3.14, if I^A exhibits a_i of high degree, then I^H has primary components of high multiplicity. These components force $m(I)$ to be large, and distinguish I^H from good ideals considered in Theorem 3.12 and related conjectures.

A major step in understanding the gap between the double exponential examples and the strong linear bounds on the regularity of many smooth

varieties was taken by Brownawell [Bro87] and Kollár [Kol88]. They discovered the beautiful and satisfying fact that if we replace membership in I by membership in $\sqrt{I}$, then there are single exponential bounds on the degrees of a_i:

Theorem 3.15 (Brownawell, Kollár) *Let k be any field, let $I = (f_1, ..., f_k) \subset k[x_1, ..., x_n]$ and let $d = \max(\deg(f_i), i = 1, \cdots, k; 3)$. If $n = 1$, replace d by $2d - 1$. If $g \in \sqrt{I}$, then there is an expression*

$$g^s = \sum_{i=1}^{k} a_i f_i$$

where $s \leq d^n$ and $\deg(a_i) \leq (1 + \deg(g))d^n$. In particular:

$$\left(\sqrt{I}\right)^{d^n} \subset I.$$

What this shows is that although the bad examples have to have primary components at infinity of high degree, nonetheless these primary ideals contain relatively small powers of $\sqrt{I^H}$. The picture you should have is that these embedded components at infinity are like strands of ivy that creep a long way out from the hyperplane at infinity, but only by clinging rather closely to the affine components.

Technical Appendix to Section 3

1. Proof of the equivalence of the conditions in Definition 3.2:

In [Mum66, pp. 99-101], it is proven that for any coherent sheaf $\mathcal{F}$ on $\mathbf{P}^n$, $H^i(\mathcal{F}(-i)) = (0)$, $i \geq 1$ implies that the same holds for $\mathcal{F}(d)$, all $d \geq 0$, and that $H^0(\mathcal{F}(d))$ is generated by $H^0(\mathcal{F}) \otimes H^0(\mathcal{O}(d))$. In particular, if you apply this to $\mathcal{F} = \mathcal{I}(m)$, the equivalence of (a) and (b) follows. (Note the diagram:

$$\begin{array}{ccc} I_d & \longrightarrow & H^0(\mathcal{I}(d)) \\ \bigcap & & \bigcap \\ k[\mathbf{x}]_d & \longrightarrow & H^0(\mathcal{O}_{\mathbf{P}^n}(d)) \end{array}$$

which shows that $I_m \to H^0(\mathcal{I}(m))$ is injective for every d). To show that (b) $\Rightarrow$ (c), first note that we may rephrase the reults in [Mum66] to say that if

$H^i(\mathcal{F}(-i)) = (0)$, $i \geq 1$, then the degrees of the minimal generators of the $k[\mathbf{x}]$-module

$$\bigoplus_{d \in \mathbf{Z}} H^0(\mathcal{F}(d))$$

are all zero or less. So we may construct the resolution in (c) inductively: at the k^{th} stage, say

$$\bigoplus_{\alpha=1}^{r_k+1} k[\mathbf{x}] \cdot e_{\alpha,k-1} \xrightarrow{\phi_{k-1}} \cdots \longrightarrow k[\mathbf{x}] \longrightarrow k[\mathbf{x}]/I \longrightarrow 0$$

has been constructed, let $M_k = \ker(\phi_{k-\ell})$ and let $\mathcal{F}_k$ be the corresponding sheaf of ideals. The induction hypothesis will say that $H^i(\mathcal{F}_k(m+k-1)) = (0)$, $i \geq 1$. Therefore M_k is generated by elements of degree $\leq m+k$, i.e., $d_\alpha = \deg e_{\alpha,k} \leq m+k$, all α. We get an exact sequence

$$0 \longrightarrow M_{k+1} \longrightarrow \bigoplus_{\alpha=1}^{r_k} (k[\mathbf{x}] \cdot e_{\alpha,k}) \longrightarrow M_k \longrightarrow 0$$

hence

$$0 \longrightarrow \mathcal{F}_{k+1} \longrightarrow \bigoplus_{\alpha=1}^{r_k} \mathcal{O}_{\mathbf{P}^n}(-d_\alpha) \longrightarrow \mathcal{F}_k \longrightarrow 0 \tag{1}$$

Therefore

$$\bigoplus_{\alpha=1}^{r_k} H^i(\mathcal{O}_{\mathbf{P}^n}(m+k-i-d_\alpha)) \longrightarrow H^i(\mathcal{F}_k(m+k-i)) \longrightarrow \tag{2}$$
$$H^{i+1}(\mathcal{F}_{k+1}(m+(k+1)-(i+1))) \longrightarrow \bigoplus_{\alpha=1}^{r_k} H^{i+1}(\mathcal{O}_{\mathbf{P}^n}(m+k-i-d_\alpha))$$

is exact. But $m+k-i-d_\alpha \geq -i$ so $H^{i+1}(\mathcal{O}_{\mathbf{P}^n}(m+k-i-d_\alpha)) = (0)$. This shows that $\mathcal{F}_{k+1}$ satisfies the induction hypothesis and we can continue. Thus (c) holds. To see that (c) $\Rightarrow$ (a), we just use the same exact sequences (1) and prove now by descending induction on k that $H^i(\mathcal{F}_k(m+k-i)) = (0)$, $i \geq 1$. Since $I = \mathcal{F}_0$, this does it. The inductive step again uses (2), since $H^i(\mathcal{O}_{\mathbf{P}^n}(m+k-i-d_\alpha)) = (0)$ too.

