FINITE FIELDS

MATH 195

Introduction

A finite field is a field (commutative ring R with $R^* = R \setminus \{0\}$) with only finitely many elements. For example, $\mathbb{Z}/p\mathbb{Z}$ with p prime is a field, which we denote $\mathbb{F}_p$.

Theorem. *We have the following:*

(a) *If q is a positive integer, then there is a field k such that $\#k = q$ if and only if $q = p^n$ for some prime number p and $n \geq 1$.*

(b) *Suppose k_0 and k_1 are two finite fields, $\#k_i = p_i^{n_i}$, p_i prime, $n_i \geq 1$. Then there is an embedding $k_0 \hookrightarrow k_1$ as a subfield if and only if $p_0 = p_1$ and $n_0 \mid n_1$ if and only if $\#k_1$ is a power of $\#k_0$. In addition, if these three statements are true, then the number of embeddings $k_0 \hookrightarrow k_1$ is equal to n_0.*

In this case, p is called the characteristic of k, and n is the degree of n if $\#k = q = p^n$.

For example, if $\#k = p^n$, then $\mathbb{F}_p$ can be embedded in k in exactly one manner. After all, we must map 0 and 1 uniquely, and this respects the addition law, so $1+1$, $1+1+1$, $\dots$ and so on up to $1+1+\dots+1 = p = 0$ will already have fixed image.

As a second consequence, we note that if k_0 and k_1 are finite fields with the same number of elements, then $k_0 \simeq k_1$: the two fields are *isomorphic*, meaning we can view them as the same field. In other words, given $q = p^n$, there is 'essentially' only one field of q elements, which we denote $\mathbb{F}_q$.

Example. The field of 3 elements is represented by the addition and multiplication table for $\mathbb{Z}/3\mathbb{Z}$.

For a finite field of $4 = 2^2$ elements, we must work harder: $\mathbb{F}_4$ is built from $\mathbb{F}_2$, so it will have the elements $0, 1, a, b$, and tables

$+$	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

and

$\cdot$	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

This is some of the material covered March 12–14, in Math 195: Cryptography, taught by Hendrik Lenstra, prepared by John Voight `jvoight@math.berkeley.edu`; Revised March 17, 2002.

By the ring axioms,

$$a + a = 1 \cdot a + 1 \cdot a = (1 + 1) \cdot a = 0 \cdot a = 0$$

for any a which contains $\mathbb{F}_2$ as a subring (with the same unit element). Hence this is true for any ring of characteristic 2.

Generally, if R is any ring containing $\mathbb{F}_p$ as a subring (with the same unit element) then for all $a \in \mathbb{R}$ one has $\underbrace{a + a + \cdots + a}_{p} = 0$.

Notice that $\mathbb{F}_4$ has an automorphism which fixes $0, 1$ and maps $a \mapsto b$, $b \mapsto a$.

Constructing $\mathbb{F}_q$

Given p and n we would like to produce a 'model' of $\mathbb{F}_{p^n}$. How?

In general, tables are not a sufficient model. For $q = p^n$, an addition table and multiplication table would take up space

$$2q^2 \left\lceil \frac{\log q}{\log 2} \right\rceil$$

which is quite bad. We would like to eliminate any powers of q.

The connection to linear algebra: in the above example, $\{a, 1\}$ are a basis for $\mathbb{F}_4$ viewed as a vector space over $\mathbb{F}_2$. From the example, $b = a + 1$, so we represent this as

$$\begin{aligned} 0 &= 0 \cdot a + 0 \cdot 1 = 00 \\ 1 &= 0 \cdot a + 1 \cdot 1 = 01 \\ a &= 1 \cdot a + 0 \cdot 1 = 10 \\ b &= 1 \cdot a + 1 \cdot 1 = 11. \end{aligned}$$

Hence each element of $\mathbb{F}_4$ can in a unique way be written as $c_1 a + c_0 1$ with $c_0, c_1 \in \mathbb{F}_2$. We have

$$c_1 c_0 + d_1 d_0 = (c_1 + d_1)(c_0 + d_0).$$

How does multiplication work in this scheme?

