Course Notes for Math 676: Computational Algebraic Geometry Spring 2009

Prof. Dan Bates
Colorado State University

February 23, 2009

Chapter 1

Foreword

These notes are being written on the fly as I teach a graduate topics class in computational algebraic geometry. We'll be covering standard symbolic sorts of computation (Gröbner bases, resultants, etc.) and newer numerical methods (homotopy methods, numerical algebraic geometry). After that, we'll have at least a few weeks left to really dig into some topics. Please note that these notes are available to anybody but were originally intended as a reference for students in this topics course.

Since some of the details I will provide in lectures will come from various copyrighted sources, some parts of these notes will be a bit sparse (to be filled in later). In general, when possible, I will point out good sources for details, extensions, or examples.

I am not intending to provide a thorough introduction to algebraic geometry. Instead, my intention is to get to computation and applications as quickly as possible. We'll spend much more time on algorithms, complexity, and adventures with software and applications than with theory. The next time around, things may change, but that's the plan for now.... Also, I am pitching the course more towards math grad students since there are no engineers/scientists in the course.

Please note that (a) there will be plenty of details missing from these notes and, much more importantly, (b) I make no claims that what I record here is entirely correct. I am often rather informal with these notes, so, for example, I am sometimes not so careful about including all necessary conditions when

stating theorems. So take everything with a grain of salt – the gist is there, but don't take the details as hard fact. That said, if you spot any errors, please let me know. I would go broke if I offered even a penny for each such report, but I will at least give you a smile and a hearty handshake....

Just two weeks into the course, I already owe a significant debt to my students for their help both with these notes and with the lectures based upon them. Thank you to all twelve of you (in alphabetical order):

- Dan Brake
- Shawn Farnell
- Beth Malmskog
- Tim McCoy
- Ken Monks
- Eric Nelson
- Matt Niemerg
- Dusty Ross
- Eric Schmidt
- Elly Smith
- Jenna Tague
- Cassie Williams

Contents

Chapter 2

Basics

Much of this chapter was motivated by the two Cox, Little, and O'Shea books, but especially the UTM book *Ideals, Varieties, and Algorithms* (which I'll call [CLO1]). Their GTM book *Using Algebraic Geometry* (which I'll call [CLO2]) is also canonical for this sort of course, but the intro is a bit more terse and points to [CLO1] a lot. Hal Schenck's book *Computational Algebraic Geometry* and Bernd Sturmfels' *Solving Polynomial Systems* are also close at hand while we go through some of these symbolic methods in the next few chapters....

2.1 Polynomials and their solutions

I assume that you know what polynomials are (along with monomials, coefficients, terms, degree, rings, and fields). If not, check out [CLO1] for a nice introduction. For now, we only allow for nonnegative exponents in the monomials (i.e., no Laurent polynomials), but that restriction may be lifted later.

Polynomials live in polynomial rings $\mathbb{F}[x_1, \ldots, x_n]$ (commutative rings with 1), with $\mathbb{F}$ a field (or occasionally a ring), typically $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$. In fact, for this course, we'll mostly think about $\mathbb{F} = \mathbb{C}$. A polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ may obviously be thought of as a function $f : \mathbb{F}^n \to \mathbb{F}$.

Definition 2.1. *A* **solution of a polynomial** $f(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ *is a point* $z \in \mathbb{F}^n$ *such that* $f(z) = 0$*, i.e., s.t.* $z \in f^{-1}(0)$.

$\mathbb{F}^n$ is called **n-dimensional affine space** for field $\mathbb{F}$. Think of $\mathbb{R}^n$. We'll get to projective space later....

If f is a set of m polynomials in n variables, then we have a function $f : \mathbb{F}^n \to \mathbb{F}^m$.

Definition 2.2. *Given a set of polynomials $f = \{f_1, \ldots, f_m\} \subset \mathbb{F}[x_1, \ldots, x_n]$ (a* **polynomial system***), a* **solution of the polynomial system** *f is a point $z \in \mathbb{F}^n$ such that $f_i(z) = 0$ for $i = 1, \ldots, m$, i.e., z is a solution for each polynomial in system f.*

This brings us to the fundamental dichotomy in algebraic geometry – polynomials are algebraic objects which have as solution sets geometric objects. The interplay between these two worlds is the focus of our course. Sometimes we will use algebraic (symbolic) manipulations of polynomials to understand the corresponding geometric objects. Sometimes, though, we will move to the geometric side to gather information about the algebraic side; this is common in the numerical methods that we'll encounter later.

We will focus on the problem of finding solutions of polynomial systems. This brings about a point that should be made clear: $f(x) = 0$ has two meanings (for some polynomial f). There is a polynomial meaning (i.e., all coefficients appearing in f are zero) and a functional meaning (i.e., f evaluates to 0 for all input x). These need not be the same – $x(x-1)$ certainly evaluates to 0 for all elements of $\mathbb{Z}^2$ but is not itself the zero polynomial.

Luckily, as proven nicely in [CLO1], the two meanings are the same for infinite fields.

2.2 Ideals

In algebra, we don't think about subsets of rings; we think about ideals. Given a set of polynomials $f_1, \ldots, f_m \subset \mathbb{F}[x_1, \ldots, x_n]$, the **ideal** $< f_1, \ldots, f_m >$ **generated by** $f_1, \ldots, f_m$ is just the collection of all polynomial combinations of the f_i. (Think of linear combinations with polynomial coefficients rather than numerical coefficients.) If you haven't seen ideals for a while, this might be a good time to crack open your favorite algebra text. To make a long story short, (a) sums of elements of an ideal and (b) products of an element of an ideal with any ring element stay in the ideal.

Example 2.1. *Let* $f = \{x^2y, y^2\} \subset \mathbb{R}[x, y]$. *Obviously,* x^2y *and* y^2 *are in* $< f >$. *So is* $3x^2y - y^2$, *as is* $(x^{40}y^{19} + 5)x^2y + (3000y^4 - x^2 + x - 1)y^2$. x *isn't, though; neither is* y *or even* x^2.

We want to study solution sets of polynomial systems $f \subset \mathbb{F}[x_1, \ldots, x_n]$. It would be handy (as we'll see later) if we could use ideals rather than sets of polynomials. A "solution" of an ideal I is just a point that satisfies all (infinitely many) polynomials in I. Fortunately, we can make this change of focus:

Proposition 2.1. *The solutions of* $f \subset \mathbb{F}[x_1, \ldots, x_n]$ *are the same as those of* $< f >$.