2. Proof of Proposition 3.6:

Look first at the case $r = 0$. Let $\mathcal{I}$ be the sheaf of ideals defined by I and let $\mathcal{I}^* \supset \mathcal{I}$ be the sheaf defined by omitting all 0-dimensional primary components of I. Consider the exact sequence:

$$0 \longrightarrow \mathcal{I}(m-1) \longrightarrow \mathcal{I}^*(m-1) \longrightarrow (\mathcal{I}^*/\mathcal{I})(m-1) \longrightarrow 0$$

This gives us:

$$H^0(\mathcal{I}^*(m-1)) \longrightarrow H^0((\mathcal{I}^*/\mathcal{I})(m-1)) \longrightarrow H^1(\mathcal{I}(m-1))$$

Now $H^1(\mathcal{I}(m-1)) = (0)$ by m-regularity, and $h^0((\mathcal{I}^*/\mathcal{I})(m-1)) = h^0(\mathcal{I}^*/\mathcal{I}) = \text{length}(\mathcal{I}^*/\mathcal{I}) = \text{arith-deg}_0(I)$ since $\mathcal{I}^*/\mathcal{I}$ has 0-dimensional support. But $H^0(\mathcal{I}^*(m-1)) \subset H^0(\mathcal{O}_{\mathbf{P}^n}(m-1))$, so

$$\begin{aligned} \text{arith-deg}_0(I) &\leq h^0(\mathcal{I}^*(m-1)) \\ &\leq h^0(\mathcal{O}_{\mathbf{P}^n}(m-1)) \\ &= \binom{m+n-1}{n} \end{aligned}$$

If $r > 0$, we can prove the Proposition by induction on r. Let H be a generic hyperplance in $\mathbf{P}^n$, given by $h = 0$. Let $I_H = (I,h)/(h) \subset k[x_0,\ldots,x_n]/(h) \cong k[x'_0,\ldots,x'_{n-1}]$ for suitable linear combinations x'_i of x_i. Then it is easy to check that:

$$\text{arith-deg}_r(I) = \text{arith-deg}_{r-1}(I_H)$$

and that I_H is also m-regular, so by induction

$$\begin{aligned} \text{arith-deg}_{r-1}(I_H) &\leq \binom{m+(n-1)-(r-1)-1}{(n-1)-(r-1)} \\ &= \binom{m+n-r-1}{n-r} \end{aligned}$$

If $r = -1$, we use the fact that

$$0 \longrightarrow I_d \longrightarrow H^0(\mathcal{I}(d)) \xleftarrow{\approx} (I^{\text{sat}})_d$$

if $d \geq m$, hence

$$\dim(I^{\text{sat}}/I) \leq \dim k[\mathbf{x}]/(x_0,\ldots,x_n)^m = \binom{m+n}{n+1}.$$

3. Proof of Proposition 3.8:

Let $I \subset k[x_0, \ldots x_n]$ and assume, after a linear change of coordinates, that x_n is not contained in any associated prime ideals of I. Let $\overline{I} \subset k[x_0, \ldots x_{n-1}]$ be the image of I. Then $d(\overline{I}) = d(I)$ and by induction we may assume

$$m(\overline{I}) \leq (2d(I))^{(n-1)!}.$$

We will prove, in fact, that

$$m(I) \leq m(\overline{I}) + \binom{m(\overline{I}) - 1 + n}{n} \tag{3}$$

and then we will be done by virtue of the elementary estimate:

$$\text{if } m^* = (2d(I))^{(n-1)!}, \text{ and } d \geq 2, \text{ then } m^* + \binom{m^* - 1 + n}{n} \leq (2d(I))^{n!}$$

To prove (3), we use the long exact sequence

$$\begin{array}{ccccccccc} 0 & \longrightarrow & (I : (x_0))_{k-1} & \xrightarrow{x_0} & I_k & \longrightarrow & \overline{I}_k & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & H^0(\mathcal{I}(k-1)) & \longrightarrow & H^0(\mathcal{I}(k-1)) & \longrightarrow & H^0(\overline{\mathcal{I}}(k-1)) & \xrightarrow{\delta} & \end{array}$$

$$\xrightarrow{\delta} \quad H^1(\mathcal{I}(k-1)) \quad \longrightarrow \quad H^1(\mathcal{I}(k)) \quad \longrightarrow \quad H^1(\overline{\mathcal{I}}(k))$$

where $(I : (x_0)) = \{\, f \,|\, x_0 f \in I \,\}$. Let $\overline{m} = m(\overline{I})$. Note that $H^i(\overline{\mathcal{I}}(k-1)) = (0)$, $i \geq 1$, $k \geq \overline{m}$, hence

$$H^i(\mathcal{I}(k-1)) \to H^i(\mathcal{I}(k))$$

is an isomorphism if $k \geq \overline{m} - 1 + 1$ and $i \geq 2$. Since $H^i(\mathcal{I}(k)) = (0)$, $k \gg 0$, this shows that $H^i(\mathcal{I}(k)) = (0)$, $i \geq 2$, $k \geq \overline{m} - i$. Moreover $\overline{I}_k \to H^0(\overline{\mathcal{I}}(k))$ is an isomorphism if $k \geq \overline{m}$, hence $\delta = 0$ if $k \geq \overline{m}$, hence $H^1(\mathcal{I}(k)) = (0)$, $k \geq \overline{m} - 1$. But now look at the surjectivity of $I_k \to H^0(\mathcal{I}(k))$. For all k, let M_k be the cokernel. Then $\oplus M_k$ is a $k[\mathbf{x}]$-module of finite dimension. Multiplication by x_0 induces a sequence:

$$0 \longrightarrow \frac{(I : (x_0))_{k-1}}{I_{k-1}} \longrightarrow M_{k-1} \xrightarrow{x_0} M_k \longrightarrow 0$$

which is exact if $k \geq \overline{m}$. But if, for one value of $k \geq \overline{m}$,

$$(I : (x_0))_k = I_k \tag{4}$$

then by Theorem 3.3, I is k-regular and (4) continues to hold for larger k, and M_k must be (0). In other words,

$$\dim M_k, \quad k \ \geq \ \overline{m} - 1$$

is non increasing and monotone decreasing to zero when $k \geq \overline{m}$. Therefore

$$\begin{aligned} m(I) &\leq \overline{m} + \dim M_{\overline{m}-1} \\ &\leq \overline{m} + \dim k[\mathbf{x}]_{\overline{m}-1} \\ &\leq \overline{m} + \binom{\overline{m} - 1 + n}{n} \end{aligned}$$

which proves (3).

