MATH 371 – Class notes/outline – September 12, 2013

Polynomials

Another source of ring theory is the study of polynomials, particularly polynomials “over a field”. For now, we’ll restrict our attention to the fields $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ of rational, real and complex numbers respectively. Later we’ll deal with other fields, and sometimes with polynomials over a ring (like $\mathbb{Z}$).

A *monomial* in the variables $x_1, \ldots, x_n$ is a product $x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ where all the exponents $\alpha_1, \ldots, \alpha_n$ are non-negative integers. The *total degree* of the monomial is $\alpha_1 + \cdots + \alpha_n$.

A *multi-index* $\alpha = (\alpha_1, \ldots, \alpha_n)$ represents all the exponents in a monomial, and we’ll often write x^α for $x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdots x_n^{\alpha_n}$. We’ll also write $|\alpha| = \alpha_1 + \cdots + \alpha_n$ for the total degree.

A *polynomial* f in $x_1, \ldots, x_n$ with coefficients in the field k is a finite linear combination of monomials. We’ll write

$$f = \sum_\alpha a_\alpha x^\alpha$$

where the $a_\alpha \in k$. The set of all polynomials in $x_1, \ldots, x_n$ with coefficients in k is denoted $k[x_1, \ldots, x_n]$. If there are only two or three variables we might write things like $\mathbb{Q}[x, y]$ or $\mathbb{R}[x, y, z]$.

Just so we’re using words the same way: a_α is called the *coefficient of the monomial* x^α. The product $a_\alpha x^\alpha$ is called a *term* of f. And the *total degree* of f is the maximum total degree of any of its terms.

Since the sum and product of polynomials is a polynomial and since addition and multiplication of polynomials satisfies all the standard properties (commutative, associative, distributive), we have that $k[x_1, \ldots, x_n]$ is a *commutative ring.*

Definition: For a field k and a positive integer n, define *n-dimensional affine space over k* to be the set

$$k^n = \{(a_1, a_2, \ldots, a_n) \mid a_1, \ldots, a_n \in k\}.$$

By analogy with $\mathbb{R}$, $\mathbb{R}^2$ etc, call $k^1 = k$ the affine line and k^2 the affine plane.

Connect polynomials to affine space by evaluation. If $f \in k[x_1, \ldots, x_n]$, then $f \ldots k^n \to k$ by substituting a_1 for x_1, a_2 for x_2, etc in each term of f and calculating the resulting value, which will clearly be an element of k.

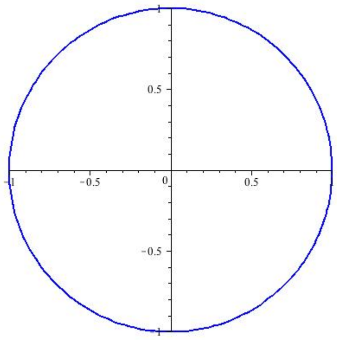
Over an *infinite* field (like $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$, a polynomial f represents the zero function if and only if it is the zero polynomial. But over finite fields ($\mathbb{Z}/p$ for p prime is a field) it’s possible that f could be the zero function even though f is not the zero polynomial (think $x^p - x$ thanks to Fermat’s little theorem). So over an infinite field, f and g represent the same function iff they are the same polynomial. (Proof of this uses that polynomial of degree n in one variable has at most n roots, which we’ll prove later).

Affine varieties: “Varieties” are the basic objects in algebraic geometry. If k is a field and $f_1, \ldots f_s$ are polynomials in $k[x_1, \ldots, x_n]$. then

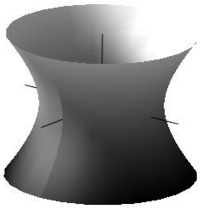
$$\mathbf{V}(f_1, \ldots, f_s) = \{(a_1, \ldots, a_n) \in k^n \mid f_i(a_1, \ldots, a_n) = 0,\ i = 1 \ldots, s\}$$

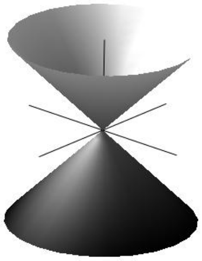
is the *affine variety* defined by $f_1, \ldots, f_s$ (the set of solutions of the simultaneous equations $f_1 = 0, \ldots, f_s = 0$).