Proof: Let $f = \{f_1, \ldots, f_m\}$. Suppose z is a solution of f, i.e., $f_i(z) = 0 \ \forall f_i \in f$. Then, for any $g \in < f >$, g is a polynomial combination of the f_i. But $f_i(z) = 0$ for each i, meaning that $g(z) = 0$. Conversely, the f_i are certainly elements of $< f >$, so solutions of the ideal must satisfy each f_i. □

Given some description of an ideal I in a ring R (without knowing any generators), we have no guarantee that there is a finite set of generators for I. The algorithms described later for "solving" polynomial systems (computing solution sets) typically put information about the generators of the ideal at hand into a matrix or cycle through them, performing operations on each (or each pair). It would be utter disaster if some ideals in polynomial rings had to have infinitely many generators! Fortunately, we have the Hilbert Basis Theorem:

Theorem 2.1. *Every ideal in a polynomial ring has a finite set of generators.*

Good news, huh? Thank you, Hilbert. There are several ways of proving this. For now, though, given the audience, we'll move on. If you haven't seen a proof before, we'll probably hit one later, after talking about Gröbner bases.

One quick note before we move to the geometric side: In practice, polynomial systems coming from applications come as exactly that – a polynomial system rather than an ideal. It is often adequate to use this polynomial system as it is handed to you (just by plugging it into some numerical or symbolic software package), but part of what makes us (as mathematicians) valuable to engineers and scientists is that we also know that there is an underlying ideal and that changing generators may be of use to us. More on that later....

2.3 Varieties

Ideals are nice and all, but we want to find solutions of polynomial systems. Here's the fundamental definition on this geometric side of the coin:

Definition 2.3. *An* **affine algebraic set** *is the solution set of a polynomial system/ideal in a polynomial ring. Given a polynomial system* $f \in \mathbb{F}[x_1, \ldots, x_n]$*, we can move to the ideal* $I =< f >$*. We denote the solution set corresponding to* I *as* $V(I)$*.*

A word about words: Some authors would also call this a variety (or, more specifically, an affine variety). Hence the "$V(I)$" notation. Other authors define varieties as *irreducible* affine algebraic sets (whatever *irreducible* might mean – we'll get to that soon). I could be very careful and use the term "affine algebraic set," but that's a lot of keystrokes for not much benefit. Just to set the record straight, **variety** for me (in these notes) means an affine (or projective, depending on the context) algebraic set. I make no assumptions about irreducibility.

So what can these varieties look like? You have seen some before:

Example 2.2. *Think about* $x \in \mathbb{R}[x]$*. The solution set of* $x = 0$ *is, well,* $x = 0 \in \mathbb{R}$*. Shocking, huh?*

OK, let's be a little more sophisticated:

Example 2.3. *A system of linear equations*

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n - b_1 &= 0 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n - b_2 &= 0 \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n - b_m &= 0, \end{aligned}$$

is also a system of (degree 1) polynomials (which generates an ideal). The solution set of this system of linear equations is thus the variety for this ideal.

Thus, linear spaces (points, lines, planes, etc.) are varieties. x *is a point in one variable, a line in two, a plane in three, and so on. As you probably know, a linear space of dimension one less than the ambient dimension (i.e., of codimension 1) is called a* hyperplane *and is given by a single linear equation.*

Example 2.4. *It turns out that we can provide a single (degree 2 or higher) polynomial and still get something of codimension 1. Think of* $y - x^2$ *or* $y^2 - x^2$ *in* $\mathbb{R}^2$*. The former is a parabola (dimension 1); the latter is a pair of lines (note that the latter factors). In* $\mathbb{R}^3$*, take* $x^2 + y^2 + z^2 - 1$*, yielding a (two-dimensional) sphere. A single (nonconstant) polynomial will* always *give you something of codimension 1.*

2.3.1 Dimension

OK, so what is the *dimension* of a variety? There is a lot to be said about this, but let's keep it simple (and mostly right) for now. Points are zero-dimensional. Lines and curves are one-dimensional. Surfaces are two-dimensional. This much is clear. The codimension of a variety is just the ambient dimension minus the dimension.

What about, for example, a line and a point together, such as you get with the system $x(x-1), x(y-1)$ in $\mathbb{R}^2$? Obviously, we think of the line as being dimension one and the point as being zero-dimensional (i.e., **isolated**). By convention, we declare that this variety (the union of the line and the point) has dimension one as this is the maximum dimension of the two bits.

Loosely speaking, dimension is the number of degrees of freedom on a variety (not including rotation). A bit more technically (and locally), it is the dimension of the tangent space at a point on the variety. For example, for a nice, smooth curve, the tangent space at a point is a line, making it one-dimensional. For a "nice" surface like a sphere, the tangent space at any point is two-dimensional, so we say that the sphere is two-dimensional.

OK, so what does "nice" mean? Think about a self-intersecting curve, i.e., some curve that has a loop. Other than at the point of self-intersection, the tangent space is one-dimensional. However, for that one special, nasty point, there are two tangent directions. What do we do about this?

Definition 2.4 (NOT CORRECT!)**.** *The dimension of a variety is the dimension of the tangent space at a "nice" point on the variety.*

By "nice," I am talking about nonsingular/regular/smooth/manifold points, i.e., points at which the Jacobian matrix has full rank. The "nasty" points are then singular (where the Jacobian does not have full rank). Unfortunately, this notion of dimension is still not entirely correct. In the example

above (a point and a line), the entire line is singular – no "nice" (nonsingular/regular/smooth/manifold) points. If we are looking for nice points, we are left with only the isolated point, which would lead us to call the variety zero-dimensional (which is incorrect).

In fact, this will *always* happen when a component of a solution set is not of the "correct" dimension, i.e., when the codimension of a component does not equal the number of polynomials in the system. Such a component will always be singular. Yikes. So how can we talk about dimension?

Well, there are at least two answers (for now), on which we may elaborate a bit later. First, a fairly standard definition of dimension is that it is the degree of the Hilbert polynomial. That will take some unwrapping (which I hope to get to later on). Second, we can move to "I(V)" (defined below) for each component – this is the ideal of all polynomials that vanish on our algebraic set. In the example above, we would then get $< x >$ for the line, so we get that all points are nonsingular on the line (since the Jacobian would be (1 0), which is full rank).

At this point, we could go on and carefully define dimension, but we already have intuition about this. For now, we are better off using that intuition and moving on. Perhaps later we will come back to Hilbert polynomials and at that point cover dimension....

One last thing before we start getting to more sophisticated connections between varieties and ideals. Given no polynomials, our solution set is (by default) the entire ambient space, e.g., $\mathbb{F}^n$. Adding one polynomial equation to our polynomial system chops the dimension by one, yielding a **hypersurface**. What if we add another? Intuition says that the dimension should again drop by one, right?

For example, think about the linear picture. Adding one nontrivial linear equation drops the dimension by one (always). Adding another nontrivial linear equation *may* decrease the dimension by one more, but not necessarily. For example, suppose the two linear equations are linearly dependent.

Unfortunately, the same holds true in the nonlinear setting. The number of equations does not directly dictate the codimension of the corresponding variety (though it does provide an upper bound on the codimension!).