4. Proof of Theorem 3.12(b):

Let Z be a smooth r-dimensional subvariety of $\mathbf{P}^n$ and $d =$ degree of Z. We first consider linear projections of Z to $\mathbf{P}^r$ and to $\mathbf{P}^{r-1}$. To get there, let $L_1 \subset \mathbf{P}^n$ be a linear subspace of dimension $n-r-1$ disjoint from Z and $L_2 \subset L_1$ a linear subspace of dimension $n-r-2$. Take these as centers of projection:

$$\begin{array}{ccccc} \mathbf{P}^n - L_1 & \supset & Z & \xrightarrow{p_2} \\ \downarrow & & \downarrow p_1 & \\ \mathbf{P}^{r+1} - \{P\} & \supset & Z_1 & \\ \downarrow & & & \\ \xrightarrow{p_2} \ \mathbf{P}^r & & & \end{array}$$

Let $x_0, \ldots x_{r+1}$ be coordinates on $\mathbf{P}^{r+1}$ so that $p = (0, \ldots, 0, 1)$, hence $x_0, \ldots x_r$ are coordinates on $\mathbf{P}^r$. Let $f(x_0, \ldots x_{r+1}) = 0$ be the equation of the hypersurface Z_1.

Now there are two ways of getting r-forms on Z: by pullback of r-forms on $\mathbf{P}^r$ and by residues of $(r+1)$-forms on $\mathbf{P}^{r-1}$ with simple poles along Z_1. The first gives us a sheaf map

$$p_2^* \, \Omega^r_{\mathbf{P}^r} \hookrightarrow \Omega^r_Z$$

whose image is $\Omega^r_Z(-B_1)$, B_1 the branch locus of p_2. Corresponding to this on divisor classes:

$$\begin{aligned} K_Z &\equiv p_2^*(K_{\mathbf{P}^r}) + B_1 \\ &\equiv -(r+1)H + B_1, \end{aligned} \tag{5}$$

where $H =$ hyperplane divisor class on Z. The second is defined by

$$a(\mathbf{x}) \cdot \frac{dx_1 \wedge \ldots \wedge dx_{r+1}}{f} \longmapsto p_1^* \left(a(\mathbf{x}) \cdot \frac{dx_1 \wedge \ldots \wedge dx_r}{\partial f / \partial x_{r+1}} \right) \tag{6}$$

and it gives us an isomorphism

$$p_1^*(\Omega^{r+1}_{\mathbf{P}^{r+1}}(Z_1)|_{Z_1}) \cong \Omega^r_Z(B_2)$$

B_2 is a divisor which can be interpreted as the *conductor* of the affine rings of Z over those of Z_1: i.e.,

$$f \in \mathcal{O}_Z(-B_2) \iff f \cdot (p_{1,*}\mathcal{O}_Z) \subset \mathcal{O}_{Z_1}.$$

In particular,

$$p_{1,*}(\mathcal{O}_Z(-B_2)) \cong \text{ sheaf of } \mathcal{O}_{Z_1} - \text{ ideals } C \text{ in } \mathcal{O}_{Z_1}. \tag{7}$$

A classical reference for these basic facts is Zariski [Zar69], Prop. 12.13 and Theorem 15.3. A modern reference is Lipman [Lip84] (apply Def. (2.1)b to p_1 and apply Cor. (13.6) to $Z_1 \subset \mathbf{P}^{r+1}$). (4) gives us the divisor class identity:

$$\begin{aligned} K_Z + B_2 &\equiv p_1^*(K_{\mathbf{P}^{r+1}} + Z_1) \\ &\equiv (d-r-2)H. \end{aligned} \tag{8}$$

(5) and (8) together tell us that

$$B_1 + B_2 \equiv (d-1)H.$$

In fact, the explicit description (6) of the residue tells us more: namely that if $y_1, \ldots, y_r$ are local coordinates on Z, then

$$\frac{\partial(x_1, \ldots, x_r)}{\partial(y_1, \ldots, y_r)} \cdot \frac{1}{\partial f / \partial x_{r+1}} \, dy_1 \wedge \ldots \wedge dy_r$$

generates $\Omega^r_Z(B_2)$ locally. But $\frac{\partial(x_1,\ldots,x_r)}{\partial(y_1,\ldots,y_r)} = 0$ is a local equation for B_1, so this means that $\partial f/\partial x_{r+1} = 0$ is a local equation for $B_1 + B_2$. But $\partial f/\partial x_{r+1} = 0$ is a global hypersurface of degree $d-1$ in $\mathbf{P}^{r+1}$, hence globally:

$$B_1 + B_2 = p_1^*(V(\frac{\partial f}{\partial x_{r_1}}))$$

(equality of divisors, not merely divisor classes). All this is standard classical material.

