CCS Discrete Math I | Professor: Padraic Bartlett

Lecture 7: More Finite Fields

Week 7 | *UCSB 2014*

> We just don't recognize life's most significant moments while they're happening. Back then I thought, "Well, there'll be other days." I didn't realize that that was the only day.
>
> Dr. Graham, Field of Dreams

1 Irreducible Polynomials

In our last class, we proved the following result:

Proposition. Suppose that $h(x)$ is an irreducible polynomial in $\mathbb{F}_p[x]$ of degree n. Then $\mathbb{F}_p[x]/h(x)$ is a field of order p^n.

Motivated by the above result, in this set of notes, we will attempt to prove the following:

Theorem. For any n and p, there is an irreducible polynomial of degree n in $\mathbb{F}_p[x]$.

This theorem, if we can prove it, will give us the following exciting corollary:

Corollary. For any prime p and positive integer n, there is a finite field of order p^n.

If this was an abstract algebra class, we would likely attempt to discuss how we can create such polynomials. Instead, because we are combinatorialists, we will attempt the simpler problem of simply **counting** the total number of irreducible polynomials! That is: if we can show that there is at least one irreducible polynomial of every degree, we will have proven our desired result (even though we might not have a great way to actually find said polynomials!)

While this might seem less useful of a result than an actual construction, this proof will have the advantages of (1) being **far** faster, (2) **far** easier to understand, and (3) actually giving you some ways to search for such irreducibles when you examine it closely!

We start with a few lemmas, that will help us understand how polynomials work in $\mathbb{F}_p$. The first of these is designed to help us understand how to factor polynomials; this will help us when we look for irreducibles, as those will be precisely the polynomials that do not factor into smaller parts!

For integers, we had some useful tools to help factor them into smaller pieces: namely, we could use the Euclidean algorithm to find the GCD of any two numbers, which was often useful for breaking larger numbers into smaller parts! We restate this process here:

Algorithm. The Euclidean algorithm is a method for taking two positive integers $a > b$ and calculating their GCD. It does so as follows.

To initialize our algorithm, set $r_1 = a, r_2 = b$. Our algorithm will create a sequence of values $r_1, r_2, r_3 \dots r_k,$, where this last value r_k will be the GCD of a, b.

1. Suppose that we have defined our sequence up to r_i, r_{i+1}, that $r_i > r_{i+1}$, and that $r_{i+1} > 0$.
2. If the remainder of r_i when divided by r_{i+1} is 0, quit our algorithm: r_{i+1} is the GCD we were looking for.
3. Otherwise, to define r_{i+2}, simply set it equal to the remainder of r_i when divided by r_{i+1}. This always gives us a number smaller than r_{i+1} that is positive, by definition.
4. Go to 1.

We claim that this process also works on polynomials! That is, consider the following process for determining the GCD of two polynomials:

Algorithm. Take any two polynomials $p(x), q(x)$, where the degree of $p(x)$ is not smaller than the degree of $q(x)$.

To initialize our algorithm, set $r_1(x) = p(x), r_2(x) = q(x)$. Our algorithm will create a sequence of polynomials $r_1(x), r_2(x), r_3(x) \dots r_k(x),$, where this last value $r_k(x)$ will be the GCD of a, b.

1. Suppose that we have defined our sequence up to $r_i(x), r_{i+1}(x)$, that the degree of $r_i(x)$ is greater than the degree of $r_{i+1}(x)$, and that $r_{i+1}(x) \neq 0$.
2. If the remainder of $r_i(x)$ when divided by $r_{i+1}(x)$ is 0, quit our algorithm: $r_{i+1}(x)$ is the GCD we were looking for.
3. Otherwise, to define $r_{i+2}(x)$, simply set it equal to the remainder of $r_i(x)$ when divided by $r_{i+1}(x)$. (Note that we are doing polynomial long division here! If you are unsure how to do this, check out Wikipedia for some background, or talk to me!)

   This always gives us a polynomial with degree smaller than $r_{i+1}(x)$, by definition.
4. Go to 1.

