

**MATH 101A: ALGEBRA I**  
**PART D: GALOIS THEORY**

This is the title page for the notes on Galois Theory.

CONTENTS

1. Basics of field extensions	1
1.1. irreducible polynomials give field extensions	1
1.2. finite extensions give irreducible polynomials	2
1.3. properties of polynomials	3
2. Basics of Galois extensions	6
2.1. separable extensions	6
2.2. Galois extensions	8
2.3. Galois group	9
3. Examples	11
3.1. roots of unity	11
3.2. Galois correspondence	11
3.3. quadratic extensions	14
3.4. cubic extensions	15
3.5. Vandermonde	16
3.6. $\alpha = \sqrt{2} + \sqrt{3}$	17
4. Reed-Solomon code	19
4.1. the code	19
4.2. decoding with no errors	19
4.3. errors	20
4.4. finding the errors	21
4.5. homogeneous linear recurrence	21
4.6. Euclidean algorithm	23
4.7. proof of uniqueness	25
5. Finite fields	29
5.1. Galois group	29
5.2. the field $\mathbb{F}_4$	29
5.3. the field $\mathbb{F}_{16}$	30
5.4. the field $\mathbb{F}_{256}$	31

## 1. BASICS OF FIELD EXTENSIONS

On the first day I talked about the relationship between algebraic field extensions and irreducible polynomials. Namely, irreducible polynomials give field extensions and elements of finite extensions give irreducible polynomials.

## 1.1. irreducible polynomials give field extensions.

**Definition 1.1.** If  $K \subseteq L$  are fields then  $L$  is called a *field extension* of  $K$  and  $K$  is called a *subfield* of  $L$ . In this case,  $L$  becomes a vector space over  $K$  and its dimension is called the *degree* of the extension and denoted:

$$|L : K| = \dim_K L.$$

**Theorem 1.2.** *If  $K$  is a field and  $p(X) \in K[X]$  is an irreducible polynomial of degree  $n$  then*

- (1)  $L = K[X]/(p(X))$  is a field containing  $K$ . (Actually,  $L$  contains a subfield isomorphic to  $K$ .)
- (2)  $|L : K| = n$
- (3)  $L$  contains a root of the polynomial  $p(X)$ .

*Proof.* (1) Since  $K[X]$  is a PID, it is also a UFD. Therefore,  $p(X)$  irreducible implies  $(p(X))$  is prime. So,  $L = K[X]/(p(X))$  is a domain. By the lemma below and (2), this implies that  $L$  is a field.

(3) The monomial  $X$  (actually it is the coset  $X + (p(X))$ ) is a root of  $p(X)$  since  $p(X) = 0$  in the quotient  $L$ .

(2) The monomials  $1, X, X^2, \dots, X^{n-1} \in L$  are linearly independent over  $K$  because any  $K$ -linear relation  $\sum c_i X^i = 0, c_i \in K$  is a polynomial in  $K[X]$  of degree  $< n$  which lies in  $(p(X))$  and is therefore 0. On the other hand, any polynomial in  $X$  is equivalent module  $p(X)$  to a polynomial of degree less than  $n$ . So, these  $n$  monomials span  $L$ . So, the dimension of  $L$  over  $K$  is  $n$ .  $\square$

**Lemma 1.3.** *Any finite dimensional domain over a field is a field.*

*Proof.* Suppose  $D$  is a domain which contains a field  $K$  so that  $D$  is finite dimensional as a vector space over  $K$ . Let  $a \neq 0 \in D$  be a nonzero element of  $D$ . Then multiplication by  $a$  gives a  $K$ -linear mapping

$$\mu_a : D \rightarrow D$$

which is a monomorphism. If  $\dim_K D = n$  then this is a  $K$ -linear map  $K^n \rightarrow K^n$ . Any such map is a monomorphism if and only if it is an isomorphism.  $\square$

The theorem can be restated as follows.

**Corollary 1.4.** *Given any field  $K$  and any irreducible polynomial  $p(X) \in K[X]$ , there is a field extension  $L$  of  $K$  which contains a root of  $p(X)$ .*

Using Zorn's Lemma we can continue to adjoin roots of irreducible polynomials until we can't add any more. Then we get the *algebraic closure*  $\overline{K}$  of  $K$ . This is a field extension of  $K$  which contains all roots of all polynomials with coefficients in  $K$ . Although the existence of the algebraic closure follows from Zorn's Lemma, the uniqueness follows from Galois theory which we will discuss later.