(7) has an important cohomological consequence: let $C^* \subset \mathcal{O}_{\mathbf{P}^{r+1}}$ be the sheaf of ideals consisting of functions whose restriction to Z_1 lies in C. Then we get an exact sequence:

$$0 \to \mathcal{O}_{\mathbf{P}^{r+1}}(-Z_1) \to C^*\,\mathcal{O}_{\mathbf{P}^{r+1}} \to C\,\mathcal{O}_{Z_1} \to 0$$

hence an exact sequence

$$0 \to \mathcal{O}_{\mathbf{P}^{r+1}}(\ell - d) \to C^*\mathcal{O}_{\mathbf{P}^{r+1}}(\ell) \to p_{1,*}(\mathcal{O}_Z(\ell H - B_2)) \to 0$$

for all integers ℓ. But $H^1(\mathcal{O}_{\mathbf{P}^{r+1}}(\ell - d)) = (0)$, hence

$$H^0(C^*\mathcal{O}_{\mathbf{P}^{r+1}}(\ell)) \to H^0(\mathcal{O}_Z(\ell H - B_2))$$

is surjective, hence

$$H^0(\mathcal{O}_Z(\ell H - B_2)) \subset \text{ Im}\left[H^0(\mathcal{O}_{\mathbf{P}^{r+1}}(\ell)) \to H^0(\mathcal{O}_Z(\ell H))\right]. \tag{9}$$

Now let us vary the projections p_1 and p_2. For each choice of L_1, we get a different B_1: call it $B_1(L_1)$, and for each choice of L_2, as different B_2: call it $B_2(L_2)$. By (5) and (8), all divisors $B_1(L_1)$ are linearly equivalent as are all divisors $B_2(L_2)$. Moreover:

$$\begin{aligned} \bigcap_{L_1} B_1(L_1) &= \emptyset \\ \bigcap_{L_2} B_2(L_2) &= \emptyset \end{aligned}$$

This is because, if $x \in Z$, then there is a choice of L_1 such that $p_1 : Z \to \mathbf{P}^r$ is unramified at y; and a choice of L_2 such that $p_2(x) \in Z_1$ is smooth, hence p_2 is an isomorphism near x. Thus

$$|\,B_1(L_1)\,| = |\,K_Z + (r+1)H\,|$$

and

$$|\, B_2(L_2)\,| = |\, K_Z + (d-r-2)H\,|$$

are base point free linear systems.

Next choose $(r+1)$ L_2's, called L_2^α, $1 \le \alpha \le r+1$, so that if $B_2^{(\alpha)} = B_2(L_2^{(\alpha)})$, then $\bigcap\limits_\alpha B_2^{(\alpha)} = \emptyset$. Look at the Koszul complex:

$$\begin{aligned} 0 \;\to\; \mathcal{O}_Z(\ell H - \textstyle\sum B_2^{(\alpha)}) \;\;&\to\;\; \cdots \\ \to \sum_{\alpha,\beta} \mathcal{O}_Z(\ell H - B_2^{(\alpha)} - B_2^{(\beta)}) \;\;&\to\;\; \sum_\alpha \mathcal{O}_Z(\ell H - B_2^{(\alpha)}) \;\to\; \mathcal{O}_Z(\ell H) \;\to\; 0. \end{aligned}$$

This is exact and diagram chasing gives the conclusion:

$$\begin{gathered} H^i(\mathcal{O}_Z(\ell H - (i+1)B_2)) = (0), \text{ all } i \ge 1 \\ \Rightarrow \sum_\alpha H^0(\mathcal{O}_Z(\ell H - B_2^{(\alpha)})) \to H^0(\mathcal{O}_Z(\ell H)) \text{ surjective} \end{gathered}$$

hence by (9)

$$H^0(\mathcal{O}_{\mathbf{P}^n}(\ell)) \to H^0(\mathcal{O}_Z(\ell H)) \text{ surjective}$$

and

$$\begin{gathered} H^{i+j}(\mathcal{O}_Z(\ell H - (i+1)B_2)) = (0), \text{ all } i \ge 0 \\ \Rightarrow \;\; H^j(\mathcal{O}_Z(\ell H) = (0). \end{gathered}$$

Now $I(Z)$ is m-regular if and only if $H^i(\mathcal{I}_Z(m-i)) = (0)$, $i \ge 1$, hence if and only if

$$\begin{gathered} H^0(\mathcal{O}_{\mathbf{P}^n}(m-1)) \to H^0(\mathcal{O}_Z(m-1)) \text{ surjective} \\ H^i(\mathcal{O}_Z(m-i-1)) = (0), \; i \ge 1. \end{gathered}$$

By the previous remark, this follows provided that

$$H^{i+j}(\mathcal{O}_Z((m-i-1)H - (j+1)B_2)) = (0), \text{ if } i,j \ge 0, \; i+j \ge 1.$$

But let us rewrite:

$$(m-i-1)H - (j+1)B_2 \;\equiv\; K_Z + jB_1 + (m-i-(j+1)(d-1)+r)H$$

using (5) and (8). Note that $jB_1 + \ell H$ is an ample divisor if $\ell \ge 1, j \ge 0$, because $|\, B_1\,|$ is base point free. Therefore by the Kodaira Vanishing Theorem,

$$H^i(\mathcal{O}_Z(K_Z + jB_1 + \ell H)) = (0), \; i,j \ge 1, \; j \ge 0$$

and provided $m = (r+1)(d-2)+2$, this gives the required vanishing.

4 Applications

From some points of view, the first main problem of algebraic geometry is to reduce the study of a general ideal I to that of prime ideals, or the study of arbitrary schemes to that of varieties. One way of doing this is to find a decomposition of the ideal into primary ideals: i.e. write it as an intersection of primary ideals. But even when non-redundancy is added, this is not unique, and usually one actually wants something less: to find its radical and perhaps write the radical as an intersection of prime ideals, or to find its top dimensional part, or to find its associated prime ideals and their multiplicities. There are really four computational problems involved here which should be treated separately: (i) eliminating the multiplicities in the ideal I, (ii) separating the pieces of different dimension, (iii) "factoring" the pieces of each dimension into irreducible components, and finally (iv) describing the original multiplicities, either numerically or by a primary ideal. Three of these four problems are the direct generalizations of the basic problems for factoring a single polynomial: we can eliminate multiple factors, getting a square-free polynomial, we can factor this into irreducible pieces and we can ask for the multiplicities with which each factor appeared in the original polynomial. There is a fifth question which arises when we work, as we always must do on a computer, over a non-algebarically closed field k: we can ask (v) for an extension field k' of k over which the irreducible components break up into absolutely irreducible components.