So in $\mathbb{R}^2$, here is $\mathbf{V}(x^2 + y^2 - 1)$:

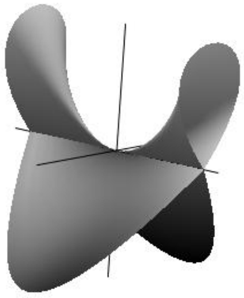


All conic sections are affine varieties, as are graphs of polynomials and graphs of rational functions (quotients of polynomials, since $y = p/q$ can be rewritten $qy - p = 0$). Examples in $\mathbb{R}^3$ include things like the hyperboloid $\mathbf{V}(x^2 + y^2 - z^2 - 1)$ and the cone $\mathbf{V}(x^2 + y^2 - z^2)$:



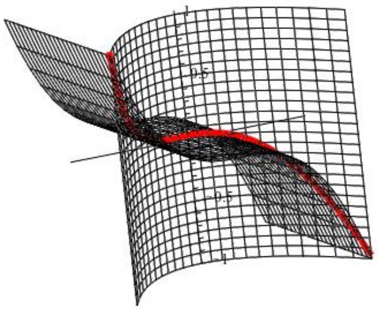


A more complicated example is provided by $\mathbf{V}(x^2 - y^2z^2 + z^3)$:

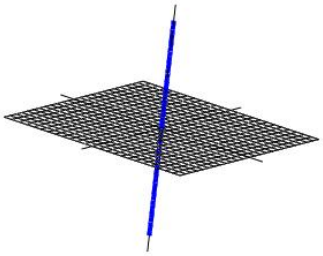


The latter two varieties show examples of "singular points" – the origin in the cone and the whole y axis where the last surface intersects itself.

Now consider a curve in $\mathbb{R}^3$ — we can obtain a curve as the intersection of two surfaces, say $\mathbf{V}(y - x^2, z - x^3)$. Here are the two surfaces $y - x^2 = 0$ and $z - x^3 = 0$ together with the curve:



It would seem that when we have one equation in $\mathbb{R}^2$ we get a curve (a 1-dimensional object); one equation in $\mathbb{R}^3$ gives a surface, which is 2-dimensional. And when we had two equations in $\mathbb{R}^3$ we got a curve, 1-dimensional. So it looks like the imposition of each equation reduces the dimension of the geometric object by one. While this is generally true, things are much more subtle. For instance, consider $\mathbf{V}(xz, yz)$:



This variety consists of two pieces, a plane (which has the "wrong" dimension) wqtogether with a line. In thinking about dimension, you might also remember how tricky it can be in linear algebra to predict in advance the dimension of the solution space to a set of linear equations in many unknowns (this is another example of an affine variety, sometimes called a *linear variety* for obvious reasons).

Sometimes the dimension can depend on the field chosen, for instance $\mathbf{V}(x^2 + y^2 + 1)$ is the empty set over $\mathbb{R}$ or $\mathbb{Q}$, but defines a variety of one (complex) dimension over $\mathbb{C}$.

A basic property: If $V, W \subset k^n$ are affine varieties, so are $V \cup W$ and $V \cap W$. In fact, if $V = \mathbf{V}(f_1, \ldots, f_s)$ and $W = \mathbf{V}(g_1, \ldots, g_t)$ then

$$\begin{aligned} V \cap W &= \mathbf{V}(f_1, \ldots, f_s, g_1, \ldots, g_t) \\ V \cup W &= \mathbf{V}(f_i g_k \,|\, 1 \leqslant i \leqslant s,\ 1 \leqslant j \leqslant t). \end{aligned}$$

Rational functions and parametrizations: A *rational function* is a quotient of polynomials f/g, where $f, g \in k[x_1, \ldots, x_n]$ and g is not the zero polynomial. The set of rational functions is denoted $k(x_1, \ldots, x_n)$.

