LECTURE 0: INTRODUCTION (V1)

1. Introduction

Computational algebraic geometry. That is the title of the course, and, on the one hand, that title is quite apt. There is a large and varied part of mathematics—more than I could reasonably cover, or even survey, in five lectures—devoted to computation inspired or informed by algebraic geometry. It ranges from the use of combinatorics and polyhedral geometry in toric and tropical geometry, through homological methods based on free resolutions, to numerical and homotopy methods for finding and certifying solutions to particular systems over $\mathbb{R}$ or $\mathbb{C}$, and much more.

However, having finished writing 95% of these notes—the remaining 5% being, well, what I am currently writing—the title also feels like a misnomer. These notes are not about a field separate from algebraic geometry or a niche subfield within algebraic geometry. Instead, I wish to suggest that, in many ways, computation is the spiritual heart of algebraic geometry. All these notes attempt to do is take that viewpoint seriously, with the modern tools at our disposal in mind.

The claim that computation is the spiritual heart of algebraic geometry may seem indefensibly bold. After all, over the past seventy-five years, the field has become known to many outsiders for its abstract and formidable machinery: schemes and sheaves, a plethora of cohomology theories, stacks and derived categories, and, more recently, even more exotic things such as shtukas, homotopical algebra, ∞-categories, etc. So let me take a moment to defend the claim with two arguments.

1.1. **Examples in Algebraic Geometry.** Algebraic geometry seeks to study the shapes cut out by systems of polynomial equations. Such shapes lurk in a rather large and dark forest. The central questions that have driven the modern field ask us to chart that forest: Which shapes exist? How do they vary? Which features are typical? Which apparent possibilities never occur? Faced with a forest this large, it may seem absurd that inspecting a single tree could tell us much. Yet, remarkably, it often can.

An enormous number of these questions—including some of the most celebrated ones—come down, in the end, to a single humble task: *compute something about a single concrete example.* We must not merely describe such an example abstractly, but build it and prove something true about it. What makes this possible is a piece of true magic worth stating up front. For many of the properties we care about, the locus of members having the property is *open* in the relevant parameter space: if one member has the property, then so does every sufficiently nearby member. Provided that parameter space is irreducible, any nonempty open locus is dense, so the property holds for what algebraic geometers call a "general" member. Showing that a property is open, while not always immediate, often seems to come down to finding the right definitions. Of course, the empty set is

open! Thus, once openness is known, to prove that a general object has the nice property we care about—and to be certain that we are not proving a vacuous theorem about nothing—it is enough to construct a specific example with that property. In this sense, one well-chosen example becomes a theorem about a dense open subset of the family. In fact, I do not think I am lying when I say that one could win a Fields Medal simply by constructing the right example.

Once you know this pattern, it is impossible to miss. To convince you that I am not lying—and perhaps to argue from authority by appealing to papers that have appeared in so-called fancy journals—let me briefly list a few of my favorite instances: Brill–Noether theory, Green's conjecture for a general curve, the Maximal Rank Conjecture, and interpolation problems. You are not expected to know these results yet; the common pattern is the point.

1.2. **Abstraction in Algebraic Geometry.** The framing of the previous paragraph deserves to be challenged rather than answered. It supposes that the machinery—sheaves, cohomology, resolutions, derived categories—sits at the opposite end of the subject from computation, so that the best one could hope to show is some historical link between the two. However, the historical record suggests this is quite far from the case.

Take the machinery at its source. Hilbert's theorem that rings of invariants are finitely generated produced no generators, and Paul Gordan—who had spent his career producing generators by hand—is supposed to have complained,"*Das ist keine Mathematik. Das ist Theologie!*" The anecdote is too tidy to be true, and, more importantly, it stops far too early. Pressed on precisely the question of effectivity, Hilbert kept going: the Basis Theorem, the Syzygy Theorem, and the Hilbert function all come out of the effort to get his hands on the invariants he had shown to exist. Faced with a computation he could not finish, he invented graded free resolutions. That is not a retreat from calculation into abstraction. It is what the calculation turned into.

Noether then did to Hilbert what Hilbert had done to Gordan. The Basis Theorem, as Hilbert proved it, is a statement about polynomial rings over a field, and its proof is entangled with the elimination-theoretic setting it came from. Noether asked what the argument had actually used. The answer was: very little. Every ideal finitely generated, or equivalently the ascending chain condition, and nothing else – so that the theorem is better stated as the assertion that this single property passes from R to $R[x]$. A hypothesis extracted from one proof became the definition of a class of rings, and Hilbert's theorem became the statement that the class is closed under adjoining a variable. The path here is the one worth naming:

$$\text{calculation} \longrightarrow \text{pattern} \longrightarrow \text{abstraction} \longrightarrow \text{better calculation},$$

and it is the last arrow that matters for us. Noetherianity is not a piece of bookkeeping we tolerate on the way to the interesting theorems. It is the reason Buchberger's algorithm terminates and the reason free resolutions are finite, which is to say it is the reason that essentially every computation performed in these notes returns an answer at all. Gordan, incidentally, later gave a proof of Hilbert's theorem whose monomial arguments read today like an anticipation of Gröbner bases; the two sides of the quarrel converged on the same object.

The Path Through These Notes. In writing these notes, I have tried—and we shall see to what extent I have succeeded—to weave together a number of threads that appear, disappear, and reappear throughout the lectures. At times, an exercise will deliberately be left hanging, a calculation left unexplained, or a theorem left unproved. This is intentional—or I did not catch it in my hurried proofreading. You will see relatively few proofs presented in the traditional theorem–proof format. Sometimes a proof lies beyond the scope of these notes and is best revisited after you have built more machinery; other times, the proof is buried in the exercises. That, too, is deliberate. If you are reading this, I hope your goal is to learn something, and, well, in my experience, no amount of my knowing something can magically make another person know it, too. (The same statement is even more true with the roles reversed.) Learning requires doing, and I hope these notes serve as a guide while you do that work for yourself. That is not to say that I intend to send you into the woods of algebraic geometry unprepared and unaccompanied. Instead, I have tried to write in the spirit of friends setting out on a hike together: I have walked enough of the trail to point toward the next turn and warn you about the loose rocks, but you still have to walk it yourself.

Concretely, the five lectures follow this route:

- **Lecture 1: Equations and Geometry.** We ask what it could possibly mean to *solve* a system of polynomial equations, a question that turns out to have at least six answers! This leads us to study algebraic varieties, ideals, coordinate rings, the operation of taking the radical of an ideal, and Hilbert's Nullstellensatz. Along the way, our four major recurring players enter the scene: the intersection of a parabola and a circle, a plane conic, the twisted cubic, and an elliptic curve. More about them later.

- **Lecture 2: Division and Hidden Equations.** One-variable division suggests a nonlinear analogue of row reduction, so we try to imitate it. Monomial orders and multivariable division show us why an arbitrary list of generators is not good enough; initial ideals and standard monomials tell us what a better list would have to accomplish; and S-polynomials isolate the one cancellation that seems to build such a list every time we try it.

- **Lecture 3: Finding Gröbner Bases.** Was that luck? Dickson's Lemma delivers finite Gröbner bases and, as a bonus, Hilbert's Basis Theorem; Buchberger's criterion turns an infinite condition into a finite test; and Buchberger's algorithm actually constructs a basis. We then pass to projective space, where a Gröbner basis for a degree-compatible monomial order lets us homogenize an affine ideal without the false components at infinity that naively homogenizing an arbitrary generating set can happily invent.

- **Lecture 4: Questions of Rationality & Interpolation.** Here we reverse the direction of travel and start building. The Elimination Theorem recovers equations for the Zariski closure of a parametrized image; standard monomials turn interpolation on a finite set into linear algebra; and rational and birational maps make precise when two apparently different varieties carry the same rational functions—and when, as for a smooth plane cubic, no

birational parametrization by $\mathbb{P}^1$ exists. Both themes named in the title are entry points into broad research programs in which basic questions remain open.

- **Lecture 5: What an Initial Ideal Remembers.** Finally, we see that what we have been doing with Gröbner bases has an entirely different geometric meaning. We read initial ideals as special fibers of degenerations, count standard monomials using Hilbert functions and Hilbert polynomials, and watch a complicated variety move to a combinatorial model. The point set, smoothness, and components may all change along the way; we ask precisely what does not, because that is what a degeneration argument is allowed to conclude.

1.3. **Using These Notes.** Each lecture builds on its predecessors, and I would not encourage you to read them backwards. At the same time, you should not start at the first page and march steadily forward, moving on only after you have finished and digested every part of every exercise. That is not how these notes are meant to be used, and it is not how I find mathematics is best learned. Instead, I encourage you to try many parts of many exercises. If you stall on an exercise, move forward, do something else, and eventually come back. Let things simmer while you move on: as new ideas appear and the bigger picture comes into focus, earlier exercises that seemed daunting may become clearer.

There are also deliberately more exercises, parts, and remarks than we can cover together. Honestly, I have included far more than I expect anyone to cover and digest deeply in five lectures, in the hope that these notes will serve as a friendly introduction to the subject. Still, if you find the notes at all interesting and continue studying algebraic geometry, I like to imagine that you might, from time to time, come back to them and revisit some old friends and familiar examples in a new light.

When it comes to computation, I encourage you both to compute by hand and to compute with *Macaulay2*. Each is a valuable skill in its own right. *Macaulay2* is free, and a browser-based interface is available, so there is nothing to install before trying it. You can, of course, also download it and run it locally, something you will definitely want to do if you continue using it. For working with *Macaulay2*, I have three suggestions:

(1) Instead of copying and pasting the snippets of code in these notes directly into a session, retype them by hand. Yes, this may be old-fashioned, but the process of typing things out often helps us remember commands, syntax, and structure more deeply than copying and pasting does. Pause over each command and ask what mathematical object it creates or computes.

(2) Consult the documentation frequently. The documentation for *Macaulay2* is written largely by research mathematicians, many of whom work in algebraic geometry, and is a treasure trove of examples, ideas, and papers. It is generally the best way to learn the language.

(3) If you enjoy using *Macaulay2*, consider getting involved! The *Macaulay2* is a vibrant and diverse group of researchers from numerous areas of mathematics. There are frequent

workshops dedicated to building and extending the functionality of the language, and they are just a lot of fun.

1.4. **Apologia.** I must end with a fairly serious warning and an apology. These notes are very, very much a first draft and an experiment. This is both the first time I have taught a course like this and the first time these notes have been used. I am certain that there are mathematical errors, typographical mistakes, idiosyncratic explanations that make sense only to me, and exercises that do not work as I had hoped. These issues and the perhaps peculiar choices scattered throughout the notes are entirely mine. The lectures were written at different times, sometimes weeks or months apart. So I apologize if the tone, voice, or tenor of the writing ever shifts in a way that feels jarring.

In the hope of correcting some of these issues and of saving any future souls who find and use these notes, I have created a Google Form through which you can report any issues you find.

https://forms.gle/eWsQYhPhvA3edpfSA

I would be genuinely grateful for any corrections, comments, or other feedback you might have. When submitting feedback, please keep in mind that it is much easier to act on specific, isolated comments. You are, of course, welcome to use the form to tell me that you hate everything about these notes and that they are useless, but I doubt I will know what to do with that. On the other hand, being told that I missed a sign on the third line of the fourth exercise in the second lecture should result in a speedy and satisfying resolution.

LECTURE 1: EQUATIONS AND GEOMETRY (V1)

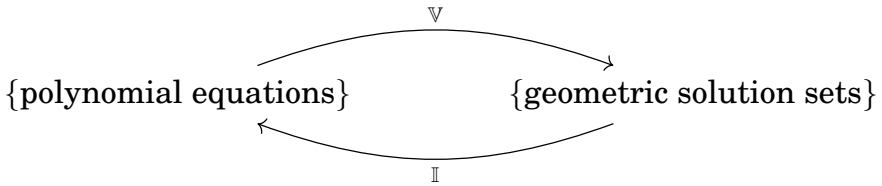
Throughout this lecture, $\mathbb{K}$ denotes a field. The algebra–geometry dictionary we shall build works over an algebraically closed field, but we will often restrict to $\mathbb{R}$ when we wish to visualize things. Further, for reasons we will explain shortly we generally do computations over $\mathbb{Q}$.

1. What Does It Mean to Solve?

The passage between algebra and geometry is familiar to us from high school. Consider the equation

$$y = x^2,$$

the solutions to which determines a subset of the plane $\mathbb{R}^2$, namely its graph. Conversely, if we are shown a graph in $\mathbb{R}^2$, we may try to recover the equation(s) which cut it out. There is something remarkable hidden in this familiar example. Formally the parabola as a subset of $\mathbb{R}^2$ consists of infinitely many points, however, the single expression $y - x^2$ encodes all of them in one piece of finite data. Algebraic geometry takes this familiar exchange seriously:



The two directions lead to different questions. The *forward problem*, which we denote $\mathbb{V}$ for reasons that will become clear shortly, asks us to find the common solution set to a given system of polynomial equations, i.e. we are asked to solve a polynomial system. The *inverse problem*, denoted by $\mathbb{I}$ for reasons that we will see later in this lecture, asks how we can find polynomial equations that cut out a given subset, i.e. we are asked to identify the graph.

The bulk of this lecture is devoted to understanding both directions of this relationship in detail, beginning with the forward problem and then considering the inverse problem. Towards this let us consider a relatively simple system of polynomial equations, which none the less will show that there are several different questions hidden in the instruction *solve the system*.

---

(1) **A Circle & A Parabola.** Consider the following system of two polynomial equations:

$$f_1 = x^2 - y = 0, \qquad f_2 = x^2 + y^2 - 1 = 0.$$

Since f_1 defines a parabola and f_2 the unit circle at the origin we are looking at the intersection of two very familiar shapes.

(a) Substitute the first equation into the second and show that every solution satisfies

$$x^4 + x^2 - 1 = 0.$$

The roots of this new polynomial are the possible x-coordinates of the intersection. What information was eliminated in passing from two equations in x, y to one equation in x? Here the equation $y = x^2$ reconstructs the discarded coordinate. Explain why every root lifts uniquely.

(b) Now work over $\mathbb{R}$. Set $u = x^2$ and solve $u^2 + u - 1 = 0$. Show that exactly one of its roots is nonnegative.

(c) Find the two real solutions of the original system and sketch the picture.

(d) Now work over $\mathbb{C}$. Show that $x^4 + x^2 - 1$ has four distinct roots and conclude that the system has four complex solutions. You need not choose a preferred square root of each complex number: an answer of the form $(\pm\sqrt{u}, u)$ is more informative.

(e) Subtract the first equation from the second to obtain a polynomial involving only y. Explain why this polynomial vanishes at every solution.

(f) Compare what is recorded by the real picture, the complex solution set, and the two one-variable elimination equations obtained in parts (a) and (e). What information does each description retain, and what does it leave out?

(g) A projection can lose information in two different ways. Over $\mathbb{C}$, every nonzero x-coordinate on the curve $y^2 = x$ has two lifts, so the discarded coordinate cannot be recovered uniquely. Now assume $\mathbb{K}$ is infinite and consider the hyperbola

$$H = \left\{(x, y) \in \mathbb{K}^2 \mid xy = 1\right\}.$$

Show that its projection to the x-axis is $\mathbb{K}^\times$. Prove that the only polynomial in $\mathbb{K}[x]$ which vanishes on this image is the zero polynomial. (Recall that a nonzero one-variable polynomial has at most as many roots as its degree.) Thus any common zero set of polynomials in $\mathbb{K}[x]$ which contains the image must be the whole line, even though the origin is not in the image. Here $y = 1/x$ does recover the discarded coordinate on the image; the failure is instead that the image cannot itself be described as a common zero set of polynomial equations. We will recover this distinction by elimination later in the lecture.

The previous exercise suggests that the instruction *solve the system* is a bit ambiguous. For a collection of polynomials $\mathscr{F} \subset \mathbb{K}[x_1, \ldots, x_n]$, there are several different geometric questions one might be asking. The following table captures several more specific versions of what one might hope to do when solving a polynomial system. this one might have when actually trying to solve a system of polynomial equations

Geometric question	**Algebraic form**	***Macaulay2***
Does a solution exist?		
Which equations hold on every solution?		
Which consequences involve only selected coordinates?		
Is the solution set finite or positive-dimensional, and how many points occur when multiplicity is remembered?		
Does the set break into irreducible pieces?		
Which equations vanish on the image of a parametrization or on the closure of that image?		

Throughout this course we will use the word *solve* in each of these senses. Each geometric question has an algebraic counterpart, and our eventual goal is to make these algebraic questions computable. We shall frequently return to this table throughout every lecture, and please keep it in mind.The model is Gaussian elimination, which answers the linear versions of these questions.

---

(2) **Linear Equations - The Gold Standard.** Work over an arbitrary field $\mathbb{K}$ and consider the following system of linear polynomials:

$$\ell_1 = x + y + z - 1, \qquad \ell_2 = 2x - y + z.$$

(a) Row reduce the system $\ell_1 = \ell_2 = 0$ and write its solutions as a one-parameter family.

(b) Show that

$$x - 2y + 1 = \ell_2 - \ell_1, \qquad 3y + z - 2 = 2\ell_1 - \ell_2.$$

Use these equations to recover your parametrization from part (a).

(c) Explain why an elementary row operation replaces a list of linear polynomials by another list generating the same ideal. Conclude, without appealing to the geometry, that row operations preserve the common zero set.

(d) Suppose a linear system in n variables is consistent and its coefficient matrix has rank r. What is the dimension of its solution set? What comparison between the coefficient matrix and the augmented matrix detects inconsistency?

(e) Consider instead the linear equations

$$m_1 = x + y, \qquad m_2 = 2x + 2y - 1.$$

Verify the identity $1 = 2m_1 - m_2$. Explain why this identity is an algebraic certificate that the system $m_1 = m_2 = 0$ has no solutions over *any* field.

(f) Start now with the parametrized line

$$L = \{(2t - 1, t, 2 - 3t) : t \in \mathbb{K}\} \subseteq \mathbb{K}^3.$$

Find two independent affine-linear polynomials which vanish on L. This is the inverse problem for a linear set.

(g) Match parts (a)–(f) with the rows of the opening table which they address. Which row has not yet appeared in a substantial way? Which steps use only ideal-preserving combinations of equations, and which use the special echelon organization of leading variables? What nonlinear replacement is still missing?

---

2. From Equations to Ideals

For systems of linear equation, replacing a subset of the equations by suitable linear combinations, does not change the underlying geometry of their solution set. Further, via these operations we may eventually move our system of equations into echelon form from which the geometry is very visible. For systems of non-linear equations, we do not have an analogous notion of echelon form, but there are natural operations, analogous to linear combinations, which preserve the solution set. Namely, we may replace the care list of equations by the ideal the list of equations generates. In partocular, if $f_1, \ldots, f_s$ vanish at a point $p \in \mathbb{K}^n$, then every expression of the form

$$h_1 f_1 + \cdots + h_s f_s$$

for $h_1, \ldots, h_s \in \mathbb{K}[x_1, \ldots, x_n]$ also vanishes at p. At a moment to ponder why the converse is also true. Thus passing from a list of equations to its ideal records all consequences obtained by these elementary algebraic operations without changing the common zero set.

