Math 75 notes, Lecture 3

P. Pollack and C. Pomerance

To review, here are some things we've learned:

1. If F is a field, then $F[x]$ denotes the set of polynomials in the variable x with coefficients in F. We can add and multiply polynomials via the usual rules you learned in high school.

2. For $f \in F[x]$ with f not the zero-polynomial, the degree of f, denoted $\deg(f)$, is the exponent on the highest power of x that appears with a nonzero coefficient. Further, as we've seen from homework, if neither of f, g are zero, then $\deg(fg) = \deg(f) + \deg(g)$.

3. For a nonzero element $M \in F[x]$ we have the notion of congruence modulo f, which is completely analogous to congruence modulo an integer m in $\mathbb{Z}$. Namely, $f \equiv g \pmod{M}$ if and only if $f - g$ is a multiple of M. Congruence modulo M is an *equivalence relation* on elements of $F[x]$.

4. Just as we create a new system $\mathbb{Z}/(m)$, we can create a new system $F[x]/(M)$. Elements are the equivalence classes of the mod M equivalence relation. The equivalence class of f is denoted $[f]$, or if we are feeling brazen and it is understood we are dealing with $F[x]/(M)$, just f. We have

$$[f] = \{g \in F[x] : g \equiv f \pmod{M}\}.$$

   As with any equivalence relation, *any* element of the equivalence class $[f]$ can stand in as a representative of the class (equivalence classes are very egalitarian!). So if $g \equiv f \pmod{M}$, then $[g] = [f]$ (and conversely).

5. Just as with $\mathbb{Z}/(m)$ we can add and multiply equivalence classes by adding and multiplying their representatives, so

$$[f] + [g] = [f + g], \quad [f][g] = [fg].$$

   There is actually something to prove here, namely that these definitions of plus and times are good definitions, and do not depend on the choice of the representative. But would you expect anything less from our egalitarian equivalence classes?

6. With our addition and multiplication on $F[x]/(M)$, we nearly have a field; the only rule that may be in doubt is multiplicative inverses for nonzero elements.

7. Just as in $\mathbb{Z}$ we have divide-with-remainder in $F[x]$. Namely, if $f, M \in F[x]$ with M nonzero, there are unique $q, r \in F[x]$ with

$$f = qM + r, \quad \deg(r) < \deg(M) \text{ or } r = 0.$$

   (See Theorem 3, day 2.)

8. Note that $[f] = [r]$ in $F[x]/(M)$ if r is the remainder after dividing M into f. Thus, though any element of $[f]$ can stand in as a representative of the class, there is one distinguished element (ok, we're leaving the egalitarioan utopia) which is either 0 or has degree smaller than $\deg(M)$. This distinguished representative is unique (that's what makes it distinguished). Thus the elements of $F[x]/(M)$ are in a natural one-to-one correspondence with the polynomials $r \in F[x]$, where $r = 0$ or $\deg(r) < \deg(M)$. (See Theorem 4, day 2.)

9. We conclude that if F is finite field with q elements, and if $M \in F[x]$ has degree d, then $F[x]/(M)$ has q^d elements. (See Theorem 4, day 2.)

As you can see, we have accomplished much already, and the analogy of $F[x]/(M)$ to $\mathbb{Z}/(m)$ has stood us in good stead. We would like to go further and decide when $F[x]/(M)$ is a field. As we have seen, $\mathbb{Z}/(m)$ is a field precisely when m is a prime number. It was easy to see that it is not a field when m is not a prime number, and we used divide with remainder to show that it is a field when m is a prime number. Well, we have divide with remainder in our new setting of $F[x]$, so maybe the same deal goes through. Yes, it does.

Theorem 1. *If F is a field and $M \in F[x]$ is not zero, then $F[x]/(M)$ is a field precisely when M is irreducible.*

And what, you might ask, does *irreducible* mean? We know you already have a good idea, but here's a precise definition. A polynomial $M \in F[x]$ is irreducible if it is not zero, not a unit, and not the product of two polynomials of smaller degree.