Checking that this works is a problem we leave for the HW! Instead, we run an example to illustrate how this works (and more generally, how polynomial arithmetic works over $\mathbb{F}_p[x]$:

Example. The GCD of $3x^5 + 4x^3 + 2x + 1$ and $4x^4 + 2x^2 + 4$ in $\mathbb{F}_5[x]$ is

Proof. We simply run the algorithm above! Set $r_1(x) = 3x^5 + 2x^3 + 2x + 3$ and $r_2(x) = 4x^4 + 2x^2 + 4$. Then, by polynomial long division, using the fact that $2 \cdot 4 \equiv 3 \mod 5$, we have

$$
\begin{array}{r|rrrr}
 & & & 2x & \\
\hline
4x^4+2x^2+3 & 3x^5 & +2x^3 & +2x & +3 \\
 & -3x^5 & -4x^3 & -3x & \\
\hline
 & & 3x^3 & +4x & +3
\end{array}
$$

and thus that $r_3(x) = 3x^3 + 4x + 3$. We repeat the process on $r_2(x), r_3(x)$, using the fact that $3 \cdot 3 \equiv 4 \mod 5$:

$$
\begin{array}{r|rrrr}
 & & & 3x & \\
\hline
3x^3+4x+3 & 4x^4 & +2x^2 & & +4 \\
 & -4x^4 & -2x^2 & -4x & \\
\hline
 & & & x & +4
\end{array}
$$

Finally, we notice that

$$
\begin{array}{r|rrrr}
 & & 3x^2 & & \\
\hline
x+4 & 3x^3 & & +4x & +3 \\
 & -3x^3 & -2x^2 & & \\
\hline
 & & 3x^2 & +4x & +3 \\
 & & -3x^2 & -2x & \\
\hline
 & & & 2x & +3 \\
 & & & -2x & -3 \\
\hline
 & & & & 0
\end{array}
$$

Therefore, we know that we're done at $r_4(x)$, and thus that the GCD of these two polynomials is $x + 4$, which is perhaps more usefully written as $x - 1$! To check, you can notice that both

$$3x^5 + 2x^3 + 2x + 3 = (x-1)(3x^4 + 3x^3 + 2),$$
$$4x^4 + 2x^2 + 4 = (x-1)(4x^3 + 4x^2 + x + 1),$$

and so this is a factor of both polynomials. (Checking that it's the largest such shared common factor I leave for the reader! It's not interesting: just try to factor the remaining polynomials above into smaller ones, and see that they don't have factors in common.) □

We will use this polynomial-Euclidean-algorithm to prove the following theorem:

Theorem. For any prime p and positive integer n, $x^{p^n} - x$ is the product of all of the monic[1] irreducible polynomials in $\mathbb{F}_p[x]$ whose degree divides n.

The reasons that we care about this theorem shouldn't be obvious yet; while it certainly tells us things about irreducibles, it's not obvious what things it tells us! Don't worry about that for now; when we get to the next section of these notes, on the inclusion-exclusion principle, this will become very clear!

For now, however, process it as a "weird fact about irreducibles that we're interested in proving."

To prove it, we use the following lemmas. The first is simple, but will help us later:

[1] A polynomial is called **monic** if the coefficient of its highest-degree term is 1. So, $x^2 + 2x + 4$ is monic, while $6x^2 - 53$ is not monic.

Lemma. Take any prime p and any integers n, d. Suppose that $n \equiv m \mod d$. Then

$$p^n \equiv p^m \mod (p^d - 1).$$

Proof. Notice that by definition,

$$p^d - 1 \equiv 0 \mod (p^d - 1),$$

and thus that

$$p^d \equiv 1 \mod (p^d - 1).$$

If $n \equiv m \mod d$, then we can find some k such that $n = kd + m$. Consequently, we can use our observation above to see that

$$p^n = p^m \cdot p^{kd} = p^m \cdot (p^d)^k \equiv p^m \cdot 1 \equiv p^m \mod (p^d - 1).$$

So we have proven our claim. □

The second lemma is not any harder, but is useful to point out for understanding how roots work in $\mathbb{F}_p[x]$:

Lemma. Let $f(x)$ be a polynomial in $\mathbb{F}[x]$, for any field $\mathbb{F}$. Then, if $f(a) = 0$ for some $a \in \mathbb{F}$, we can write $f(x) = (x - a) \cdot g(x)$, for some other polynomial $g(x) \in \mathbb{F}[x]$.

Proof. On the HW! (Basically: this is obvious when $a = 0$, as $f(0) = 0$ just means that the constant term of $f(x)$ is 0, which means we can factor out an x from all remaining terms. When $a \neq 0$, try looking at the polynomial $h(x) = f(x - a)$. How can a factorization of $h(x)$ help you factor $f(x)$?) □

The third is much tricker, but will help us understand the idea of "factors of $x^{p^n} - x$."

Lemma. Take any prime p and any integers n, d. Then in $\mathbb{F}_p[x]$, we have that

$$gcd\left(x^{p^n} - x, x^{p^d} - x\right) = x^{p^{gcd(n,d)}}.$$

(When we talk about the GCD of two polynomials, we mean the highest-degree factor that divides both polynomials. For example, the GCD of $x^2 - 2x + 1$ and $x^2 - 1$ is $(x - 1)$, as it is a factor of both polynomials.)

