Math 75 NOTES on finite fields

C. Pomerance

Suppose F is a finite field with q elements and F is a subfield of L, which as an F-vector space, has dimension d. We say that d is the degree of the field extension L over F, and we write $[L : F] = d$.

We see that this occurs if $M \in F[x]$ is irreducible of degree d and $L = F[x]/(M)$. In this case we identify the field F with the subfield of L consisting of constant polynomials.

Suppose that F is a subfield of K and K is a subfield of L. Then F is a subfield of L. Assume that $[L : F] = d$. Then K is a finite dimensional F-vector space, in fact it is a subspace of L. Say $[K : F] = j$. Since a basis for L over F also spans L if we allow the scalar field to increase to K, we have that L is a finite dimensional vector space over K, say $[L : K] = i$. We have

$$|L| = |F|^d, \quad |K| = |F|^j, \quad |L| = |K|^i.$$

Substituting the middle equation into the third, we get with the first equation that

$$|F|^d = \left(|F|^j\right)^i = |F|^{ji}.$$

Thus, $d = ji$, that is,

$$[L : F] = [L : K][K : F].$$

Suppose again we have F a subfield of L, say $|F| = q$, and let $\alpha \in F$. Let K be the subset $F[\alpha]$ of L. This is the set of all polynomials with coefficients in F, each one evaluated at α. Since $\alpha \in L$, it follows that $F[\alpha]$ is a subring of L. That is, we have closure under addition, multiplication, additive inverses, and we have the usual rules, like commutative, etc. So, is $F[\alpha]$ a subfield of L? That is, does it have multiplicative inverses for its nonzero elements?

To answer this question, we first note that there is some polynomial $f \in F[x]$ that is nonzero and satisfies $f(\alpha) = 0$. Indeed, if we look at the elements $1, \alpha, \alpha^2, \ldots, \alpha^d$ in $F[\alpha] \subset L$, we have $d + 1$ elements in the d-dimensional vector space L, so there is a linear dependency among them, that is, there are scalars $c_0, c_1, \ldots, c_d$ in F, not all 0, with $c_0 \cdot 1 + c_1 \cdot \alpha + \ldots c_d \cdot \alpha^d = 0$. So we have our $f(x)$. It is

$$c_d x^d + \cdots + c_1 x + c_0.$$

Now replace f with the monic polynomial M of minimal degree in $F[x]$ which has α as a root. We've learned that M is irreducible, and so $F[x]/(M)$ is a field. Let us compare this field with the ring $F[\alpha]$. I claim we do arithmetic exactly the same way. First, in $F[\alpha]$ we can represent elements uniquely as $b_0 + b_1\alpha + \cdots + b_{d-1}\alpha^{d-1}$, since whenever you have a power of α with exponent at least d, we can use the fact that $M(\alpha) = 0$ to solve for α^d as a linear combination of smaller powers of α. But this is exactly the same arithmetic as dividing a polynomial by $M(x)$ and replacing it with the remainder. So, identifying the "x" in $F[x]/(M)$ with the "α" in $F[\alpha]$, we see that we have an isomorphism. And since $F[x]/(M)$ is a field, so is $F[\alpha]$.

To recap, when we have F a subfield of L and $[L : F] = d$, then for every $\alpha \in L$, we have an intermediate field $K = F[\alpha]$ with $[K : F]$ isomorphic to $F[x]/(M)$ for M the monic irreducible polynomial in $F[x]$ with $M(\alpha) = 0$. Moreover, if j is the degree of M, then $[F[\alpha] : F] = j$ and $j \mid d$.

These, fairly easy facts go a long way towards our task of classifying and understanding finite fields!