To prove the theorem, we have to consider the case when M is not irreducible and the case when M is irreducible. The first case is eaay. First, $F[x]/(M)$ is not defined when $M = 0$ (or if you want to be a stickler by saying $f \equiv g \pmod 0$ if and only if $f = g$, which is the correct viewpoint from abstract algebra, then $F[x]/(0)$ is really the same as $F[x]$, and we've seen that $F[x]$ is not a field). Further, if M is a unit, then $f \equiv g \pmod M$ for *every* $f, g \in F[x]$, so that $F[x]/(M)$ has just 1 element and is not a field. Finally say $M = uv$ where $\deg(u), \deg(v) < \deg(M)$. Then $uv = 0$ in the arithmetic of $F[x]/(M)$, but neither u nor v is 0. But, u cannot have an inverse, for if $wu = 1$, then

$$0 = w \cdot 0 = wuv = 1 \cdot v = v \neq 0.$$

Thus, $F[x]/(M)$ is not a field in this case.

So, it remains to show that M irreducible forces $F[x]/(M)$ to be a field. To see this we take a slight but necessary detour and copy over the idea of the extended Euclic algorithm to $F[x]$. Namely, if $f, g \in F[x]$ are not both 0, and h is the highest common factor of f, g (the same as the greatest common divisor, namely the common divisor of highest degree and with leading coefficient 1), then there are polynomials $u, v \in F[x]$ with $uf + vg = h$. The proof is exactly the same as in $\mathbb{Z}$, and there is no need to repeat it. In addition, the proof is *constructive*. Namely, it actually gives us a reasonable way to find choices for u, v.

Now suppose that $M \in F[x]$ is not zero and that $f \in F[x]$ is *relatively prime* to M, meaning that their highest common factor is 1. Then, there are $u, v \in F[x]$ with $uf + vM = 1$. Thus, in $F[x]/(M)$ we see that f has an inverse, namely u, that is, $uf = 1$. We'll try an example in a moment, but let's first finish the proof of the theorem.

So, we have M irreducible, and f is some nonzero element of $F[x]/(M)$. Recall, we can think of f as a nonzero polynomial of degree smaller than $\deg(M)$. Now what might the highest common factor h of f and M be? Since $h \mid M$, we have $M = hk$ for some polynomial k. Since $\deg(h) + \deg(k) = \deg(M)$, if $0 < \deg(h) < \deg(M)$, then $\deg(k)$ satisfies the same double inequality, and so M is the product of two polynomials of smaller degree. Aha! This cannot occur when M is irreducible. So it must be that $\deg(h) = 0$ or $\deg(h) = \deg(M)$. But $\deg(f) < \deg(M)$ and also $h \mid f$, so we cannot have $\deg(h) = \deg(M)$, which leaves only the possibility that $\deg(h) = 0$, that is, $h = 1$ (since we have decided that to stand as the highest common factor, the leading coefficient is to be 1).

Well, now we're laughing. We have seen that this is precisely the condition that implies that f has an inverse in $F[x]/(M)$. Since we started with an arbitrary nonzero element f, this shows that every nonzero element has an inverse, and that $F[x]/(M)$ is a field.

Some examples.

Let us take as an example $F = \mathbb{Z}/(7)$ and $M(x) = x^2 + 1$. We know that F is a field, since 7 is a prime number. We know that $F[x]/(M)$ has $7^2 = 49$ elements, but do we know that $F[x]/(M)$ is a field? Using Theorem 1 above, the question can be restated: do we know that M is irreducible?

Quadratics and cubics are usually easy to detect for irreducibility: they are irreducible if and only if they do not have a root. So, let us check to see if $x^2 + 1$ has a root in $F = \mathbb{Z}/(7)$. There are just 7 possibilities to try for x, and really just 4 calculations need to be tried, since r and $-r$ can be tried together. We have in $\mathbb{Z}/(7)$:

$$0^2 + 1 = 1, \quad (\pm 1)^2 + 1 = 2, \quad (\pm 2)^2 + 1 = 5, \quad (\pm 3)^2 + 1 = 3,$$

so we see that $x^2 + 1$ has no root, and so is irreducible. Thus, $F[x]/(M)$ is a finite field with 49 elements.

These 49 elements are succinctly written as $a + bx$, where $a, b \in \{0, 1, \ldots, 6\}$, and the arithmetic uses $x^2 = -1 = 6$. For example (you should check this)

$$(x + 1)(x + 2) = 3x + 1.$$

If we wanted to, we could write down the full 7×7 multiplication table for this field, and if you are feeling a bit unsure about things, it may be good for you to try this. From such a table you could read off an element's inverse. For example, the inverse of $x + 1$ is $\ldots$. Yes, that's one way to find it, but we have a better way.