Proof. For convenience of notation, I will replace our $p(x) \equiv_{h(x)} q(x)$ notation with $p(x) \equiv q(x) \mod h(x)$, as this will let me use different $h(x)$'s with more ease / make the subscripts less messy.

First, notice that for any k, we can write $x^{p^n} - x$ as the following sum/product:

$$x^{p^n} - x = \left(\sum_{i=1}^{k} x^{p^n - i(p^d - 1) - 1}\right) \cdot \left(x^{p^d} - x\right) + \left(x^{p^n - k(p^d - 1)} - x\right).$$

This is a standard telescoping-trick thing, that we verify here for the skeptical:

$$\begin{aligned}
&\left(\sum_{i=1}^{k} x^{p^n - i(p^d-1)-1}\right) \cdot \left(x^{p^d} - x\right) + \left(x^{p^n - k(p^d-1)} - x\right) \\
=&\left(\sum_{i=1}^{k} x^{p^n - i(p^d-1)-1+p^d}\right) - \left(\sum_{i=1}^{k} x^{p^n - i(p^d-1)-1+1}\right) + \left(x^{p^n - k(p^d-1)} - x\right) \\
=&\left(\sum_{i=1}^{k} x^{p^n - (i-1)(p^d-1)}\right) - \left(\sum_{i=1}^{k} x^{p^n - i(p^d-1)}\right) + \left(x^{p^n - k(p^d-1)} - x\right) \\
=&\left(\sum_{i=0}^{k-1} x^{p^n - i(p^d-1)}\right) - \left(\sum_{i=1}^{k} x^{p^n - i(p^d-1)}\right) + \left(x^{p^n - k(p^d-1)} - x\right) \\
=&\left(x^{p^n - 0(p^d-1)} - x^{p^n - k(p^d-1)}\right) + \left(x^{p^n - k(p^d-1)} - x\right) \\
=&x^{p^n} - x.
\end{aligned}$$

Consequently, if we use this identity, we can see that

$$x^{p^n} - x \equiv \left(x^{p^n - k(p^d-1)} - x\right) \mod x^{p^d} - 1.$$

We showed earlier that if $n \equiv m \mod d$, then

$$p^n \equiv p^m \mod (p^d - 1);$$

in other words, we can find some k such that

$$p^n - k(p^d - 1) = p^m.$$

But (if we set m to be the remainder of n divided by d, i.e. $\mathrm{rem}(n/d)$) this means that we have

$$x^{p^n} - x \equiv \left(x^{p^{\mathrm{rem}(n/d)}} - x\right) \mod x^{p^d} - 1.$$

Notice that because adding copies of $x^{p^d} - 1$ to $x^{p^n \bmod d} - x$ will result in a polynomial of degree strictly larger than $p^{n \bmod d}$, we have actually shown that the remainder of $x^{p^n} - x$ on division by $x^{p^d} - 1$ **is** $x^{p^{n \bmod d}} - x$. In other words, we performed the first step of our Euclidean algorithm on polynomials!

For notational convenience, let's set

$$\begin{aligned}
f_1(x) &= x^{p^n} - x = x^{p^{r_1}} - x, \\
f_2(x) &= x^{p^d} - x = x^{p^{r_2}} - x, \\
f_3(x) &= x^{p^{\mathrm{rem}(r_1/r_2)}} - x = x^{p^{r_3}} - x.
\end{aligned}$$

Then, if we simply apply our above logic to $x^{p^{r_2}} - x, x^{p^{r_3}} - x$, we can get that

$$f_4(x) = x^{p^{\text{rem}(r_2/r_3)}} - x = x^{p^{r_4}} - x,$$

and in general get that

$$f_k(x) = x^{p^{\text{rem}(r_{k-2}/r_{k-1})}} - x = x^{p^{r_k}} - x.$$

When do we terminate our algorithm? When we get to some $f_k(x) = 0$; that is, some $r_k = 0$, as $x^{p^0} - x = x - x = 0$.

So it suffices to examine the sequence of the r_i's! Look at what we've formed here:

- $r_1 = n, r_2 = d$,
- $r_k =$ the remainder of r_{k-2} on division by r_{k-2},

This is just the Euclidean algorithm for integers applied to n, d! So we know the last nonzero value of this sequence is $gcd(n, d)$; consequently, we also know that the last term of our f_k sequences is

$$x^{p^{gcd(n,d)}} - x.$$

In other words, we've proven our claim: that

$$gcd\left(x^{p^n} - x, x^{p^d} - x\right) = x^{p^{gcd(n,d)}} - x.$$

□

Think about this for a while; it is almost surely the trickiest proof we have done in this class so far. There is nothing really crazy here — to find the GCD of polynomials, we used the Euclidean algorithm, and noticed a pattern that made it work like the Euclidean algorithm on integers if you just looked at the exponents! But there were a lot of weird superscripts and it may be hard to keep track of it all.