Definition 1. Let $\mathscr{F} \subseteq \mathbb{K}[x_1, \ldots, x_n]$ be an arbitrary subset. The *vanishing set* of $\mathscr{F}$ over $\mathbb{K}$ is

$$\mathbb{V}_{\mathbb{K}}(\mathscr{F}) = \{p \in \mathbb{K}^n \mid f(p) = 0 \text{ for every } f \in \mathscr{F}\}.$$

When the ground field $\mathbb{K}$ is clear, we simply write $\mathbb{V}(\mathscr{F})$. A subset of $\mathbb{K}^n$ of the form $\mathbb{V}(\mathscr{F})$ is called an *affine algebraic set* or an *affine algebraic variety*.

Out of economy, or perhaps laziness, we often drop adjectives and simply say *algebraic set*, *algebraic variety*, or just *variety* when it seems little confusion will arise. As a slight warning, there are dueling notations throughout the algebraic geometry over the use of the word variety. Some authors reserve the term algebraic variety for an algebraic set, which is also irreducible. (A topological property we shall see later in this exercise.) Others adopt the slightly broader definition we have taken where algebraic set and variety are synonymous and we shall stress the adjective irreducible when appropriate.

Notice the collection $\mathscr{F}$ in the above definition is allowed to be infinite. Nothing, so far, rules out an algebraic set which genuinely requires infinitely many equations. One of the central results of these lectures, Hilbert's Basis Theorem, which we will prove in Lecture 3, says that this cannot occur: every affine algebraic variety can be cut out by finitely many equations.

The notation $\mathbb{V}(\mathscr{F})$ admits two useful mnemonics: $\mathbb{V}$ for the common *vanishing* locus and $\mathbb{V}$ for the *variety* it defines. The historical origin of the this notation is a bit mysterious, however, the mathematical word *variety* comes to use from various translations of Riemann's *Mannigfaltigkeit*. While in modern usage this word translates as manifold, Riemann's use of the word was much more general than the current usage of a smooth or even topological manifold. The German adjective *mannigfaltig*, which is a root of Mannigfaltigkeit, roughly translates as plural, many, or perhaps various, which seems to be in roughly line with Riemann's work. As such Italian mathematicians translated Mannigfaltigkeit as *varietà*, and many French mathematicians translated it as *variété*. Note other French mathematicians translated it as *multiplicité* or *continuum*. When this word began solely referring to algebraic solution sets is unclear.

---

(3) **From Equations to Ideals.** Let $\mathscr{F} \subseteq \mathbb{K}[x_1, \ldots, x_n]$ and let $I, J \subseteq \mathbb{K}[x_1, \ldots, x_n]$ be ideals.

(a) Prove that $\mathbb{V}(\mathscr{F}) = \mathbb{V}(\langle \mathscr{F} \rangle)$. Thus the common zero set remembers the ideal generated by the equations, not merely the displayed list.

(b) Prove that $\mathbb{V}$ reverses inclusions: if $I \subseteq J$, then $\mathbb{V}(J) \subseteq \mathbb{V}(I)$.

(c) Describe the following varieties, and when possible, draw the corresponding picture. Over $\mathbb{C}$, a literal sketch is likely difficult, so you may need to be creative.

$$\mathbb{V}_{\mathbb{R}}(x^2 + y^2 - 1), \quad \mathbb{V}_{\mathbb{R}}(x^2 - y^2), \quad \mathbb{V}_{\mathbb{R}}(x^2 + y^2),$$
$$\mathbb{V}_{\mathbb{C}}(x^2 + y^2 - 1), \quad \mathbb{V}_{\mathbb{C}}(x^2 - y^2), \quad \mathbb{V}_{\mathbb{C}}(x^2 + y^2).$$

An *affine plane conic* is the zero set of a polynomial of degree two in two variables, and thus, the six varieties above are all examples of plane conics.

(d) Which of the defining polynomials in the previous part are irreducible over $\mathbb{R}$, but reducible over $\mathbb{C}$? For these examples, how do the varieties over $\mathbb{R}$ and $\mathbb{C}$ differ?

(e) Prove that for any polynomials $f, g \in \mathbb{K}[x_1, \ldots, x_n]$ the following identity holds:

$$\mathbb{V}(fg) = \mathbb{V}(f) \cup \mathbb{V}(g).$$

Thus factorization is an algebraic signal that the corresponding variety can be broken into (simpler) pieces.

(f) A nonempty algebraic variety $X \subset \mathbb{K}^n$ is *reducible* if and only if there exists proper algebraic varieties $X_1, X_2 \subset \mathbb{K}^n$, with $X_i \neq X$ for $i = 1, 2$, such that $X = X_1 \cup X_2$. A variety is

irreducible if it is not reducible. Prove that a nonempty algebraic variety $X \subset \mathbb{K}^n$ is irreducible if and only if given proper algebraic varieties $X_1, X_2 \subset \mathbb{K}^n$ such that $X = X_1 \cup X_2$ then $X_i = X$ for some i. (Note here proper means a proper nonempty subset of $\mathbb{K}^n$.)

(g) Note being reducible or irreducible very much depends on the ground field we are working over. For example, prove the following:

(i) The variety $\mathbb{V}_{\mathbb{K}}(xy)$ is reducible over all fields $\mathbb{K}$.

(ii) The variety $\mathbb{V}_{\mathbb{K}}(x^2 - y^2)$ is reducible if and only if $\mathbb{K}$ does not have characterisitc two. What happens when $\mathbb{K}$ is characterisitc two?

(iii) The variety $\mathbb{V}_{\mathbb{K}}(x^2 + y^2)$ is irreducible when $\mathbb{K} = \mathbb{Q}$ or $\mathbb{R}$, but is reducible when $\mathbb{K} = \mathbb{C}$.

The Nullstellensatz

The same algebraic set can be presented by many different ideals. For example, over every field we have that $\mathbb{V}(x) = \mathbb{V}(x^2) = \mathbb{V}(x^3)$. Geometrically, the equations $x = 0$, $x^2 = 0$, and $x^3 = 0$ all impose the same conditions on a point. Algebraically, however, the ideals $\langle x \rangle$, $\langle x^2 \rangle$, and $\langle x^3 \rangle$ are all different. (Be sure you know why.) More generally, if $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$ is an ideal and $f^k \in I$ for some $k \in \mathbb{Z}_{\geq 1}$, then for any point $p \in \mathbb{V}(I)$ we have

$$(f(p))^k = f^k(p) = 0.$$

Since $\mathbb{K}$ is a field, it has no nonzero nilpotents, so this forces $f(p) = 0$. Thus every equation whose power belongs to I already vanishes on $\mathbb{V}(I)$, even if the equation itself does not belong to I. This leads to the following operation.

Definition 2. Let R be a commutative ring. The *radical* of an ideal $I \subseteq R$ is

$$\sqrt{I} \coloneqq \left\{ f \in R \mid f^k \in I \text{ for some } k \geq 1 \right\}.$$

An ideal is *radical* if $I = \sqrt{I}$.

As argued above, and as you will formalize shortly, $\mathbb{V}(I)$ and $\mathbb{V}(\sqrt{I})$ are equal. Hence, taking radicals is one explanation for why the operation $I \mapsto \mathbb{V}(I)$ loses at least some information: it cannot distinguish an ideal from its radical. Amazingly, when $\mathbb{K}$ is algebraically closed, this is the only information that is lost in the process. This is the first of a series of results broadly know as 'Hilbert's Nullstellensatz", each of which highlights a different perspective of what it might mean to "solve" a system of polynomial equations.

Theorem 3 (Hilbert's Nullstellensatz I (Geometric Form))**.** *Let $\mathbb{K}$ be an algebraically closed field. The map $I \mapsto \mathbb{V}(I)$ gives a bijection:*

$$\left\{ \begin{matrix} \textit{radical ideals} \\ \textit{in } \mathbb{K}[x_1, \ldots, x_n] \end{matrix} \right\} \xrightarrow{\quad \mathbb{V} \quad} \left\{ \begin{matrix} \textit{algebraic varieties} \\ \textit{in } \mathbb{K}^n \end{matrix} \right\}.$$

It is worth pausing to see how this mirrors the relationship between equations and solution sets recalled at the beginning of this lecture. A complete proof of the Nullstellensatz is somewhat beyond the scope of this course, and so we will not give one here. However, we will frequently make use of it and will try to refine our understanding of it.

The theorem underlying this modern geometric formulation was proved, for homogeneous polynomials, in Section 3 of Hilbert's 1893 paper *Ueber die vollen Invariantensysteme*. It appeared three years after his 1890 paper *Ueber die Theorie der algebraischen Formen*, which contained what we now call Hilbert's Basis Theorem and applied it to the studying the question of when the ring of invariants is finitely generated. Famously, Hilbert's 1890 argument was highly non-constructive, leading his contempary Paul Gordan to exclaim, "Das ist keine Mathematik, das ist Theologie", "This is not mathematics, this is theology". While the provienance of this remark is a bit shakey, the first record of I could find it in-print is from Max Noether in 1914, Hilbert's 1893 paper also concered the problem of studying invariants, however from a more constructive perpsective. Interesting, Gordan subsequently gave his own proof of Hilbert's Basis Theorem, and it is his approach we shall roughly follow in Lecture 3.

---

(4) **Radical Ideals.** Let R be a ring and $I \subseteq R$ an ideal.

(a) Show that $I \subseteq \sqrt{I}$ and that $\sqrt{\sqrt{I}} = \sqrt{I}$.

(b) Suppose $f, g \in \sqrt{I}$, with $f^r \in I$ and $g^s \in I$. Use the binomial theorem to show that $(f + g)^{r+s} \in I$. Deduce that $\sqrt{I}$ is closed under addition.

(c) Show that if $f \in \sqrt{I}$ and $h \in R$, then $hf \in \sqrt{I}$. Conclude that $\sqrt{I}$ is an ideal.

(d) Let $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$ be an ideal. Prove that $\mathbb{V}(I) = \mathbb{V}(\sqrt{I})$. In other words, passing to the radical does not change its zero set.

(e) Compute $\sqrt{\langle x^3y^2 \rangle} \subseteq \mathbb{K}[x, y]$. Compare the zero sets of $\langle x^3y^2 \rangle$ and its radical.

(f) Give three different ideals in $\mathbb{K}[x, y]$ which define the x-axis. Which, if any, of your ideals are radical?

The geometric form of Hilbert's Nullstellensatz I asserts that $\mathbb{V}$ has an inverse, but it does not yet tell us what that inverse is. This returns us to the inverse problem with which we began: given a set of points, which polynomial equations hold on all of them? The natural answer is to collect every such equation.

Definition 4. For any subset $X \subseteq \mathbb{K}^n$, its *vanishing ideal* is

$$\mathbb{I}(X) \coloneqq \{f \in \mathbb{K}[x_1, \ldots, x_n] \mid f(p) = 0 \text{ for every } p \in X\}.$$

Thus $\mathbb{I}(X)$ records every polynomial vanishing on all of X, rather than one particular list of equations describing X. Notice this definition makes sense whether or not X is a variety. When X is a variety we often call $\mathbb{I}(X)$ the *defining ideal* of X. As we shall see shortly, it only really makes sense to think of $\mathbb{I}(X)$ as defining X when X is itself an algebraic variety.

In general the operations $\mathbb{V}$ and $\mathbb{I}$ are not literal inverses, even when we restrict our attention to only algebraic varieties. They reverse inclusions and satisfy

$$X \subseteq \mathbb{V}(\mathbb{I}(X)) \qquad \text{and} \qquad I \subseteq \mathbb{I}(\mathbb{V}(I)).$$

The first enlargement replaces an arbitrary set by the smallest algebraic set containing it; the second enlarges an ideal to include every polynomial which vanishes on its zero set. Exercise 5 asks you to make the first statement precise.

Two words of warning are useful at this moment: i) A finite list of equations whose common zero set is X need not generate the full ideal $\mathbb{I}(X)$. ii) A priori the ideal $\mathbb{I}(X)$ need not have a finite generating set. This claim is the content of Hilbert's Basis Theorem.

---

(5) **The Inverse Problem & Zariski Closure.** Let $X, Y \subseteq \mathbb{K}^n$ be arbitrary subsets.

(a) Prove that $\mathbb{I}(X)$ is a radical ideal. Hint: First check it is an ideal and then check that it is equal to its radical.

(b) Prove that if $X \subseteq Y$, then $\mathbb{I}(Y) \subseteq \mathbb{I}(X)$.

(c) Show that $X \subseteq \mathbb{V}(\mathbb{I}(X))$.

(d) Show that $\mathbb{V}(\mathbb{I}(X))$ is the smallest algebraic subset of $\mathbb{K}^n$ containing X: if $X \subseteq Z$ and Z is algebraic, then $\mathbb{V}(\mathbb{I}(X)) \subseteq Z$. We call this set the *Zariski closure* of X.

(e) Conclude that X is an algebraic set if and only if $X = \mathbb{V}(\mathbb{I}(X))$.

(f) Assume now that $\mathbb{K}$ is infinite and let $X = \{(t, t^2) \mid t \in \mathbb{K}^\times\} \subseteq \mathbb{K}^2$. If $f \in \mathbb{I}(X)$, consider the one-variable polynomial $f(t, t^2)$. Use the root bound recalled in Exercise 1(g) to show that it is identically zero, and hence that $f(0,0) = 0$. Conclude that $(0,0) \in \mathbb{V}(\mathbb{I}(X))$ even though $(0,0) \notin X$. In particular, a parabola with one point removed is not an algebraic set.

(g) In the previous part, where does the assumption that $\mathbb{K}$ is infinite enter? Over the finite field $\mathbb{F}_q$, set $X_q = \{(t, t^2) \mid t \in \mathbb{F}_q^\times\}$. Show that $x^{q-1} - 1 \in \mathbb{I}(X_q)$ but does not vanish at $(0,0)$, so $(0,0) \notin \mathbb{V}(\mathbb{I}(X_q))$. Also explain why $t^q - t$ is a nonzero polynomial in $\mathbb{F}_q[t]$ even though it vanishes at every element of $\mathbb{F}_q$. Identify the inference from part (f) which fails over a finite field.

Assume now that $\mathbb{K}$ is algebraically closed. Exercises 4 and 5 allow us to identify the inverse promised by Hilbert's Nullstellensatz I. Given $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$, the ideals $\sqrt{I}$ and $\mathbb{I}(\mathbb{V}(I))$ are both

radical, and those exercises show that they have the same zero set:

$$\mathbb{V}\left(\sqrt{I}\right) = \mathbb{V}(I) = \mathbb{V}(\mathbb{I}(\mathbb{V}(I))).$$

The injectivity in Theorem 3 therefore forces them to be equal. The set $\mathbb{V}(\mathbb{I}(X))$ is the smallest algebraic set containing X; in particular, if X is already algebraic, then $\mathbb{V}(\mathbb{I}(X)) = X$. The other composition, $\mathbb{I}(\mathbb{V}(I))$, contains the deeper content of the Nullstellensatz, and the equality $\mathbb{I}(\mathbb{V}(I)) = \sqrt{I}$ is false over a general field. Under our algebraically closed hypothesis, this brings us to our second version of the Nullstellensatz.

Theorem 5 (Hilbert's Nullstellensatz II (Recovery Form))**.** *Let $\mathbb{K}$ be algebraically closed and let $\mathscr{S} \subseteq \mathbb{K}[x_1, \ldots, x_n]$. If $V = \mathbb{V}(\mathscr{S}) \subseteq \mathbb{K}^n$, then*

$$\mathbb{I}(V) = \sqrt{\langle \mathscr{S} \rangle}.$$

An immediate consequence of Hilbert's Nullstellensatz II is that if I is radical, then $\mathbb{I}(\mathbb{V}(I)) = I$. Thus the vanishing-ideal operation $\mathbb{I}$ is precisely the inverse promised by the bijection $I \mapsto \mathbb{V}(I)$ in Hilbert's Nullstellensatz I. Let us record this conclusion one final time, as this will be the most memorable and used form of the Nullstellensatz.

Theorem 6 (Hilbert's Nullstellensatz III (Affine Algebra–Geometry Dictionary))**.** *Let $\mathbb{K}$ be an algebraically closed field. The maps $\mathbb{V}$ and $\mathbb{I}$ induce inverse, inclusion-reversing bijections:*

$$\left\{\begin{matrix}\textit{radical ideals} \\ \textit{in } \mathbb{K}[x_1, \ldots, x_n]\end{matrix}\right\} \underset{\mathbb{I}}{\overset{\mathbb{V}}{\rightleftarrows}} \left\{\begin{matrix}\textit{algebraic varieties} \\ \textit{in } \mathbb{K}^n\end{matrix}\right\}.$$

In many ways much of algebraic geometry can be view as understanding how various properties can and cannot be passed back and forth thru these correspondences. In this form, the Nullstellensatz solves the inverse problem completely: over an algebraically closed field, an algebraic set determines its radical ideal of equations, and that ideal determines the algebraic set.

A particularly important special case asks which ideal corresponds to the empty set. Since every polynomial vanishes vacuously on the empty set, $\mathbb{I}(\varnothing) = \mathbb{K}[x_1, \ldots, x_n]$. Hilbert's Nullstellensatz II therefore gives the following.

Theorem 7 (Hilbert's Weak Nullstellensatz)**.** *If $\mathbb{K}$ is algebraically closed and $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$ is an ideal, then $\mathbb{V}(I) = \varnothing$ if and only if $I = \langle 1 \rangle$.*

Notice the reverse implication that $1 \in I$ implies $\mathbb{V}(I) = \varnothing$ is immediate since as a polynomial 1 never vanishes. This is true over any field since in a field $1 \neq 0$. Hence the difficult part of the Weak Nullstellensatz is the forward implication that if $\mathbb{V}(I) = \varnothing$ then $1 \in I$. This implication is where the hypothesis that $\mathbb{K}$ be algebraically closed is crucial. (Consider $\mathbb{V}_{\mathbb{R}}(x^2 + 1)$.)

Although we have obtained the weak form here as a consequence of Hilbert's Nullstellensatz II, the usual proof proceeds in the opposite direction. One first proves the weak Nullstellensatz and then uses Rabinowitsch's trick to deduce the full recovery statement $\mathbb{I}(\mathbb{V}(I)) = \sqrt{I}$. While we shall see Rabinowitsch's trick shortly, it will be of interest to us mainly for its computational consequences.