Classical algorithms for all of these of these rely heavily on making explicit projections of $V(I)$ to lower dimensional projective spaces. This can be done either by multi-variable resultants if you want only the set-theoretic projection, or by Gröbner bases with respect to the lexicographic order or an elimination order, to get the full ideal $I \cap k[X_0, \cdots, X_m]$. Recent treatments of multi-variable resultants can be found in [Can89], [Cha91], and a recent treatment of the basis method can be found in [GTZ88]. There is no evidence that either of these is an efficient method, however, and taking Gröbner bases in the lexicographical order or an elimination order is often quite slow, certainly slow in the worst case. The general experience is that taking projections can be very time consuming. One reason is that the degree of the generators may go up substantially and that sparse defining polynomials may be replaced by more or less generic polynomials. A specific

example is given by principally polarized abelian varieties of dimension r: they are defined by quadratic polynomials in $(4^r - 1)$-space, but their degree here (hence the degree of their generic projection to $\mathbf{P}^{r+1}$) is $4^r r!$ [Mum70a]. In fact, any variety is defined purely by quadratic relations in a suitable embedding [Mum70b].

Instead of using real computational experience, the fundamental method in theoretical computer science for analyzing complexity of algorithms is to count operations. For algebraic algorithms, the natural measure of complexity is not the number of bit operations, but the number of field operations, addition, subtraction, multiplication and (possibly) division that are used. In this sense, any methods that involve taking Gröbner bases for any order on monomials will have a worst-case behavior whose complexity goes up with the regularity of the ideal hence will take "double exponential time". However, it appears that this worst-case behavior may in fact only concern problem (iv) – finding the primary ideals – and that problems (i), (ii) and (iii) may be solvable in "single exponential time". The idea that such algorithms should exist for finding $V(I)$ set-theoretically was proposed in the 1984 lecture on which this article is based, but turned out, in fact, to have been already proven by Chistov and Grigoriev, cf. their unpublished 1983 note [CG83]. Their line of research led, in some sense, to the work of Brownawell and Kollár, showing the single exponential bound $\left(\sqrt{I}\right)^m \subset I$ for $m = d^n$, where $d = \max($ degrees of generators of $I)$.

Based on this work, Giusti and Heintz [GH91] give a singly exponentially bounded algorithm for computing ideals $\mathbf{q}_i$ such that $V(\mathbf{q}_i)$ are the irreducible components of $V(I)$ (over the ground field k). The method depends on computing what is essentially the Chow form of each component, and leads to an ideal defining this variety but not its full ideal. In fact, their $\mathbf{q}_i$ may be guaranteed to be prime except for possible embedded components.

A direct approach to constructing both $\sqrt{I}$ and the intersection of the top-dimensional primary components of I, denoted $\mathrm{Top}(I)$, is given in a recent paper by Eisenbud, Huneke and Vasconcelos [EHV92]. Their construction of the radical uses the Jacobian ideals, i.e. the ideals of minors of various sizes of the Jacobian matrix of generators of I. This is certainly the most direct approach, but, again they have trouble with possible embedded components, and must resort to ideal quotients, hence they need a Gröbner basis of I

in the reverse lexicographic order. They compute $\mathrm{Top}(I)$ as the annihilator of $\mathrm{Ext}^{\mathrm{codim}(I)}(k[X_0,\cdots,X_n]/I, k[X_0,\cdots,X_n])$, which is readily found from a full resolution using Gröbner bases. Their algorithm appears to be practical in some cases of interest, but still has double exponential time worst-case behavior.

It may turn out to be most effective in practice to combine these ideas. Often an ideal under study has regularity far smaller than the geometric degree of its top dimensional components; projecting these components to a hypersurface requires computing in degrees up to the geometric degree, which is wasteful. On the other hand, methods such as those in [EHV92] work better in low codimensions, if only because there are fewer minors to consider in the Jacobian matrix. Thus, projecting an arbitary scheme down to low codimension and then switching to direct methods may work best of all.

This still does not settle the issue of the complexity of calculating $\sqrt{I}$, or, for that matter, calculating the full prime ideal of any subvariety of codimension greater than one. Chow form type methods give you an effective method of defining the set $V(I)$ but only of generating I up to possible embedded components. For this reason, the two schools of research, one based on the algebra of I, the other based on subsets of $\mathbf{P}^n$ have diverged. If we knew, as discussed in the previous section, that the regularity of a reduced ideal could be bounded singly exponentially, then we could bound the degrees of the generators of $\sqrt{I}$, and, using Brownawell-Kollár, we could determine $\sqrt{I}$ up to these degrees and get the whole ideal. But without such a bound, it is still not clear whether only $V(I)$ and not $\sqrt{I}$ can be found in worst-case single exponential time.

Let's look at problem (iii). Assume you have found a reduced equidimensional I. To study splitting it into irreducible or absolutely irreducible pieces, we shall assume initially it is a hypersurface, i.e. $I = (f)$. Computationally, there may often be advantages to not projecting a general I to a hypersurface, and we will discuss one such approach below. Geometrically, there is nothing very natural about irreducible but not absolutely irreducible varieties: from the standpoint of their properties, they behave like reducible varieties, except that, being conjugate over k, their components have very similar properties. If the ground field k gets bigger or smaller, the set of absolutely irreducible components gets partitioned in finer or coarser ways

into the k-components. If one has never done any calculations, one would therefore be inclined to say – let's extend k as far as needed to split our algebraic set up into absolutely irreducible components. *This is a very bad idea!* Unless this extension k' happens to be something simple like a quadratic or cyclotomic extension of k, the splitting field k' is usually gigantic. This is what happens if one component of $V(I)$ is defined over an extension field k_1 of k of degree e, and the Galois group of k_1/k is the full symmetric group, a very common occurence. Then $V(I)$ only splits completely over the Galois closure of k_1/k and this has degree $e!$. The moral is: never factor unless you have to.