It also may be hard to see why we care! That comes in when we look at the following promised theorem:

Theorem. For any prime p and positive integer n, $x^{p^n} - x$ is the product of all of the monic irreducible polynomials in $\mathbb{F}_p[x]$ whose degree divides n.

Proof. Take any monic irreducible polynomial $\pi(x)$ of degree d, where $d|n$. We will show that $\pi(x)$ divides $x^{p^n} - x$!

To see why: notice that because $\pi(x)$ is irreducible, then $\mathbb{F}_p[x]/\pi(x)$ is a field of order p^d, as shown in class!

Therefore, if you look at all of the nonzero elements in this field, you get a multiplicative group of order $p^d - 1$. The order of any element in this group, by Lagrange's theorem, must divide the size of this group; consequently we know that for any a in our field, we must have $a^{p^d-1} = 1$, and therefore that $a^{p^d} - a = 0$ for any a in our field!

In other words: the expression $a^{p^d} - a$ is equal to 0 for every element a in our field $\mathbb{F}_p[x]/\pi(x)$ (as plugging in 0 to this expression also yields 0!)

But what does this mean? Well: consider two cases.

1. The degree of $\pi(x)$ is greater than 1. This means that x is an element of $\mathbb{F}_p[x]/\pi(x)$, as it has smaller degree than $\pi(x)$.

   We just showed that plugging in any element a of our field into the expression $a^{p^d} - a$ yields 0; so, let's do this with x! We get that

   $$x^{p^d} - x = 0;$$

   in other words, $x^{p^d} - x$ is a multiple of $\pi(x)$. By our lemma earlier, this means that $x^{p^n} - x$ is also a multiple of $\pi(x)$, as claimed!

2. The degree of $\pi(x) = 1$. This we can verify directly: write $\pi(x) = x + c$ for some constant $c \in \mathbb{F}_p$ (any monic irreducible polynomial is in this form!), and note that the typical telescoping-sum trick gives us

   $$x^{p^n} - x = (x+c)\Big(x^{p^n-1} + (-c)x^{p^n-2} + (-c)^2x^{p^n-3} + (-c)^3x^{p^n-4} + (-c)^4x^{p^n-5} + \dots \\ \dots + (-c)^{p^n-3}x^2 + (-c)^{p^n-2}x^1 + (-c)^{p^n-1}\Big),$$

   because

   $$\begin{aligned}(x+c)\left(\sum_{k=1}^{p^n-1} x^{p^n-k}(-c)^{k-1}\right) &= \left(\sum_{k=1}^{p^n-1} x^{p^n-k+1}(-c)^{k-1}\right) - \left(\sum_{k=1}^{p^n-1} x^{p^n-k+1}(-c)^{k}\right) \\ &= \left(\sum_{k=0}^{p^n-2} x^{p^n-k}(-c)^{k}\right) - \left(\sum_{k=1}^{p^n-1} x^{p^n-k+1}(-c)^{k}\right) \\ &= x^{p^n}(-c)^0 \pm (\text{terms that all cancel out}) - x^1(-c)^{p^n-1} \\ &= x^{p^n} - x.\end{aligned}$$

   (The last step here is justified by the observation that $(-c)^{p-1} = 1$ by Fermat's little theorem, and therefore that $(-c)^{p^n-1}$ is also 1, because $p^n - 1 = (p-1)(p^{n-1} + p^{n-2} + \dots + p + 1)$.)

So we have shown that $\pi(x)$ divides $x^{p^n} - x$, whenever $\pi(x)$ is an irreducible polynomial of degree d with $d|n$.

We now do something very silly. Observe that the derivative of $x^{p^n} - x$ is just $p^n \cdot x^{p^n-1} - 1 = -1$ in $\mathbb{F}_p[x]$. Now, consider any polynomial of the form $f(x))^2g(x)$ in $\mathbb{F}_p[x]$; the derivative of this polynomial is just $2f(x)(f'(x)g(x) + f(x)g'(x))$, which is a multiple of $f(x)$! Finally, observe that -1 is not a multiple of any nonconstant polynomial. What can you conclude?

No factor of $x^{p^n} - x$ occurs more than once!

This is a fun trick, and it comes up a lot! Derivatives: they come in handy.