2.4 Ideals and varieties

2.4.1 Combining ideals and varieties

Given two varieties, i.e., $V(f_1, \ldots, f_n)$ and $V(g_1, \ldots, g_k)$, what can we say about the intersection and union of these two varieties? In particular, are they also varieties, i.e., can we cook up polynomials that define them? Before reading on, try it.

Take the case of two hypersurfaces, each given by one polynomial, f and g, say. Then the intersection is given by the system $f = g = 0$. In general, we just merge the two polynomial systems together to get the intersection of the two varieties.

As for the union, we want one set of polynomials to hold or the other. In the case of two polynomials as in the last paragraph, $fg = 0$ suffices. (Check this.) What if we have the union of $V(f_1, f_2)$ and $V(g_1, g_2)$? Then take all four combinations. (Check this, too.) This case is a little harder to prove than the intersection case above.

2.4.2 Moving between ideals and varieties

Now we can finally get to the main set of ideas – the "dictionary" as it is called in [CLO1]. We already know that we have a map from ideals to varieties, i.e., $V()$. There is also a map in the other direction. Indeed, given some set of points $V \subset \mathbb{F}^n$, we define $I(V) = \{f \in \mathbb{F}[x_1, \ldots, x_n] : f(z) = 0 \ \forall z \in V\}$, i.e., $I(V)$ is just the set of all polynomials that vanish at all points of V.

Using these two mappings, we can now move back and forth between ideals and varieties. What if we go over and back – will we come back to where we came from?

Let's start with a variety V. $I(V)$ is all polynomials that vanish on V. Now what is the common zero set ("locus") of the polynomials of $I(V)$? Well, it's just V, fairly obviously. This isn't a formal proof, but one can prove that $V = V(I(V))$ if V is a variety to begin with. This seems nice. (What if V is not a variety to begin with???)

How about the other direction? Take an ideal I and move to the variety $V(I)$. Now take all polynomials that vanish on $V(I)$ to get $I(V(I))$. Does

$I = I(V(I))$ necessarily?

Think about x^2 in $\mathbb{R}$. It vanishes at $x = 0$, so $I(V(I))$ includes $x^2, x^3, \ldots$ and x. x is not in $< x^2 >$, so we have moved to something bigger. In fact, with some work, we could prove the Hilbert Nullstellensatz:

Theorem 2.2. *$I(V(I)) = \sqrt{I}$ if we are working over an algebraically closed field.*

What is $\sqrt{I}$? It is the radical of the ideal I and is defined by $\sqrt{I} = \{f \in \mathbb{F}[x_1, \ldots, x_n] : f^m \in I \text{ for some } m\}$. Clearly, $I \subset \sqrt{I}$. Recall that an ideal is said to be **radical** if $I = \sqrt{I}$.

Suppose we take two ideals $I_1 \subset I_2$. Since I_2 is larger, it places more restrictions on the corresponding variety, so we get that $V(I_2) \subset V(I_1)$. The same actually holds true in the opposite direction, too. (Check this!)

So, to sum up where we are so far, we have inclusion-reversing correspondences between ideals and varieties that are not bijective.

2.4.3 Going a bit further

We can say more if we don't think about general ideals and varieties. For example, suppose we restrict to the case of radical ideals (over an algebraically closed field) and their corresponding varieties. Then $I = \sqrt{I}$, so we have $I = I(V(I))$ and $V = V(I(V))$, so we have a bijection!

More can be said along these lines. When we look back to our example with a point and a line, we want to say that there are two components – a line and a point. How can we make that precise?

Definition 2.5. *A variety is* **irreducible** *if it cannot be written as the finite union of strictly smaller varieties.*

Great, but how do we *know* that something is irreducible? For example, what keeps us from writing the y-axis as the union of the positive, zero, and negative parts? Couldn't those be varieties? We need to go back to algebra:

Definition 2.6. *An ideal I is* **prime** *if either $f \in I$ or $g \in I$ anytime $fg \in I$.*

Proposition 2.2. *An affine variety V is irreducible if and only if $I(V)$ is prime.*

Proof sketch (straight from [CLO1]): Assume V is irreducible. Let $fg \in I(V)$, and let $V_1 = V \cap V(f)$ and $V_2 = V \cap V(g)$ (which are varieties since intersections of varieties are varieties). Since $fg \in I(V)$, we have that $fg(z) = 0$ for all $z \in V$. So, for any point $z \in V$, we have that $z \in V_1 \cup V_2$ (since $fg(z) = 0$ means that $f(z) = 0$ or $g(z) = 0$. Conversely, $V \subset V_1 \cup V_2$ by definition of those two varieties, so that $V = V_1 \cup V_2$. V is irreducible, though, so, WLOG, assume that $V = V_1$. Then f vanishes throughout V, meaning that $f \in I(V)$.

In the other direction, assume $I(V)$ is prime and let $V = V_1 \cup V_2$. WLOG, assume that $V \neq V_1$. WWTS that $V = V_2$, so we'll show that $I(V) = I(V_2)$ (since, then, we can apply the $V()$ mapping). $V_2 \subset V$ (by definition), and V_1 is a proper subset of V (meaning that $I(V)$ is a proper subset of $I(V_1)$. Choose $f \in I(V_1) - I(V)$ and $g \in I(V_2)$. $V = V_1 \cup V_2$ tells us that fg vanishes on all of V. But I prime forces f or g to be in $I(V)$. By definition, f is not in $I(V)$, so $g \in I(V)$, meaning that $I(V) = I(V_2)$. □

Every prime ideal is also radical, so this proposition gives us that there is a one-to-one correspondence between prime ideals and irreducible varieties.

But wait, there's more! Every variety can be decomposed into a union of irreducible components. Even better:

Proposition 2.3. *Every variety has a* minimal *irreducible decomposition (where no variety may be contained within another) which is unique (up to ordering of the union).*

Because of our propositions above, we also have the following:

Proposition 2.4. *(Over an algebraically closed field) every radical ideal of a polynomial ring has a unique decomposition as an intersection of prime ideals, no one of which is contained in another. All of these statements are proven in Chapter 4 of [CLO1] in case you are curious.*

The previous proposition holds *only* for radical ideals. What about general ideals? Can we decompose them into an intersection of primes?

Unfortunately, the intersection of primes is always radical, so no such luck. We need to go one step further and define *primary ideals*:

Definition 2.7. *An ideal I is* primary *if $fg \in I$ implies that either f or g^m is in I, for some m.*

Primes are primary, and, if I is primary, then $\sqrt{I} = P$ is prime and is the smallest prime containing I. In that case, we call I **P-primary.**

Proposition 2.5. *Every ideal can be written as a finite intersection of primary ideals.*

The previous proposition holds for all Noetherian rings, so it includes all polynomial rings over any field, including finite fields. There is some sense of minimality, too, but this is enough for us. If you toss in minimality, you get the Lasker-Noether Theorem.