As an interesting historical aside, Rabinowitsch's trick appeared in a one-page note by J. L. Rabinowitsch. The note, *Zum Hilbertschen Nullstellensatz*, was received in 1929 and published on page 520 of volume 102 of *Mathematische Annalen* in 1930. The page identifies its author only by the initials J. L. and the location Moscow. A colorful story, often repeated to identify J. L. Rabinowitsch with George Yuri Rainich, goes as follows. Rainich supposedly used the trick during a lecture and was interrupted by someone who indignantly accused him of presenting the famous Rabinowitsch trick as his own. Without replying, Rainich turned to the blackboard, wrote RABINOWITSCH, erased the letters B, O, W, T, S to leave RAINICH, and then continued his lecture. It is a wonderful story, but almost certainly not history. The current best evidence points towards a different Moscow based mathematician: Yuli Lazarevich Rabinovich, about whom little seems to be know.

---

(6) **Consequences and Certificates.** Assume $\mathbb{K}$ is algebraically closed unless otherwise stated.

(a) Let $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$ be an ideal. Use Theorem 5 to prove that a polynomial h vanishes at every point of $\mathbb{V}(I)$ if and only if $h^k \in I$ for some $k \geq 1$.

(b) Let $I = \langle x, xy - 1 \rangle \subseteq \mathbb{K}[x, y]$. Find an explicit identity expressing 1 as a polynomial combination of the displayed generators. Explain what this proves about $\mathbb{V}(I)$.

(c) Why does the fact that $\mathbb{V}_{\mathbb{R}}(x^2 + 1) = \varnothing$ not contradict Hilbert's Nullstellensatz?

(d) Return to the circle–parabola system in Exercise 1 given by $f_1 = x^2 - y$ and $f_2 = x^2 + y^2 - 1$. Prove that the polynomial $h = y^2 + y - 1$ vanishes on the common zeros of $f_1 = f_2 = 0$ by explicitly demonstrating that $h \in \langle f_1, f_2 \rangle$. Hint: Express h in terms of f_1 and f_2.

(e) *Further: Rabinowitsch's trick.* Let $R = \mathbb{K}[x_1, \ldots, x_n]$ and let $I \subseteq S$ be an ideal. Suppose $g \in S$ vanishes on $\mathbb{V}(I)$. Introducing a new variable z consider the polynomial ring $R[z] \cong \mathbb{K}[x_1, \ldots, x_n, z]$ and set $J = \langle I, 1 - zg \rangle \subset R[z]$. Notice that variety $\mathbb{V}(J)$ is a subset of $\mathbb{K}^{n+1}$.

(i) Prove that $\mathbb{V}(J) = \varnothing$. Use the weak Nullstellensatz to conclude that $1 \in J$.

(ii) More generally, prove that $g \in \sqrt{I}$ then $1 \in \langle I, 1 - zg \rangle$. (Hint: If $g^k \in I$ use the geometric series identity for $1 - (zg)^k$.)

The converse is also true, however, it is much more difficult to prove. Put together we get what is know as Rabinowitsch's trick: $g \in \sqrt{I}$ if and only if $1 \in \langle I, 1 - zg \rangle$. We shall use this statement in Lecture 3 to turn a membership test for the ideal on the right hand side into an algorithm to check membership in the radical ideal on the left hand side.

3. Parametrizations and Kernels

Up to this point, we have mostly focused on describing a solution set (i.e. a variety) implicitly, through its defining equations. A second perspective to describing a variety is by parametrizing it. Roughly, one should think of a parametrization as a convenient recipe for producing points of the solution set without having to solve the system of equations defining it For example, the parabola from the beginning of the lecture can be described implicitly as $\mathbb{V}(y-x^2)$, but it can also be described parametrically as the set $\{(t,t^2) \mid t \in \mathbb{K}\}$. These two description are useful for different purposes: the implicit equation $y - x^2$ makes it exceptionally easy to test whether a point $p \in \mathbb{K}^2$ lies on the parabola, while the parametric description makes it easy to construct points on the curve.

More generally, a collection of polynomials $f_1, \ldots, f_n \in \mathbb{K}[t_1, \ldots, t_m]$ determines a polynomial map

$$F : \mathbb{K}^m \longrightarrow \mathbb{K}^n \qquad \text{given by} \qquad a \longmapsto \big(f_1(a), \ldots, f_n(a)\big).$$

We call this a *polynomial parametrization* of its image. Note the parametrization F need not be injective, and the image of F image need not be algebraic, i.e., the image might not be a variety). The process of finding the vanishing ideal of the image of F from the polynomials $f_1, \ldots, f_n$ defining F is called *implicitization*. Moregenerally, we might define rational parameterizations by replacing the $f_1, \ldots, f_n$ by rational functions. Understanding which varieties can be rationally parameterized, and when such a parametrization is generically one-to-one, has be of great interest in algebraic geometry since the 19th century. These questions lead us to the notions of unirational and rational varieties, and more generally birational geometry, which we shall see in Lecture 4.

For the moment, we wish to see how algebra translates the map F into equations. Notice that we may use the same polynomials $g_1, \ldots, g_n$ defining F to construct a ring homomorphism

$$\begin{aligned} F^* : \mathbb{K}[x_1, \ldots, x_n] &\longrightarrow \mathbb{K}[t_1, \ldots, t_m], \\ f &\longmapsto f(g_1, \ldots, g_n), \end{aligned}$$

which we call the *substitution map* for $g_1, \ldots, g_n$. A polynomial f belongs to $\ker(F^*)$ exactly when $f(g_1, \ldots, g_n)$ is the zero polynomial in $\mathbb{K}[t_1, \ldots, t_m]$. Any such f certainly vanishes at every point of the image, since $f(F(a)) = (F^* f)(a) = 0$ for all $a \in \mathbb{K}^m$. Hence we always have

$$\begin{aligned} \ker(F^*) &\subseteq \mathbb{I}\left(F(\mathbb{K}^m)\right), \\ F(\mathbb{K}^m) &\subseteq \mathbb{V}\left(\ker(F^*)\right). \end{aligned}$$

The relationship among these sets depends both on the field and on whether the image is algebraic. Suppose $\mathbb{K}$ is infinite. A polynomial in $\mathbb{K}[t_1, \ldots, t_m]$ which vanishes at every point of $\mathbb{K}^m$ is the zero polynomial; this follows by induction on m from the corresponding fact for one-variable polynomials. Consequently, $f \in \mathbb{I}(F(\mathbb{K}^m))$ forces $F^*(f) = 0$, and therefore $\ker(F^*) = \mathbb{I}\big(F(\mathbb{K}^m)\big)$. In this case, $\mathbb{V}(\ker F^*) = \mathbb{V}\big(\mathbb{I}(F(\mathbb{K}^m))\big)$ is precisely the Zariski closure of the parametrized image, and it cuts out the image itself when that image is already algebraic. For a general field, $\ker(F^*)$ can be strictly smaller than $\mathbb{I}(F(\mathbb{K}^m))$, as we shall see in the exercise below.

There are several closely related but distinct problems concerning a parameterization F, which are worth keeping separate in our minds:

(i) Deciding whether the image of F is algebraic.

(ii) Computing the kernel of the substitution map F^*.

(iii) Computing the vanishing ideal of $F(\mathbb{K}^m)$, i.e., computing $\mathbb{I}(F(\mathbb{K}^m))$.

Answering (i) means proving there exists an ideal $J \subset \mathbb{K}[x_1, \ldots, x_n]$ such that $\mathbb{V}(J) = F(\mathbb{K}^m)$. If we find such an ideal J it need not be equal to $\mathbb{I}(F(\mathbb{K}^m))$, instead over an arbitrary field all we know is that $J \subset \mathbb{I}(F(\mathbb{K}^m))$. Moreover, although when $\mathbb{K}$ is infinite we have that $\ker(F^*) = \mathbb{I}(F(\mathbb{K}^m))$, this can fail over a finite field. Further, this abstract equality is not yet an explicit answer to the problem of implicitization, as we need a way to find explicit generators for $\ker(F^*)$. At the moment, it is not even clear that such a *finite* generating set exists, yet alone how we might construct one. This will be the focus of Lecture 3 where we will see Gröbner bases and Hilbert's Basis Theorem. The following exercise shows how the three questions above can differ in a relatively simple example.

---

(7) **From a Parametrization to Equations: The Affine Twisted Cubic.** Consider the polynomial map

$$F : \mathbb{K} \longrightarrow \mathbb{K}^3 \qquad \text{given by} \qquad t \longmapsto (t, t^2, t^3),$$

and let C denote the image of F. This curve is called the *affine twisted cubic*; the reason for the name will become clearer when we introduce projective space. Substitution gives the ring homomorphism

$$\begin{aligned} \mathbb{K}[x, y, z] &\xrightarrow{\quad F^* \quad} \mathbb{K}[t] \\ x &\longmapsto t \\ y &\longmapsto t^2 \\ z &\longmapsto t^3 \end{aligned}$$

which is the identity on $\mathbb{K}$.

(a) Verify that $y - x^2$ and $z - xy$ belong to $\ker(F^*)$. Let $J = \langle y - x^2, z - xy \rangle$. Explain why this proves that $J \subseteq \ker(F^*)$ and $C \subseteq \mathbb{V}(J)$.

(b) Conversely, let $(a, b, c) \in \mathbb{V}(J)$. Show that $b = a^2$ and $c = a^3$, and hence that $(a, b, c) = F(a)$.

(c) Conclude that $C = \mathbb{V}(J)$. Hence, in this example, the image of the parametrization is already algebraic. This equality holds over every field.

(d) At this point we know that $J \subseteq \ker(F^*)$ and

$$\mathbb{V}_{\mathbb{K}}(J) = \mathbb{V}_{\mathbb{K}}(\ker(F^*)) = C.$$

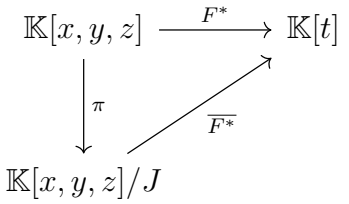
However, as we have seen two ideals defining the same variety does not imply the ideals are equal. Further, over an arbitrary field $\mathbb{V}(I) = \mathbb{V}(J)$ does not imply that $\sqrt{I} = \sqrt{J}$. (What assumption are we missing?) The remainder of the exercise proves $\ker(F^*) = J$. For $\alpha, \beta, \gamma \in \mathbb{Z}_{\geq 0}$, show that

$$x^\alpha y^\beta z^\gamma \equiv x^{\alpha+2\beta+3\gamma} \pmod{J}.$$

Hint: First use the generators of J to show that $y \equiv x^2$ and $z \equiv x^3 \pmod{J}$.

(e) Deduce that every polynomial in $\mathbb{K}[x, y, z]$ is congruent modulo J to a polynomial in x alone.

(f) Since $J \subseteq \ker(F^*)$, the map F^* factors as follows:



where π is the natural quotient map. Show that $\overline{F^*}$ is surjective.

(g) Prove that $\overline{F^*}$ is injective. Then apply the first isomorphism theorem to conclude that $\ker(F^*) = J$. *Hint:* Represent a class in the quotient $\mathbb{K}[x, y, z]/J$ by a polynomial $r(x) \in \mathbb{K}[x, y, z]$ only involving x. Using part (e), then show that $\overline{F^*}(r(x)) = r(t)$.

(h) We always have $\ker(F^*) \subseteq \mathbb{I}(C)$. The reverse inclusion requires more care. Suppose that $\mathbb{K}$ is infinite, and let $h \in \mathbb{I}(C)$. Show that $h(t, t^2, t^3) \in \mathbb{K}[t]$ vanishes at every $t \in \mathbb{K}$, and hence is the zero polynomial. Conclude that $\mathbb{I}(C) = \ker(F^*) = J$.

(i) Explain why the preceding conclusion requires an infinite field. If $\mathbb{K} = \mathbb{F}_q$, show that $x^q - x \in \mathbb{I}(C)$ even though $F^*(x^q - x) = t^q - t \neq 0$ as an element of $\mathbb{F}_q[t]$; the latter is a nonzero polynomial which vanishes as a function on $\mathbb{F}_q$. Thus $\ker(F^*) \subsetneq \mathbb{I}(C)$.

(j) Show that $\langle y - x^2, z - xy \rangle$ is equal to $\langle y - x^2, z - x^3 \rangle$. What does this tell us about the nonuniqueness of a chosen list of equations, even though the ideal of all polynomial relations is uniquely determined?

(k) Verify the further relation

$$xz - y^2 = y(x^2 - y) - x(xy - z) \in J.$$

Record it for now. Lecture 2 will reveal why the two displayed generators do not yet form the kind of organized generating set needed for systematic computation.

As we saw in the previous exercise, the affine twisted cubic admits a polynomial parametrization in a single parameter. Not every curve admits such a parametrization. As a preview of Lecture 4,

consider the affine cubic curve

$$\mathbb{V}_{\mathbb{C}}(y^2 - x^3 + x) \subset \mathbb{K}^2,$$

whose projective closure is the elliptic curve $\mathbb{V}_{\mathbb{P}^2}(y^2z = x^3 - xz^2)$. (Projective space and projective closure will appear in Lecture 3.) We shall see that unlike the twisted cubic, this curve cannot be rationally parameterized.

We can now carry out the same computations in *Macaulay2* (M2), an open-source computer algebra system created by Daniel Grayson and Michael Stillman for algebraic geometry and commutative algebra. It grew out of its predecessor *Macaulay*, written by Dave Bayer and Stillman; David Eisenbud has been a longtime collaborator, and a broad international community now develops the system and its packages. We will use *Macaulay2* through the browser, so nothing needs to be installed. Open either server on the official

"Try it out" page.

In *Macaulay2* one defines something with the = symbol. During a session all defined variables, rings, polynomials, ideals, etc. are remembered. In order to clear these things from memory we use the command `restart`, which starts a fresh session. The *Macaulay2* syntax for multiplication is `*`, powers are written with `^`, and additional and subtraction are `+` and `-` respectively. One can surpress the output of a line by placing a semicolon `;` at the end of the line. Once ran all inputs and outputs `i1`, `o1`, and so on; these labels are not part of the commands. For help with a command, one can enter, for example, `help ideal`.

Since computers cannot store infinite decimal expansions, computations with real or complex numbers generally require floating-point approximations. Although dealing with these numerical issues is itself a beautiful area of mathematics, we will avoid them completely. Instead, we will generally compute over the rational numbers, denoted `QQ` in *Macaulay2*, or over a finite field, where arithmetic is exact. Since our equations have rational coefficients, one may show that any polynomial identity proved over $\mathbb{Q}$ remains true if we view the same identity over $\mathbb{C}$ (The fancy way to say this is "after extending scalars to $\mathbb{C}$.) So for example, a polynomial that factors over $\mathbb{Q}$ also factors over $\mathbb{C}$. However the converse is not true, and identities proven over $\mathbb{C}$ may fail to be true over $\mathbb{Q}$. In general it is good to develop a sense of what precisely a computation done with *Macaulay2* shows, and then think about what theoretical results can be used to generalize it.

One beauty of *Macaulay2* is that much of its syntax mimics mathematical notation; what you type as code will often look like what you would write on the blackboard. For example, `QQ` denotes $\mathbb{Q}$, `ZZ` denotes $\mathbb{Z}$, and, when p is prime, `ZZ/p` denotes $\mathbb{Z}/p\mathbb{Z}$. When q is a prime power, the finite field $\mathbb{F}_q$ is constructed with `GF(q)`. As noted above, there are also `RR` and `CC`, but they use finite-precision approximations rather than exact copies of $\mathbb{R}$ and $\mathbb{C}$. A polynomial ring is specified by listing the coefficient ring followed by variables, separated by commas, in brackets `[]`. For example, the command `R = QQ[x,y,z]` creates the polynomial ring $\mathbb{Q}[x, y, z]$ and gives it the name `R`.

For now, we will treat commands such as `ker` and `eliminate` as black boxes whose outputs we can interpret. Lectures 2 and 3 will develop the algorithms that make these commands possible; both rely on Gröbner basis computations.

The namesake of both *Macaulay2* and the original *Macaulay* is Francis Sowerby Macaulay (1862–1937), an early pioneer in commutative algebra. Macaulay's career was filled with contrasts. He spent most of it as a schoolmaster at St Paul's School in London, yet was elected a fellow of the Royal Society in 1928. His work received little attention from many contemporary algebraic geometers in England but was appreciated by Emmy Noether and those around her who were building the modern language of commutative algebra. His 1916 tract *The Algebraic Theory of Modular Systems* likewise received little attention after his death but became highly influential in the second half of the twentieth century. In Macaulay's terminology, "modular systems" are what we now call ideals and modules. Much of the book is recognizably computational—elimination, resultants, and explicit systems of equations. The name of the software is therefore apt: as Eisenbud and Gray put it, it *"recognizes that much of his work is based on examples created by refined computation."* _

(8) **First *Macaulay2* Session.** Let us return to the circle–parabola equations, which is defined by the ideal $I = \langle x^2 - y, x^2 + y^2 - 1 \rangle \subset \mathbb{K}[x, y]$ Enter the commands one line at a time and compare every output with Exercise 1.

```
restart
R = QQ[x,y];
f1 = x^2-y;
f2 = x^2+y^2-1;
I = ideal(f1,f2); -- encode the intersection as an ideal

R
gens I
f2-f1 -- eliminate x
```

The command `R = QQ[x,y]` creates the polynomial ring $R = \mathbb{Q}[x, y]$. After running this line, we can use the variables x and y to construct polynomials in R. For example `f1 = x^2-y` constructs the polynomial $f_1 = x^2 - y$ and *Macaulay2* knows this is a polynomial in R. The command `I = ideal(f1,f2)` creates the ideal $I = \langle f_1, f_2 \rangle \subset R$. In order to see the generators of I we use the command `gens I`, which displays the chosen generators as a row matrix.

(a) After running the above code explain the three displayed outputs.

(b) Write code that determines whether $y^2 + y - 1$ lies in the ideal I. What does this mean geometrically? *Hint:* One direct test is `isSubset(ideal(y^2+y-1),I)`, but can you find a more direct method?

(9) **A Second *Macaulay2* Session.** Now consider the parametrization of the affine twisted cubic. Before running the block, predict the two zero outputs and at least one generator of the kernel.

```
restart
A = QQ[x,y,z];
T = QQ[t];
phi = map(T,A,{t,t^2,t^3}); -- encode t |-> (t,t^2,t^3)
phi(y-x^2)
phi(z-x*y)
K = ker phi; -- compute the full ideal of relations
trim K
```

Here `ker phi` computes an ideal, while `trim K` asks for a smaller list of generators.

(a) *Macaulay2* lists the target of a ring map first and its source second. Explain why the line `map(T,A,{t,t^2,t^3})` above represents the map $A \to T$ used in Exercise 7.

(b) Explain the two zero outputs. Compare the generators in `trim K` with the generators found by hand. Why do the displayed lists not need to agree term for term?

(c) Use `K == ideal(y-x^2,z-x*y)` to check equality of the two ideals. What mathematical statement is being verified by the output `true`?