We are nearly done. We know that if $d|n$ and $\pi(x)$ is an irreducible polynomial of degree d, it shows up exactly once in the factorization of $x^{p^n} - x$ into irreducibles. To finish our proof, it suffices to show that no other irreducible polynomials show up in this factorization.

We prove this by induction on n. For $n = 1$, we know that $x^p - x$ satisfies this property, by our "factoring out linear roots" result earlier: because $a^p - a = 0$ for any $a \in \mathbb{F}_p$ by Fermat's little theorem, we can factor out p linear roots from $x^p - x$, and can see that there are no more because this is a polynomial of degree p.

For our induction step: take any irreducible polynomial $\pi(x)$ that divides $x^{p^n} - x$. Suppose that the degree of $\pi(x)$ is d: then π divides $x^{p^d} - x$ by our work above. Consequently, $\pi(x)$ divides the GCD of $x^{p^n} - x, x^{p^d} - x$, which we showed was $x^{p^{gcd(n,d)}} - x$.

If the GCD of n and d is smaller than d, then by induction we know that it is impossible for $\pi(x)$ to divide $x^{p^{gcd(n,d)}} - x$! But $x^{p^{gcd(n,d)}} - x$ divides $x^{p^n} - x$ by our earlier arguments; so this is a contradiction.

So we must have the GCD of n and d equal to d; in other words, d divides n. □

Woo! Hardest proof of the quarter.

. . . what can we **do** with it?

2 Inclusion-Exclusion

The principle of **inclusion-exclusion** is a relatively simple method for determining the number of things in a collection, akin to our counting rules from earlier.

Theorem. (**Principle of Inclusion-Exclusion**.) Suppose that we have two sets A, B, and we want to count the total number of elements in $A \cup B$. We can express this quantity in terms of the sets A, B and $A \cap B$ as follows:

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

The reasoning behind this is simple: if we want to count all of the elements in $A \cup B$, we can just count the elements in A, and add to this the number of elements in B. However, doing this double-counts everything in their overlap; so we need to subtract off $|A \cap B|$ to insure that we've counted every element in $A \cup B$ exactly once.

We can extend this to three sets as well, by the same reasoning:

$$|A \cup B \cup C| = |A| + |B| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|.$$

This is because counting $A \cup B \cup C$ can be done by the followingL

- First, count all of the elements in A, B, C separately and add these quantities.
- This process "double-counted" all of the elements in the overlap of those sets. So, subtract off $|A \cap B|, |A \cap C|$ and $|B \cap C|$ to fix this.
- Now, think about elements in $A \cap B \cap C$. They were counted in each of A, B, C, positively and $|A \cap B|, |A \cap C|$ and $|B \cap C|$ negatively; so at the moment we're not counting them at all! Fix this by adding back in $|A \cap B \cap C|$.

- We are now counting each element in $A \cup B \cup C$ once!

In general, we can express the size of the union of n sets $A_1, \ldots A_n$ by a similar process:

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n} |A_i| \; - \sum_{1 \le i < j \le n} |A_i \cap A_j| \; + \sum_{1 \le i < j < k \le n} |A_i \cap A_j \cap A_k| \; - \; \ldots \; + \; (-1)^{n-1} |A_1 \cap \cdots \cap A_n| \, .$$

We can express this compactly as follows, if you like:

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{k=1}^{n} (-1)^{k+1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n} |A_{i_1} \cap \cdots \cap A_{i_k}| \right) .$$

The justification for this formula is similar to the above; we leave it for the homework!

Instead, we focus on what we can accomplish with this idea: clever counting tricks! We start with a warmup:

Question 1. *Take the collection of all numbers from 1 to 1000. How many are multiples of* $2, 3$ *or* 5*?*

Answer. We simply apply the principle of inclusion-exclusion to find all of the multiples of 2, 3 or 5. We can "over-count" by simply finding all multiples of 2, adding this to all multiples of 3, and then adding this to all multiples of 5:

$$\left\lfloor \frac{1000}{2} \right\rfloor + \left\lfloor \frac{1000}{3} \right\rfloor + \left\lfloor \frac{1000}{5} \right\rfloor$$

However, this "overcounts" by counting multiples of 6, 10 and 15 twice, as all of these show up in two of the quantities above. So fix this by subtracting one copy of all such numbers off:

$$\left\lfloor \frac{1000}{2} \right\rfloor + \left\lfloor \frac{1000}{3} \right\rfloor + \left\lfloor \frac{1000}{5} \right\rfloor - \left\lfloor \frac{1000}{6} \right\rfloor - \left\lfloor \frac{1000}{10} \right\rfloor - \left\lfloor \frac{1000}{15} \right\rfloor$$