A *parametrization* of an algebraic variety $V \in k^n$ is a set of equations

$$\begin{aligned} x_1 &= \varphi_1(t_1, \ldots, t_m) \\ x_2 &= \varphi_2(t_1, \ldots, t_m) \\ \vdots \ &\ \vdots \quad \vdots \\ x_n &= \varphi_n(t_1, \ldots, t_m) \end{aligned}$$

where for each $(t_1, \ldots, t_m) \in k^m$, for which $\varphi_1, \ldots, \varphi_n$ are defined, the point $(x_1, \ldots, x_n)$ lies in V. It is usually also required that V be the "smallest" affine variety containing these points. If all the functions $\varphi_1, \ldots, \varphi_n$ are of the same type (polynomial, rational, etc.) the parametrization gets that label (a rational parametrization, etc.).

For example

$$x = \frac{1-t^2}{1+t^2}\,, \quad y = \frac{2t}{1+t^2}$$

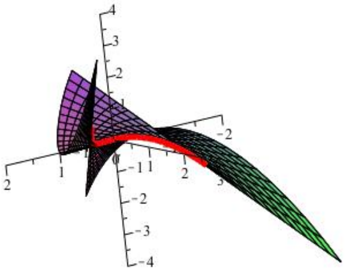
is a rational parametrization of the circle $\mathbf{V}(x^2+y^2-1)$ (it hits all the points except $(-1,0)$). You can derive this parametrization by considering lines through the point $(-1,0)$. Except for the vertical line, each line through $(-1,0)$ will intersect the circle at another point (x,y), and will also intersect the y-axis at the point $(0,t)$.

As another example, the equations

$$x = t(u^2-t^2) \quad y = u \quad z = u^2 - t^2$$

parametrize the surface $\mathbf{V}(x^2 - y^2z^2 + z^3)$ given above.

For yet another example, it is easy to see that the parametric equations $x = t$, $y = t^2$, $z = t^3$ parametrize the curve $\mathbf{V}(y-x^2, z-x^3)$ discussed above (it is usually called the *twisted cubic.* Let's now consider the surface consisting of all the tangent lines to the twisted cubic. For a specific value of t, the tangent line at the point (t,t^2,t^3) has direction vector $(1,2t,3t^2)$ and so it can be parametrized as $x = t+u$, $y = t^2+2tu$, $z = t^3+3t^2u$. And letting t vary in these equations as well, we get a parametrization of the tangent surface to the twisted cubic. Here is a picture of this surface.



To show that the tangent surface of the twisted cubic is actually an affine variety, we note that it is $\mathbf{V}(4x^3z - 3x^2y^2 + 4y^3 - 6xyz + z^2)$.

These examples naturally lead so two questions: (1) Does every affine variety have a rational parametrization? (Unfortunately, the answer to this question is "No".) (2) Given a rational parametrization of an affine variety, can we recover the defining (implicit) equations? The answer to this question is, perhaps surprisingly, "Yes", and there are algorithms for this. We need to develop a little bit of machinery to understand them.

Ideals

The concept of an ideal in a ring is fundamental. An *ideal* in the polynomial ring $k[x_1,\ldots,x_n]$ is a subset $I \subset k[x_1,\ldots,x_n]$ that satisfies:

1. the zero polynomial is in I.
2. If f and g are in I then $f + g \in I$.
3. If $f \in I$ and h is *any* polynomial in $k[x_1, \ldots, x_n]$ then $hf \in I$.

The ideal generated by a finite set: Let $f_1, \ldots, f_s$ be polynomials in $k[x_1, \ldots, x_n]$. Then

$$\langle f_1, \ldots . f_s \rangle = \left\{ \sum_{i=1}^{s} h_i f_i \mid h_1, \ldots, h_s \in k[x_1, \ldots, x_n] \right\}$$

is an ideal, called *the ideal generated by* $f_1, \ldots, f_s$.

The ideal $\langle f_1, \ldots, f_s \rangle$ gives the left-hand sides of all the "polynomial consequences" of the equations $f_1 = 0$, $f_2 = 0, \ldots, f_s = 0$.