**1.2. finite extensions give irreducible polynomials.** Suppose that  $L$  is a *finite extension* of  $K$ , i.e., a field extension of finite degree:  $|L : K| < \infty$ . Then, for any  $a \in L$  the evaluation map

$$ev_a : K[X] \rightarrow L$$

sending  $f(X)$  to  $f(a)$  cannot be a monomorphism since  $K[X]$  is infinite dimensional and  $L$  is finite dimensional by assumption. Since  $L$  is a domain, the kernel of  $ev_a$  must be a prime ideal:

$$\ker(ev_a) = (p(X))$$

which is necessarily generated by an irreducible polynomial  $p(X)$ . If we take  $p(X)$  to be *monic* (i.e., with leading coefficient 1) then it is uniquely determined by  $a$  and we write:

$$p(X) = irr(a, K).$$

**Definition 1.5.**  $a \in L$  is *algebraic* over  $K$  if  $p(a) = 0$  for some polynomial  $p(X) \in K[X]$ . If every element of  $L$  is algebraic over  $K$  then  $L$  is called an *algebraic extension* of  $K$ . If  $a \in L$  is not algebraic over  $K$  it is called *transcendental* and  $L$  is called a *transcendental extension* of  $K$ .

**Example 1.6.** The *field of rational functions*  $L = K(X) = Q(K[X])$  is a transcendental extension of  $K$  with  $X$  being a transcendental element. This is the quotient field of  $K[X]$ . So, elements are fractions

$$\frac{f(X)}{g(X)}$$

where  $f(X), g(X) \in K[X]$  with  $g(X) \neq 0$ . These fractions are called *rational functions*.

**1.3. properties of polynomials.** We went over several properties of polynomials, some without proof. Suppose that  $f(X) \in K[X]$  is a monic polynomial. I.e.,

$$f(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0.$$

(a)  $f(a) = 0$  if and only if  $X - a$  divides  $f(X)$ . This is because

$$f(X) = (X - a)g(X) + c$$

where the remainder  $c \in K$  must be equal to  $f(a)$  since  $(X - a) = 0$  when  $X = a$ .

(b)  $f(X) = \prod_{i=1}^n (X - r_i)$  where  $r_i \in \overline{K}$  and  $\overline{K}$  is the algebraic closure of  $K$ . Thus  $X - r_i \in \overline{K}[X]$ .

(c)  $f(X)$  has multiple roots, i.e., the roots  $r_i$  in (b) are not distinct, if and only if  $f$  and its derivative have a common factor:  $(f(X), f'(X)) \neq 1$ . The *derivative* is given formally by

$$f'(X) = \sum k a_k X^{k-1} = nX^{n-1} + a_{n-1}(n-1)X^{n-2} + \cdots + a_1.$$

Using the product rule  $(fg)' = f'g + fg'$  we get

$$f'(X) = \left( \prod (X - r_i) \right)' = \sum_{j=1}^n \prod_{i \neq j} (X - r_i)$$

The linear factor  $X - r_j$  divides the right hand side if and only if  $r_j$  is a multiple root of  $f(X)$ . If  $f(X)$  is irreducible, this does not appear to be possible since the greatest common divisor  $(f(X), f'(X))$  is a polynomial of degree  $< n$  which divides  $f(X)$ . The only way to avoid a contradiction is if  $f'(X) = 0$ .

**Example 1.7.** Suppose that  $f(X) = X^p - b$  and  $\text{char } K = p$ . Then  $f'(X) = pX^{p-1} = 0$ . In this case, all roots of  $f(X)$  are equal. To see this suppose that  $r$  is a root, i.e.,  $r^p = b$ . Then

$$(X - r)^p = X^p - r^p = X^p - b.$$

If  $r \notin K$  then  $K(r)$  is called a *purely inseparable extension* of  $K$ .