There are two other very important facts. First, if F is a finite field and $|F| = q$, then $a^q = a$ for all $a \in F$. If you've studied algebra, you know this almost instantly from Lagrange's theorem (if G is a finite group with n elements and $g \in G$, then g^n is the group identity; this is applied to the multiplicative group $F^* = F \setminus \{0\}$ of the finite field F). But lets prove it from first principles. Let $a \in F$ and first note that $a^q = a$ holds if $a = 0$. So assume that $a \neq 0$. Let P be the product of all of the nonzero elements of F, that is

$$P = \prod_{b \in F^*} b.$$

Now the function $F \to F$ that takes an element b to ab is 1-to-1, since if $ab_1 = ab_2$, then multiplying both sides by a^{-1}, we get $b_1 = b_2$. Further, the function takes F^* to F^*, a set of size $q - 1$. A 1-to-1 function on a finite set to itself must be onto, so the list of elements ab as b runs over F^* is exactly the same as the list of all elements b in F^*, just in some permuted order. But multiplication doesn't care about the order of the factors, so we also have

$$P = \prod_{b \in F^*} (ab).$$

We have here $q - 1$ factors, each having a factor a, so we can re-write this last equation as

$$P = a^{q-1} \prod_{b \in F^*} b = a^{q-1} P.$$

Since $P \neq 0$, we have $1 = a^{q-1}$. Multiplying both sides by a proves our result. Namely, $a^q = a$ for all $a \in F$.

Another way of thinking of this is that the polynomial $x^q - x \in F[x]$ factors completely into q different linear factors:

$$x^q - x = \prod_{a \in F} (x - a).$$

We know this since in general, a is a root of $f(x)$ if and only if $x - a$ is a factor. Can you prove this last assertion?

Our next fact tells us something profound about the size q of a finite field F: it must be a prime number or a power of a prime. Here's why: We know that if we keep adding 1 to itself inside of F, we will eventually get a repeat, since there are just q elements in F, so subtracting off the shorter sum of 1's from the longer repeated sum of 1's, we get a sum of 1's equal to 0. Say k is the length of the shortest sum of 1's equal to 0. And say that k is a composite number

mn with $1 < m, n < k$. By the distributive law and $1 \cdot 1 = 1$, we have that m ones added together and then multiplied by the sum of n ones, gives $mn = k$ ones added together, which is 0. For example, say $k = 6 = 2 \cdot 3$. Then

$$0 = 1 + 1 + 1 + 1 + 1 + 1 = (1 + 1)(1 + 1 + 1).$$

But if the product of two field elements is 0, a factor must be 0. So, this violates the minimality of k, and we deduce that k must not be composite. We have $k \neq 1$ (since $1 \neq 0$), so k must be prime. Denote it by p, so it looks more "prime" like. This prime number is called the characteristic of the field F.

The characteristic p of the finite field F has the following nice properties. The first is that we see that F contains a copy of $\mathbb{Z}/(p)$ as a subfield. Second, if *any* element of F is added to itself p times, we get 0. We saw this with the element 1. Now let a be any element of F. We write it as $1 \cdot a$, then add it to istelf p times, and then factor out a. We see 1 added to itself p times, which is 0. So a added to itself p times is 0 times a, which is 0.

Next, we have for any $a, b \in F$:

$$(a + b)^p = a^p + b^p.$$

This follows since all of the intermediate terms from the binomial theorem are of the form

$$\binom{p}{j} a^j b^{p-j}, \quad \binom{p}{j} = \frac{p!}{j!(p-j)!}.$$

When $1 \leq j \leq p - 1$, the binomial coefficient $\binom{p}{j}$, which is an integer, has a factor p in the numerator, but no factors p in the denominator, so when it is reduced, it is a multiple of p. Thus, this intermediate term involves repeatedly adding something to itself a multiple of p times, so it must be 0. Thus, it's perfectly legal to ignore all those nasty intermediate terms when working out $(a + b)^p$, so long as p is the characteristic of the field.

By mathematical induction, we also have

$$(a + b)^{p^n} = a^{p^n} + b^{p^n}$$

for all a, b in the finite field F of characteristic p and for all positive integers n.