If V is an affine variety, then the set of polynomials $f \in k[x_1, \ldots, x_n]$ such that $f(a_1, \ldots, a_n) = 0$ for all $(a_1, \ldots, a_n) \in V$ is an ideal. Call this ideal $\mathbf{I}(V)$ (the *ideal of* V). It's an easy proposition that if the ideals $\langle f_1, \ldots, f_s \rangle$ and $\langle g_1, \ldots, g_r \rangle$ are the same (two different bases for the same ideal) then $\mathbf{V}(f_1, \ldots, f_s) = \mathbf{V}(g_1, \ldots, g_r)$.

Example: The ideal of the twisted cubic: Recall that the twisted cubic is $V = \mathbf{V}(y - x^2, z - x^3)$. We'll show that $\mathbf{I}(V) = \langle y - x^2, z - x^3 \rangle$. It's easy to see that $\langle y - x^2, z - x^3 \rangle \subset \mathbf{I}(V)$. To show the reverse inclusion, we first show that any polynomial $f \in \mathbb{R}[x, y, z]$ can be expressed as

$$f = q_1(x, y, z)(y - x^2) + q_2(x, y, z)(z - x^3) + r(x)$$

where $q_1, q_2, r \in \mathbb{R}[x, y, z]$ and r is a function of x alone. First consider the case where $f = x^a y^b z^c$ is a monomial. We can then write:

$$\begin{aligned} x^a y^b z^c &= x^a (x^2 + (y - x^2))^b (x^3 + (z - x^3))^c \\ &= x^a (x^{2b} + \text{terms divisible by } (y - x^2))(x^{3c} + \text{terms divisible by } (z - x^3)). \end{aligned}$$

Multiplying this out gives

$$x^a y^b z^c = q_1(x, y, z)(y - x^2) + q_2(x, y, z)(z - x^3) + x^{a+2b+3c}$$

so the claim is true for monomials. But then it's true for all polynomials, which after all are sums of monomials.

Now we can prove that $\mathbf{I}(V) \subset \langle y - x^2, z - x^3 \rangle$. Suppose $f \in \mathbf{I}(V)$. Using the parametrization $x = t$, $y = t^2$, $z = t^3$ of V we see that we must have $f(t, t^2, t^3) = 0$. But then

$$0 = f(t, t^2, t^3) = h_1(t, t^2, t^3)(t^2 - t^2) + h_2(t, t^2, t^3)(t^3 - t^3) + r(t) = 0 + 0 + r(t).$$

Therefore $r(x) = 0$ and so f must be in $\langle y - x^2, z - x^3 \rangle$.

As a corollary, we get that $f(x, y, z) \in \langle y - x^2, z - x^3 \rangle$ if and only if $f(t, t^2, t^3) = 0$.

Observe what happened here. We started with some polynomials $f_1, \ldots, f_s \in k[x_1, \ldots, x_n]$, which determine the affine variety $\mathbf{V}(f_1, \ldots, f_s)$, which in turn determines the ideal $\mathbf{I}(\mathbf{V}(f_1, \ldots, f_s))$. In the example of the twisted cubic, we had $\mathbf{I}(\mathbf{V}(f_1,\dots,f_s)) = \langle f_1,\dots,f_s\rangle$, but in general we only have

$$\langle f_1,\dots,f_s\rangle \subset \mathbf{I}(\mathbf{V}(f_1,\dots,f_s)).$$

For instance $\langle x^2, y^2\rangle \subset \mathbb{R}[x,y]$ is a proper subset of $\langle x,y\rangle = \mathbf{IV}(x^2,y^2)$. More about this later (the precise relation is given by a theorem called Hilbert's Nullstellensatz).

For now, we note the inverse relationship: For affine varieties $V, W \subset k[x_1,\dots,x_n]$, we have $V \subset W$ if and only if $\mathbf{I}(W) \subset \mathbf{I}(V)$, and $V = W$ if and only if $\mathbf{I}(V) = \mathbf{I}(W)$.