This last correction has left us without any multiples of 30, which we need to count! So add those back in:

$$\left\lfloor \frac{1000}{2} \right\rfloor + \left\lfloor \frac{1000}{3} \right\rfloor + \left\lfloor \frac{1000}{5} \right\rfloor - \left\lfloor \frac{1000}{6} \right\rfloor - \left\lfloor \frac{1000}{10} \right\rfloor - \left\lfloor \frac{1000}{15} \right\rfloor + \left\lfloor \frac{1000}{30} \right\rfloor .$$

This is the size of our set! For fun, let's evaluate it:

$$\begin{aligned} &\left\lfloor \frac{1000}{2} \right\rfloor + \left\lfloor \frac{1000}{3} \right\rfloor + \left\lfloor \frac{1000}{5} \right\rfloor - \left\lfloor \frac{1000}{6} \right\rfloor - \left\lfloor \frac{1000}{10} \right\rfloor - \left\lfloor \frac{1000}{15} \right\rfloor + \left\lfloor \frac{1000}{30} \right\rfloor \\ =&500 + 333 + 200 - 166 - 100 - 66 + 33 \\ =&734. \end{aligned}$$

A second, trickier example of inclusion-exclusion is the following puzzle:

Example. Make a $100 \times 200 \times 300$ box out of $1 \times 1 \times 1$ cubes, and place this box in $\mathbb{R}^3$ so that one corner is at the origin and the other is at $(60, 140, 210)$. Draw the diagonal connecting $(0, 0, 0)$ to $(60, 140, 210)$. How many $1 \times 1 \times 1$ cubes does this cross?

Proof. Our diagonal crosses over from one small cube to another precisely when it passes through a face, edge or vertex of one of our cubes. This happens precisely at places where at least one of the coordinates of our line has integer coördinates! So it suffices to count all of these points, as doing this will tell us the total number of cubes that we enter. (Note that we don't want to count the point $(60, 140, 210)$ in any such count, as at this point we don't "cross over" into a new cube.)

How can we do this? Well: our diagonal line's coördinates are all points of the form

$$(60t, 140t, 210t),$$

for $t \in [0, 1]$. When is this integral and not $(60, 140, 210)$?

Well: the first coördinate is integral for precisely 60 values of t, namely $t = \frac{0}{60}, \frac{1}{60}, \frac{2}{60}, \ldots \frac{59}{60}$. Similarly, the second coördinate is integral for precisely 140 values of t, and the third coordinate is integral for 210 values of t.

However, this process overcounts some points! For example, points with their first two coördinates as integers are counted twice above; however, we only want to count such a point once! So we need to remove one copy of all of the points with two integer coördinates. This is not hard to do: $(60t, 140t)$ are both integral iff $t = \frac{k}{gcd(60,140)} = \frac{k}{20}$, of which there are twenty possible values. Similarly, $(60t, 210t)$ is integral at $gcd(60, 210) = 30$ values, while $(140t, 210t)$ is integral at $gcd(140, 210) = 70$ values.

Finally, we've over-subtracted some points in this last step: namely, those for which all three coordinates are integral! Those happen at values of t such that $(60t, 140t, 210t)$ are all integral; i.e. points of the form $\frac{k}{gcd(60,140,210)} = \frac{k}{10}$, of which there are 10.

So, in total, we have that there are

$$60 + 140 + 210 - 20 - 30 - 70 + 10 = 300$$

many such cubes intersected by our diagonal! □

2.1 Applying inclusion-exclusion to irreducibles.

Why mention this counting technique here? The answer is that it lets us answer the problem we started this lecture with: namely, how to count irreducible polynomials!

In specific, we have the following result:

Theorem. Let $M_n(p)$ denote the number of monic irreducible polynomials of degree n in $\mathbb{F}_p[x]$. Then we have the following equation:

$$M_n(p) = \frac{1}{n} \sum_{d|n} \mu(d) p^{n/d},$$

where $\mu(d)$ is the **Möbius function**, defined as follows:

$$\mu(d) := \begin{cases} 1, & d \text{ is squarefree and has an even number of prime factors,} \\ -1, & d \text{ is squarefree and has an odd number of prime factors,} \\ 0, & d \text{ has a squared prime factor.} \end{cases}$$

Proof. We proved earlier in the notes the following result: for any prime p, natural number n,

$$x^{p^n} - x = \prod_{\substack{\pi(x) \text{ monic and} \\ \text{irreducible,} \\ \text{s.t. } \deg(\pi(x)) \mid n}} \pi(x).$$

So: if we just think about the degrees of all of the irreducibles in the right-hand-side product, we can see that the sum of their degrees must be p^n, because the product of their polynomials is a polynomial of degree d^n! Consequently, we have

$$p^n = \sum_{\substack{\pi(x) \text{ monic and} \\ \text{irreducible,} \\ \text{s.t. } \deg(\pi(x)) \mid n}} \deg(\pi(x)).$$

If we group irreducibles by their degrees, we can further refine this statement to the equation

$$p^n = \sum_{d|n} d \cdot M_d(p),$$

which we have proven holds for all primes p, natural numbers n!