In fact, unless you need to deal simultaneously with more than one of its irreducible components, you can proceed as follows: the function field $K = k[X_0, \cdots, X_n]/(f)$ contains as a subfield an isomorphic copy of k_1: you find that field as an extension $k_1 = k[y]/(p(y))$ of k, and solve for the equation of one irreducible component $f_1 \in k_1[X_0, \cdots, X_n]$ by the formula $\mathrm{Norm}_{k_1/k}(f_1) = f$.

Pursuing this point, why should one even factor the defining equation f over k? Factoring, although it takes polynomial time [LLL82], is often very slow in real time, and, unless the geometry dictates that the components be treated separately, why not leave them alone. In some situations, for instance, [DD84] one may have an ideal, module or other algebraic structure defined by polynomials or matrices of polynomials over a *ground ring* $D = k[y]/(p(y))$, where p is a square-free polynomial. Thus D is a direct sum of extension fields, but there is no need to factor p or split up D until the calculations take different turns with the structures over different pieces of $Spec(D)$.

The standard methods of factoring in computer algebra all depend on (i) writing the polynomial over a ring, finitely generated over $\mathbf{Z}$, and reducing modulo a maximal ideal $\mathbf{m}$ in that ring, obtaining a polynomial over a finite field; and (ii) restricting to a line L, i.e. substituting $X_i = a_i X_0 + b_i, i \geq 1$ for all but one variable, obtaining a polynomial in one variable over a finite field. This is then factored and then, using Hensel's lemma, one lifts this factorization modulo higher powers of $\mathbf{m}$ and of the linear space L. One then checks whether a coarsened version of this factorization works for f. This is all really the arithmetic of various small fields. Geometrically, every polynomial in one variable factors over a suitable extension field and the question of counting the absolutely irreducible components of a variety is really more

elementary: it is fundamentally topological and not arithmetic. One should, therefore, expect there to be direct geometric ways of counting these components and separating them. Assuming I is a reduced, equi-r-dimensional ideal, the direct way should be to use Serre duality, computing the cohomology $H^r(\Omega^r_{V(I)})$, where $\Omega^r_{V(I)} \subset \omega_{V(I)}$ is the subsheaf of the top-dimensional dualizing sheaf of $V(I)$ of absolutely regular r-forms. Its dimension will be the number of absolutely irreducible components into which $V(I)$ splits. Calculating this cohomology involves two things: algebraically resolving the ideal I and geometrically resolving the singularities of $V(I)$ far enough to work out $\Omega^r_{V(I)}$. Classically, when $I = (f)$ was principal, $\Omega^r_{V(I)}$ was called its ideal of "subadjoint" polynomials.

There is one case where this is quite elementary and has been carried out: this is for plane curves. One can see immediately what is happening by remarking that a non-singular plane curve is automatically absolutely irreducible, hence one should expect that its singularities control its decomposition into absolutely irreducible pieces. Indeed, if $\mathbf{C} \subset k[X_0, X_1, X_2]/(f)$ is the conductor ideal, then $\Omega^1_{V(f)}$ is given by the homogeneous ideal $\mathbf{C}$, but with degree 0 being shifted to be polynomials of degree $d-3$, d the degree of f. To calculate H^1, assume X_0 is not zero at any singularity of $V(f)$ and look at the finite-dimensional vector space of all functions $k[X_1/X_0, X_2/X_0]/(\mathbf{C} + (f))$ modulo the restrictions $g/(X_0^{d-3})$ for all homogeneous polynomials g of degree $d-3$. This will be canonically the space of functions on the set of components of $V(f)$ with sum 0. In particular, it is (0) if and only if $V(f)$ is absolutely irreducible. This follows from standard exact sequences and duality theory. It was known classically as the Cayley-Bacharach theorem, for the special case where $V(f)$ was smooth except for a finite number of ordinary double points. It states that $V(f)$ is absolutely irreducible if and only if for every double point P, there is a curve of degree $d-3$ passing through all the double points except P.

This example gives one instance where a deeper computational analysis of varieties requires a computation of its resolution of singularities. We believe that there will be many instances where practical problems will require such an analysis. In many ways, resolution theorems look quite algorithmic, and, for instance, Abhyankar and his school have been approaching the problem in this way [Abh82], as have Bierstone and Milman [BM91]. However, the only case of resolution of singularities to be fully analyzed in the sense of

computational complexity is that of plane curves. This has been done by Teitelbaum [Tei89], [Tei90]. His analysis is notable in various ways: he is extremely careful about not making unnecessary factorizations, let alone taking unnecessary field extensions, and uses the "D" formalism discussed above. He describes his algorithm so precisely that it would be trivial to convert it to code and, as a result, he gives excellent bounds on its complexity.

References

[Abh82] S. S. Abhyankar, *Weighted expansions for canonical desingularization*, Lecture Notes in Math., vol. 910, Springer-Verlag, 1982.

[Art76] M. Artin, *Lectures on deformations of singularities*, Tata Institute on Fundamental Research, Bombay, 1976.

[Bay82] Dave Bayer, *The division algorithm and the Hilbert scheme*, Ph.D. thesis, Harvard University, Department of Mathematics, June 1982, order number 82-22588, University Microfilms International, 300 N. Zeeb Rd., Ann Arbor, MI 48106.