In general, if  $f(X)$  is irreducible and has multiple roots then it must be of the form

$$f(X) = h(X^p)$$

for some polynomial  $h(X)$  and  $p$  must be the characteristic of  $K$ .

(d) If  $K$  has characteristic 0 and  $f(X)$  is irreducible then the roots of  $f(X)$  are all distinct.

(e) If  $K$  is finite then  $|K| = p^n$  is a power of a prime and  $K = GF(p^n) = GF(q) = \mathbb{F}_q$  where  $q = p^n$ .

**Theorem 1.8.**  $\mathbb{F}_q^\times$  is cyclic of order  $q - 1$ .

*Proof.* Since  $\mathbb{F}_q^\times$  is a group of order  $q - 1$ , all of its elements satisfy the equation  $X^{q-1} = 1$ . If the group is not cyclic then it has exponent  $n < q - 1$  which would mean that the elements of  $\mathbb{F}_q^\times$  are all roots of the polynomial  $X^n - 1$ . However, a polynomial of degree  $n$  can have at most  $n$  roots. So, we would get a contradiction.  $\square$

(f) I stated the following lemma without proof.

**Lemma 1.9.** *Let  $R$  be a PID and  $F = Q(R)$  the quotient field. Then  $f(X) \in R[X]$  is irreducible in  $Q(R)[X]$  if and only if  $\frac{1}{d}f(X)$  is irreducible in  $R[X]$  where  $d \in R$  is the greatest common divisor of the coefficients of  $F$ .*

*Proof.* If  $f(X)$  is irreducible in  $Q(R)[X]$  then it is certainly irreducible in  $R[X]$  excepts possibly for factoring out a common divisor from the coefficients. Conversely, suppose that  $d = 1$  and  $f(X)$  is irreducible in  $R[X]$ . If it is not irreducible in  $Q(R)[X]$  then it factors as a product of two polynomials with coefficients in the fraction field  $Q(R)$ . But we can clear the denominators and obtain a factorization of the form:

$$af(X) = bg(X)h(X)$$

where  $g(X), h(X) \in R[X]$  having no common divisor of their coefficients and  $a, b \in R$  are relatively prime. If  $a$  is a unit in  $R$  we get a contradiction to the assumption that  $f(X)$  is irreducible. Therefore there is an irreducible element  $p \in R$  so that  $p$  divides  $a$  (and therefore does not divide  $b$ ). But then we can pass to the quotient domain  $R/(p)$  and we get

$$0 = b\bar{g}(X)\bar{h}(X).$$

Since  $R/(p)[X]$  is a domain, we must have either  $\bar{g}(X) = 0$  or  $\bar{h}(X) = 0$ . I.e.,  $p$  divides the coefficients of  $g$  or of  $h$ . This is a contradiction.  $\square$

(e) This implies the following important irreducibility criterion.

**Theorem 1.10** (Eisenstein). *Suppose that  $R$  is a PID and*

$$f(X) = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0 \in R[X]$$

*If there is an irreducible  $p \in R$  so that  $p$  divides  $a_0, a_1, \dots, a_{n-1}$ ,  $p$  doesn't divide  $a_n$  and  $p^2$  does not divide  $a_0$  then  $f(X)$  is irreducible  $Q(R)[X]$ .*

*Proof.* Let  $d$  be the greatest common divisor of the coefficients of  $f(X)$ . Then dividing by  $d$  will not change the validity of the condition. So, we may assume that  $d = 1$  and irreducibility in  $Q(R)[X]$  is the same as irreducibility in  $R[X]$  by the lemma. Suppose that  $f(X) = g(X)h(X)$

where  $g(X), h(X) \in R[X]$ . Then, we pass to the field of fractions  $Q(R/(p))$  of the quotient domain  $R/(p)$ :

$$R[X] \twoheadrightarrow R/(p)[X] \hookrightarrow Q(R/(p))[X]$$

Since  $p$  does not divide  $a_n$ , the image  $\overline{a_n}$  of  $a_n$  in  $Q(R/(p))$  is nonzero. Therefore, the image  $f(X)$  in  $Q(R/(p))[X]$  is the monomial

$$\overline{f}(X) = \overline{a_n}X^n.$$

Since  $Q(R/(p))[X]$  is a PID and thus a UFD, the images  $\overline{g}(X), \overline{h}(X)$  must also be constants times monomials:

$$\overline{g}(X) = bX^k, \quad \overline{h}(X) = cX^{n-k}.$$

This implies that the original polynomials  $g(X), h(X)$  must have the form

$$\begin{aligned} g(X) &= b_k X^k + b_{k-1} X^{k-1} + \cdots + b_0 \\ h(X) &= c_{n-k} X^{n-k} + \cdots + c_0 \end{aligned}$$

where  $p$  divides  $b_0, \dots, b_{k-1}, c_0, \dots, c_{n-k-1}$ . But then  $p^2$  divides  $a_0 = b_0 c_0$  which is a contradiction.  $\square$

## 2. BASICS OF GALOIS EXTENSIONS

- (1) separable extensions
- (2) Galois (normal) extensions
- (3) Galois group

**2.1. separable extensions.** I gave an unusual definition of “separable extension” which I replaced with a standard definition in these notes and the usual definition of separable element. Then I matched the two definitions using basic properties of field extensions from the first day.

**Definition 2.1.** Suppose that  $L, M$  are two field extensions of  $K$ . Then an *embedding* of  $L$  into  $M$  over  $K$  is defined to be a homomorphism

$$\phi : L \rightarrow M$$

so that  $\phi|_K = id_K$ .

*Remark 2.2.* Any homomorphism  $\phi : L \rightarrow R$  of a field  $L$  into a ring  $R$  (with  $1 \neq 0$ ) is necessarily an embedding since the only ideal in  $L$  is 0 and  $\ker \phi$  cannot be all of  $L$  since  $\phi(1) = 1 \neq 0$ .

**Lemma 2.3.** *Suppose that  $f(X)$  is an irreducible polynomial over  $K$  with  $n$  distinct roots  $a_1, a_2, \dots, a_n \in \overline{K}$ . Then there are  $n$  embeddings*

$$\phi_i : L = K[X]/(f(X)) \rightarrow \overline{K}$$

*over  $K$  so that  $\phi_i(X) = a_i$ . Furthermore, there are no other embeddings of  $L$  into  $\overline{K}$  over  $K$ .*

*Remark 2.4.* This is the key lemma which I used many times in the lecture. I also used a version which appears to be more general, but, by a “slight-of-hand,” can be seen to be the same. The more general statement is that the  $n$  embedding can be taken to be equal to some fixed embedding  $\phi : K \rightarrow \overline{K}$  which is not the inclusion map. But this is really the same thing because *we can replace  $K$  with its image in  $\overline{K}$* . Call this image  $K'$ . Replace all the coefficients of  $f(X)$  by their images in  $K'$ . This gives  $\phi(f)$ . Then the embeddings of  $K'[X]/(\phi(f))$  into  $\overline{K}$  over  $K'$  are equivalent to the extensions of the embedding  $\phi$  to  $K[X]/(f)$ .

*Proof.* Let

$$\psi_i = ev_{a_i} : K[X] \rightarrow \overline{K}$$

be the evaluation map  $\psi_i(h) = h(a_i)$ . Then  $\psi_i(f) = f(a_i) = 0$ . So,  $\psi_i$  induces a unique homomorphism

$$\phi_i : L = K[X]/(f) \rightarrow \overline{K}$$

so that  $\phi_i(h) = h(a_i)$ . In particular,  $\phi_i(X) = a_i$ . The uniqueness is obvious. Any embedding  $\phi : L \rightarrow \overline{K}$  must map  $X$  to a root of  $f(X)$ . So, it must map  $X$  to some  $a_i$ . Since  $\phi$  is assumed to be the identity on  $K$ , its value is determined on all of  $K[X]$ . So, it must be the one we already have.  $\square$

**Definition 2.5.** An extension  $L$  of  $K$  is called *separable* if

- (1)  $L$  is an algebraic extension of  $K$  and
- (2) for any  $a \in L$  with  $[K(a) : K] = n$ , there are embeddings

$$\phi_1, \phi_2, \dots, \phi_n : L \rightarrow \overline{K}$$

over  $K$  so that  $\phi_i(a)$  are all distinct.