To finish the algebraic picture, an ideal is **maximal** if it is proper but not contained in any larger proper ideal. These are nice:

Proposition 2.6. *Over an algebraically closed field, maximal ideals always have the form $< x_1 - a_1, \ldots, x_n - a_n >$ (and vice versa), i.e., maximal ideals correspond to points. Only the opposite direction holds if the field is not algebraically closed.*

To see that this doesn't hold over $\mathbb{R}$, for example, consider the ideal generated by $x^2 + 1$. Over $\mathbb{R}$, the variety is empty so it seems believable that this ideal could be maximal. Moreover, think of it over $\mathbb{C}$: it is contained in the maximals $< x + i >$ and $< x - i >$, which obviously don't exist over $\mathbb{R}$.

So the picture is that our ring (sitting at the bottom of some tree which we will describe) is contained in (and is equal to the intersection of) a set of primary ideals. The radical of each of these primaries is a so-called associated prime. Some of these could be maximal, though that isn't necessary.

Proposition 2.7. *The set of primes associated to ideal I is uniquely determined, i.e., it doesn't depend on a choice of primary decomposition.*

There is nothing saying that associated primes cannot sit inside one another. In fact, we have the following definition:

Definition 2.8. *If associated prime P contains no other associated primes, it is considered a* **minimal prime**. *Otherwise, it is an* **embedded prime**. *Moreover, we define the primary component associated to a minimal/embedded prime to be minimal/embedded itself.*

There is a lot to keep straight here. In our polynomial ring, we have general ideals, some of which are primary. Among the primaries, we have primes. Among those primes, some are maximal and, if we are considering

the decomposition of a particular ideal, some are minimal while others are embedded. The terms *minimal* and *maximal* both come from their role on the algebraic side (obviously). Since embedded primes contain other associated primes, the variety of an embedded prime (or primary) is contained in the variety of some other component, i.e., it is embedded. Thus, the term *embedded* reflects the geometric interpretation.

Example 2.5. *Think about the ideal* $< x > \subset \mathbb{F}[x, y]$. *It is pretty clearly prime but not maximal. Each point* $(0, a)$ *on the y-axis has maximal ideal* $< x, y - a >$, *which clearly contains the ideal* $< x >$. *However, these are not* embedded *primes since they are not* associated.

Alright, so primes pick out our irreducible components for us. What extra information do we get from looking at primaries rather than primes? To answer this, we need schemes.

2.4.4 Did you say schemes?

Yep. But don't worry - we won't go too far down this path. In fact, let's just say that primary ideals contain "derivative information" that is missed by primes ideals. Also, let's take a look at an example or two (straight out of Eisenbud's nice book on Commutative Algebra).

Think about the variety $I =< x, y >$ in $\mathbb{R}[x, y] = R$. This is just the origin – not so exciting. In fact, I is prime (since it's maximal), so we really do get *just* the origin. What primary ideals are associated to I? Here are a few, along with a brief (and incomplete) geometric interpretation:

Example 2.6. $J =< x^2, y >$. *For some reason (we don't have time to get into it, at least for now), we look at* R/J, *i.e., the quotient ring. A general element* $f \in R$ *has all monomials in* x *and* y *with coefficients in* $\mathbb{R}$. *The class of* f *in* R/I, *though, is of the form* $a + bx$ *since all other monomials are killed by* J *(check this). Note that* $a = f(0, 0)$ *and* $b = \frac{\partial f}{\partial x}(0, 0)$. *Thus, we are led to say that although the set-theoretic geometric object for* J *is just the origin, the scheme-theoretic geometric object is the origin plus a smidge in the* x *direction (corresopnding to the function value and the derivative with respect to* x, *respectively). The picture we would draw is the origin with a tiny arrow pointing in the positive* x *direction.*

Example 2.7. $J =< x^2, xy, y^2 >$. *Using the same kind of analysis as above, we get the first partial in both the x and y directions. We could draw an arrow in each direction, but, in this case, we usually just draw a fuzzy little dot around the origin. This is referred to as the first infinitessimal neighborhood at the origin.*

Example 2.8. *Let J be given by all 11th degree monomials in x and y. Then we get all partials up to order 10 and draw a bigger dot at the origin (this seems silly, I know). This is the 10th infinitessimal neighborhood.*

Now let's move to $I =< x >$ in the same ring. This is prime and gives us the y-axis as the variety.

Example 2.9. $J =< x^2 >$. *This is just like the first example above, except that we get the tangent in the x direction for all points on the line, i.e., a fuzzy line.*

Now what if we choose two different primary decompositions for the same ideal?

Example 2.10. $I =< x^2, xy >=< x > \cap < x^2, xy, y^2 >=< x > \cap < x^2, y >$. *Either way, we get the y-axis plus the first infinitessimal neighborhood at the origin. Check this....*

That's all we're going to say about this, at least for now. Would you like to know more? If so, Eisenbud's *Commutative Algebra with a View toward Algebraic Geometry* is very easy to read and informative. The two Cox, Little, &O'Shea books are nice, too, though they do not delve quite as deep. Decker & Schreyer also discuss these ideas at the beginning of their book. If you really like commutative algebra, Atiyah–MacDonald is canonical, as is Matsumura.

2.5 Big outstanding questions

Gröbner basis methods are going to let us answer some big questions that would otherwise be more difficult (if possible at all). Here are a few:

1. Given some ideal, what is the corresponding variety? This is a big one, and it will be essential during much of the course.

2. How can we tell whether some polynomial is in a given ideal?

3. How can we tell whether two ideals are identical?

Chapter 3

Development of Gröbner Bases

As we will see shortly, the ideal membership problem (our focus for the next short while) is solved via simple division in the univariate setting. It would be great if that were the case in the multivariate setting, too, though there are a couple confounding factors. We'll hit those on our way to Gröbner bases, one of the main goals for this course. In the next chapter, we'll take a look at how to compute (and recognize) Gröbner bases. Cox, Little, and O'Shea give one of the best treatments of the development of Gröbner bases that I have seen, so I will just give the highlights here. For a nice treatment related to this chapter and the next couple, check out Chapter 2 of [CLO1]!

3.1 Univariate division

Welcome back to high school; today, we'll be doing long division of univariate polynomials. Formatting division is a headache, so I'll just talk you through an example.

Example 3.1. *Let's divide* x^3+3x^2-8 *by* $x+5$*. First, notice that we order the monomials by decreasing degree. Now, to start dividing, we compare leading terms and figure out what the first term of the quotient must be. In this case, it's* x^2*. Notice that this depended only on the leading terms. Now we multiply* x^2 *by* $x+5$ *and subtract this from* x^3+3x^2-8*, yielding* $-2x^2-8$ *(notice again that we have the monomials in a particular order). Again, we check leading terms.... In the end, we have that* $x^3+3x^2-8 = (x^2-2x+10)(x+5)-58$*,*

i.e., the remainder is 58.