To find the inverse of $x+1$, lets perform the division algorithm with $x+1$ and x^2+1. First, the quotient is $x - 1$ and the remainder is 2. That is,

$$x^2 + 1 = (x - 1)(x + 1) + 2.$$

If we divide 2 into $x+1$, the remainder will be 0 since we recognize 2 as a unit in $F[x]$. So this division can be avoided. Actually, we need the multiplicative inverse of 2 in F. This can be done by a further Euclid algorithm in $\mathbb{Z}$ with 2 and 7, or we could just note that $2 \cdot 4 = 1$. Multiplying the equation displayed above by 4, we have

$$4(x^2+1) = 4(x-1)(x+1) + 1,$$

and reading this in $F[x]/(M)$ we see that

$$-4(x-1)(x+1) = 1.$$

Thus, the inverse of $x+1$ is $-4x+4 = 3x+4$.

OK, lets try another example. Take $F = \mathbb{Z}/(2)$ and $M(x) = x^3 + x^2 + 1$. In this example there are just two possible roots to try, 0 and 1, and we see that neither works. Since M has degree 3, this means that M is irreducible and that $F[x]/(M)$ is a field with $2^3 = 8$ elements. Lets try for the inverse of x^2+1. The first step of the division algorithm gives quotient $x+1$ and the remainder is x; that is,

$$x^3 + x^2 + 1 = (x+1)(x^2+1) + x.$$

Next, we divide x into x^2+1 getting quotient x and remainder 1:

$$x^2 + 1 = x \cdot x + 1.$$

We're now ready to assemble the inverse of x^2+1. Perhaps it's easier to do this starting from the last equation, since then we can keep our goal of getting something equal to 1 in front of us:

$$1 \;=\; (x^2+1) + x \cdot x \;=\; (x^2+1) + \big((x^3+x^2+1) + (x+1)(x^2+1)\big)x.$$

(One reason that this may be confusing is that when $F = \mathbb{Z}/(2)$ there is never any call for writing minus signs!) Do you see where the big parentheses comes from? It is actually using the first divide-with-remainder to substitute for the factor x in the previous expression. So, keeping in mind that we are dealing with x^2+1 and x^3+x^2+1, we have

$$1 \;=\; x(x^3+x^2+1) + \big(1 + (x+1)x\big)(x^2+1).$$

Reading this as an equation in $F[x]/(M)$, and simplifying the big parentheses, we get

$$1 \;=\; (x^2+x+1)(x^2+1).$$

That is, the inverse of x^2+1 is x^2+x+1. Do you believe it? You know of course there is plenty of opportunity to make a careless error. So, lets check it. Doing the multiplication, we get x^4+x^3+x+1. Now $x^4 = x \cdot x^3 = x(x^2+1) = x^3+x$, and substituting this into x^4+x^3+x+1, we do indeed get 1. It checks!

The take-home message here is that it is perfectly possible to find multiplicative inverses if you carefully follow the rules. Yes, you will make mistakes, so it is a good idea to check your work.

Where to now?

We have seen how to construct new fields from old, the problem that was posed on day 2. Namely, given a field F, we can construct a new field if we come up with an irreducible polynomial in $F[x]$. Further, if F is a finite field, the new field will also be a finite field. So, here are some natural questions that we will be addressing (not necessarily in the order given here):

1. If F is a finite field and M is a polynomial in $F[x]$, is there an efficient way to tell if M is irreducible?

2. If F is a finite field and d is a positive integer, must there always be at least one irreducible polynomial in $F[x]$ of degree d? If so, how many of them are there? And if so, how might we find one of them?

3. A ubiquitous viewpoint in algebra, and in fact all of mathematics: it is not just the objects that are important, but the functions between the objects. In this case, we'd like some criterion for when 2 finite fields are *isomorphic*. (The technical definition: there is a bijection between the fields that preserves addition and multiplication. More on this later.) For example, both x^3+x^2+1 and x^3+x+1 are irreducible in $F[x]$ when $F = \mathbb{Z}/(2)$, and so we can construct two finite fields with 8 elements. Are they really the same; that is, are they isomorphic?

4. We've learned one way of constructing finite fields, namely start with $F = \mathbb{Z}/(p)$ with p a prime and then find an irreducible polynomial M in $F[x]$. Up to isomorphism, do all finite fields arise in this way?