$$(c_1 a + c_0)(d_1 a + d_0) = (c_1 d_1) a^2 + (c_0 d_1 + c_1 d_0) a + c_0 d_0.$$

But we need to know what a^2 is for this to work. This is 1 entry in the table: $a^2 = b = a + 1$, then substitute again to get the formula in terms of the basis $a, 1$. This says that $a^2 + a + 1 = 0$, and we have

$$(c_1 a + c_0)(d_1 a + d_0) = (c_1 d_1 + c_1 d_0 + c_0 d_1) a + (c_0 d_0 + c_1 d_1) 1.$$

In other words, the arithmetic in the field F_4 is summarized by the equation $a^2 + a + 1 = 0$.

Polynomials over $\mathbb{F}_p$

Let p be a prime number, and $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. A *polynomial over* $\mathbb{F}_p$ is an expression of the form $c_n X^n + c_{n-1} X^{n-1} + \cdots + c_1 X + c_0$, where X is just a symbol and each $c_i \in \mathbb{F}_p$.

Usually, we resrict to the case where the highest degree term has a nonzero coefficient, $c_n \neq 0$. Then n is called the *degree* of f, written $n = \deg f$. The set of all polynomials over $\mathbb{F}_p$ is denoted by $\mathbb{F}_p[X]$, and it is a commutative ring containing $\mathbb{F}_p$ as a subring. This is far from being a finite field as this ring is infinite.

Under addition,

$$(c_nX^n + \cdots + c_1X + c_0) + (d_mX^m + \cdots + d_1X + d_0) = (c_n + d_n)X^n + \cdots + (c_0 + d_0)$$

(adding in a few zero terms we assume $m = n$). We multiply as usual, subject to the conditions that the coefficients are in $\mathbb{F}_p$: this looks like

$$(c_nX^n + \cdots + c_0)(d_mX^m + \cdots + d_0) = \sum_{k=0}^{m+n} \sum_{i+j=k} c_id_jX^k.$$

The typical name for a polynomial is $f = f(X)$. Notice that we have $\deg(fg) = \deg f + \deg g$. If $c_n = 1$, then f is *monic*.

There are extended analogies between $\mathbb{Z}$ and $\mathbb{F}_p[X]$. If we do computations with polynomials, we may write $f = c_nc_{n-1}\dots c_0$, for $c_i \in \mathbb{F}_p$.

Example. For $p = 2$, $1 \cdot X^3 + 1 \cdot X + 1$ would be written 1011.

This is not always a more compact representation: the polynomial $X^{2^{10}} = \underbrace{100\cdot 0}_{2^{10}\text{zeros}}$.

Addition of polynomials can be done digitwise modulo 2: $(X^3 + X + 1) + (X^4 + X^2 + 1)$ becomes $1011 + 10101 = 11110$, $X^4 + X^3 + X^2 + X$, with no carries. To multiply these, we write

$$\begin{array}{ccccccc}
 & & & 1 & 0 & 1 & 1 \\
 & & 1 & 0 & 1 & 0 & 1 \\
\hline
 & & & 1 & 0 & 1 & 1 \\
 & 1 & 0 & 1 & 1 & & \\
1 & 0 & 1 & 1 & & & \\
\hline
1 & 0 & 0 & 1 & 0 & 1 & 1 & 1
\end{array}$$

represents the multiplication $(X^3 + X + 1)(X^4 + X^2 + 1) = X^7 + X^4 + X^2 + X + 1$, just like multiplying integers, the second representing the repeated "shifts" of the first to be added together.