**Definition 2.6.** Let  $L$  be an algebraic extension of  $K$ . Then an element  $a \in L$  is called *separable* if the irreducible=minimal polynomial  $f(X) = \text{irr}(a, K)$  does not have multiple roots, i.e., if  $f(X)$  and its derivative  $f'(X)$  are relatively prime.

Going slightly out of order, I explained at this point that the minimal polynomial is irreducible:

**Lemma 2.7.** Suppose that  $a \in L$  is algebraic over  $K$  and  $f(X)$  is the minimal polynomial of  $a$ , i.e., the polynomial in  $K[X]$  of smallest degree so that  $f(a) = 0$ . Then  $f(X)$  is irreducible and divides any polynomial  $g(X)$  for which  $a$  is a root.

*Proof.* The set of all polynomials  $g(X) \in K[X]$  so that  $g(a) = 0$  forms an ideal. This ideal is principal since  $K[X]$  is a PID. So it is generated by some  $f(X)$ . The only thing is to show that  $f(X)$  is irreducible. This is by contradiction. If not then  $f(X) = g(X)h(X)$  and  $f(a) = 0 = g(a)h(a)$ . So, either  $g(a) = 0$  or  $h(a) = 0$  which means that one of these lies in the ideal. So, the ideal is prime and its generator is irreducible.  $\square$

**Theorem 2.8.** An algebraic extension  $L$  of  $K$  is separable if and only if every element is separable over  $K$ .

*Proof.* ( $\Leftarrow$ ) Suppose that every element of  $L$  is separable over  $K$ . To show that  $L$  is a separable extension, take any  $a \in L$ ,  $a \notin K$ . Let  $f(X)$  be the minimal polynomial of  $a$  over  $K$ . Then  $\deg f = n \geq 2$  with distinct roots  $a_1, a_2, \dots, a_n \in \overline{K}$  and

$$K(a) \cong K[X]/(f)$$

which, by Lemma 2.3, has  $n$  embeddings  $\phi_i : K(a) \rightarrow \overline{K}$  sending  $a$  to  $n$  different elements  $a_i$ . So, the  $\phi_i$  “separate”  $a$ . We would be finished if  $L = K(a)$ . Otherwise we need the lemma below.

( $\Rightarrow$ ) Suppose that  $L$  is a separable extension and  $a \in L, a \notin K$ . Then the images of the homomorphisms  $\phi_i : L \rightarrow \overline{K}$  are  $n$  distinct roots of the irreducible polynomial of  $a$ .  $\square$

**Lemma 2.9.** *Suppose that  $K \subseteq E \subseteq L$  and  $L$  is algebraic over  $K$ . Then any embedding  $\phi : E \rightarrow \overline{K}$  over  $K$  extends to an embedding of  $L$  into  $\overline{K}$ .*

*Proof.* This is a typical Zorn’s Lemma argument. Take the partially ordered set of all pairs  $(F_\alpha, \phi_\alpha)$  where  $F_\alpha$  is an intermediate field  $E \subseteq F_\alpha \subseteq L$  and  $\phi_\alpha : F_\alpha \rightarrow \overline{K}$  is an extension of  $\phi$  and say that  $(F_\alpha, \phi_\alpha) \leq (F_\beta, \phi_\beta)$  if  $F_\alpha \subseteq F_\beta$  and  $\phi_\alpha = \phi_\beta|_{F_\alpha}$ . If we have a tower  $\{(F_\alpha, \phi_\alpha)\}$ , then we get an upper bound by taking  $F_\infty = \bigcup F_\alpha$  and  $\phi_\infty = \bigcup \phi_\alpha$ . So, by Zorn’s Lemma, there exists a maximal element  $(F_\infty, \phi_\infty)$ . If  $F_\infty = L$  we are done. Otherwise, we get a contradiction using Lemma 2.3 (and Remark 2.4) as follows.