For now, we leave this situation with three questions:

1. (*Ideal Description*) Can every ideal $I \subset k[x_1,\dots x_n]$ be written as $\langle f_1,\dots,f_s\rangle$ for some $f_1,\dots,f_s \in k[x_1,\dots,x_n]$? In particular, is every ideal $I \subset k[x_1,\dots,x_n]$ *finitely generated*?
2. (*Ideal Membership*) Given $f_1,\dots,f_s \in k[x_1,\dots,x_n]$, is there a way to decide whether a given $f \in k[x_1,\dots,x_n]$ is in the ideal $\langle f_1,\dots,f_s\rangle$?
3. (*Nullstellensatz*) Given $\langle f_1,\dots,f_s\rangle \in k[x_1,\dots,x_n]$, what is the precise relation between $\langle f_1,\dots,f_s\rangle$ and $\mathbf{I}(\mathbf{V}(f_1,\dots,f_s))$?

Polynomials in one variable over a field

The case of $k[x]$, polynomials in a single variable over a field, is special but particularly instructive. We can deal with all three of the questions above and get reasonably satisfying answers. The main reason for this is that there is a division with remainder algorithm for polynomials in $k[x]$.

Some jargon first: We'll just say the *degree* of a polynomial in one variable instead of total degree, and we'll write $\deg(f)$ for the degree of the polynomial f. There can be only one term of degree $\deg(f)$, and if $m = \deg(f)$ the polynomial looks like

$$f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0 \quad \text{with } a_m \neq 0.$$

The term $a_m x^m$ is the *leading term* of f, and we'll write $\mathrm{LT}(f)$, and a_m is the *leading coefficient* of f.

Because k is a field, we have that $\mathrm{LT}(f) \mid \mathrm{LT}(g)$ if and only if $\deg(g) \geqslant \deg(f)$.

The division algorithm: Suppose $f, g \in k[x]$, and g is not the zero polynomial. Then there are unique polynomials $q, r \in k[x]$ with $r = 0$ or $deg(r) < deg(g)$ such that

$$f = qg + r.$$

Moreover, there is an algorithm for finding q and r.

Uniqueness first: If $f = q'g + r'$ as well, then we have $0 = (q'-q)g + (r'-r)$. If $q'-q \neq 0$ then $r'-r \neq 0$ in which case $g \mid r'-r$. But $\deg(r) < \deg(g)$ so this is impossible unless $r'-r = 0$, a contradiction.

Existence of q and r is provided by the algorithm itself — it is the familiar long division algorithm from high school algebra: Given f and g, we set $q = 0$ and $r = f$. If $\deg(f) \geqslant \deg(r)$ we then let $z = \mathrm{LT}(f)/\mathrm{LT}(g)$, replace q by $q + z$ and replace r by $r - zg$. Note that the degree of r must have decreased in this process. We repeat this process until either $r = 0$ or $\deg(r) < \deg(g)$. The resulting q and r are the ones guaranteed by the statement above.

We know that the algorithm will terminate (and work!) because the degree of r keeps decreasing, and it can't decrease below zero.

The first corollary of the division algorithm is the long-awaited result about the number of roots of a polynomial in one variable:

Corollary: If k is a field and f is a non-zero polynomial in $k[x]$ then f has at most $\deg(f)$ roots in k.

Proof is by induction on $\deg(f)$. If a satisfies $f(a) = 0$, then write $f(x) = (x - a)q(x) + r$ where $\deg(r) < \deg(x - a)$, so r is a constant. We must have $r = 0$ as can be seen by substituting $x = a$ into this last equation. So now $f = (x - a)q$ where $\deg(q) = \deg(f) - 1$. But by induction the result is true for polynomials of degree $\deg(f) - 1$, so we're done.

Next, we can learn something about the structure of ideals in $k[x]$.

Corollary: If k is a field then every ideal of $k[x]$ can be written in the form $\langle f\rangle$ for some $f \in k[x]$. Moreover, f is unique up to multiplication by a non-zero constant.