The point of this example is to notice our assumptions – we have a term order, we pay attention to leading terms only, and we end up with a quotient and a remainder. Moreover, we can write $f = qg + r$ for any polynomials f and g, where $r = 0$ or $deg(r) \leq deg(g)$.

Even better, this division method is an *algorithm*: It is correct (gets the right answer) and terminates in finite time (since the degrees of the subresults keep dropping). Also, it is *deterministic* since multiple runs with the same input will certainly guarantee the same output (there are no choices to be made anywhere). Finally, there is a *certificate* that it is correct – we can just check the answer at the end by simple arithmetic.

What happens in the multivariate case? We'll get to that soon, though the basic requirements are similar (with one extra wrinkle about remainders...). First, though, let's see what division does for us in terms of our big questions (membership and solving).

3.2 Univariate ideal membership

First, notice that univariate polynomial rings are principal ideal domains (i.e., each ideal is principal, i.e., each ideal can be generated by a single generator). Indeed, for any (nonempty) ideal I, let $g \in I$ have minimal degree (among polynomials in I). Then, for any $f \in I$, we can write $f = qg + r$ with $r = f - qg \in I$ of degree smaller than g. But g was chosen with minimal degree, so $r = 0$, meaning that $g = qg \in < g >$.

So, to decide if a polynomial f is in some ideal I, just find a generator g s.t. $I =< g >$ and see if $g|f$ evenly (a remainder of zero). Indeed, $f \in I =< g >$ iff $f = qg$ for some polynomial q.

Now all that remains is to cook up a generator for any given ideal I.

The GCD h of polynomials f and g is characterized by the facts that (1) h divides both f and g and (2) any other divisor of both must also divide h.

Proposition 3.1. *GCD(f,g) exists, is unique up to a constant, can be calculated somehow, and generates* $< f, g >= I$.

Most of this should be fairly clear. The Euclidean algorithm yields the

GCD. As for generating I, we know that there is a polynomial h (more than one, actually) such that $I =< h >$. The claim is that $h = GCD(f, g)$. As proof, notice that h divides both f and g, so that condition (1) of the definition of the definition of the GCD is met. Suppose p is some other common divisor, so that $f = q_1 p$ and $g = q_2 p$. Since $h \in I =< f, g >$, $h = c_1 f + c_2 g = (c_1 q_1 + c_2 q_2)p$, so p divides h. □

All of these concepts extend up to multiple univariate polynomials. In particular, the GCD of a set of (more than two polynomials) can be found be finding the GCD of any 2, then the GCD of that GCD and another polynomial, then the GCD of *that* GCD and another polynomial, etc.

Example 3.2. *Is* $p = x^4 - 8x^3 + 22 * x^2 - 24 * x + 9 \in I =< x^3 - 5x^2 + 7x - 3, x^2 - 7x^2 + 15x - 9, x^3 - 6x^2 + 11x - 6 >$*? To decide this, we first find the GCD of the generators of the ideal. In Maple, just include the* `Algebraic` *package and use the command* `Gcd(Gcd(f,g),h);` *(where* f*,* g*, and* h *are the generators of* I*). That command yields* $x^2 - 4x + 3$*, and division of* p *by this GCD yields a remainder of 0 (via* `Remainder(f,gcd1)`*). Note that a nonzero remainder* guarantees *that the polynomial is not in the ideal.*

3.3 Solving univariate polynomial "systems"

Ideal membership is interesting and has its uses, but, in the real world, solving polynomial systems is arguably more important. How do we solve univariate polynomial systems?

Obviously, we can move from any given set of generators to a single generator via the GCD, as described in the previous section. So we need to solve single, univariate polynomials. You probably know some things about this:

- $f = 0$: The solution set is $x \in \mathbb{F}$.
- $deg(f) = 0$ and $f \neq 0$: No solution.
- $deg(f) = 1$: $f = x - a$, so $x - a$ is the unique solution.
- $deg(f) = 2$: Quadratic formula.

- $deg(f) = 3, 4$: There are formulas. The MacTutor History of Math page (at St. Andrew's in Scotland) has a great write-up about the characters who first discovered these formulas.

- $deg(f) \geq 5$: Out of luck (symbolically). Abel supposedly gave the first proof of this, though (according to MacTutor) Ruffini certainly played a role....

So what do we do about solving general univariate polynomials? Note first of all that there are no more than d distinct complex (or, therefore, real) solutions. Thus, if we find d, we know that we are done.

Method 1: Newton's method (for real roots)

We've all seen this one before. Just grab a host of starting points and plug them all into the Newton machine. We'll talk more about Newton's method when we get to numerical methods, but there are a few things to notice. First, not all starting points converge. In fact, try to find the roots of $x^2 - 1$ starting from 0.5, 3, and 12. You will see that the first converges to 1 quickly (quadratically!), the second converges immediately (exactly, in one step) to -1, and 12 diverges. Well, the last is at 10^{24} after 5 steps. Also, note that singular solutions kill the quadratic convergence of Newton's method (but we'll worry about that sort of statement later...).

Unfortunately, this doesn't cut it for us. Without knowing *a priori* how many real solutions a polynomial has, there is no guarantee that we can find all of them. Indeed, if there are less than the degree, we will not know that and will (*ad infinitum*) choose more and more starting points from which to try Newton's method....

Method 2: Eigenvalues of companion matrices

Let $f = a_d x^d + a_{d-1} x^{d-1} + \ldots + a_0$. The following matrix is called the *companion matrix* of f, denoted C_f:

$$\begin{pmatrix} 0 & 0 & \ldots & 0 & -\frac{a_0}{a_d} \\ 1 & 0 & \ldots & 0 & -\frac{a_1}{a_d} \\ 0 & 1 & \ldots & 0 & -\frac{a_2}{a_d} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -\frac{a_{d-1}}{a_d} \end{pmatrix}$$

It's then fairly easy to prove that the solutions of f are precisely the eigenvalues of C_f (and vice versa). Check out Sturmfels' CBMS book for more on this....

Method 3: Sturm sequences and bisection (for real roots)

This one is also out of Sturmfels' book and is probably not as well known (judging from my experience). Define $p_0 = f$, $p_1 = \frac{\partial f}{\partial x}$, and $p_i = -\text{rem}(p_{i-2}, p_{i-1})$ for $i \geq 2$. This is the *Sturm sequence* for polynomial f.

Theorem 3.1. *Let $[a, b]$ be an interval in $\mathbb{R}$ such that $f(a) \neq 0$ and $f(b) \neq 0$. Then the number of real roots in the interval is $N(a) - N(b)$ where $N(x)$ is the number of sign changes in the Sturm sequence when evaluated at x.*

Thus, given an interval, we can count the number of real roots in the interval. Bisecting the interval then tells us how many roots lie in each half-interval. From there, we can continue bisecting and dropping any intervals that contain no roots. Eventually, the diameter of the intervals will be less than ϵ (as long as $\epsilon > 0$).