If  $F_\infty \neq L$  there exists some  $a \in L, a \notin F_\infty$ . But  $a$  satisfies an irreducible polynomial  $f(X) \in K[X]$  which factors as a product of irreducible polynomials

$$f(X) = \prod g_i(X)$$

$g_i(X) \in F_\infty[X]$  and  $a$  is a root of one of the factors, say  $g_0(X)$ . Since the embedding  $\phi_\infty$  is the identity on  $K$ , it fixes  $f(X)$  (acting on coefficients). So:

$$f(X) = \prod \phi_\infty(g_i)(X).$$

Let  $b \in \overline{K}$  be a root of the polynomial  $\phi_\infty(g_0)(X)$ . Then we get an embedding

$$F_\infty(a) \rightarrow \overline{K}$$

extending  $\phi_\infty$  by sending  $a$  to  $b$ .  $\square$

## 2.2. Galois extensions.

**Definition 2.10.** A separable algebraic extension  $L$  of  $K$  is called *normal* or *Galois* if the embeddings of  $L$  into  $\overline{K}$  over  $K$  all have the same image.

**Example 2.11.** Take  $K = \mathbb{Q}$  and  $L = \mathbb{Q}(\zeta)$  where  $\zeta = e^{2\pi i/5}$  is a primitive 5th root of unity. The irreducible polynomial of  $\zeta$  is

$$f(X) = X^4 + X^3 + X^2 + X + 1.$$

(One way to see this is to put  $X = Y + 1$ . Then

$$f(X) = f(Y + 1) = Y^4 + 5Y^3 + 10Y^2 + 10Y + 5$$

is irreducible by Eisenstein.) The roots of  $f(X)$  are  $\zeta, \zeta^2, \zeta^3, \zeta^4$  all of which are elements of  $L = \mathbb{Q}(\zeta)$ . Therefore, the embeddings  $\phi_i : L \rightarrow \overline{\mathbb{Q}}$  have the same image  $\mathbb{Q}(\zeta^i) = L$ . So,  $\mathbb{Q}(\zeta)$  is a normal extension of  $\mathbb{Q}$ .

By the same argument, we have the general statement:

**Proposition 2.12.** *Suppose that  $L = K(a)$  is a separable extension of  $K$ . Then  $L$  is normal if and only if it contains all the conjugates of  $a$ .*

**Definition 2.13.** If  $a$  is algebraic over  $K$  with minimal polynomial  $f(X)$ , the *conjugates* of  $a$  are defined to be the roots of  $f(X)$ .

**Definition 2.14.**  $L$  is called the *splitting field* of a polynomial  $f(X) \in K[X]$  if it is equal to  $K$  adjoin all of the roots of  $f(X)$ .

**Proposition 2.15.** *A finite separable extension is normal if and only if it is the splitting field of some polynomial.*

*Proof.* This is more or less obvious. You just take some elements which generate  $L$  as a field extension. Then  $L$  must be the splitting field of the product of the corresponding irreducible polynomials.

Conversely, if  $L$  is the splitting field of some polynomial, then any embedding into  $\overline{K}$  must map the roots of this polynomial to other roots of the same polynomial. So, all these embeddings will have image contained in  $L$ . Since  $L$  is a finite extension, the images must equal  $L$ .  $\square$

**2.3. Galois group.** If  $L$  is a normal extension of  $K$  then the embeddings have the same image and therefore form a group (after we identify  $L$  with one particular subfield of  $\overline{K}$ ). This group is called the *Galois group* of the extension  $L/K$  and denoted  $Gal(L/K)$ .

**Theorem 2.16.** *Suppose that  $L \subseteq \overline{K}$  is a finite normal extension of  $K$  of degree  $|L : K| = n$ . Then  $Gal(L/K)$  is a finite group of order  $n$ .*

This follows immediately from the following lemma.

**Lemma 2.17.** *Suppose that  $L$  is a finite separable extension of  $K$  of degree  $n$ . Then there are exactly  $n$  embeddings of  $L$  into  $\overline{K}$  over  $K$ .*

*Remark 2.18.* This would follow immediately from Lemma 2.3 if we had proved the theorem that any finite separable extension is generated by a single element.

*Proof.* This is by induction on  $n$  using Lemma 2.9 and the formula:

$$|L : K| = |L : E| \cdot |E : K|$$

Let  $E = K(a)$  for some  $a \in L, a \notin K$ . Suppose  $a$  has degree  $m$  over  $K$ . Then  $|K(a) : K| = m$  and  $|L : K(a)| = n/m < n$ . By Lemma 2.3, there are exactly  $m$  embeddings  $\phi_i$  of  $E = K(a)$  into  $\overline{K}$  over  $K$ . By induction on  $n$  each of these embeddings has exactly  $n/m$  extensions to  $L$ . This uses the “slight-of-hand” argument (Remark 2.4).  $\square$

## 3. EXAMPLES

I did some examples and explained the theory at the same time.