Proof. Consider the ideal $I \subset k[x]$. If I is the zero ideal, we're done since $I = \langle 0\rangle$. Otherwise, let f be the non-zero polynomial of minimal degree in I, and claim that $\langle f\rangle = I$. Certainly $\langle f\rangle \subset I$. Now suppose $g \in I$ and write $g = qf + r$ where either $r = 0$ or $\deg(r) < \deg(f)$. Then $r = g - qf \in I$ so we must have $r = 0$ to avoid a contradiction. Therefore $g = qf \in \langle f\rangle$, so $I \subset \langle f\rangle$ and we are done. (Uniqueness is easy, since if $\langle f\rangle = \langle g\rangle$ then $f = pg$ and $g = qf$ so $\deg(g) = \deg(f)$ and p and q are constants.

Definition: An ideal (in any ring, more later) generated by a single element is called a *principal ideal*, and a ring like $k[x]$ where every ideal is principal is called a *principal ideal domain.* (We actually need a bit more to justify the word "domain" but we'll come back to this later). The ring of integers $\mathbb{Z}$ is also a principal ideal domain.

How do you find the generator of an ideal in $k[x]$ if the ideal isn't presented in this form? For instance, what is the generator of the ideal $\langle x^4 - 1, x^6 - 1\rangle$? The answer is the same as in the integers:

Definition: A *greatest common divisor* of polynomials $f, g \in k[x]$ is a polynomial h such that $h \mid f$ and $h \mid g$ and such that if p is another polynomial that divides both f and g, then $g \mid h$. We write $\gcd(f, g)$ for such a polynomial.

Properties of the gcd: (1) $\gcd(f, g)$ exists and is unique up to multiplication by a non-zero constant, (2) $\gcd(f, g)$ generates the ideal $\langle f, g\rangle$, and (3) There is an algorithm for calculating $\gcd(f, g)$.

Proof: (1) and (2): The ideal $\langle f, g\rangle$ is principal, say $\langle f, g\rangle = \langle h\rangle$, and claim that $h = \gcd(f, g)$. Certainly h divides f and g since they are in the ideal generated by h, and conversely $h = \lambda f + \mu g$ for some polynomials $\lambda, \mu \in k[x]$, since h is in the ideal generated by f and g (sound familiar?). So if p is another polynomial that divides both f and g then $f = ap$ and $g = bp$ so $h = \lambda ap + \mu bp = (\lambda a + \mu b)h$ so $p \mid h$ as well, so h is a gcd for f and g. By the above corollary, h is unique up to multiplication by a constant.

Not surprisingly perhaps, the algorithm for computing $\gcd(f, g)$ is the Euclidean algorithm. First note that $\gcd(f, g) = \gcd(f - qg, g)$ for any polynomial q. Then the Euclidean algorithm for polynomials works pretty much the same way as it does for integers. There's even an extended Euclidean algorithm so you can find the λ and μ so that $\lambda f + \mu g = \gcd(f, g)$.

Just for practice, we'll run the extended Euclidean algorithm on the pair $f = x^4 - 1$, $g = x^6 - 1$:

i	-1	0	1	2
r_i	$x^6 - 1$	$x^4 - 1$	$x^2 - 1$	0
q_i	$-$	$-$	x^2	$x^2 + 1$
λ_i	1	0	1	$-$
μ_i	0	1	$-x^2$	$-$

From this we conclude that $\gcd(x^4 - 1, x^6 - 1) = x^2 - 1$ and that $1(x^6 - 1) - x^2(x^4 - 1) = x^2 - 1$.

What we have done so far gives us a complete answer to questions 1 and 2 above. Since $k[x]$ is a principal ideal domain, the answer to the ideal description problem is "yes" — and in fact every ideal $I \subset k[x]$ can be written as $\langle f \rangle$ for some $f \in k[x]$. To solve the ideal membership problem, given $f_1, \ldots f_s \in k[x]$, we need only calculate their gcd (by successively calculating $\gcd(f_1, \gcd(f_2, \gcd(f_3, \ldots, f_s) \cdots)))$ and then to decide whether a given f is in $\langle f_1, \ldots, f_s \rangle$, we just have to see whether $\gcd(f_1, \ldots, f_s) \mid f$. The Nullstellensatz question is a little more subtle and we'll take that up later.