We have seen, then, the following analogies between $\mathbb{Z}$ and $\mathbb{F}_p[X]$:

In $\mathbb{Z}$, we may write numbers in base $b = 10$, whereas we think of polynomials as "base" X. The degree function on $\mathbb{F}_p[X]$ corresponds roughly then to $\log |n|$ for $\mathbb{Z}$. The unit groups are $\mathbb{Z}^* = \{\pm 1\}$, $\mathbb{F}_p[X]^* = \mathbb{F}_p^*$. Note then that every integer can be written as a positive integer times a unit (± 1), and similarly, every polynomial in $\mathbb{F}_p[X]$ can be written as a monic polynomial times the leading coefficient $c_n \neq 0$, hence a unit.

This analogy is limited: for example, if you add two integers which are positive, you again get a positive integer, but this is not true if you add two monic polynomials.

As another example, $(X^3 + X + 1)^2$ can be computed as

$$\begin{array}{ccccccc}
 & & & 1 & 0 & 1 & 1 \\
 & & & 1 & 0 & 1 & 1 \\
\hline
 & & & 1 & 0 & 1 & 1 \\
 & & 1 & 0 & 1 & 1 & \\
1 & 0 & 1 & 1 & & & \\
\hline
1 & 0 & 0 & 0 & 1 & 0 & 1
\end{array}$$

so that $(X^3 + X + 1)^2 = X^6 + X^2 + 1$.

This is a more general phenomenon, called *Freshperson's dream*: $(a+b)^2 = a^2+b^2$ in any commutative ring (e.g. $\mathbb{F}_2[X]$) containing $\mathbb{F}_2$. (For a proof, consider the missing term $2ab$.) Note that

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

so if $3 = 0$ in the ring, i.e. if $\mathbb{F}_3$ is contained in your ring, then this becomes just $a^3 + b^3$. More generally:

Claim (Freshperson's dream). If a commutative ring R contains $\mathbb{F}_p$, then $(a+b)^p = a^p + b^p$ in R.

From this, and the trivial facts that $(ab)^p = a^pb^p$ and $1^p = 1$, we see that the map $a \mapsto a^p$ for a ring R containing $\mathbb{F}_p$ is called the *Frobenius map* is in fact a ring homomorphism. Notice that there is no analogy of the Frobenius map for $\mathbb{Z}$.

Recall that the integers have division with remainder: for any positive integers a, b, there exists $q > 0$ and $0 \le r \le b-1$ such that

$$a = qb + r$$

where q is the quotient and r the remainder. There is an analogous statement for $\mathbb{F}_p[X]$.

Claim. If $f, g \in \mathbb{F}_p[X]$, $g \neq 0$, then there exists unique $q, r \in \mathbb{F}_p[X]$ such that

$$f = qg + r$$

where $\deg r < \deg g$ or $r = 0$.

Example. For $p = 3$. Take $f = X^4 - X^3$, $g = X^3 - X - 1$, with coefficients $\mathbb{F}_3 = \{0, 1, -1\}$. We perform long division (synthetic division just subject to the arithmetic in $\mathbb{F}_p$), and obtain $q = X - 1$, $r = X^2 - 1$.

Here is another schematic way to compute this:

Example. For $p = 2$, $g = X^3 + X + 1 = 1011$, $f = X^{10} + X^5 + X^2 = 10000100100$, so we compute:

$$\begin{array}{ccccccccccc}
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 & & & & & & & \\
 & & 1 & 0 & 1 & 1 & & & & & \\
 & & & 1 & 0 & 1 & 1 & & & & \\
 & & & & 1 & 0 & 1 & 1 & & & \\
 & & & & & 1 & 0 & 1 & 1 & &
\end{array}$$

We repeatedly add 1011 so that the columns add to to the top. This says that, in fact, $r = 0$.

We say that $\mathbb{Z}$ and $\mathbb{F}_p[X]$ are *Euclidean domains* since they have this division with remainder.

Theorem (Unique factorization)**.** *Every monic polynomial in $\mathbb{F}_p[X]$ can in a unique way (up to ordering) be written as a product of monic* irreducible *polynomials.*

A polynomial $f \in \mathbb{F}_p[X]$ is called *irreducible* if $\deg f > 0$ and there do not exist $g, h \in \mathbb{F}_p[X]$ such that $f = gh$, $\deg g, \deg h < \deg f$.