We want to solve the equations above for $M_n(d)$. To do this, we can simply use inclusion-exclusion!

Specifically: notice that in a sense, p^n "counts" $M_n(n)$ in its sum, as we have

$$p^n = n \cdot M_n(n) + \sum_{d|n, d<n} d \cdot M_d(p).$$

The only issue is that we've inadvertently also counted the $M_d(p)$'s at the same time! Conveniently, however, we know that we can count these with p^d's as well, for appropriate values of d!

To get a feel for how this works, let's consider some sample values of n. For $n = 1$, for example, we don't even need to think about double-counting:

$$p^1 = \sum_{d|1} d \cdot M_d(p) = M_1(p).$$

So the number of degree-1 monic irreducibles is p! This makes sense; we have that $x - c$ is a monic irreducible of degree 1 for every $c \in \mathbb{F}_p$, as these terms can't be broken down into smaller terms.

Let's consider $n = q$, for any prime q. Then, we have

$$p^n = p^q = \sum_{d|n} d \cdot M_d(p) = M_1(p) + qM_q(p).$$

$$\Rightarrow qM_q(p) = p^q - M_1(p) = p^q - p.$$

In this case, we were able to "correct" for $M_q(p)$'s overcounting by $M_1(p)$ by just subtracting this term off!

Let's go a bit further, and consider $n = qr$, for any two primes q, r. Then, we have two possibilities: either $q = r$, in which case we have

$$p^n = \sum_{d|n} d \cdot M_d(p) = M_1(p) + qM_q(p) + q^2 M_{q^2}(p),$$

or $q \neq r$, in which case we have

$$p^n = \sum_{d|n} d \cdot M_d(p) = M_1(p) + qM_q(p) + rM_r(p) + qrM_{qr}(p).$$

These are different cases! In the first, we would want to correct by subtracting off two terms;

$$q^2 M_{q^2}(p) = p^{q^2} - qM_q(p) - M_1(p).$$

Here, we could simply use our earlier knowledge of what $M_q(p)$ is for any primes q, p, but this is actually not the most useful observation to make here. Instead, notice that we showed that p^q counts $qM_q(p)$, with an overcount by $M_1(p)$: therefore, we have

$$q^2 M_{q^2}(p) = p^{q^2} - qM_q(p) - M_1(p) = p^{q^2} - (p^q - M_1(p)) - M_1(p) = p^{q^2} - p^q.$$

Similarly, in the second case, we could apply this knowledge as well: both $qM_q(p)$ and $rM_r(p)$ are counted by p^q, p^r respectively, with each overcounting by $M_1(p)$. Therefore, we have

$$qrM_{qr}(p) = p^{qr} - (p^q - M_1(p)) - (p^r - M_1(p)) + M_1(p) = p^{qr} - p^q - p^r + p.$$

So far, this lines up with our claim: that

$$M_n(p) = \frac{1}{n} \sum_{d|n} \mu(d) p^{n/d},$$

where $\mu(d)$ is the **Möbius function**, defined as follows:

$$\mu(d) := \begin{cases} 1, & d \text{ is squarefree and has an even number of prime factors,} \\ -1, & d \text{ is squarefree and has an odd number of prime factors,} \\ 0, & d \text{ has a squared prime factor.} \end{cases}$$

In particular, we've proven this claim for all values of n with at most two prime factors!

We will prove this claim by induction on the number k of prime factors of n. We have already accomplished our base cases above; now, let's assume that our claim works for all integers with k prime factors, and take any n with $k + 1$ prime factors. Label them $q_1, \ldots q_{k+1}$ for convenience.

As proven earlier in class, we know that

$$p^n = \sum_{d|n} d \cdot M_d(p).$$

Solving for $M_n(p)$ gives us

$$nM_n(p) = p^n - \sum_{d|n, d<n} d \cdot M_d(p).$$

Because $d|n$ and $d < n$, we know that each value of d in the sum above has k or fewer factors! Therefore, we can apply our inductive claim to it, and get

$$nM_n(p) = p^n - \sum_{d|n, d<n} \left(\sum_{c|d} \mu(c) p^{d/c} \right)$$

What do the individual terms of the sum above look like? Well: take any factor b of n. In how many ways can b occur as some expression of the form d/c? Well:

- If b has k prime factors, then there is only one way in which this happens: when $d = b, c = 1$. This is because d is restricted to terms that are less than n, and therefore terms that have at most k of n's prime factors.