(d) Testing one polynomial at a time is already becoming tedious. A *Macaulay2* list is enclosed in braces. Enter

```
candidates = {y-x^2,z-x*y,x*z-y^2,x};
apply(candidates,g -> phi(g))
```

The arrow expression on that line defines a function, in this case, taking something named g to $\phi(g)$, i.e., evaluating the map ϕ defined above at the element g. The command `apply` applies a function to every element of a list or a sequence before returning the outputs as a new list or new sequence. (Note your original list remains unchanged.) Before running the second line, predict the list it should return.

(e) The second useful command to know is `scan`, which also applies a function to every entry of a list or sequence. However, `scan` discards the returned values, as such it useful when the function itself returns an output, such as printing. Run the following code.

```
scan(candidates,g -> print (g,phi(g) == 0))
```

What is the meaning of the output? How does the differ from the output you got when you ran the `apply` in the previous part?

(f) Add polynomial polynomial of your choosing to the list `candidates`, predict whether it belongs to $\ker(\phi)$, and use the same commands to test your prediction.

(g) *Further:* Also verify membership in K using `(x*z-y^2) % gb K`. Explain the mathematical meaning of each test.

In addition to being open source, much of the specialized functionality of *Macaulay2* is developed by researchers in commutative algebra and algebraic geometry through *packages*: collections of commands and documentation devoted to a particular area of mathematics or type of computation. These packages are written and maintained by members of the *Macaulay2* community, allowing the program to grow alongside current research. A list of the packages distributed with *Macaulay2*, organized by subject, is available in the package index.

Before using a package, we load it with the command `needsPackage`. For example, the command `needsPackage "Elimination"` loads the package `Elimination`. Loading a package makes all of the commands in the package available for the current session, however, we must load the package each time we restart if we wish to continue using it. Every *Macaualy2* package comes with extensive documentation, which includes an overview of the package, a list of functions, methods, and objects, and a host of examples. You access the package documentation via the command `viewHelp`, for example, `viewHelp "Elimination"`, which will open the documentation as an HTML file. I would encourage you to frequently consult and use the *Macaulay2* documentation; one rarely memorizes every command, and it often important to carefully understand exactly what a command is computing.

In the following exercise we will use the command `dim I`, which computes the Krull dimension of R/I for an ideal $I \subseteq R$. The precise definition of Krull dimension is beyond this course, but you should informally think of it as the number of locally independent parameters of the corresponding affine variety. For example, an ideal defining a non-empty finite set has Krull dimension zero, an ideal defining a curve has Krull dimension 1, and an ideal defining a surface has Krull dimension 2. (Note we define the dimension relative the the ground field $\mathbb{K}$, and so we think of $\mathbb{C}$ as being one-dimensional while $\mathbb{C}^2$ is two-dimensional. Although we should be careful not to confuse Krull dimension with vector space dimension. We will also use the command `degree I`, which computes the degree of R/I. When $\mathbb{V}(I)$ is zero-dimensional, `degree I` is equal to the $\mathbb{Q}$-vector-space dimension of R/I. As we will see, in this setting, we should roughly think of degree as counting the number of points in $\mathbb{V}(I)$ if we count the points with the correct multiplicity. Exercises 12 and 13 will make this precise by comparing the degree of an ideal defining four distinct points with the degree of an ideal defining a tangential double point.

---

(10) **A Third *Macaulay2* Session.** Let us return to the circle–parabola ideal and start a fresh *Macaulay2* session. Before running the block below, predict the two elimination generators, the dimension, and the degree.

```
restart
needsPackage "Elimination";
R = QQ[x,y];
I = ideal(x^2-y,x^2+y^2-1);

eliminate({y},I) -- algebraic equations involving only x
eliminate({x},I) -- algebraic equations involving only y

dim I == 0  -- check that the intersection is a finite set of points
degree I -- count (with multiplicity) the number of points in the
   intersection
```

(a) Match the two elimination ideals with the hand calculations in Exercise 1. Notice that eliminating y leaves equations in x, and eliminating x leaves equations in y.

(b) Give a geometric interpretation of the outputs of `dim I` and `degree I`. (Note there is something weird going on here; how many points does $x^2 - y$ and $x^2 + y^2 - 1$ intersect in over $\mathbb{Q}$, which is the field we are working over.)

(c) *Further: Closure Rather Than Image.* Set `H = ideal(x*y-1)` and eliminate y with the following command:

```
eliminate({y},H)
```

Compare the output with Exercise 1(g): which point has elimination added, and why?

(d) *Further: Components.* The following computation returns to the factorization examples from Exercise 3. Predict the visible components before asking the computer.

```
restart
needsPackage "MinimalPrimes";
R = QQ[x,y];

axes = ideal(x*y);
decompose axes -- the two irreducible components

twoLines = ideal((x-y)*(x+y));
decompose twoLines -- factorization becomes component decomposition
```

Compare the outputs with the decompositions found by hand. In Lecture 5, we will return to primary decomposition and to nonreduced equations, where they have a geometric job.

---

4. Coordinate Rings and Multiplicity

We now return to a distinction already encountered in the equality $\mathbb{V}(x) = \mathbb{V}(x^2)$. A polynomial $f \in \mathbb{K}[x_1, \ldots, x_n]$ restricts to a function $X \to \mathbb{K}$ on any subset $X \subseteq \mathbb{K}^n$. We say that a function $\phi : X \to \mathbb{K}$ is a *polynomial function on* X if there exists a polynomial $f \in \mathbb{K}[x_1, \ldots, x_n]$ such that $\phi = f|_X$. Note sometimes it is possible for polynomial functions to hide, and not look like polynomials. Consider for example the function $(x, y) \mapsto e^{x^2+y^2}$. This is a well-definined function on all of $\mathbb{R}^2$, but it is certainly not a polynomial function on $\mathbb{R}^2$. However, if we restrict this an consider it as a function on the unit circle, $\mathbb{V}_{\mathbb{R}}(x^2 + y^1 - 1)$, it is a polynomial function. Can you find the polynomial f it agrees with on $\mathbb{V}_{\mathbb{R}}(x^2 + y^2 - 1)$?

Since function addition, subtraction, and multiplication are defined point-wse, the set of all functions $X \to \mathbb{K}$ forms a ring. (You should think about this carefully if it is not obvious.) Further, more, since restriction commutes with these operations the subset of polynomial functions on X forms a subring. This motivates the following definition.

Definition 8. The *coordinate ring* of an affine algebraic variety $X \subseteq \mathbb{K}^n$, which we denote by $\mathbb{K}[X]$, is the ring of polynomial functions on X under point-wise addition and multiplication.

While the above definition makes it seem very abstract, the coordinate ring of a variety X, has a beautiful and simple description. Namely, by definition there is a surjective map of rings:

$$\mathbb{K}[x_1, \ldots, x_m] \xrightarrow{\rho} \mathbb{K}[X] f \longmapsto f|_X$$

where $f|_X$ is the polynomial function on X determined by f, i.e. $p \mapsto f(p)$. A polynomial $f \in \mathbb{K}[x_1, \ldots, x_n]$ restricts to give the zero function on X if and only if $f(p) = 0$ for all $p \in X$. Hence a polynomial f is in the kernel of ρ if and only if f vanishing on all of X, i.e. if and only if $f \in \mathbb{I}(X)$. Hence the kernel of ρ is precisely $\mathbb{I}(X)$. Using everyone's favorite isomorphism theorem we get:

$$\mathbb{K}[X] \coloneqq \mathbb{K}[x_1, \ldots, x_n]/\mathbb{I}(X).$$

Since $\mathbb{I}(X)$ is a radical ideal the coordinate ring of X is always reduced, i.e., it has no nonzero nilpotent elements Further, if $F : X \to Y$ is a map of varieties that is given in coordinates by polynomials and $\phi : Y \to \mathbb{K}$ is a polynomial function then the composition $\phi \circ F : X \to \mathbb{K}$ is a polynomial function. Hence F "pullsback" polynomial functions and induces a ring homomorphism

$$F^* : \mathbb{K}[Y] \longrightarrow \mathbb{K}[X], \qquad \text{given by} \qquad [h] \longmapsto [h \circ F].$$

Notice this ring homomorphism goes in the opposite direction to the map of algebraic varieties.

Now set $S = \mathbb{K}[x_1, \ldots, x_n]$ and suppose that $X = \mathbb{V}(I)$ for a specified ideal $I \subseteq S$. We wish to distinguish the coordinate ring $\mathbb{K}[X]$ from the quotient S/I attached to the ideal with which we began. The inclusion $I \subseteq \mathbb{I}(\mathbb{V}(I))$ induces a canonical surjection

$$S/I \longrightarrow \mathbb{K}[\mathbb{V}(I)] = S/\mathbb{I}(\mathbb{V}(I)).$$

Its kernel is $\mathbb{I}(\mathbb{V}(I))/I$. When $\mathbb{K}$ is algebraically closed, the Nullstellensatz identifies this map with

$$S/I \longrightarrow S/\sqrt{I},$$

whose kernel $\sqrt{I}/I$ consists precisely of the nilpotent elements of S/I. Indeed, the class of f in S/I is nilpotent exactly when $f^m \in I$ for some $m \geq 1$.

This gives a precise meaning to the statement that nilpotents cannot be detected by points. Every $p \in \mathbb{V}(I)$ determines an evaluation homomorphism

$$\mathrm{ev}_p : S/I \longrightarrow \mathbb{K}, \qquad \overline{f} \longmapsto f(p).$$

Every nilpotent is sent to zero by every such homomorphism, since a field has no nonzero nilpotents. Conversely, when $\mathbb{K}$ is algebraically closed, the Nullstellensatz says that an element annihilated by every point evaluation must be nilpotent. Thus nilpotents are not additional points hidden from the picture; they are algebraic information retained by S/I but lost when we pass to the underlying solution set. In the finite intersections below, this information records repeated vanishing and tangency.

The algebraically closed hypothesis in the converse is essential. For example, $\langle x^2 + y^2 \rangle$ is radical in $\mathbb{R}[x, y]$, but its real zero set is only the origin, whose vanishing ideal is $\langle x, y \rangle$. Over a non-algebraically closed field, passing to the set of $\mathbb{K}$-valued solutions may therefore forget more than nilpotents.

---

(11) **Coordinate Rings, Nilpotents, and Points.**

Let $S = \mathbb{K}[x_1, \ldots, x_n]$.

(a) Let $I \subseteq S$. Prove that S/I is an integral domain if and only if I is prime. Show that the nilpotent elements of S/I are precisely the classes $f + I$ with $f \in \sqrt{I}$. Deduce that S/I is reduced if and only if I is radical.

(b) Let X be a nonempty algebraic set. Using $\mathbb{V}(fg) = \mathbb{V}(f) \cup \mathbb{V}(g)$, prove that

$$\begin{aligned} X \text{ is irreducible} &\iff \mathbb{I}(X) \text{ is prime} \\ &\iff \mathbb{K}[X] \text{ is an integral domain.} \end{aligned}$$

Notice that this equivalence does not require $\mathbb{K}$ to be algebraically closed.

Hint: For the reverse implication, if $X = X_1 \cup X_2$ with both $X_i \subsetneq X$, choose $f_i \in \mathbb{I}(X_i) \setminus \mathbb{I}(X)$.

(c) Let $I \subseteq S$ and suppose that $X = \mathbb{V}(I)$ is nonempty. Show that the natural map $q_I : S/I \longrightarrow \mathbb{K}[X]$ has kernel $\mathbb{I}(X)/I$.

(d) Show that every $p \in X$ induces an evaluation homomorphism $\mathrm{ev}_p : S/I \longrightarrow \mathbb{K}$ and that

$$\ker(q_I) = \bigcap_{p \in X} \ker(\mathrm{ev}_p).$$

(e) Assuming that $\mathbb{K}$ is algebraically closed, use the Nullstellensatz to conclude that this intersection is precisely the set of nilpotent elements of S/I.

(12) **A Double Point from Tangency.** Let $A_1 = \mathbb{K}[x]/\langle x\rangle$ and $A_2 = \mathbb{K}[x]/\langle x^2\rangle$. Notice that both $\langle x\rangle$ and $\langle x^2\rangle$ define the single point $\{0\} \subseteq \mathbb{K}$.

(a) Prove that $\overline{x} \neq 0$ in A_2 while $\overline{x}^2 = 0$. Identify the kernel of the natural map $A_2 \longrightarrow A_1 \cong \mathbb{K}[\{0\}]$, and explain why evaluation at the unique point cannot detect $\overline{x}$.

(b) The preceding algebra occurs geometrically when curves meet tangentially. The parabola $y = x^2$ is tangent to the line $y = 0$ at the origin. Set

$$I_{\tan} = \langle x^2 - y, y\rangle \subseteq \mathbb{C}[x, y].$$

Show that $\mathbb{C}[x, y]/I_{\tan} \cong \mathbb{C}[x]/\langle x^2\rangle$.

(c) Compute $\mathbb{V}(I_{\tan})$, $\sqrt{I_{\tan}}$, and $\mathbb{C}[\mathbb{V}(I_{\tan})]$.

(d) Find a basis for $\mathbb{C}[x, y]/I_{\tan}$ as a $\mathbb{C}$-vector space and identify its nonzero nilpotent classes.

(e) Explain why the underlying intersection consists of only the origin, while the quotient ring has dimension 2 as a $\mathbb{C}$-vector space. Since this is the only intersection point, this dimension is its intersection multiplicity: the origin is therefore *counted twice*.

(13) **Four Distinct Intersection Points.** Consider the intersection of the parabola $y = x^2$ with the circle $x^2 + y^2 = 1$, and set

$$I_{\text{circ}} = \langle x^2 - y, x^2 + y^2 - 1\rangle \subseteq \mathbb{C}[x, y].$$

(a) Show that $\mathbb{C}[x, y]/I_{\text{circ}} \cong \mathbb{C}[x]/\langle x^4 + x^2 - 1\rangle$.

(b) Deduce that the classes of $1, x, x^2, x^3$ form a $\mathbb{C}$-basis for this quotient, and use $y = x^2$ in the quotient to represent this basis by the classes of $1, x, y, xy$.

(c) Recall from Exercise 1 that $x^4 + x^2 - 1$ has four distinct complex roots $\alpha_1, \ldots, \alpha_4$. Use the Chinese remainder theorem to prove that evaluation induces an isomorphism

$$\mathbb{C}[x]/\langle x^4 + x^2 - 1\rangle \longrightarrow \mathbb{C}^4, \qquad r(x) \longmapsto (r(\alpha_1), \ldots, r(\alpha_4)).$$

(d) Conclude that $\mathbb{C}[x, y]/I_{\text{circ}}$ is reduced and that the intersection consists of the four distinct points (α_i, α_i^2) for $i = 1, 2, 3, 4$.

(e) Compare this intersection with the tangent intersection in Exercise 12. Explain precisely why the tangent intersection is described as *one point counted twice*, whereas the circle–parabola intersection consists of *four distinct points*. Your answer should compare the underlying point sets, the vector-space dimensions of the quotient rings, and the presence or absence of nonzero nilpotents.

Exercises 12 and 13 are two halves of one classical statement called Bézout's theorem. Roughly, Bézout's theorem states that two plane curves of degrees d and e with no common component intersect in exactly de points. Thus, a conic and a line intersect in exactly $2 \cdots 1 = 2$ points, while two conics meet in $2 \cdot 2 = 4$ points. Note this can be thought of as a generalization of the fact you might have learned in high school: the degree of a polynomial is equal to the number of intersection points of the graph of the polynomial with a generic horizontal line.

As the previous exercises highlight, in order to make bëzout's theorem true three refinements are needed to how we could intersection points: (i) we must work over an algebraically closed field, (ii) we must count points with multiplicity, and (iii) we must work retain points that inersect "at infiniy", which means working in projective space. We will see more about projective space in the later lectures. While this theorem is normally credit to Bézout's workin in 1779, the result is also stated, without proof, in Newton's Principia in 1687

---

(14) **The Affine Dictionary.** Assuming $\mathbb{K}$ is algebraically closed, return to the opening table and complete the algebraic column using $1 \in I$, $\sqrt{I}$, elimination ideals, finite quotient dimension, primality of $\sqrt{I}$, and kernels. Fill in the *Macaulay2* entries learned so far: `eliminate`, `dim`, `degree`, and `ker`; if you completed the *Further: Components* part, also include `decompose`. Mark the rows where an output may describe a Zariski closure rather than the exact image. We will return to the remaining entries in the subsequent lectures.

LECTURE 2: DIVISION AND HIDDEN EQUATIONS (V1)

Throughout this lecture, $\mathbb{K}$ denotes a field. In Lecture 1 we studied the affine twisted cubic $C \subset \mathbb{K}^3$ in detail and saw that it was given by $\mathbb{V}(I)$, where $I = \langle x^2 - y, xy - z \rangle$. While studying this variety we found the following identity

$$xz - y^2 = y(x^2 - y) - x(xy - z),$$

which certifies that $xz - y^2 \in I$. Recall geometrically this means that the polynomial $xz - y^2$ vanishes along the twisted cubic. That said, how could we have found this certificate, i.e., the coefficients y and $-x$, systematically? More importantly, if no such identity exists, how could we ever prove nonmembership?

This is the *ideal-membership problem*: Given an ideal $I = \langle g_1, \ldots, g_s \rangle \subseteq \mathbb{K}[x_1, \ldots, x_n]$ and a polynomial $f \in \mathbb{K}[x_1, \ldots, x_n]$, determine whether $f \in I$. By definition, $f \in I$ if and only if there exist polynomials $q_1, \ldots, q_s \in \mathbb{K}[x_1, \ldots, x_n]$ such that

$$f = q_1 g_1 + \cdots + q_s g_s.$$

Thus an explicit representation of the above form *certifies* or is a *certificate of* membership of f in I. Notice given the q_i checking such an identity is true is easy as all we have to do is multiple and add everything out and see it gives f. Hence the difficulty behind the ideal-membership problem is that the polynomials q_i are unknown, and it seems like the definition of I gives no obvious finite search process. What we need is a some form of constructive algorithm that terminates and produces either a certificate as above (proving $f \in I$) or a conclusive obstruction showing that $f \notin I$.

We have already seen two cases where we have essentially such an algorithm: (i) For systems of linear equations, Gaussian elimination isolates the leading variables, and eventually reduces them to echelon form ii) For polynomials in one-variable the trusty polynomial long division reduces the question of whether the remainder is zero is not. We will solve the ideal-membership problem in general by generalizing the division algorithm, by asking what makes one-variable division work, imitated it as literally as possible, seeing why these imitations fail before making the needed changes.

(1) **What Would an Ideal-Membership Algorithm Decide?** Return to the table near the beginning of Lecture 1.

   (a) Which rows can be approached through membership tests? Identify separately the roles played by $1 \in I$, $f \in I$, and $f \in \sqrt{I}$.

   (b) Explain why $f \in I$ always implies that f vanishes on $\mathbb{V}_{\mathbb{K}}(I)$. When $\mathbb{K}$ is algebraically closed, what does Hilbert's Nullstellensatz say if f vanishes on $\mathbb{V}_{\mathbb{K}}(I)$? Why does its conclusion place f in $\sqrt{I}$, rather than necessarily in I?