Finally, since $\mathbb{Z}/(p)$ is a subfield of F, we have F a finite dimensional vector space over the scalar field $\mathbb{Z}/(p)$, say it has dimension d. So $|F| = p^d$.

Lets put these ideas together to prove another astounding property of finite fields. We have seen that if F is a subfield of L and $[L : F] = d$, then for each $\alpha \in L$, we have an intermediate field $K = F[\alpha]$, say $[K : F] = j$. Also, $j \mid d$. Here's the astounding fact: There is at most one intermediate field K between F and L of degree j over F. Here's why. Say K_1, K_2 are intermediate fields between F and L both of degree j over F. If F has q elements, we've learned that K_1, K_2 both have q^j elements. Now consider the polynomial $x^{q^j} - x \in F[x]$. We've learned that this polynomial factors in $K_1[x]$ as

$$x^{q^j} - x = \prod_{a \in K_1} (x - a)$$

and it factors in $K_2[x]$ as

$$x^{q^j} - x = \prod_{b \in K_2} (x - b).$$

But both of these factorizations are living in $L[x]$, since both K_1, K_2 are subfields of L. Then, by unique factorization into irreducibles in $L[x]$, we must have these as the same factorization, so $K_1 = K_2$. (Another, equivalent way to see this: count roots of $x^{q^j} - x$ in L. It has at most q^j roots, since no polynomial over a field can have more roots than its degree. But it has q^j roots coming from the field K_1, and it has q^j roots coming from the field K_2, so these sets of roots must be identical.)

We're almost done with our classification. Our goal is to show that for every prime power there is a field with that number of elements and any two fields with that number of elements are isomorphic. For this it is very helpful to show that the polynomial $x^{q^d} - x$ over F, a finite field with d elements, is squarefree. This means that in the factorization of $x^{q^d} - x$ into monic irreducibles, there are no repeated factors.

Consider the function D on $F[x]$ which sends a polynomial to its derivative. In calculus class it was drilled into you that a derivative involves a limit. But the formulas for polynomials are quite simple. Here we want to forget the limit deal and just use the formulas. The question is if we can prove the familiar rules like the product rule. First note that D is an F-linear operator, that is, if $f, g \in F[x]$ and $a \in F$, then $D(f+g) = D(f)+D(g)$ and $D(af) = aD(f)$. Since every polynomial is built up from monomials, it would suffice to prove this for monomials. Well, if $f(x) = ux^j, g(x) = vx^k$, then $D(f) = jux^{j-1}$, $D(g) = kvx^{k-1}$, in the case when $\deg f$ and $\deg g$ are positive. Note that j is an integer in the exponent, but when it comes down in front to be "multiplied" by u, it is now thought of as j ones added together, which is the element j in $\mathbb{Z}/(p)$, where p is the characteristic of F. We see that $D(af) = D(aux^j) = jaux^{j-1} = aD(f)$. And similarly we get $D(f+g) = D(f)+D(g)$. The cases where f, g, or $f+g$ are constant work out similarly. So, to prove product rule, $D(fg) = fD(g) + gD(f)$, it would suffice to prove it in the case that f, g are monomials. And this is easy:

$$\begin{aligned} D(x^j x^k) &= D(x^{j+k}) = (j+k)x^{j+k-1} = jx^{j+k-1} + kx^{j+k-1} = jx^{j-1}x^k + kx^{k-1}x^j \\ &= D(x^j)x^k + D(x^k)x^j. \end{aligned}$$

Good. One corollary is that if $f(x)^2 \mid g(x)$ in $F[x]$, then $f(x) \mid D(g(x))$. Indeed, if $g = f^2 h$, then

$$D(g) = D(f^2 h) = f^2 D(h) + hD(f^2) = f^2 D(h) + h(fD(f) + fD(f)) = f(fD(h) + 2hD(f)).$$

So if $f^2 \mid x^{q^d} - x$, then we would have $f \mid D(x^{q^d} - x)$. This last derivative is $q^d x^{q^d - 1} - 1 = -1$, since the characteristic p divides q^d. But -1 cannot be divisible by any irreducible polynomial, so we get that $x^{q^d} - x$ is squarefree.