  In this case, we have that the only term containing a q^b is the $\mu(1)q^b = q^b$ term.

- If b has $k-1$ prime factors, then we can write $bq_1q_2 = n$, for two primes q_1, q_2. We again look at all of the ways to write $b = d/c$:

  - When $d = b, c = 1$. $\mu(c)$ is 1 here.
  - When $d = b \cdot q_i, c = q_i$, for one of the two prime factors q_1, q_2. $\mu(c)$ is -1 here.

  So, in total, we have $1 - 2 = -1$ ways in which this happens, if we scale by the $\mu(c)$ terms.

- In general, what happens? Well: suppose that n/b consists of l **distinct** prime factors. Call them $q_1, \ldots q_l$. In this case, we can have $b = d/c$ in several different ways:

  - Once, when $d = b$, $c = 1$. $\mu(c)$ is 1 here.
  - We can have this happen l different ways, when $d = bq_j, c = q_j$ for any of the l distinct factors q_j of n/b. $\mu(c)$ is -1 here.
  - We can have this happen $\binom{l}{2}$ different ways, when $d = bq_iq_j, c = q_iq_j$. $\mu(c)$ is 1 here.
  - We can have this happen $\binom{l}{3}$ different ways, when $d = bq_iq_jq_m, c = q_iq_jq_m$. $\mu(c)$ is -1 here.
  - $\ldots$

– We can have this happen $\binom{l}{l-1}$ different ways, when $d = \frac{n}{q_j}, c = \frac{q_1 \cdot \ldots \cdot q_{k-l}}{q_j}$. $\mu(c) = (-1)^{l-1}$ here.

So, in total we have

$$\sum_{i=0}^{l-1} (-1)^i \binom{l}{i}$$

ways to make b, if we scale by the $\mu(c)$'s.

What is this sum? I claim it is more recognizable if we throw in an additional term: namely, notice that

$$\sum_{i=0}^{l} (-1)^i \binom{l}{i} = (1-1)^l,$$

by the binomial theorem! We know that the RHS above is just 0; therefore, the LHS is 0 as well, and therefore the first $l-1$ terms of the LHS are equal to the last term; that is,

$$\sum_{i=0}^{l-1} (-1)^i \binom{l}{i} = \binom{l}{l} \cdot (-1)^l = (-1)^l.$$

So the number of ways to make b is just $(-1)^l$, if n/b consists of l distinct prime factors.

- Finally, let's consider what happens if n/b has a repeated prime q in it. In this situation, notice that we can "pair" ways to write $b = d/c$ as follows:
  - If $b = d/c$ and c has no factors of q in it, pair this to the way of expressing b as $(dq)/(cq)$. Note that $\mu(c) = -\mu(cq)$, as we've gained exactly one new prime factor in this way.
  - This pairs off all pairs d/c where c contains no q's or one q. Any other c's contain at least 2 factors of q, and thus have $\mu(c) = 0$.

  So, in total, we have 0 such ways to make this b, if we scale by the $\mu(c)$'s.

But what does this mean? We have shown that a term q^b shows up in our sum iff b has no repeated factors, and moreover that the sign with which it shows up in our sum is given by the number of distinct prime factors in n/b! In other words, if we set $b = d$, we've proven our claim: that

$$M_n(p) = \frac{1}{n} \sum_{d|n} \mu(d) p^{n/d}.$$

□

As a nice corollary, we have the following:

Corollary. There are irreducible polynomials of degree n in $\mathbb{F}_p[x]$ for any n, p.

Proof. It suffices to show that $M_n(p) \geq 1$ for any n, p.

Doing this is not hard: if we have

$$nM_n(p) = \sum_{d|n} \mu(d) p^{n/d},$$

simply observe that the sum on the right is strictly larger than the sum

$$p^n - \sum_{k=0}^{n-1} p^k.$$

But for any integer $p \geq 2$, this is always greater than 1, which we can observe by induction; $p - 1 > 1$, and if

$$p^n - \sum_{k=0}^{n-1} p^k > 1,$$

we can multiply both sides by p to get

$$p^{n+1} - \sum_{k=0}^{n-1} p^{k+1} > p \Rightarrow p^{n+1} - \sum_{k=0}^{n-1} p^{k+1} - 1 > p - 1 \Rightarrow p^{n+1} - \sum_{k=0}^{n} p^k \quad > p - 1 > 1.$$

So we've proven our claim! □