[BEL91] Aaron Bertram, Lawrence Ein, and Robert Lazarsfeld, *Vanishing theorems, a theorem of Severi, and the equations defining projective varieties*, J. Amer. Math. Soc. **4** (1991), 587–602.

[Ber78] G. M. Bergman, *The diamond lemma for ring theory*, Adv. in Math. **29** (1978), 178–218.

[BM88] Dave Bayer and Ian Morrison, *Standard bases and geometric invariant theory I. Initial ideals and state polytopes*, J. Symb. Comput. **6** (1988), no. 2–3, 209–217, reprinted in [Rob89].

[BM91] E. Bierstone and P. Milman, *A simple constructive proof of canonical resolution of singularities*, Effective methods in algebraic geometry (Castiglioncello, 1990), Progr. Math., vol. 94, Birkhauser Boston, 1991, pp. 11–30.

[BCR91] A. M. Bigatti, M. Caboara, and L. Robbiano, *On the computation of Hilbert-Poincare series*, Applicable Algebra in Engineering, Communications, and Computing **2** (1991), 21–33.

[Bri73] J. Briancon, *Weierstrass prepare a la Hironaka*, Astérisque **7,8** (1973), 67–73.

[Bro87] W. D. Brownawell, *Bounds for the degrees in the Nullstellensatz*, Ann. of Math. (2) **126** (1987), 577–591.

[BS87a] Dave Bayer and Mike Stillman, *A criterion for detecting m-regularity*, Invent. Math. **87** (1987), 1–11.

[BS87b] Dave Bayer and Mike Stillman, *A theorem on refining division orders by the reverse lexicographic order*, Duke Math. J. **55** (1987), no. 2, 321–328.

[BS88] Dave Bayer and Mike Stillman, *On the complexity of computing syzygies*, J. Symb. Comput. **6** (1988), 135–147.

[BS92a] Dave Bayer and Mike Stillman, *Macaulay: A system for computation in algebraic geometry and commutative algebra*, 1982–1992, computer software available via anonymous ftp from zariski.harvard.edu.

[BS92b] Dave Bayer and Mike Stillman, *Computation of Hilbert functions*, J. Symb. Comput. **6** (1992), 31-50.

[Buc65] B. Buchberger, Ph.D. thesis, Univ. Innsbrück, 1965.

[Buc76] B. Buchberger, *A theoretical basis for the reduction of polynomials to canonical forms*, ACM SIGSAM Bull. **39** (1976), 19–29.

[Buc79] B. Buchberger, *A criterion for detecting unnecessary reductions in the construction of Gröbner bases*, Symbolic and Algebraic Computation (Proceedings of EUROSAM 79), Lecture Notes in Computer Science, vol. 72, Springer-Verlag, 1979, pp. 3–21.

[Can89] J. Canny, *Generalized characteristic polynomials*, Symbolic and Algebraic Computation (Proceedings of ISSAC 88), Lecture Notes in Computer Science, vol. 358, Springer-Verlag, 1989, pp. 293–299.

[CG83] A. L. Chistov and D. Yu. Grigoriev, *Subexponential-time solving systems of algebraic equations I, II*, Steklov Mathematical Institute, Leningrad department, LOMI Preprints E-9-93, 0E-10-c83, 1983.

[Cha91] Marc Chardin, *Un algorithme pour le calcu des résultants*, Effective methods in algebraic geometry (Castiglioncello, 1990), Progr. Math., vol. 94, Birkhauser Boston, 1991, pp. 47–62.

[DD84] C. Dicrescenzo and D. Duval, *Computations on curves*, Lecture Notes in Computer Science, vol. 174, Springer-Verlag, 1984.

[EG84] David Eisenbud and Shiro Goto, *Linear free resolutions and minimal multiplicity*, J. Algebra **88** (1984), no. 1, 89–133.

[EHV92] David Eisenbud, Craig Huneke, and Wolmer Vasconcelos, *Direct methods for primary decomposition*, Invent. Math. (1992), to appear.

[Eis92] David Eisenbud, *Commutative algebra with a view toward algebraic geometry*, 1992, in preparation.

[Gal74] A. Galligo, *A propos du theoreme de preparation de Weierstrass*, Fonctions de Plusieurs Variables Complexes, Lecture Notes in Math., vol. 409, Springer-Verlag, 1974, pp. 543–579.

[Gal79] A. Galligo, *Theoreme de division et stabilite en geometrie analytique locale*, Ann. Inst. Fourier (Grenoble) **29** (1979), 107–184.

[GH91] Marc Giusti and Joos Heintz, *Algorithmes—disons rapides—pour la decomposition d'une variete algebrique en composantes irreductibles et equidimensionnelles ["Fast" algorithms for the decomposition of an algebraic variety into irreducible and equidimensional components]*, Effective methods in algebraic geometry (Castiglioncello, 1990), Progr. Math., vol. 94, Birkhauser Boston, 1991, pp. 169–194.

[Giu84] Marc Giusti, *Some effectivity problems in polynomial ideal theory*, EUROSAM 84), Lecture Notes in Computer Science, vol. 204, Springer-Verlag, 1984, pp. 159–171.

[GLP83] L. Gruson, R. Lazarsfeld, and C. Peskine, *On a theorem of Castelnuovo, and the equations defining space curves*, Invent. Math. **72** (1983), 491–506.

[GTZ88] P. Gianni, B. Trager, and G. Zacharias, *Gröbner bases and primary decomposition of polynomial ideals*, J. Symb. Comput. **6** (1988), no. 2–3, 149–167, reprinted in [Rob89].

[Her26] Grete Hermann, *Die Frage der endlich vielen Schritte in der Theorie der Polynomideale*, Math. Ann. **95** (1926), 736–788.