(c) Explain why $1 \in I$ always forces $\mathbb{V}_{\mathbb{K}}(I) = \varnothing$. Which hypothesis is required for the converse?

(d) Suppose that you search for polynomials q_i satisfying $f = \sum_i q_i g_i$ and do not find them. Why does this not prove that $f \notin I$? What additional output would turn the failed search into a finite certificate of nonmembership? At this stage, you are not expected to name such an output; describe only what it would need to establish.

The preceding exercise exposes the computational difficulty behind the dictionary. A successful search for $q_1, \ldots, q_s$ produces a checkable certificate of membership, but an unsuccessful search says nothing: we may simply not have searched far enough. We need a systematic way to conduct our search. To begin building such an approach, let us consider a far simpler special case of the problem: principal ideals. Let $\langle g_1 \rangle$ be a principal ideal. Given an element $f \in \mathbb{K}[x_1, \ldots, x_n]$, we know that $f \in \langle g_1 \rangle$ if and only if $f = q_1 g_1$ for some $q_1 \in \mathbb{K}[x_1, \ldots, x_n]$, i.e., if and only if f is divisible by g_1. Hence, for principal ideals, the ideal-membership problem really amounts to finding a systematic way to perform division.

Luckily for us, we are all familiar with at least one special case in which we understand division: polynomials in one variable. Even better, since $\mathbb{K}[x]$ is a principal ideal domain (PID), every ideal is principal. As we will see shortly, division gives a beautiful answer to the ideal-membership question in $\mathbb{K}[x]$. We will then attempt to generalize this to $\mathbb{K}[x_1, \ldots, x_n]$, where ideals need no longer be principal and we do not (yet) have a familiar notion of polynomial division. This leads us naturally to monomial orders, Gröbner bases, normal forms, and beyond.

However, let us not get ahead of ourselves. We begin by considering the one-variable case in depth, adopting the convention that $\deg(0) = -\infty$.

Theorem 1 (Polynomial Division Algorithm)**.** *Let $g \in \mathbb{K}[x]$ be nonzero. For every $f \in \mathbb{K}[x]$ there are unique polynomials $q, r \in \mathbb{K}[x]$ such that*

$$f = qg + r \qquad \textit{and} \qquad \deg(r) < \deg(g).$$

The usual proof already contains an algorithm. For $0 \neq h$, let $\mathrm{LT}(h)$ denote its term of highest degree. Define two new polynomials, the 'working" polynomials p and the quotient polynomial q, which we shall update at each step of our algorithm. Initially define them by $p = f$ and $q = 0$. Now if $\deg(p) \geq \deg(g)$, then define

$$m = \frac{\mathrm{LT}(p)}{\mathrm{LT}(g)}.$$

Subtracting mg cancels the leading term of p, so the new working polynomial is either zero or has smaller degree. We repeat this until either $p = 0$ or $\deg(p) > \deg(g)$. We may write this as an algorithm as the following.

Algorithm: One-Variable Division
Input: $f, g \in \mathbb{K}[x]$ with $g \neq 0$.
Initialize: $q \leftarrow 0$ and $p \leftarrow f$.
Repeat: While $p \neq 0$ and $\deg(p) \geq \deg(g)$, set

$$m \leftarrow \frac{\mathrm{LT}(p)}{\mathrm{LT}(g)}, \qquad q \leftarrow q + m, \qquad p \leftarrow p - mg.$$

Output: q and $r \leftarrow p$.

Notice that at every stage of the algorithm the bookkeeping preserves the identity

$$f = qg + p,$$

and after each update the degree of p strictly decreases. Since the degree of a polynomial is a finite positive number, the well-ordering prinicipal for the natural numbers implies that eventually the algorithm must stop. Exercise 2 asks you to carefully check both of these claims, while Exercise 3(a) isolates what makes the remainder output by the algorithm unique.

The division algorithm also gives us a unique way to represent classes in the quotient ring $\mathbb{K}[x]/\langle g \rangle$. In particular, given $f \in \mathbb{K}[x]$, if we write $f = qg + r$ as in the division algorithm, then the remainder r is the unique representative of the class of f in $\mathbb{K}[x]/\langle g \rangle$ whose degree is smaller than $\deg(g)$.

This gives the remainder a geometric interpretation. Every class in $\mathbb{K}[x]/\langle g \rangle$ determines a polynomial function on $\mathbb{V}(g)$, because every element of $\langle g \rangle$ vanishes there. The phrase "polynomial function" requires some care: distinct classes may determine the same function unless $\langle g \rangle = \mathbb{I}(\mathbb{V}(g))$. The equality $f \equiv r \pmod{\langle g \rangle}$ means that the classes of f and r determine the same function on $\mathbb{V}(g)$. For example, if $g = x^2 - 1$ and $\mathrm{char}(\mathbb{K}) \neq 2$, then $\langle g \rangle = \mathbb{I}(\{-1, 1\})$.

---

(2) **Division and Polynomial Functions on Two Points.** Carefully apply the division algorithm to divide $f = x^5 - 1$ by $g = x^2 - 1$.

(a) Make a table recording the values of p and q at each stage. Verify that at every stage of the algorithm we have the identity $f = qg + p$.

(b) Find the final quotient q and remainder r.

(c) Explain why the degree of p strictly decreases after every pass through the loop.

(d) Assume $\mathrm{char}(\mathbb{K}) \neq 2$. Evaluate both f and the remainder r at $x = -1$ and $x = 1$. Why do both f and r determine the same function on $\mathbb{V}(x^2 - 1)$ even though they are different polynomials?

(e) *Further: Characteristic* 2. Using the fact that $x^2 - 1 = (x - 1)^2$, prove that the class of $x - 1$ in $\mathbb{K}[x]/\langle x^2 - 1 \rangle$ is nonzero and nilpotent. How does this recover the distinction between the

coordinate ring of $\{1\}$ and the quotient ring $\mathbb{K}[x]/\langle x^2-1\rangle$? How does this comapre to the tangent-double point exercise in Lecture 1?

(3) **Ideal Membership in One Variable.** Let $g \in \mathbb{K}[x]$ be nonzero and $f \in \mathbb{K}[x]$.

(a) Prove the uniqueness of the remainder in the division algorithm: if $f = qg + r = q'g + r'$ and each remainder is either zero or has degree strictly less than $\deg(g)$, show that $q = q'$ and $r = r'$.

(b) Carefully prove that $f \in \langle g \rangle$ if and only if the remainder of f divided by g is zero.

(c) Assume that an ideal of $\mathbb{K}[x]$ is presented by a generator g. Describe an algorithm that decides whether a given f lies in $\langle g \rangle$. Why does the fact that $\mathbb{K}[x]$ is a PID mean that this method applies to every ideal once a generator has been found?

(d) *Further: Ideals in* $\mathbb{K}[x]$. Let $I \subseteq \mathbb{K}[x]$ be a nonzero ideal and choose a nonzero element $g \in I$ of least degree. Divide an arbitrary $f \in I$ by g and prove that $I = \langle g \rangle$.

(4) **One-Variable Division in *Macaulay2*.** Return to the browser-based *Macaulay2* session from Lecture 1 and start again with `restart`. In one variable, `//` returns the quotient and `%` returns the remainder. Predict both outputs from Exercise 2 before entering the following block.

```
restart
R = QQ[x];
f = x^5-1;
g = x^2-1;

q = f//g -- quotient in the identity f=qg+r
r = f%g  -- normal form of f modulo the ideal (g)

f == q*g+r -- verify the division identity
```

(a) Compare the outputs with your hand computation. What does the final output `true` verify?

(b) Replace f and g by $x^{1000}-1$ and x^2-1. Predict the remainder before asking the computer. Interpret the result as a statement about the values of $x^{1000}-1$ on $\mathbb{V}_{\mathbb{C}}(x^2-1)$.

(c) Explain why the test `f%g == 0` decides membership in $\langle g \rangle \subseteq \mathbb{K}[x]$. Which part of Theorem 1 is being used?

(d) *Further: The Factor Theorem and a Bound on Roots.* For $a \in \mathbb{K}$, use division by $x-a$ to prove that $f(a)$ is the remainder and deduce that $a \in \mathbb{K}$ is a root of f if and only if f is divisible by $(x-a)$. Then prove that a nonzero polynomial in $\mathbb{K}[x]$ has at most $\deg(f)$ roots in $\mathbb{K}$. This is the one-variable fact used in Lecture 1 to show that a polynomial vanishing at every point of an infinite field is zero; induction on the number of variables gives the corresponding statement for $\mathbb{K}^m$.

Who would have guessed that polynomial division would be so useful? I certainly did not when I first learned about it in high school. We have just seen that the polynomial division algorithm in $\mathbb{K}[x]$ (i) answers the ideal-membership problem for $\mathbb{K}[x]$, (ii) gives us a canonical way to represent classes in quotient rings $\mathbb{K}[x]/\langle g\rangle$, and (iii) proves the factor theorem, among other things. The natural question is whether these ideas extend to several variables.

What prevents us from repeating this calculation in several variables? Let us again return to our running example of the affine twisted cubic, defined by the ideal $I = \langle x^2 - y, xy - z\rangle \subseteq \mathbb{K}[x, y, z]$, and the relation

$$xz - y^2 = y(x^2 - y) - x(xy - z),$$

certifying that $xz - y^2 \in I$. If we wished to obtain this relation from a division algorithm, we would encounter several issues. First, the ideal I is not principal, so we need a notion of division that allows us to divide $xz - y^2$ by two polynomials simultaneously—namely, $x^2 - y$ and $xy - z$—rather than by a single polynomial. Second, if we are to invent such a division algorithm, we need a way to define the leading term of $xz - y^2$. Both monomials in this polynomial have degree two! It is therefore unclear which one the division algorithm should try to cancel first.

We therefore need to make two choices that were invisible in one variable: an order on the monomials and an ordered list of divisors. These choices are not merely bookkeeping. Different ways of declaring a term to be largest can expose different consequences of the same equations, just as different elimination steps can expose different features of a linear system.

Let us turn to the first issue: ordering the monomials in a way that lets us choose between the two terms of $xz - y^2$.

Definition 2. A *monomial order* on $\mathbb{K}[x_1, \ldots, x_n]$ is a relation $<$ on the set of monomials satisfying the following conditions:

(i) $<$ is a strict total order;

(ii) if $m_1 < m_2$ and m is any monomial, then $mm_1 < mm_2$; and

(iii) $<$ is a well-order.

Each condition does a specific job. The first says that any two monomials can be compared: for all monomials m_1 and m_2, exactly one of $m_1 < m_2$, $m_1 = m_2$, $m_2 < m_1$ holds. In particular, every nonzero polynomial has a unique largest term, which is what we set out to obtain. The second says that an inequality survives multiplication by a common monomial; since division proceeds by subtracting monomial multiples of the divisors, this is what keeps a division step from disturbing the order. The third says that every nonempty collection of monomials has a smallest element, or equivalently that there is no infinite strictly decreasing chain $m_1 > m_2 > m_3 > \cdots$. It will replace the fact, used in one-variable division, that a decreasing sequence of nonnegative integer degrees must terminate, and it is what will make our division algorithm stop.

These three conditions are not independent. Given (i) and (ii), condition (iii) is equivalent to requiring that $1 < m$ for every monomial $m \neq 1$. Indeed, if $m < 1$, then multiplying repeatedly by m produces the infinite descending chain $\cdots < m^3 < m^2 < m < 1$; the converse implication rests on Dickson's Lemma, which we shall encounter in the next lecture.

---

(5) **Non-Principal Ideals.**

(a) Prove that the ideal $\langle x, y\rangle \subseteq \mathbb{K}[x, y]$ is not principal. Hint: if $\langle x, y\rangle = \langle g\rangle$, then g divides both x and y. What are their common divisors? Why can the ideal not be the whole ring?

(b) Show directly that over every field $\mathbb{I}(\{(0, 0)\}) = \langle x, y\rangle$. Explain why part (a) says that the ideal of *all* equations vanishing at the origin cannot be generated by one equation.

(c) Why does this not, by itself, rule out the possibility that a single polynomial might have the origin as its set of $\mathbb{K}$-valued zeros? Hint: Describe $\mathbb{V}_{\mathbb{R}}(x^2 + y^2)$.

(6) **Implications of Monomial Orders.**

(a) Deduce from the well-ordering axiom that there is no infinite strictly decreasing sequence of monomials.

(b) Prove that 1 is the smallest monomial in every monomial order. Hint: if $x^\alpha < 1$, use compatibility with multiplication and part (a) to compare $1, x^\alpha, x^{2\alpha}, \ldots$.

(c) Show that ordering monomials in $\mathbb{K}[x, y]$ by total degree is not a total order.

(d) Compare part (a) with the degree drop in one-variable division. What property of a monomial order will replace "the degree strictly decreases" in the proof of termination?

The definition above is appealing because it is intrinsic: it orders the monomials themselves, without first identifying them with anything else. To construct such orders, however, it is far easier to work with exponent vectors. Throughout, we write $\mathscr{M}_n$ for the set of monomials of $\mathbb{K}[x_1, \ldots, x_n]$, which is a commutative monoid under multiplication, just as $\mathbb{Z}^n_{\geq 0}$ is one under addition.

Definition 3. A *monomial order* on $\mathbb{Z}^n_{\geq 0}$ is a relation $\prec$ on $\mathbb{Z}^n_{\geq 0}$ satisfying the analogous conditions:

(i) $\prec$ is a strict total order;

(ii) if $\alpha \prec \beta$ and $\gamma \in \mathbb{Z}^n_{\geq 0}$ then $\alpha + \gamma \prec \beta + \gamma$; and

(iii) $\prec$ is a well-order.

This definition should look very similar to the preceding one. In fact, it is the same definition transported along a dictionary between $\mathscr{M}_n$ and $\mathbb{Z}^n_{\geq 0}$. Let us make that precise. Let $\Phi\colon \mathscr{M}_n \to \mathbb{Z}^n_{\geq 0}$ be a multiplicative bijection, meaning that $\Phi(mm') = \Phi(m) + \Phi(m')$. Given a monomial order $\prec$

on $\mathbb{Z}^n_{\geq 0}$, define a relation $<$ on $\mathscr{M}_n$ by $m_1 < m_2$ if and only if $\Phi(m_1) \prec \Phi(m_2)$. One checks that $<$ is indeed a monomial order.

Thus, specifying a monomial order in this way amounts to specifying two pieces of data: the isomorphism Φ and an order $\prec$ on $\mathbb{Z}^n_{\geq 0}$. The first piece is much simpler than it looks. Since every monomial is a product of variables and Φ is multiplicative, Φ is determined by its values on the variables. Further, the bijection requires Φ to take irreducible monomials to irreducible vectors, i.e., vectors that cannot be written as the sum of two nonzero vectors in $\mathbb{Z}^n_{\geq 0}$. Hence $\Phi(x_i)$ must equal e_j for some j.

Choosing Φ is therefore nothing more than choosing which variable is recorded in which coordinate. This is why a monomial order is customarily specified in two steps: first an ordering of the variables, say $x > y > z$, and then an ordering of $\mathbb{Z}^n_{\geq 0}$. Unless we say otherwise, we use the tautological identification $\Phi(x^\alpha) = \alpha$. We write $\alpha \succ \beta$ for $\beta \prec \alpha$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n$ for the total degree.

There are numerous monomial orders. For example, *Macaulay2* has at least twelve families of them, each of which can highlight a slightly different algebraic or geometric feature of a problem, as we shall see later. In this course we will largely restrict ourselves to the three most commonly used monomial orders. We informally refer to $|\alpha|$ as the (total) degree of α. (Think about why this is.)

(a) *Lexicographic order (Lex).* For distinct $\alpha, \beta \in \mathbb{Z}^n_{\geq 0}$, we declare $\alpha >_{\text{lex}} \beta$ if the leftmost nonzero entry of $\alpha - \beta$ is positive. In other words, in lex we compare α_1 and β_1; if they agree, we compare α_2 and β_2, and so on until the first difference appears.

(b) *Graded lexicographic order (GrLex).* We declare $\alpha >_{\text{grlex}} \beta$ if $|\alpha| > |\beta|$, or if $|\alpha| = |\beta|$ and $\alpha >_{\text{lex}} \beta$. Thus graded lex first compares total degree and uses lex to break ties.

(c) *Graded reverse lexicographic order (GRevLex).* We declare $\alpha >_{\text{grevlex}} \beta$ if $|\alpha| > |\beta|$, or if $|\alpha| = |\beta|$ and the rightmost nonzero entry of $\alpha - \beta$ is negative. Thus, when the total degrees tie, we scan the vectors from the right; at the first place where they differ, the vector with the *smaller* entry is declared larger.

The last two compare total degree first, so both refine the grading; they differ only in how they break a tie. Graded lexicographic order then prefers the monomial with the larger power of the earliest variable, while graded reverse lexicographic order prefers the one with the *smaller* power of the latest variable. Two warnings about the name are in order. Graded reverse lexicographic order is not the reverse of graded lexicographic order, nor is it obtained by reversing the variables; the sign condition and the scanning direction are both flipped, and for $n \leq 2$ the two orders in fact coincide. Notation also varies: grlex means graded lexicographic order in some sources and graded reverse lexicographic order in others, and deglex and $\mathrm{degrevlex}$ are common alternatives.

Each of these is a monomial order on $\mathbb{Z}^n_{\geq 0}$, and all three order the standard basis vectors as

$$e_1 > e_2 > \cdots > e_n.$$

What differs is how that comparison is extended to vectors of higher degree. To obtain monomial orders on $\mathbb{K}[x_1, \ldots, x_n]$, we must specify a multiplicative bijection $\Phi\colon \mathscr{M}_n \to \mathbb{Z}^n_{\geq 0}$ as discussed above. In practice, mathematicians encode Φ by naming a variable order and simply say, for example, lex with $x > y > z$ or lex with $y > x > z$. Let us look at a few examples to make this clearer.

Consider the monomial orders on $\mathbb{K}[x, y, z]$ arising from the variable orders $x > y > z$, $z > y > x$, and $y > z > x$, for lex, graded lex, and graded reverse lexicographic order.

- $x > y > z$: This variable order specifies the familiar bijection $\Phi(x^a y^b z^c) = (a, b, c)$, so that $\Phi(x) = (1, 0, 0)$, $\Phi(y) = (0, 1, 0)$, and $\Phi(z) = (0, 0, 1)$.
- $z > y > x$: Here $\Phi(x^a y^b z^c) = (c, b, a)$, so $\Phi(x) = (0, 0, 1)$, $\Phi(y) = (0, 1, 0)$, and $\Phi(z) = (1, 0, 0)$.
- $y > z > x$: Here $\Phi(x^a y^b z^c) = (b, c, a)$, so $\Phi(x) = (0, 0, 1)$, $\Phi(y) = (1, 0, 0)$, and $\Phi(z) = (0, 1, 0)$.