Our next step: If F is a finite field with q elements, we'll show that if $f(x) \in F[x]$ is an irreducible factor of $x^{q^d} - x$ of degree j, then $j \mid d$. Here's why. Using the division algorithm

in $\mathbb{Z}$, let $u, v \in \mathbb{Z}$ be such that $d = uj + v$, where $0 \le v \le j-1$. Also, let $K = F[x]/(f)$. We know that every element $a \in K$ satisfies $a^{q^j} = a$. Now the special element $x \in K$, lets call it α so that we don't get too confused, is a root of f, so it's a root of $x^{q^d} - x$, since $f \mid x^{q^d} - x$. Thus, $\alpha^{q^d} = \alpha$. But then every element in $a \in K$ has $a^{q^d} = a$. Indeed, by raising each side to, say, the ith power, we have $(\alpha^i)^{q^d} = \alpha^i$, for $i = 1, 2, \ldots$. And since members c of F have $c^q = c$, we also have $c^{q^d} = c$, for $i = 1, 2, \ldots$. So, since $(a+b)^{q^d} = a^{q^d} + b^{q^d}$ (why?), we have that every polynomial with F coefficients evaluated at α will give a value a with $a^{q^d} = a$. But this is everything in K. Now recall that we have $d = uj + v$. So, each $a \in K$ has

$$a = a^{q^d} = a^{q^{uj+v}} = (a^{q^j})^{q^{(u-1)j+v}} = a^{q^{(u-1)j+v}} = \cdots = a^{q^v}.$$

Thus, the polynomial $x^{q^v} - x$ has q^j roots, and $0 \le v < j$. So, if $v > 0$, we have too many roots for this polynomial, so we must have $v = 0$ and the polynomial $x^{q^v} - x$ reduces to the 0-polynomial. This proves that $j \mid d$.

For $j \mid d$, let $N_q(j)$ denote the number of irreducible factors of $x^{q^d} - x$ over $F[x]$ of degree j. Then

$$\sum_{j|d} jN_q(j) = q^d, \tag{1}$$

since $x^{q^d} - x$ is squarefree and thus the sum of the degrees of the irreducible divisors is the degree q^d of $x^{q^d} - x$. At this point we introduce the Möbius function $\mu(n)$ from elementary number theory and combinatorics. It is defined for positive integers n as follows: If n is the product of k different primes, then $\mu(n) = (-1)^k$, while if n is divisible by the square of a prime (i.e., it is not squarefree), then $\mu(n) = 0$. For example,

$$\mu(1) = 1,\ \mu(2) = \mu(3) = -1,\ \mu(4) = 0,\ \mu(5) = -1,\ \mu(6) = 1, \ldots$$

This function has the following cool identity:

$$\sum_{j|n} \mu(j) = \begin{cases} 1, & n = 1, \\ 0, & n > 1. \end{cases}$$

(We may discuss a proof; it is standard in elementary number theory.) Using this, we have by (1)

$$\begin{aligned} \sum_{j|d} \mu(d/j) q^j &= \sum_{j|d} \mu(d/j) \sum_{i|j} iN_q(i) = \sum_{i|d} iN_q(i) \sum_{j:\, i|j,\, j|d} \mu(d/j) = \sum_{i|d} iN_q(i) \sum_{k|d/i} \mu(k) \\ &= dN_q(d). \end{aligned}$$

This argument was a little tricky. We interchanged the order of summation, and then we introduced a new variable k equal to d/j in the next-to-last expression. In any event, we have what we want, a formula for $N_q(d)$. It is

$$N_q(d) = \frac{1}{d} \sum_{j|d} \mu(d/j) q^j. \tag{2}$$

This is an important formula and we shall make good use of it.