[Hir64] H. Hironaka, *Resolution of singularities of an algebraic variety over a field of characteristic zero: I, II*, Ann. of Math. (2) **79** (1964), 109–326.

[Kol88] János Kollár, *Sharp effective Nullstellensatz*, J. Amer. Math. Soc. **1** (1988), no. 4, 963–975.

[Lak91] Y. N. Lakshman, *A simple exponential bound on the complexity of computing gröbner bases of zero-dimensional ideals*, Effective methods in algebraic geometry (Castiglioncello, 1990), Progr. Math., vol. 94, Birkhauser Boston, 1991, pp. 227–234.

[Laz87] Robert Lazarsfeld, *A sharp Castelnuovo bound for smooth surfaces*, Duke Math. J. **55** (1987), 423–429.

[Len92] H. W. Lenstra, Jr., *Algorithms in algebraic number theory*, Bull. Amer. Math. Soc. (N.S.) **26** (1992), no. 2, 211–244.

[Lip84] Joseph Lipman, *Dualizing sheaves, differentials and residues on algebraic varieties*, Astérisque, vol. 117, 1984.

[LL91] Y. N. Lakshman and D. Lazard, *On the complexity of zero-dimensional algebraic systems*, Effective methods in algebraic geometry (Castiglioncello, 1990), Progr. Math., vol. 94, Birkhauser Boston, 1991, pp. 217–225.

[LLL82] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász, *Factoring polynomials with rational coefficients*, Math. Ann. **261** (1982), 515–534.

[Mac27] F. S. Macaulay, *Some properties of enumeration in the theory of modular systems*, Proc. London Math. Soc. **26** (1927), 531–555.

[MM82] Ernst W. Mayr and Albert R. Meyer, *The complexity of the word problem for commutative semigroups and polynomial ideals*, Adv. in Math. **46** (1982), 305–329.

[MM83] H. Michael Möller and Ferdinando Mora, *Upper and lower bounds for the degree of Gröbner bases*, Computer Algebra (EUROCAL 83), Lecture Notes in Computer Science, vol. 162, Springer-Verlag, 1983, pp. 157–167.

[MM86] H. Michael Möller and Ferdinando Mora, *New constructive methods in classical ideal theory*, J. Algebra **100** (1986), no. 1, 138–178.

[MR88] T. Mora and L. Robbiano, *The Gröbner fan of an ideal*, J. Symb. Comput. **6** (1988), no. 2–3, 183–208, reprinted in [Rob89].

[Mum66] David Mumford, *Lectures on curves on an algebraic surface*, Princeton University Press, Princeton, New Jersey, 1966.

[Mum70a] David Mumford, *Abelian varieties*, Oxford University Press, Oxford, 1970.

[Mum70b] David Mumford, *Varieties defined by quadratic equations*, Questions on Algebraic Varieties, Centro Internationale Matematica Estivo, Cremonese, Rome, 1970, pp. 29–100.

[MW83] D. W. Masser and G. Wüstholz, *Fields of large transcendence degree generated by values of elliptic functions*, Invent. Math. **72** (1983), 407–464.

[Pin86] Henry C. Pinkham, *A Castelnuovo bound for smooth surfaces*, Invent. Math. **83** (1986), 491–506.

[Ran90] Ziv Ran, *Local differential geometry and generic projections of threefolds*, J. Differential Geom. **32** (1990), 131–137.

[Rav90] M. S. Ravi, *Regularity of ideals and their radicals*, Manuscripta Math. **68** (1990), 77–87.

[Ric74] F. Richman, *Constructive aspects of Noetherian rings*, Proc. Amer. Math. Soc. **44** (1974), 436–441.

[Rob85] L. Robbiano, *Term orderings on the polynomial ring*, Proceedings of EUROCAL '85 (Linz), Lecture Notes in Computer Science, vol. 204, Springer-Verlag, 1985, pp. 513–517.

[Rob89] Lorenzo Robbiano (ed.), *Computational aspects of commutative algebra*, Academic Press, 1989, ISBN 0-12-589590-9.

[Sch80] Frank-Olaf Schreyer, *Die Berechnung von Syzygien mit dem verallgemeinerten Weierstrass'schen Divisionssatz*, Diplomarbeit am Fachbereich Mathematik der Universität Hamburg, 1980.

[Sch91] Frank-Olaf Schreyer, *A standard basis approach to syzygies of canonical curves*, J. Reine Angew. Math. **421** (1991), 83–123.

[Ser55] J.-P. Serre, *Faisceaux algebrique coherents*, Ann. of Math. (2) **61** (1955), 197–278.

[Spe77] D. Spear, *A constructive approach to commutative ring theory*, Proceedings of the 1977 MACSYMA Users' Conference, NASA CP-2012, 1977, pp. 369–376.

[Tei89] Jeremy Teitelbaum, *On the computational complexity of the resolution of plane curve singularities*, Symbolic and algebraic computation (Rome, 1988), Lecture Notes in Computer Science, vol. 358, Springer, 1989, pp. 285–292.

[Tei90] Jeremy Teitelbaum, *The computational complexity of the resolution of plane curve singularities*, Math. Comp. **54** (1990), no. 190, 797–837.

[Tri78] W. Trinks, *Über B. Buchberger's Verfahren, Systeme algebraischer Gleichungen zu lösen*, J. Number Theory **10** (1978), 475–488.

[Zac78] G. Zacharias, Bachelor's thesis, Mass. Inst. of Technology, 1978.

[Zar69] Oscar Zariski, *An introduction to the theory of algebraic surfaces*, Lecture Notes in Math., vol. 83, Springer-Verlag, 1969.

[ZS76] Oscar Zariski and Pierre Samuel, *Commutative algebra, Volumes I, II*, Graduate texts in mathematics, vol. 28–29, Springer-Verlag, 1975–1976.