Now that we understand the three bijections, let us compare the corresponding monomial orders by arranging all monomials of degree at most 2 in the nine possible combinations. Within each cell, the line breaks are purely typographical: every new line continues the same decreasing chain.

	Variable order		
Monomial order	$x > y > z$	$z > y > x$	$y > z > x$
Lex	$x^2 > xy > xz > x$ $y^2 > yz > y$ $z^2 > z > 1$	$z^2 > yz > xz > z$ $y^2 > xy > y$ $x^2 > x > 1$	$y^2 > yz > xy > y$ $z^2 > xz > z$ $x^2 > x > 1$
GrLex	$x^2 > xy > xz$ $y^2 > yz > z^2$ $x > y > z > 1$	$z^2 > yz > xz$ $y^2 > xy > x^2$ $z > y > x > 1$	$y^2 > yz > xy$ $z^2 > xz > x^2$ $y > z > x > 1$
GRevLex	$x^2 > xy > y^2$ $xz > yz > z^2$ $x > y > z > 1$	$z^2 > yz > y^2$ $xz > xy > x^2$ $z > y > x > 1$	$y^2 > yz > z^2$ $xy > xz > x^2$ $y > z > x > 1$

---

(7) **Reading Monomial Orders.** Work in $\mathbb{K}[x, y, z]$ with $x > y > z$.

(a) Arrange the following monomials from largest to smallest in lex, graded lex, and graded reverse lexicographic order:

$$x^2, \quad xy, \quad xz, \quad y^2, \quad y^3, \quad z^2.$$

(b) In each order above, compare x^2 and y^3. Why does lex privilege the first variable while the graded orders first privilege total degree?

(c) Arrange all monomials of total degree at most 3 in $\mathbb{K}[x, y]$ in lex order. Use the definition to explain why $y^k < x$ for every $k \geq 1$. Thus every monomial involving only y is smaller than every monomial involving x. Why might this make lex useful for exposing consequences that eliminate x?

(d) Verify directly from the definitions that lex, graded lex, and graded reverse lexicographic order are compatible with multiplication.

(e) *Further: Verifying That These Are Monomial Orders.* Show that graded lexicographic and graded reverse lexicographic orders are well-orders by first choosing the least total degree occurring in a nonempty set of monomials and using the fact that only finitely many monomials have that degree. For lex, argue by induction on the number of variables: choose the least first coordinate which occurs and then apply the induction hypothesis to the remaining coordinates. Where does the proof use that every exponent is nonnegative?

Our goal in introducing monomial orders is to give each nonzero polynomial a unique distinguished leading term, just as we were able to do in one variable. This is the analogue of choosing a pivot in a row of a matrix. Toward this end, we introduce the following definition.

Definition 4. Let $<$ be a monomial order on $\mathbb{K}[x_1, \ldots, x_n]$. Let $0 \neq f \in \mathbb{K}[x_1, \ldots, x_n]$ and write $f = \sum_{i=1}^{t} c_i x^{\alpha_i}$, where $c_i \in \mathbb{K}^\times$ and the monomials x^{α_i} are distinct. Suppose that x^{α_j} is the largest monomial appearing in f with respect to $<$. The *leading monomial*, *leading coefficient*, and *leading term* of f with respect to $<$ are

$$\mathrm{LM}_<(f) = x^{\alpha_j}, \qquad \mathrm{LC}_<(f) = c_j, \qquad \mathrm{LT}_<(f) = c_j x^{\alpha_j}.$$

When the intended monomial order is clear, we often write $\mathrm{LT}(f)$, $\mathrm{LM}(f)$, and $\mathrm{LC}(f)$ for $\mathrm{LT}_<(f)$, $\mathrm{LM}_<(f)$, and $\mathrm{LC}_<(f)$.

---

(8) **Examples of Leading Terms.** Consider $f = x - y^2$ and $h = xz - y^2$ as elements of $\mathbb{K}[x, y, z]$.

(a) Compute the leading term, leading coefficient, and leading monomial of both f and h with respect to lex, graded lex, and graded reverse lexicographic order, with $x > y > z$.

(b) Run the following in a new *Macaulay2* session. After running the code, replace `Lex` first by `GLex` and then by `GRevLex`, restarting each time. Compare the three outputs with your answers to part (a).

```
restart
R = QQ[x,y,z,MonomialOrder=>Lex];
f = x-y^2;    -- the equation y^2=x in the xy-plane
h = x*z-y^2; -- the hidden twisted cubic relation

{leadMonomial f,leadCoefficient f,leadTerm f}
leadMonomial h
```

(c) In *Macaulay2*, we give a one-input function a name using the arrow notation

```
name = input -> output.
```

A parenthesized block may contain several statements separated by semicolons; the final expression without a semicolon is the value returned by the function. For example, the following function returns the list $\{\mathrm{LM}(p), \mathrm{LC}(p), \mathrm{LT}(p)\}$:

```
leadData = p -> (
    LMp := leadMonomial p;
    LCp := leadCoefficient p;
    LTp := leadTerm p;
    {LMp, LCp, LTp}
)
leadData f
apply({f,h},p -> leadData p)
```

Predict both outputs before running the code, then explain how the final line applies the function to every entry of the list `{f,h}`.

(9) **Working with Leading Terms.** Let $u, v \in \mathbb{K}[x_1, \ldots, x_n]$ be nonzero polynomials and fix a monomial order $<$.

(a) Prove that $\mathrm{LM}(uv) = \mathrm{LM}(u)\,\mathrm{LM}(v)$ and $\mathrm{LT}(uv) = \mathrm{LT}(u)\,\mathrm{LT}(v)$.

(b) Give an example with $u + v \neq 0$ showing that $\mathrm{LM}(u + v)$ need not equal either $\mathrm{LM}(u)$ or $\mathrm{LM}(v)$.

We are finally ready to return to the plan that led us to monomial orders in the first place: imitating long division in several variables. One problem remains. We no longer wish to divide f by a single polynomial g, but instead want to make sense of dividing f by a list of polynomials $G = (g_1, \ldots, g_s)$. We will always require that G be treated as an *ordered* list, and, just as in ordinary division, it is important that none of the g_i be zero. Although we might not yet have much intuition for this goal, let us try to mimic the one-variable algorithm, noting the adjustments we make along the way.

Suppose that we wish to divide a polynomial f by an ordered list of nonzero polynomials $G = (g_1, \ldots, g_s)$. Start by initializing a working polynomial $p = f$ and quotient polynomials $q_1 = \cdots = q_s = 0$, which we shall update as we go. Unlike before, we also keep a separate remainder polynomial $r = 0$, which will be updated at each step. Scanning through the list G, look for the first g_i such that $\mathrm{LM}(g_i)$ divides $\mathrm{LM}(p)$. If one exists, then, just as before, the term

$$\frac{\mathrm{LT}(p)}{\mathrm{LT}(g_i)}$$

is exactly the term by which we must multiply g_i in order to cancel the leading term of p. Here is one new wrinkle. If none of the g_i can cancel $\mathrm{LT}(p)$, we cannot simply stop: one of the smaller terms of p might still be divisible by a leading monomial of one of the g_i. Instead, we move the leading term of p into a separate remainder and continue. Thus, in addition to choosing a monomial order, we must choose an ordering of the divisors. Once both choices have been made, the division procedure is completely determined.

Algorithm: Multivariable Division by $G = (g_1, \ldots, g_s)$
Input: f; nonzero divisors $g_1, \ldots, g_s$ in a fixed order; and a fixed monomial order.
Initialize: $q_1, \ldots, q_s \leftarrow 0$, $r \leftarrow 0$, and $p \leftarrow f$.
Repeat while $p \neq 0$*:*

(1) Scan $g_1, \ldots, g_s$ in order.

(2) If g_i is the first divisor in the list for which $\mathrm{LM}(g_i) \mid \mathrm{LM}(p)$, set

$$m \leftarrow \frac{\mathrm{LT}(p)}{\mathrm{LT}(g_i)}, \qquad q_i \leftarrow q_i + m, \qquad p \leftarrow p - mg_i,$$

and return to the beginning of the list.

(3) If no such g_i exists, set

$$r \leftarrow r + \mathrm{LT}(p), \qquad p \leftarrow p - \mathrm{LT}(p).$$

Output: $q_1, \ldots, q_s, r$.

Notice that just as with one-variable division, we have chosen our bookkeeping so that the identity

$$f = q_1 g_1 + \cdots + q_s g_s + p + r$$

is preserved after every update. This algorithm will terminate since at each stage the lead term of the working polynomial p strictly decreases with respect to the chosen monomial order $<$. Hence termination follows from the well-ordering condition in the definition of a monomial order. When the algorithm does stop, no monomial of the remainder is divisible by any $\mathrm{LM}(g_i)$. Importantly, if the final remainder is zero then at the end the algorithm has found an identity

$$f = q_1 g_1 + \cdots + q_s g_s,$$

which is a certificate that f is in the ideal generated by $g_1, \ldots, g_s$. More generally, the preserved division identity always gives

$$f \equiv r \pmod{\langle g_1, \ldots, g_s \rangle}.$$

This means that f and r represent the same class in the quotient ring $\mathbb{K}[x_1, \ldots, x_n]/\langle g_1, \ldots, g_s \rangle$. As a warning, at the present moment we do not know that r is uniquely determined and so this representation might not be canonical. Even more worryingly while reminder zero proves membership it is not clear the converse – nonzero remainder implies nonmembership – is true. In fact it is not!

---

(10) **Dividing the Circle–Parabola Elimination Equation.** Work in $\mathbb{K}[x, y]$ with lex order $x > y$. Let us return to the circle–parabola intersection from Lecture 1. Let $f = x^4 + x^2 - 1$ and $G = (g_1, g_2) = (x^2 - y, y^2 + y - 1)$. Here g_1 is the parabola equation, g_2 is the y-elimination equation from the circle–parabola problem that opened Lecture 1, and f is the corresponding x-elimination equation. Notice that the divisors have been organized in a special way: Exercise 14 will explain where the second divisor comes from, and Exercise 20 will eventually prove that the resulting list is a Gröbner basis.

(a) Divide f by the ordered list G, recording the values of p, q_1, q_2, r after every update.

(b) Verify after every update that the identity $f = q_1 g_1 + q_2 g_2 + p + r$ remains true.

(c) Show that the final remainder is zero. Record the resulting identity

$$x^4 + x^2 - 1 = q_1(x^2 - y) + q_2(y^2 + y - 1).$$

Why does this identity serve as a certificate for $f \in \langle g_1, g_2 \rangle$?

(d) Viewing $g_1 = g_2 = 0$ as equations, use $y = x^2$ to verify directly that every common zero satisfies $f = 0$.

(e) Why is the ideal-membership certificate from part (c) stronger than this pointwise verification in part (d)? If $\mathbb{K}$ is algebraically closed, what converse follows from the Nullstellensatz, and why does it place f only in the radical of the ideal?

(f) Show that $\langle g_1, g_2 \rangle = \langle x^2 - y, x^2 + y^2 - 1 \rangle$. Conclude that the two displayed systems have the same common zero set, and interpret the certificate from part (c) as an algebraic version of the earlier elimination-by-substitution computation.

(11) **Why the Algorithm Terminates.** Fix a monomial order $<$ on $\mathbb{K}[x_1, \ldots, x_n]$. Let $f, g_1, \ldots, g_s \in \mathbb{K}[x_1, \ldots, x_n]$, with $g_i \neq 0$ for all i. Regard $G = (g_1, \ldots, g_s)$ as an ordered list.

(a) Prove that at every stage of the multivariable division algorithm the identity

$$f = q_1 g_1 + \cdots + q_s g_s + p + r$$

is preserved. Hint: See how the identity changes with both kinds of update in the algorithm.

(b) Suppose some g_i is used to cancel $\mathrm{LT}(p)$. Prove that the new working polynomial p is either zero or has a leading monomial strictly smaller with respect to $<$ than it had at the previous step.

(c) Suppose instead that $\mathrm{LT}(p)$ is moved to the remainder. Prove the same conclusion.

(d) Use Exercise 6(a) to prove that the algorithm terminates.

(e) Prove that on termination no monomial of r is divisible by $\mathrm{LM}(g_i)$ for any i.

(f) Prove that if the remainder is zero, then $f \in \langle g_1, \ldots, g_s \rangle$. Hence, a zero remainder certifies ideal membership. Notice that we have not shown the converse!

(g) Prove that the output satisfies the following identities:

$$\mathrm{LM}(q_i g_i) \leq \mathrm{LM}(f) \quad \text{for every } i \text{ with } q_i \neq 0, \qquad \mathrm{LM}(r) \leq \mathrm{LM}(f) \quad \text{if } r \neq 0.$$

Hint: Every subtraction cancels the current leading monomial, and leading monomials never increase. A representation $f = \sum_i q_i g_i$ satisfying the first set of inequalities is called a *standard representation*.

So far this looks remarkably like one-variable division: the process stops, the remainder contains no further cancellable monomials, and a zero remainder produces a certificate of membership. It is tempting to declare victory. Unfortunately, there are two separate problems. First, the remainder depends on the order in which we list the divisors. Second, and much more seriously, a polynomial lying in the ideal can have a nonzero remainder. Thus the implication

$$\text{the remainder is zero} \quad \Longrightarrow \quad f \in \langle g_1, \ldots, g_s \rangle$$

always holds, but the converse can fail. The affine twisted cubic already shows us both failures without requiring any new geometry.

---

(12) **The Remainder Is Not Intrinsic: The Twisted Cubic Again.** Work in $\mathbb{K}[x, y, z]$ with lex order $x > y > z$, and consider the affine twisted cubic. Set $g_1 = x^2 - y$, $g_2 = xy - z$, and $I = \langle g_1, g_2 \rangle$. Assuming $\mathbb{K}$ is infinite, this is the defining ideal of the parametrized curve $t \mapsto (t, t^2, t^3)$ that we called the affine twisted cubic in Lecture 1. (Remember that we showed this, in part, by computing the kernel of the corresponding substitution map.) We have changed the signs of the displayed generators so that their leading coefficients are 1.

(a) Divide $x^2 y$ by (g_1, g_2) and show that the remainder is y^2.

(b) Divide the same polynomial by (g_2, g_1) and show that the remainder is xz.

(c) Verify the identity $y g_1 - x g_2 = xz - y^2$, which appeared at the beginning of the lecture. Notice that it gives a representation of $xz - y^2$ in terms of g_1, g_2, but this is not a standard representation: compare $\mathrm{LM}(y g_1) = \mathrm{LM}(x g_2) = x^2 y$ with $\mathrm{LM}(xz - y^2) = xz$.

(d) Recall that the substitution map sends $x \mapsto t$, $y \mapsto t^2$, and $z \mapsto t^3$. Check that both remainders y^2 and xz map to t^4. Explain why they represent the same polynomial function on the

twisted cubic even though division has produced different representatives. Verify directly that their difference is the relation from part (c).

(e) Compute $\mathrm{LM}(xz-y^2)$. Divide $xz-y^2$ by either displayed list and explain why the algorithm returns the same nonzero polynomial, even though it lies in I.

(f) A zero remainder proves ideal membership. Does a nonzero remainder prove nonmembership? Use part (e) to answer, and explain why this failure is more serious than the mere nonuniqueness in parts (a) and (b).

(g) *Further: Finite Fields.* Over $\mathbb{K} = \mathbb{F}_q$, equality as functions on the finite set is coarser than equality modulo I: the class of $x^q - x$ is nonzero in $\mathbb{K}[x,y,z]/I \cong \mathbb{K}[t]$, but it vanishes at every point of the parametrized curve. Explain why this does not contradict $I = \ker(F^*)$, and compare with the twisted cubic exercise in Lecture 1, where polynomial identities differed from equality as functions over a finite field.

The twisted cubic calculation tells us exactly what the original list failed to see. Its leading monomials were x^2 and xy, but cancellation produced the relation $xz - y^2 \in I$, whose leading monomial is xz. Thus the leading monomials of the displayed generators need not account for the leading monomials of all their polynomial consequences. This suggests collecting the leading terms of every element of the ideal.

Definition 5. Let $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$ be an ideal and fix a monomial order $<$. The *initial ideal* of I with respect to $<$ is

$$\mathrm{in}_<(I) = \langle \mathrm{LT}(f) \mid 0 \neq f \in I \rangle .$$

Because $\mathbb{K}$ is a field, the initial ideal of I is also equal to the ideal generated by the leading monomials $\mathrm{LM}(f)$. The zero ideal requires a bit of careful handling, so we set $\mathrm{in}_<(\langle 0 \rangle) = \langle 0 \rangle$.

If $I = \langle g_1, \ldots, g_s \rangle$, then

$$\langle \mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_s) \rangle \subseteq \mathrm{in}_<(I).$$

The reverse inclusion can fail, so equality need not hold. The twisted cubic has already shown us this: $\langle \mathrm{LT}(g_1), \mathrm{LT}(g_2) \rangle = \langle x^2, xy \rangle$, while $f = xz - y^2 \in \langle g_1, g_2 \rangle$ but $\mathrm{LT}(f) = xz \notin \langle x^2, xy \rangle$.

The issue causing the failure of the other inclusion is cancellation: polynomial combinations of the g_i may have leading monomials that are not divisible by any of the displayed $\mathrm{LM}(g_i)$. This is because the operation of "taking a leading term" does not necessarily commute with taking polynomial combinations. We saw this previously when we found examples in which $\mathrm{LM}(u+v)$ was not determined by $\mathrm{LM}(u)$ and $\mathrm{LM}(v)$.

This suggests a definition designed precisely to prevent the failure.

Definition 6. Let $I \subseteq \mathbb{K}[x_1, \ldots, x_n]$ be an ideal and fix a monomial order $<$. A finite list $G = (g_1, \ldots, g_s)$ of nonzero elements of I is a *Gröbner basis* of I with respect to this order if

$$\mathrm{in}_<(I) = \langle \mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_s) \rangle.$$

For the zero ideal we allow the empty list, whose leading terms generate $\langle 0 \rangle$.

There is something slightly surprising about this definition. We require that each g_i lie in I and that their leading terms generate the initial ideal of I, but we have not assumed that the g_i generate I itself. Additionally, it is not at all clear from the definition that Gröbner bases even exist. In Exercise 18 you will prove that the leading-term condition forces $I = \langle g_1, \ldots, g_s \rangle$ and is exactly what is needed for division to decide ideal membership. Before proving this formally, we will use the circle–parabola and twisted cubic examples to see concretely what the leading-term condition repairs. The proof of the existence of Gröbner bases will be one of the main goals of Lecture 3, where we will seek to prove another celebrated result of Hilbert: Hilbert's Basis Theorem.