Example. We compute the factorization of $f = X^{10} + X^9 + X^5 + X^2$ in $\mathbb{F}_2$. We note that this has a double root $X = 0$, giving us $X^{10} + X^9 + X^5 + X^2 = X^2(X^8 + X^6 + X^3 + 1)$. This second polynomial has $X = 1$ as a root, so we compute (using long division) that $X^8 + X^6 + X^3 + 1 = (X+1)(X^7 + X^4 + X^3 + X^2 + X + 1)$. This again has $X = 1$ as a root, and we are left with the factor $X^6 + X^5 + X^4 + X^2 + 1$. This has no roots. If it has an (irreducible) factor, it will be degree 2 or degree 3. The only monic quadratic with a root is $X^2 + X + 1$, and a cubic is irreducible if and only if it does not have a root, hence we have only $X^3 + X^2 + 1$ and $X^3 + X + 1$. ($X = 1$ is a root if and only if there is an even number of termss.) Dividing each of these into our degree 6 factor we see that there is a remainder, so a complete factorization is given by

$$X^{10} + X^9 + X^5 + X^2 = X^2(X+1)^2(X^6 + X^5 + X^4 + X^2 + 1).$$

Recall that

$$a_n(p) = \#\{f \in \mathbb{F}_p[X] : \deg f = n, f \text{ monic irreducible}\}.$$

We have $a_1(2) = 2$, $a_2(2) = 1$ (the unique irreducible is $X^2 + X + 1$) and $a_2(3) = 2$.

Since any monic polynomial of degree 1 is of the form $X - a$, and each of these is irreducible, we have $a_1(p) = p$.

We count that there are p^2 monic polynomials of degree 2 ($X^2 + aX + b$, p choices for each of a and b). If it factors, it does so as $X^2 + aX + b = (X - c)(X - d)$: if $c \neq d$, there are $\binom{p}{2}$ choices, and if $c = d$, total p, for a total of $(p+1)p/2$. Therefore there are a total of

$$a_2(p) = p^2 - \frac{p(p+1)}{2} = \frac{1}{2}(p^2 - p).$$

This reasoning will continue: if a cubic factors, it does so as the product of a linear and irreducible quadratic or as the product of three linear factors. This is a bit of a headache, hence a homework problem:

$$a_3(p) = \frac{1}{3}(p^3 - p).$$

If you continue in this way, you find

$$a_4(p) = \frac{1}{4}(p^4 - p^2)$$

and

$$a_5(p) = \frac{1}{5}(p^5 - p),$$

with an amazing amount of cancellation. The next term becomes

$$a_6(p) = \frac{1}{6}(p^6 - p^3 - p^2 + p),$$

which is much more complicated.

We have the following analogue to the prime number theorem:

Claim. For all $n \geq 1$ and all primes p, one has

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

Example. For example, with $n = 2$, since $d = 1$ or $d = 2$ we have

$$\sum_{d|n} da_d(p) = p^2 = a_1(p) + 2a_2(p),$$

so

$$a_2(p) = \frac{1}{2}(p^2 - a_1(p)) = \frac{1}{2}(p^2 - p).$$

Similarly,

$$6a_6(p) = p^6 - 3a_3(p) - 2a_2(p) - a_1(p) = p^6 - p^3 - p^2 + p.$$

This gives $a_n(p) \leq p^n/n$ as an upper bound, and for a lower bound,

$$na_n(p) = p^n - \sum_{\substack{d|n \\ d \leq \lfloor n/2 \rfloor}} da_d(p) \geq p^n - \sum_{d=1}^{\lfloor n/2 \rfloor} p^d > p^n - 2p^{\lfloor n/2 \rfloor}.$$

Therefore

$$\frac{p^n - 2p^{\lfloor n/2 \rfloor}}{n} < a_n(p) \leq \frac{p^n}{n}.$$