Example 3.3. *Consider $x^2 - 1$ on $[-2, 2]$. At $x = -2$, the Sturm sequence evaluates to 3,-4,2,0 ($N(-2) = 2$) and to 3,4,2,0 at $x = 2$ ($N(2) = 0$), so there are $2 - 0 = 2$ roots in the interval.*

Before moving on to multivariate considerations, it is worth recalling Descartes' Rule of Signs. In particular, the number of positive real roots of polynomial f is bounded (above) by the number of sign changes among its coefficients. Plugging in $-x$ for x, we get the same result for negative real roots, and zero can of course also be a root. Putting those three facts together, we have that a degree d univariate polynomial with m monomials will have no more than d distinct real roots (thanks to the fundamental theorem of algebra) and no more than $2m-1$ real roots (thanks to Descartes' Rule of Signs). As we may see later, these bounds in some way generalize up to multivariate polynomials in the form of the Bézout bound and fewnomial bounds.... On to multivariate stuff....

3.4 Multivariate division, round 1

Is $f = x^2y + xy^2 + y^2 \in \mathbb{R}[x, y]$ in the ideal generated by $g_1 = y^2 - 1$ and $g_2 = xy - 1$ (thanks, again, [CLO1])? Using the univariate case as our

precedent, let's divide. Wait! How do you divide f by two polynomials? I suppose you could divide by one and then the other, but what if the result of this is then again divisible by g_1? For that matter, even if we do decide in what order to divide, how do we choose leading terms (since division involves comparing leading terms)?

We need to choose a reasonable order for the monomials in our polynomial ring. But what would make a term order reasonable? Well, here are a few things:

1. Our order should compare any two monomials, i.e., we should never be unsure which term of our polynomial is the leading term because two terms are incomparable.
2. We'll be multiplying monomials together during division, so we may want orders to be maintained by multiplication. In other words, if $x^\alpha < x^\beta$ (where α and β are n-tuples of nonnegative integers), then we want $x^\alpha x^\gamma < x^\beta x^\gamma$ for any n-tuple of nonnegative integers γ that we choose.
3. For various algorithmic reasons (i.e., proofs of termination), we'll want any set of monomials to have a minimal element.

Now we can define what is called a monomial or term order.

3.5 Monomial orders

Definition 3.1. *A* monomial order *(or* term order*) for the monomials of* $\mathbb{F}[x_1 \ldots, x_n]$ *must satisfy the three conditions given above, i.e., it must be a total ordering, it must respect multiplication by monomials, and it must be a well-ordering.*

Actually, the third requirement is a consequence of the other two (though you don't necessarily see that *a priori*), so it is enough for an order to satisfy the first two conditions. Also, [CLO1] covers this much more nicely and in more detail....

Also, please note that a monomial order requires a variable order, too. A monomial order using the variable order (x,y) could (will?) be different than the same monomial order with the variable order reversed....

Example 3.4. *The lex term order (lexicographic) is perhaps the easiest to understand. To compare two monomials, first look at the exponents of the first variable. If they are different, then the monomial with the larger first exponent is larger. If there is a tie, move to the second variable and compare those exponents. If all exponents are the same, then the monomials are the same! For example, we have* $x^2y > xy^2$ *(assuming that we set* $x > y$*) because the exponents for* x *are* $2 > 1$*. If we flip the variable order, then the conclusion on this example flips, too. As another example,* $xy < xy^2$ *because the exponents on* x *are the same but those for* y *are* $1 < 2$*.*

Example 3.5. *grlex (graded lex, or degree lex) is probably the second easiest to understand. To compare two monomials, you first compare their total degrees (the sum of all exponents). If they differ, then you draw the obvious conclusion. Otherwise, you use lex as a tiebreaker. For grlex (with* $x > y$*),* $x^2y > xy^2$ *since the degrees are the same and lex orders them this way. However, lex would set* $x^2y > xy^8$ *whereas grlex goes the other way since* $3 < 9$*.*

Example 3.6. *grevlex (graded reverse lex) is the nastiest of the standard three monomial orders. Again, the first comparison is by degree, just like with grlex. BUT the tiebreaker is reverse lex – you look at the last variable (instead of the first) and favor the smaller exponent (rather than the larger). So it's actually graded double reverse lex (sounds like figureskating!). For example,* $x^2y > xy^2$ *because they have the same total degree and the* y *variable has exponents* $1 < 2$ *(and the monomial order flips this exponent order). Keep in mind, though, that grading (degree) is always the first thing to check.*

3.6 Multivariate division, round 2

Now that we have monomial orders under our belts, we can try dividing again. First, though, let's fix some terminology:

Definition 3.2. *Let* $f = \sum a_\alpha x^\alpha \in \mathbb{F}[x_1, \ldots, x_n]$ *with a fixed monomial order. Then the* multidegree *of* f*, multideg*(f)*, is just the largest exponent appearing in* f *(an* n*-tuple). The* leading coefficient *(*$LC(f)$*),* leading monomial *(*$LM(f)$*), and* leading term *(*$LT(f)$*) of* f *are the coefficient, monomial, and term (coefficient times monomial) that go with the term having the largest degree (i.e., the multidegree), respectively.*

OK, let's divide something. Let's divide $xy^2 + 1$ by $xy + 1$ and $y + 1$ (in that order, using the lex order with $x > y$). To write this down, just stack the two divisors where you usually put the divisor. Also, there will be two quotients – one for each divisor – that sit atop the dividend (where the quotient usually goes). We'll proceed as usual, checking whether the leading term of a divisor divides the leading term of the dividend (beginning with the first divisor and working our way down). In this case, $xy|xy^2$, so we put y in the first quotient, multiply, and subtract (like usual). This leaves us with $-y + 1$. $xy = LT(xy + 1)$ doesn't divide $-y$, but $y = LT(y + 1)$ does, so we put -1 in the second quotient, multiply, and subtract, leaving us with 2. Obviously, neither leading term divides 2, so we conclude that $xy^2 + 1 = y(xy + 1) - (y + 1) + 2$.

If we knew that multivariate division solved the ideal membership problem, we'd be golden. Unfortunately, we don't (yet).

Let's try another one (for a good reason – not just torture). Let's divide $x^2y + xy^2 + y^2$ by $xy - 1$ and $y^2 - 1$. (This is one of many places where you are far better looking in [CLO1]!!). After the first two steps, we are left with the partial dividend $x + y^2 + y$. Neither xy nor y^2 divide x, so it is tempting to say that we are done. HOWEVER, y^2 (the leading term of the second polynomial) divides y^2 (the second term of the partial dividend), so we should take advantage of that. Thus, we set up somewhere a list of things in the remainder and add x to it, leaving behind $y^2 + y$. We can that add a 1 to the second quotient, multiply, and subtract, leading us to the next partial dividend, $y + 1$. Now, neither leading term of a divisor divides either of the terms of $y + 1$, so we officially say that we are done and end up with $x^2y + xy^2 + y^2 = (x + y)(xy - 1) + 1(y^2 - 1) + (x + y + 1)$.