A warning: one should not confuse I and $\mathrm{in}_<(I)$; their algebraic and geometric properties can be very different. In general, one should view the passage from an ideal I to its initial ideal $\mathrm{in}_<(I)$ as a noncanonical, lossy process. It is noncanonical because it depends on the choice of monomial order $<$, and lossy because one cannot generally recover I from $\mathrm{in}_<(I)$; much of the geometry and algebra may be lost. For example, if $I = \langle y - x^2 \rangle \subseteq \mathbb{K}[x, y]$ is considered with lex order $x > y$, then $\mathrm{in}_<(I) = \langle x^2 \rangle$. Thus $\mathbb{V}(I)$ is the familiar parabola, while the zero set $\mathbb{V}(\mathrm{in}_<(I))$ is just the y-axis. Lecture 5 will center on the question: which pieces of the algebra and geometry of I does $\mathrm{in}_<(I)$ remember?

One indication of why passing to initial ideals loses information is that initial ideals are a very special type of ideal. In particular, every leading term is a nonzero scalar multiple of a monomial. Since $\mathbb{K}$ is a field, $\mathrm{in}_<(I)$ is generated equally well by the leading monomials, and hence is a monomial ideal.

Definition 7. An ideal $M \subseteq \mathbb{K}[x_1, \ldots, x_n]$ is a *monomial ideal* if it is generated by some collection of monomials.

Notice that we did not require the collection of monomial generators in the definition above to be finite. Lecture 3 will show that a finite collection always suffices.

Monomial ideals are exceptionally nice. For example, a monomial ideal is entirely determined by the set of monomials belonging to it. Since monomials in $\mathbb{K}[x_1, \ldots, x_n]$ are in bijection with $\mathbb{Z}^n_{\geq 0}$, this means that we can—and often do—think of monomial ideals as corresponding to subsets of $\mathbb{Z}^n_{\geq 0}$. Moreover, ideal membership for monomial ideals is particularly simple. We will frequently use the following result.

Here an important question is hiding in plain sight. The definition of $\mathrm{in}_<(I)$ used a leading monomial from every nonzero polynomial in I, so it appears to require infinitely much information. A

monomial ideal, however, looks like an upward-closed staircase. If such a staircase has only finitely many corners, then those corner monomials generate the whole ideal. We will prove in Lecture 3 that this always happens; for now the pictures below should make the claim feel plausible.

Lemma 8. *Let $I = \langle m_1, \ldots, m_s \rangle \subseteq \mathbb{K}[x_1, \ldots, x_n]$ be a monomial ideal, with each m_i a monomial. If $m \in \mathbb{K}[x_1, \ldots, x_n]$ is a monomial, then $m \in I$ if and only if m is divisible by m_i for some i.*

This result motivates the following definition.

Definition 9. If M is a monomial ideal, a monomial not contained in M is called a *standard monomial* of M.

Many of the ideas we have seen thus far were developed in the 1960s. For example, Gröbner bases were introduced by Bruno Buchberger in his 1965 dissertation, *An Algorithm for Finding the Basis Elements of the Residue Class Ring of a Zero-Dimensional Polynomial Ideal*. He named these bases for his advisor, Wolfgang Gröbner. That said, the concept of Gröbner bases had been in the mathematical air. For example, in his celebrated 1964 proof of the existence of resolutions of singularities in characteristic zero—work that would earn him the 1970 Fields Medal—Heisuke Hironaka introduced the closely related notion of *standard bases*.

---

(13) **Monomial Ideals.** Let $M = \langle m_1, \ldots, m_s \rangle \subseteq \mathbb{K}[x_1, \ldots, x_n]$ be an ideal generated by monomials.

- (a) Prove that a monomial m lies in M if and only if some m_i divides m. For the less immediate direction, write $m = \sum_i h_i m_i$ and use the linear independence of distinct monomials.
- (b) Deduce that a polynomial lies in M if and only if each monomial appearing in it lies in M.
- (c) Conclude that $\mathrm{in}_<(M) = M$ for every monomial order.
- (d) In two variables, draw $x^a y^b$ at the lattice point (a, b). Explain why every point weakly above and to the right of a point of M also lies in M.
- (e) In your picture from the previous part, draw the standard monomials.
- (f) Draw pictures of the following monomial ideals and list their standard monomials:

$$\langle x^2, y^2 \rangle \qquad \text{and} \qquad \langle x^2, y \rangle.$$

(14) **A Hidden Consequence of the Circle–Parabola Equations.** Work in $\mathbb{K}[x, y]$ with lex order $x > y$ and set $f_1 = x^2 - y$, $f_2 = x^2 + y^2 - 1$, and $I = \langle f_1, f_2 \rangle$.

- (a) Compute $\mathrm{LM}(f_1)$ and $\mathrm{LM}(f_2)$. What is the ideal generated by these two leading monomials?
- (b) Consider the difference $h = f_2 - f_1 = y^2 + y - 1$. Show that $h \in I$ and compute $\mathrm{LM}(h)$.
- (c) Conclude that $\langle \mathrm{LM}(f_1), \mathrm{LM}(f_2) \rangle$ is a proper subset of $\mathrm{in}_<(I)$. Which leading monomial was invisible in the original list?

(d) Divide h by the ordered list (f_1, f_2). Can either $\mathrm{LM}(f_1)$ or $\mathrm{LM}(f_2)$ be used to cancel the leading term of h at the first step? What remainder does the algorithm return, and why is this apparently uneventful calculation a decisive failure of the membership test?

(e) Geometrically, what does $h(y) = 0$ tell us about the possible y-coordinates of the intersection points? Why is it natural to call h an elimination equation?

The cancellation has exposed the elimination equation which the original leading terms could not see. The natural repair is to replace f_2 by this consequence h. Let us separate two facts about this replacement: it does not change the ideal, and it gives a particularly simple staircase.

---

(15) **The Circle–Parabola Staircase.** Continue with the notation above and set $G' = (f_1, h)$.

(a) Use $h = f_2 - f_1$ to prove that $\langle f_1, f_2 \rangle = \langle f_1, h \rangle$. Thus replacing f_2 by the elimination equation changes the organization of the generators without changing the ideal or the circle–parabola intersection.

(b) Return to Exercise 10 and notice that G' was exactly the ordered list used there.

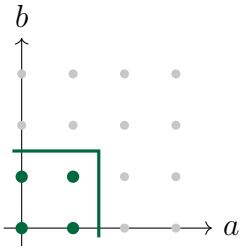
(c) Which monomials are divisible by neither $\mathrm{LM}(f_1) = x^2$ nor $\mathrm{LM}(h) = y^2$? List all of them without imposing a degree bound. Compare the resulting list $1, x, y, xy$ with the basis $1, x, x^2, x^3$ found for the circle–parabola quotient in Lecture 1.

(d) Use $x^2 \equiv y$ and $x^3 \equiv xy$ to explain why the two lists describe the same four-dimensional quotient.

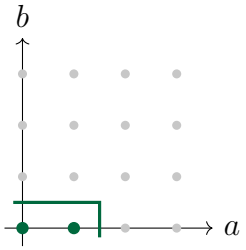
This explains why our first multivariable division worked so well. We had already replaced the original second equation by precisely the consequence whose leading monomial was missing. The ideal and its common zero set did not change; only its presentation became better adapted to division with our chosen monomial order.

The staircase gives us a picture of this improved presentation. In the diagrams below, green lattice points are standard monomials and gray points lie in the monomial ideal. Only the first few exponent vectors are shown.

This picture also returns us to F. S. Macaulay. In studying Hilbert functions—which we have not yet defined—Macaulay organized quotients by counting the monomials that survive modulo a monomial ideal. We are seeing the simplest version of that idea here: the points beneath the staircase will become a basis of the quotient.



$\langle x^2, y^2 \rangle$: $1, x, y, xy$



$\langle x^2, y \rangle$: $1, x$

The monomial quotient $\mathbb{K}[x, y]/\langle x^2, y^2 \rangle$ has $\mathbb{K}$-vector-space dimension 4, while its zero set consists only of the origin. For the finite-dimensional quotient rings considered here, we use *length* to mean this vector-space dimension; thus this quotient has length 4. Its four standard monomials $1, x, y, xy$ also index a basis of the original circle–parabola quotient. Over $\mathbb{C}$, the latter length-4 quotient is instead realized by four reduced intersection points, just as the $2 \cdot 2 = 4$ count from Bézout's theorem predicts. For $\langle x^2, y \rangle$, only 1 and x survive. This quotient is a doubled point of length 2, and the nonzero nilpotent class of x records its tangent direction.

Notice that a staircase is not a drawing of the original algebraic set. It is a picture of monomial divisibility. Its complement will give us representatives in the quotient, and it is the quotient—not merely the set of visible points—that remembers multiplicity.

---

(16) **The Hidden Twisted Cubic Relation.** Return to the ideal $I = \langle x^2 - y, xy - z \rangle \subseteq \mathbb{K}[x, y, z]$ from Exercise 12, which defines the affine twisted cubic. Still working with lex order $x > y > z$, we already found one leading monomial missing from the original list. Let us see what happens after we add the corresponding relation.

(a) Using Exercise 12(c) and (e), conclude that $\langle x^2, xy \rangle \neq \mathrm{in}_<(I)$. Which cancellation produced the invisible monomial?

(b) Set $g_3 = xz - y^2$. The leading monomial of g_3 is xz, and it can cancel against a multiple of $\mathrm{LM}(g_2) = xy$. Form the second cancellation

$$zg_2 - yg_3 = y^3 - z^2.$$

Verify the identity, show that $y^3 - z^2 \in I$, and compute its leading monomial. Why is this leading monomial also absent from $\langle x^2, xy, xz \rangle$?

(c) At this point our candidate for the initial ideal is $M = \langle x^2, xy, xz, y^3 \rangle$. Show that the standard monomials for M are

$$\{x\} \sqcup \{y^b z^c : 0 \leq b \leq 2,\ c \geq 0\}.$$

(d) Under the substitution $x \mapsto t$, $y \mapsto t^2$, $z \mapsto t^3$, determine the power of t represented by each of the standard monomials in the previous part. Show that t^d occurs exactly once for every $d \geq 0$. Hint: Separate d according to its residue modulo 3, treating $d = 1$ using x.

(e) Why is the pattern from the previous part suggested by the ring isomorphism $\mathbb{K}[x,y,z]/I \cong \mathbb{K}[t]$ proved in Lecture 1?

(f) Return for a moment to the three relations g_1, g_2, g_3. Arrange them, up to sign, as the 2×2 minors of

$$\begin{pmatrix} 1 & x & y \\ x & y & z \end{pmatrix}.$$

Thus the cancellation in Exercise 12 did not produce an accidental relation. Which minor had been omitted by the original two-generator list?

(g) *Further: The Same Ideal in a Different Order.* Use graded reverse lexicographic order with $x > y > z$. Show that the three minors have leading monomials x^2, xy, y^2. List the standard monomials of $\langle x^2, xy, y^2 \rangle$ as

$$z^c, \quad xz^c, \quad yz^c \qquad (c \geq 0),$$

and show that they again map bijectively onto the powers of t. Record this calculation; Exercise 20 will turn it into a proof that the three minors already form a Gröbner basis in this order. Which order gives the shorter basis? Which order exposes the equation $y^3 - z^2$ involving only y, z?

(h) The relation $xz - y^2$ has now appeared both as an element of the kernel computed in Lecture 1 and as the polynomial whose leading monomial was missing from the original list. Explain what it tells us in each role. Notice that no affine equation has been lost: $xz - y^2$ is already a polynomial combination of the two displayed generators.

Let us pause over what just happened. In Lecture 1 we chose $x^2 - y$ and $xy - z$ because they make the points of the twisted cubic easy to recognize: once x is known, the other coordinates must be x^2 and x^3. For division in lex order, however, this convenient geometric presentation was badly organized. Both $xz - y^2$ and $y^3 - z^2$ were already in the ideal, but their leading monomials were invisible in the original list. Adding these equations did not change the ideal or the curve. It only made information already present in the ideal visible to the division algorithm. Keep the relation $xz - y^2$ in mind. At the end of Lecture 3 it will become essential when we pass from the affine twisted cubic to its projective closure.

The staircase gives rather convincing evidence that we have now found all the missing leading terms. Its proposed standard monomials map to $1, t, t^2, \ldots$, each exactly once, just as $\mathbb{K}[x,y,z]/I \cong \mathbb{K}[t]$ predicts. Still, evidence is not proof: perhaps some further cancellation produces another leading monomial and shrinks the staircase. We will shortly use the quotient description from Lecture 1 to rule this out without assuming what we are trying to prove.

For this small example we could discover the missing relations by hand. In larger examples, this quickly becomes unreasonable, so let us now ask *Macaulay2* to perform the same reorganization and compare its output with the geometry and algebra we already understand.

---

(17) **Hidden Leading Terms in *Macaulay2*.** Let us return to the circle–parabola ideal via the *Macaualy2* code below, which will compute a Gröbner basis for the ideal $I = \langle x^2 - y, x^2 + y^2 - 1 \rangle$ with respect to the lex monomial order with $x > y > z$. Further, the code snippet below will compute the leading terms of the computed Gröbner basis, and the initial ideal of I.

```
restart
R = QQ[x,y,MonomialOrder=>Lex];
I = ideal(x^2-y, x^2+y^2-1); -- circle--parabola ideal

gens I
shownLeadingIdeal = ideal leadTerm gens I -- only the displayed
   generators

B = gb I;
G = gens B
fullLeadingIdeal = ideal leadTerm G -- leading terms after all
   cancellations
directLeadingIdeal = leadTerm I     -- an ideal input triggers a GB
   computation
directLeadingIdeal == fullLeadingIdeal

h = y^2+y-1;
h % B -- reduce to verify that the elimination equation lies in I
```

In the code above the command `gens I` returns a list of generators for I stored as a row vector. Note this will generally be the generating set you used to initial describe I. The command `gb I` asks *Macaulay2* to compute a Gröbner basis for I. To access the elements in the computed Gröbner basis we use the command `gens gb I`. There is a subtle but important distinction in the code above. Doing `leadTerm I` asks *Macaulay2* to compute the initial ideal of I, which it does by first computing a Gröbner basis for I and then returning the full initial ideal. In contrast, running `ideal leadTerm gens I` computes the ideal generated by the leading terms of a generating set for I. As we have seen these two ideals can be very different.

(a) Explain why `shownLeadingIdeal` uses only the leading terms of the displayed generators, whereas `fullLeadingIdeal` uses the basis computed from the whole ideal. Which missing leading monomial distinguishes the outputs?

(b) The object B stores a special basis computed from I. Compare `G` with `F`. Which new equation makes the missing leading monomial visible?

(c) Compare the zero remainder `h % B` with the hand division in Exercise 14(d). Explain why the two computations use different lists of divisors.

(d) A function with two inputs is defined by placing the inputs in parentheses. Write a function named `hasZeroRemainder` which takes a polynomial p and a Gröbner basis object C and returns `true` precisely when `p % C` is zero. Use it with B and `apply` to test the three polynomials

$$h, \qquad x, \qquad x^4 + x^2 - 1$$

in a single command, predicting the three Boolean outputs before asking the computer. *Hint:* Inside `apply` use a one-input function of the form `p -> hasZeroRemainder(p,B)`. What ideal-membership principle do these outputs suggest? The next exercise will prove that this interpretation is valid.

(e) *Write the code yourself.* Start a fresh session with lex order $y > x$; remember that variables are ordered as they are declared in the ring. Compute a Gröbner basis and verify that its initial ideal is $\langle y, x^4 \rangle$, with standard monomials $1, x, x^2, x^3$. For lex $x > y$, the standard monomials were $1, x, y, xy$. Use $y \equiv x^2$ and $xy \equiv x^3$ to pass between the two lists in the quotient. Why do both orders record quotient dimension 4? Which order reveals an equation in y, and which reveals one in x?

(f) *Further: The Twisted Cubic.* Repeat this exercise for the affine twisted cubic in lex order with $x > y > z$. Locate the hidden monomial xz in the full initial ideal. Verify that the additional leading monomial is y^3, as predicted in Exercise 16. Compare the standard monomials with the powers of t found there.

The computer has not changed our equations in any geometric sense. It has replaced one generating set by another for the same ideal, just as Gaussian elimination replaces a linear system by row-echelon form. The new list may look different, but it has exactly the same polynomial consequences, and its leading terms are organized so that division can see them.

For the circle–parabola ideal, the orders $x > y$ and $y > x$ select the two bases

$$1, x, y, xy \qquad \text{and} \qquad 1, x, x^2, x^3$$

of the same four-dimensional quotient. The remainder chosen by division depends on the monomial order, just as coordinates depend on a choice of basis, but the algebraic intersection itself has not changed.

We have now met the same obstruction by hand and on the computer. The next exercise proves that our definition of a Gröbner basis repairs it. Suppose that $f \in I$ leaves a remainder r. What

does the division identity tell us about r, and what could its leading monomial look like? This is precisely where the equality

$$\mathrm{in}_<(I) = \langle \mathrm{LM}(g_1), \ldots, \mathrm{LM}(g_s) \rangle$$

enters. This is why the definition is the right one: once no leading term can hide, a remainder can no longer falsely claim that an element of the ideal lies outside it.

---

(18) **When Does Division Decide Membership?** Let $S = \mathbb{K}[x_1, \ldots, x_n]$, let $I \subseteq S$ be an ideal, and fix a monomial order. Let $G = (g_1, \ldots, g_s)$ be a finite list of nonzero elements of I.

(a) Suppose the list is a Gröbner basis, meaning that

$$\mathrm{in}_<(I) = \langle \mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_s) \rangle = \langle \mathrm{LM}(g_1), \ldots, \mathrm{LM}(g_s) \rangle.$$

Let $f \in I$ and divide f by $(g_1, \ldots, g_s)$, obtaining remainder r. Show that $r \in I$.

(b) If $r \neq 0$, then $\mathrm{LM}(r) \in \mathrm{in}_<(I)$. Use the monomial-ideal criterion proved above to explain why the displayed equality forces some $\mathrm{LM}(g_i)$ to divide $\mathrm{LM}(r)$.

(c) On the other hand, no monomial of a remainder is divisible by any $\mathrm{LM}(g_i)$. Explain why this contradicts part (b), and conclude that $r = 0$.

(d) For this fixed ordered list G, write $\mathrm{rem}_G(f)$ for the remainder. Conclude that

$$f \in I \quad \Longleftrightarrow \quad \mathrm{rem}_G(f) = 0.$$

For the reverse implication, use $\langle g_1, \ldots, g_s \rangle \subseteq I$.

(e) Since every $f \in I$ has zero remainder, use the division identity to conclude that $I = \langle g_1, \ldots, g_s \rangle$. Thus a Gröbner basis automatically generates its ideal, although this was not required in the definition.

(f) Suppose division by G returns zero for every $f \in I$. Let $0 \neq f \in I$. Explain why, at the first stage of the algorithm, some $\mathrm{LM}(g_i)$ must divide $\mathrm{LM}(f)$: otherwise $\mathrm{LT}(f)$ would be moved permanently into a nonzero remainder. Indeed, the algorithm never removes a term already placed in the remainder, and every later leading monomial is strictly smaller. Deduce that

$$\mathrm{in}_<(I) = \langle \mathrm{LM}(g_1), \ldots, \mathrm{LM}(g_s) \rangle.$$

Be sure to use $g_i \in I$ for the reverse containment.