All told, in dividing f by some set of polynomials g_i, we end up with $f = \sum q_i g_i + r$ where the remainder r is 0 or is built from monomials, none of which are divisible by a leading term of a g_i. That's the good news.

The bad news is that if you swap the order of the divisors in that last example, you get a different remainder – $2x + 1$! So, as things stand, the division algorithm doesn't solve the ideal membership problem. Bummer.

Wouldn't it be nice if we could cook up a "nice" set of generators for an ideal so that division by these generators would yield unique remainders (and thereby solve the ideal membership problem for multivariate polynomial rings)?

3.7 Monomial ideals

Definition 3.3. *An ideal is a* monomial ideal *if it has a generating set consisting only of (possibly infinitely many) monomials.*

Not a shocking definition, but it will be a useful concept. Note that a monomial is in a monomial ideal iff some generator divides it evenly.

Note, too, that although the Hilbert Basis Theorem (discussed in Chapter 2) guarantees the existence of some finite generating set, there is no guarantee from that theorem that there is a finite generating set consisting only of monomials! More on that in a moment....

Ideal membership in a monomial ideal is a piece of cake. In particular, for a monomial ideal I, $f \in I$ iff every term of f is in I iff f can be written as a linear combination (i.e., the coefficients are numbers from the field!) of ideal elements. So, to check whether $f \in I$, just look at each monomial! But wait, there is no guarantee that we have a finite generating set consisting only of monomials, so this criterion is perhaps not so practical, right?

Wrong. Dickson's Lemma says that monomial ideas are finitely generated by monomials. In fact, in [CLO1], they first prove Dickson's Lemma and then prove the Hilbert Basis Theorem as a corollary. So, we now have an ideal membership *algorithm* for monomial ideals: $f \in I$ iff $Rem(f|g_i) = 0$ for some generating set g_i of I.

Fine, but not all ideals are monomial. In fact, most interesting ones aren't. What do we do in general??

3.8 Leading term ideals and Gröbner Bases

Definition 3.4. *Let I be a nonzero ideal in a multivariate polynomial ring over a field. $LT(I)$ is the set of all leading terms of polynomials in I, and $< LT(I) >$ is the ideal that they generate.*

Notice that $L =< LT(I) >$ is a monomial ideal. That means that it has a finite generating set of monomials. HOWEVER, be careful – the leading terms of a provided set of generators of I need not generate L! Indeed, let $f = x^3 - 2xy$ and $g = x^2y - 2y^2 + x$. Then $xg - yf = x^2$ is in I, meaning that x^2 is in L, but x^2 is not divisible by either $LT(f)$ or $LT(g)$! We'll always

have that the ideal generated by the leading terms of the generators of I is contained in L, but the reverse containment is not guaranteed.

OK, so not every set of generators for I can lead to a generating set of L (by taking leading terms). But there must be at least one set of generators that can (from the last section), so let's give them a name:

Definition 3.5. *Under a fixed monomial order, a set of generators of an ideal is a* Gröbner basis *for the ideal if the leading terms of the generators generate L.*

That's it. That's all that a Gröbner basis is – just a nice set of generators for an ideal. Granted, we don't know how to cook one up for a given ideal. We don't even know how to check that a set of generators is a Gröbner basis . We'll get to those points in the next section. Before we go, though, note that every nonzero ideal in a multivariate polynomial ring has a Gröbner basis .

3.9 The ideal membership problem

Gröbner bases are the bases that we were talking about back in §3.6. They are the nice bases which give us unique remainders when dividing.

In a multivariate polynomial ring $R = \mathbb{F}[x_1, \ldots, x_n]$, given a Gröbner basis $G = \{g_1, \ldots, g_k\}$ of an ideal $I \subset R$ and a polynomial $f \in R$, we know that there is a unique polynomial $r \in R$ having no term divisible by the leading terms of the g_i and there is some $g \in I$ such that $f = g + r$. This is pretty straightforward to prove – g is what it needs to be, and r is unique because otherwise, the leading term of the difference of any two (distinct) remainders is divisible by some $LT(g_i)$ (since the difference is in I). But we assumed that no term of any remainder is divisible by any of these $LT(g_i)$, so the difference is 0.

Since the remainder upon dividing a polynomial f by a Gröbner basis is unique, it is reasonable to name it – we call it the *normal form* for f. Note, however, that the quotients when performing such a division need *not* be unique. Of course, this doesn't matter – we don't care *which* linear combination of the elements of a Gröbner basis yield $f - r$. In fact, we will typically only care whether the remainder is zero. Since this normal form will come up from time to time, it makes sense to have some notation for it:

Definition 3.6. *$R(f, G)$ is the remainder of f when divided by a Gröbner basis G. Note that this differs from the notation in [CLO1].*

Finally, we can solve the ideal membership problem for multivariate polynomial rings: $f \in I$ iff the remainder when dividing f by a Gröbner basis for I is zero (i.e., if $R(f, G) = 0$. The proof is pretty obvious, given the uniqueness of the remainder.

The problem is that we don't know how to compute a Gröbner basis for a given ideal. We don't even know how to check whether a set of generators is a Gröbner basisẆe'll cover those topics (and some complexity concerns) next.

Chapter 4

Computation of Gröbner bases

As we'll see soon, Gröbner bases are handy computational tools. In fact, for a while (and still in some settings), they were/are the only option for computation. As a result, there has been a fair bit of thought put into their computation, though, like everything, there are some limitations imposed on their effectiveness due to the complexity needed to compute them.

In this chapter, we will talk about the standard way to compute a Gröbner basis some of the related computational concerns, and some of the responses that have been cooked up in response to these concerns. As above, we'll follow [CLO1] as far as we can with this, though we will start branching out at the end of this chapter.

4.1 Recognizing Gröbner bases

Given some ideal I and a potential Gröbner basis H, how do we check whether H is indeed a Gröbner basis for I? Well, what does it mean to be a Gröbner basis? It means that the leading terms of H generate the ideal generated by all leading terms of polynomials of I. How can we miss things?

Well, one obvious way (as we saw in §3.8) is that the leading terms can be cancelled out in some two term polynomial combination. In particular, it should be easy to cook up α, β, a, and b so that the leading terms of h_1 and h_2 cancel in the polynomial combination $ax^\alpha h_1 - bx^\beta h_2$. For example, in §3.8, we had that $x^2 = xh_1 - yh_2$, where $h_1 = x^2y - 2y^2 + x$ and $h_2 =$

$x^3 - 2xy$. (Note that this is *exactly* Ex. 2 of 2.5 in [CLO1].) In this example, $a = b = \alpha = \beta = 1$. x^2 is clearly the leading term of x^2, and it is certainly not in the ideal generated by the leading terms of h_1 and h_2 (at least using lex!). So, h_1 and h_2 do not constitute a Gröbner basis for the ideal that they generate.

In fact, the a, b, α, β to get leading terms to cancel is *so* straightforward, let's write them down. The following is known as the *S-polynomial* for two other polynomials f and g:

$$S(f,g) = \frac{x^\gamma}{LT(f)} f - \frac{x^\gamma}{LT(g)} g,$$

where $\gamma \in \mathbb{Z}^n_{\geq 0}$ is given by $\gamma_i = max(\text{multideg}(f)_i, \text{multideg}(g)_i)$ and $LT(\cdot)$ includes the leading coefficient (as usual). By the way, the "S" stands for "syzygy," a fancy word for alignment (particularly in astronomy and other algebraic geometry).

Again, let's check out an example from [CLO1]: Consider the polynomials $f = x^3y^2 - x^2y^3 + x$ and $g = x^4y + y^2$ using the graded lex order. $\gamma = (4, 2)$, so we get:

$$S(f,g) = \frac{x^4y^2}{x^3y^2} f - \frac{x^4y^2}{x^4y} g = -x^3y^3 + x^2 - \frac{1}{3}y^3,$$

meaning that we don't have a Gröbner basis.

[CLO1] nicely proves that this is the only way to have leading terms cancel out. Given some set of polynomials $h_1, \ldots, h_k$ all with the same multidegree, they prove that any polynomial combination of the h_i with lesser multidegree *must be* a linear combination of polynomials of the form $S(h_i, h_j)$.

Finally, based on this, we get the criterion originally provided by Buchberger for checking whether some basis is indeed a Gröbner basis. In particular, a potential generating set H for an ideal I is indeed a Gröbner basis for I iff $R(S(h_i, h_j), H) = 0$ for all $i \neq j$! The forward direction is obvious – S-polynomials are in the ideal, so division by a Gröbner basis is automatially zero (as described above). In the opposite direction, we need to take $f \in < h_1, \ldots, h_k >$ and show that $LT(f) \in < LT(h_1), \ldots, LT(h_k) >$. f is a linear combination of the h_i, and, in looking at the multidegree of f and the multidegrees of the summands of the linear combination, there is either equality (at which point division to a zero remainder is possible) or there is

inequality. In the latter case, there is a lot of work to do...this might make a good "extra topic" by me or, even better, a talk by one of the students....

For kicks, try out $< x - z, y - z >$ with lex(x,y,z). Is it a Gröbner basis for the ideal it generates? What about $< y - x^2, z - x^3 >$ with lex(y,z,x) (NOTE THE ORDER!)? What about with lex(x,y,z)?

4.2 Computing a Gröbner basis from scratch

Given Buchberger's S-criterion above, Buchberger's algorithm for computing a Gröbner basis is fairly straightforward. In particular, given an ideal $I =< f_1, \ldots, f_k >$:

1. Let $G = f_1, \ldots, f_k$.
2. Set NUMNEW to 0.
3. Form *all* S-polynomials $S(g_i, g_j)$ for all pairs of polynomials in G.
4. Compute $R(S(g_i, g_j), G)$ for each S-polynomial.
5. Increment NUMNEW for each remainder that is nonzero and throw it into G.
6. If NUMNEW is 0, then G is a Gröbner basis for I. Otherwise, go to 2.

Try this on $x^3 - 2xy$ and $x^2y - 2y^2 + x$ with graded lex. Remember that $-x^2$ was the first remainder when dividing an S-polynomial by the existing partial basis. You should end up with five generators....

So does this algorithm actually compute a Gröbner basis for I?? There are really two main considerations here – whether it is correct (and does produce a Gröbner basis for I) and whether is terminates on all input (no infinite loops). For the former, it should be clear, thanks to Buchberger's S-criterion above. As for termination in finite time, first note that each trip through the loop is clearly finite – it's the number of loops that is in question. The key here is to notice that we add fewer new S-polynomials each time (this takes a little reasoning) so that, considering the ideal J_k of S-polynomials added to G at stage k, we end up with an ascending chain of ideals (larger

and larger ideals). Noetherian rings satisfy the ascending chain condition (that all ascending chains of ideals are of finite length), so we eventually stop adding more to G, i.e., the algorithm terminates. The connection between the ascending chain condition and the Hilbert Basis Theorem could also be the subject of a talk....

Though it rather belongs in the next section, it should be noted that a Gröbner basis as computed with the previous algorithm might be *too big*, meaning that some of the generators are redundant. In fact, if $g \in G$ with $LT(G) \in < LT(G - p) >$, then $G - p$ is also a Gröbner basis for I.

Dumping all extra generators and making all polynomials monic (we are working over a field here!), we get a so-called *minimal Gröbner bases*. Unfortunately, this does not yield a unique Gröbner basis for I. Page 89 of [CLO1] has a nice example of this.

If we also require that, for all polynomials $g \in G$, no monomial of g is in $< LT(G - g) >$, then we get a *reduced Gröbner basis*. These *are* unique (up to term order).

4.3 Some computational concerns

Complexity (still unknown, though people have a good idea...), coefficient blowup (finite fields!).

4.4 Some responses

Chapter 5

Solving polynomial systems exactly

Chapter 6

Real-world polynomial systems

6.1 Kinematics

6.2 Game theory

6.3 Optimal control

6.4 Others

Chapter 7

Basics of numerical methods

Chapter 8

Homotopy continuation, zero-dimensional solving

Chapter 9

Basic numerical algebraic geometry

Chapter 10

Advanced topics in numerical algebraic geometry

Chapter 11

Potential further topics

11.1 Hilbert polynomials and dimension

11.2 Fewnomials and bounds on the number of real solutions

11.3 Toric deformations

11.4 Tropical geometry

11.5 Basic convex geometry – Bernstein and such

Appendix A

Advanced topics for talks

Here are a few ideas for the 20–25 minute talks I am asking you to give at the end of the semester. The list is not complete, and I will generally be open to suggestions (as long as they are at least tangentially connected to the material in the course). Please let me know what topic you pick so that we can avoid redundancy. Also, feel free to swing by my office if you are having trouble deciding.

- Solving a class of polynomial systems in various ways/packages
- Details about complexity of Gröbner bases, from Mayr/Meyer
- Faugere's F4 and F5 algorithms (rough idea)
- Computation on elliptic curves (basic or advanced)
- Syzygies, free resolutions, Betti numbers (definitions and such)
- Parallel computation of Gröbner bases?
- More advanced numerical methods
- Basic toric geometry
- Basic tropical geometry
- CoCoA demo

- Demo of that number theory package that Rachel Pries mentioned – one letter?
- Some other application of algebraic geometry
- Proof of Buchberger's criterion and/or Buchberger's algorithm (as in [CLO1])
- Connect the ACC (ascending chain condition) to what we have discussed