(g) Suppose $I = \mathbb{I}(X)$ for an algebraic set X. Interpret this membership test geometrically as deciding whether the hypersurface $\mathbb{V}(f)$ contains X. For example, $xz - y^2 \in \mathbb{I}(C)$ says that the corresponding quadric contains the twisted cubic. Why must this interpretation be qualified when I is merely an ideal cutting out X, rather than the full vanishing ideal?

We have now reached the goal stated at the beginning of the lecture. For an arbitrary finite list of divisors, a zero remainder always gives a certificate of membership. While the converse is true, we know that if the list happens to be a Gröbner basis then a nonzero remainder is a finite certificate for nonmembership in the ideal. This begs a number of questions: How does one check if a given finite list of polynomials is a Gröbner basis? Do finite Gröbner bases exist for every ideal? If Gröbner bases do exist how do we find them? Lecture 3 will be dedicated to answering all of these questions.

There is also one other closely related finiteness question floating around, which remains unresolved. We proved that *if* an ideal has a finite Gröbner basis, then that basis generates the ideal. However, we have not yet shown that every ideal is finitely generated, yet along finitely generated by a Gröbner basis. By applying this result to $\mathbb{I}(X)$ for an affine algebraic variety X would imply that every affine variety can be cut out by finitely many equations. This is the classical field case of Hilbert's Basis Theorem, and it will also be one of our goals in Lecture 3.

However, recall that in one-variable, division did more than just answer the ideal membership with a finite certificate. That finite certificate gave us a canonical representative for every class in the corresponding quotient ring. We will now show that our multivariable division algorithm does this as well. A *standard monomial* for an ideal I is a monomial not contained in $\mathrm{in}_<(I($. The standard monomials will play the role that $1, x, \ldots, x^{\deg(g)-1}$ played in the one-variable case, i.e., the will form a basis for $\mathbb{C}[x_1, \ldots, x_n]/I$.

---

(19) **Canonical Normal Forms and Quotient Bases.** Continue with the notation from Exercise 18, and suppose that $G = (g_1, \ldots, g_s)$ is a Gröbner basis for I. Let $f \in \mathbb{K}[x_1, \ldots, x_n]$. We will show that the remainder of f divided by G exists, is unique, and can be expressed in terms of the standard monomials of I.

(a) Suppose that $r, r' \in \mathbb{K}[x_1, \ldots, x_n]$ such that $f - r \in I$ and $f - r' \in I$. Prove that $r - r' \in I$.

(b) Assume that $r - r' \neq 0$. Prove that if every monomial appearing in r or every monomial appearing in r' is a standard monomial for I then $\mathrm{LM}(r - r') \in \mathrm{in}_<(I)$.

(c) Assume that $r - r' \neq 0$. Prove that every monomial in $r - r'$ is a standard monomial. Deduce that $r = r'$.

(d) Divide f by G obtaining the following output division identity:

$$f = q_1 g_1 + \cdots q_s g_s + r$$

where no monomial appearing in r is divisible by any of $\mathrm{LM}(g_1), \ldots, \mathrm{LM}(g_s)$. Prove that every monomial in r is standard. Hint: Since G is a Gröbner basis we know that $\mathrm{in}_<(I) =$

$\langle(\mathrm{LM}(g_1),\ldots,\mathrm{LM}(g_s)\rangle$. Now use the ideal membership criteria for monomial ideals given in Exercise 13(a).

(e) Continuing the previous part, conclude that every class in $\mathbb{K}[x_1,\ldots,x_n]/I$ can be represented uniquely in terms of the classes of standard monomials. (Note you must check that the remainder r is independent of any choices you might have made.)

(f) Deduce that $\mathbb{K}[x_1,\ldots,x_n]/I$ is spanned as a $\mathbb{K}$-vector space by the classes of standard monomials.

(g) Prove that the classes of the standard monomials in $\mathbb{K}[x_1,\ldots,x_n]/I$ are linearly independent over $\mathbb{K}$. Hint: Suppose that $\sum_i c_i m_i \in I$ where $c_i \in \mathbb{K}^\times$ and m_i is standard for all i. Apply the leading monomial argument used above to get a contradiction.

(h) Conclude that the classes of standard monomials form a $\mathbb{K}$-basis for $\mathbb{K}[x_1,\ldots,x_n]/I$.

Notice that the staircase is not merely a helpful picture. Its lattice points index an honest $\mathbb{K}$-basis of the quotient, and division writes every residue class uniquely in that basis. When the staircase is finite, counting its points computes the vector-space dimension of the quotient and hence the lengths we saw in Lecture 1. When it is infinite, the pattern of the surviving monomials can still reveal useful algebraic structure. However, since the underlying vector spaces are infinite-dimensional, we must often move from thinking about vector spaces to thinking about modules.

The unique representative of a class whose monomials are all standard is called its *normal form*. For a fixed monomial order and Gröbner basis G, we denote it by $\mathrm{NF}_G(f)$. Thus we have proved

$$f \in I \qquad \Longleftrightarrow \qquad \mathrm{NF}_G(f) = 0.$$

Although the notation displays G, the normal form depends only on I and the chosen monomial order: every Gröbner basis of I produces the same standard representative.

Accepting for the moment the Gröbner basis computed by *Macaulay2*, the circle–parabola example predicts very concretely that every residue class has a unique representative of the form $a + bx + cy + dxy$. The four green points beneath the staircase therefore do more than suggest the correct dimension: they index another basis of the four-dimensional quotient studied in Lecture 1, and a, b, c, d are the coordinates of a class in this basis.

There is one piece of unfinished business. For the circle–parabola and twisted cubic we have proposed particular initial ideals, guided by hand calculations and by *Macaulay2*. We must be careful not to argue in a circle: we cannot assume our proposed standard monomials form a quotient basis in order to prove that we have found the correct initial ideal. Instead, the next exercise uses quotient descriptions proved independently in Lecture 1 to show directly that every possible remainder of an element of I is zero.

---

(20) **Certifying the Candidate Staircases.** We now replace our computational evidence by proofs that work over an arbitrary field. The strategy is simple: divide an arbitrary $f \in I$ by the candidate basis, and then use a quotient description from Lecture 1 to prove that the resulting remainder must be zero.

(a) Return to the circle–parabola ideal $I = \langle x^2 - y, x^2 + y^2 - 1\rangle \subseteq \mathbb{K}[x, y]$, equipped with lex order $x > y$. Set $G = (x^2 - y, y^2 + y - 1)$. Divide an arbitrary $f \in I$ by G. Explain why its remainder has the form

$$r = a + bx + cy + dxy.$$

(b) Use the division identity and the fact that each element of G lies in I to show that $r \in I$. Under the isomorphism induced by $y \mapsto x^2$,

$$\mathbb{K}[x, y]/I \cong \mathbb{K}[x]/\langle x^4 + x^2 - 1\rangle,$$

which polynomial of degree at most 3 represents the class of r?

(c) Since $r \in I$, this polynomial represents zero in the quotient and is therefore a multiple of the monic polynomial $x^4 + x^2 - 1$. Why must it therefore be zero? Deduce that $r = 0$.

(d) Apply Exercise 18(f) to conclude, over every field, that

$$\mathrm{in}_<(I) = \langle x^2, y^2\rangle.$$

Why does the original list $(x^2 - y, x^2 + y^2 - 1)$ nevertheless fail to be a Gröbner basis?

(e) *Further: The Twisted Cubic in Lex Order.* Return to the affine twisted cubic $I = \langle x^2 - y, xy - z\rangle \subseteq \mathbb{K}[x, y, z]$ with lex order $x > y > z$. Divide an arbitrary $f \in I$ by $G = (x^2 - y, xy - z, xz - y^2, y^3 - z^2)$. Its remainder is a linear combination of the standard monomials listed in Exercise 16(c). As in part (b), the division identity shows that this remainder lies in I. Apply the substitution $x \mapsto t$, $y \mapsto t^2$, $z \mapsto t^3$. Since the standard monomials map to distinct powers of t, prove that the remainder is zero. Deduce

$$\mathrm{in}_<(I) = \langle x^2, xy, xz, y^3\rangle$$

over every field.

(f) *Further: Graded Reverse Lexicographic Order.* Repeat part (e) with graded reverse lexicographic order and the three minors and standard monomials from Exercise 16(f)–(g). Use the monomials z^c, xz^c, yz^c to prove

$$\mathrm{in}_<(I) = \langle x^2, xy, y^2\rangle.$$

Thus lex requires a fourth basis element but exposes an elimination equation, while graded reverse lexicographic order gives the shorter determinantal basis.

The important point is that this proof did not rely on the computer or use Buchberger's criterion, which we have not yet proved. After division, the circle–parabola remainder becomes a polynomial

of degree at most 3 in x. It is simply too small to be a nonzero multiple of $x^4 + x^2 - 1$. If you completed the *Further* parts, the same idea certifies the twisted cubic bases: the candidate standard monomials map to distinct powers of t, so no nonzero linear combination of them can vanish in $\mathbb{K}[t]$.

The two orders for the twisted cubic also give our first clear example of the tradeoff involved in choosing an order. Lex requires one more basis element, but rewards us with the elimination equation $y^3 - z^2$. Graded reverse lexicographic order gives a shorter basis, but does not expose that elimination equation. The quotient remains $\mathbb{K}[t]$ in either order. The computation changes the coordinates in which we see it, not the underlying answer.

The circle–parabola staircase was finite because its quotient was finite-dimensional. The twisted cubic gave an infinite staircase, but its parametrization made the pattern unusually transparent. Let us finish with an equation-defined curve where an infinite staircase has an equally concrete geometric meaning.

---

(21) **An Infinite Staircase: A Cubic Curve.** Assume that $\mathbb{K}$ is algebraically closed of characteristic different from 2, work in $\mathbb{K}[x, y]$ with lex order $y > x$, and set $f = y^2 - x^3 + x$, $I = \langle f \rangle$, and $E = \mathbb{V}(f)$. This is the smooth affine cubic previewed at the end of Lecture 1, which will eventually give an elliptic curve. We are not yet asking whether it can be parametrized. For now we ask a more modest question: what does division tell us about its coordinate ring and about the projection onto the x-axis?

(a) Use Exercise 9(a) to prove, more generally, that the singleton list (f) is a Gröbner basis of the principal ideal generated by every nonzero polynomial f. Deduce that

$$\mathrm{in}_<(I) = \langle y^2 \rangle.$$

Thus division by one polynomial behaves just as well in several variables as it did in one variable. The hidden-leading-term problem only appears when several divisors can cancel against one another.

(b) Draw the staircase of $\langle y^2 \rangle$ and list its standard monomials. Why is the infinite staircase what one should expect for a curve rather than a finite intersection?

(c) Use the normal-form theorem to deduce that every class in $\mathbb{K}[x, y]/I$ has a unique representative $a(x) + b(x)y$, and conclude that this quotient is a free $\mathbb{K}[x]$-module of rank 2 with basis $1, y$.

(d) *Further: What the Staircase Forgets.* Repeat parts (a)–(c) for $y^2 - x^2(x + 1)$, whose curve has a node at the origin. Its initial ideal, staircase, and normal forms are identical. Which geometric difference between the two curves is therefore invisible to the initial ideal?

The description of the quotient $\mathbb{K}[x,y]/I$ found above has a direct geometric meaning. Fixing $x = a$ asks which points of E lie above a; we call this the *fiber* of the projection over a. Let us compare the visible points in a fiber with the algebra that remembers their multiplicities.

---

(22) **Fibers and Multiplicity.** Assume that $\mathbb{K}$ remains algebraically closed of characteristic different from 2. Continue with the cubic curve E and the projection $E \to \mathbb{K}$, $(x,y) \mapsto x$.

(a) Show that the fiber over $a \in \mathbb{K}$ is obtained by solving $y^2 = a^3 - a$. Deduce that it consists of two points when $a \notin \{-1, 0, 1\}$ and one point when $a \in \{-1, 0, 1\}$.

(b) For $a \in \{-1, 0, 1\}$, impose the additional equation $x = a$ and compute the fiber ring

$$\mathbb{K}[x,y]/\langle y^2 - x^3 + x,\ x - a\rangle.$$

Show that it is isomorphic to $\mathbb{K}[y]/\langle y^2\rangle$.

(c) Each exceptional fiber has one visible point but length 2. Explain how the nilpotent class of y records the two points coming together and why this agrees with the rank-2 module in Exercise 21(c).

The infinite staircase has now told us something geometric. The standard monomials x^a and $x^a y$ do much more than say that the quotient is infinite-dimensional: they organize it into two infinite rows. The algebraic way of saying this is that the quotient is a free $\mathbb{K}[x]$-module of rank 2. Correspondingly, the x-projection has fibers of length 2. Usually this length is realized by two distinct points; over each of -1, 0, and 1 it is realized by one double point. The staircase retains exactly the multiplicity needed to make this count constant.

If you completed the *Further* part, the nodal cubic gives an equally important warning. It has the same initial ideal and the same staircase, even though one curve is smooth and the other is singular. A staircase can remember a basis, a module rank, and fiber lengths while forgetting other geometric features. Gröbner methods replace complicated equations by simpler monomial data, so we must always ask which information survived the replacement. With this warning in mind, let us update the affine dictionary and state exactly what we have gained.

---

(23) **Updating the Affine Dictionary.** Consult the table near the beginning of Lecture 1. We can now replace several existence questions by actual finite computations.

(a) For a Gröbner basis G of I, record the test

$$f \in I \quad \Longleftrightarrow \quad \mathrm{NF}_G(f) = 0$$

and the *Macaulay2* commands `gb` and `%`. Why does this decide membership in the specified ideal but not, by itself, decide whether f vanishes on every point of $\mathbb{V}(I)$? When do those questions agree?

(b) Explain why $\mathrm{NF}_G(f)$ is a canonical representative of a class in S/I. Under what hypothesis may it also be interpreted as a canonical representative of a polynomial function on $\mathbb{V}(I)$?

(c) The monomial order changes the standard-monomial basis and hence the chosen representative, but not the truth of $f \in I$. Explain. Why should you now expect the black-box computations from Lecture 1—including `eliminate`, `ker`, and ideal equality—to be implemented using Gröbner bases?

We can finally answer the ideal-membership question with which we began the lecture. Once a Gröbner basis has been found and certified, division terminates with either a membership certificate or a nonzero normal form certifying nonmembership. At the same time, normal forms give concrete coordinates on the quotient S/I.

Before we finish, let us be precise about what this says geometrically. A normal form represents a class in S/I, and therefore restricts to the same polynomial function on $\mathbb{V}(I)$ as the original polynomial. But distinct classes in S/I determine distinct polynomial functions precisely when $I = \mathbb{I}(\mathbb{V}(I))$. Over an algebraically closed field, the Nullstellensatz says that this is equivalent to I being radical. A monomial order can organize an ideal for computation, but no choice of order erases the distinction between an ideal and the vanishing ideal of its visible points.

There remains one central question. The definition of a Gröbner basis refers to the leading terms of every element of an ideal, an infinite collection we cannot inspect one polynomial at a time. How could we ever find such a basis? Our examples suggest an answer: every missing term arose in exactly the same way. We multiplied two equations until their leading terms matched and then subtracted to make those terms cancel. The final exercise isolates this move and turns it into the input for Lecture 3.

---

(24) **The Shape of a Cancellation.** For the final exercise, bring together our two running examples: the circle–parabola ideal and the affine twisted cubic. In $\mathbb{K}[x, y]$, use lex order $x > y$ and set $f_1 = x^2 - y$ and $f_2 = x^2 + y^2 - 1$, the equations defining the parabola and circle, respectively. In $\mathbb{K}[x, y, z]$, use lex order $x > y > z$ and set $g_1 = x^2 - y$, $g_2 = xy - z$, and $g_3 = xz - y^2$. We have previously encountered the three identities

$$f_2 - f_1 = y^2 + y - 1, \qquad yg_1 - xg_2 = xz - y^2, \qquad zg_2 - yg_3 = y^3 - z^2.$$

In each identity, the leading terms of the two polynomial multiples on the left cancel. Our goal is to understand the common pattern behind these cancellations.

(a) Consider, in order, the three pairs (f_2, f_1), (g_1, g_2), and (g_2, g_3). For each pair (u, v), compute $\mathrm{LM}(u)$ and $\mathrm{LM}(v)$, and then compute the least common multiple $L = \mathrm{lcm}(\mathrm{LM}(u), \mathrm{LM}(v))$.

(b) For each pair (u, v) in part (a), check that the monomials multiplying u and v in the corresponding identity above are given by

$$\frac{L}{\mathrm{LM}(u)} \qquad \text{and} \qquad \frac{L}{\mathrm{LM}(v)}.$$

(c) For any two nonzero polynomials u and v, define

$$S(u, v) \coloneqq \frac{\mathrm{lcm}\,(\mathrm{LM}(u), \mathrm{LM}(v))}{\mathrm{LT}(u)} u - \frac{\mathrm{lcm}\,(\mathrm{LM}(u), \mathrm{LM}(v))}{\mathrm{LT}(v)} v.$$

Explain why $S(u, v)$ is itself a polynomial with coefficients in $\mathbb{K}$. Compute the leading terms of the two summands in the definition of $S(u, v)$, and show that these leading terms are equal and therefore cancel.

(d) Compute $S(f_2, f_1)$, $S(g_1, g_2)$, and $S(g_2, g_3)$. Show how this computation recovers the three identities above.

(e) *Further: Uniqueness of the Basic Cancellation.* Let u and v be nonzero polynomials and suppose that h_1 and h_2 are nonzero terms such that the leading terms cancel in $h_1 u - h_2 v$, i.e., suppose $\mathrm{LT}(h_1 u) = \mathrm{LT}(h_2 v)$. Prove that their common leading monomial must be divisible by $L = \mathrm{lcm}(\mathrm{LM}(u), \mathrm{LM}(v))$. Deduce that there is a scalar $c \in \mathbb{K}^\times$ and a monomial m such that

$$h_1 u - h_2 v = c\, m\, S(u, v).$$

Thus every cancellation between two term multiples of u and v is obtained from $S(u, v)$ by multiplying by a nonzero scalar and a monomial. Hint: Write the common leading term as cP. Since P is divisible by both $\mathrm{LM}(u)$ and $\mathrm{LM}(v)$, write $P = mL$. Now use $\mathrm{LT}(h_1 u) = \mathrm{LT}(h_2 v) = cmL$ to solve for h_1 and h_2.

The polynomials defined in part (c) are called *S-polynomials*. We now know what the next algorithm must do: form these basic cancellations between pairs, divide the results by the current list, and adjoin any nonzero remainders. Lecture 3 will make this precise, prove that these pairwise cancellations detect every missing leading term, and show that repeating the process eventually stops.