**3.1. roots of unity.** Let  $L = \mathbb{Q}(\zeta)$  where  $\zeta = e^{2\pi i/5}$  is a primitive 5th root of unity.

**Theorem 3.1.** For any prime  $p$ ,

$$f(X) = X^{p-1} + X^{p-2} + \cdots + X + 1$$

is irreducible over  $\mathbb{Q}$ .

*Proof.* When you plug in  $X = Y + 1$  you get

$$f(Y+1) = \frac{(Y+1)^p - 1}{Y} = Y^{p-1} + pY^{p-2} + \binom{p}{2}Y^{p-3} + \cdots + \binom{p}{2}Y + p$$

which is irreducible by Eisenstein.  $\square$

As I pointed out earlier,  $\mathbb{Q}(\zeta)$  is the splitting field of the polynomial  $f(X)$  for  $p = 5$  since it contains all of the roots (conjugates of  $\zeta$ ):  $\zeta, \zeta^2, \zeta^3, \zeta^4$ . Since the Galois group  $G = \text{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$  is a finite group of order equal to the degree of the extension, it has order 4. So, it is either  $\mathbb{Z}/4$  or  $\mathbb{Z}/2 \oplus \mathbb{Z}/2$ .

We also know by Lemma 2.3 (and Lemma 2.9 in more general cases where we need to adjoin more than one root to get the splitting field) that the Galois group acts transitively on the roots. So, there is an element  $\sigma \in G$  so that  $\sigma(\zeta) = \zeta^2$ . But,  $\sigma$  is an automorphism of the field. So,  $\sigma(\zeta^2) = \sigma(\zeta)^2 = \zeta^4$ ,  $\sigma(\zeta^3) = (\zeta^2)^3 = \zeta$  and  $\sigma(\zeta^4) = \zeta^3$ . As a permutation,  $\sigma = (1243)$  is a 4-cycle. So, the Galois group is cyclic of order 4.

The group  $\mathbb{Z}/4$  has exactly one nonzero proper subgroup  $2\mathbb{Z}/4$  which has 2 elements. These elements are  $1, \sigma^2$ . The automorphism  $\sigma^2$  is complex conjugation. It switches  $\zeta, \zeta^4$  and  $\zeta^2, \zeta^3$ . According to the Galois correspondence, this subgroup corresponds to the subfield fixed by complex conjugation, namely the real subfield. So the unique intermediate fields is

$$\mathbb{Q}(\zeta) \cap \mathbb{R} = \mathbb{Q}(\zeta + \zeta^4).$$

**3.2. Galois correspondence.** Using the theorems that we already proved, the Galois correspondence is easy to derive. The first step is to show that intermediate fields give Galois extensions.

**Lemma 3.2.** If  $L$  is a separable extension of  $K$  and  $E$  is an intermediate field, i.e.,  $K \subseteq E \subseteq L$ , then  $L$  is separable over  $E$ .

*Proof.* Let  $a \in L$ ,  $a \notin E$ ,  $f(X) = \text{irr}(a, E)$ . We need to show that  $f(X)$  has no multiple roots. Let  $g(X) = \text{irr}(a, K)$ . Then  $g(X) \in K[X] \subseteq E[X]$  with  $g(a) = 0$ . Therefore,  $g(X) \in (f(X))$  since this is the ideal of all polynomials in  $E[X]$  of which  $a$  is a root. Therefore,  $g(X) = f(X)h(X)$  for some  $h(X) \in E[X]$ . The assumption that  $L/K$  is separable implies that  $g$  has distinct roots. But the roots of  $f$  are roots of  $g$ . So, they are also distinct.  $\square$

**Theorem 3.3.** *If  $L$  is a Galois extension of  $K$  and  $E$  is an intermediate field then  $L$  is Galois over  $E$ .*

*Proof.* We just showed that  $L/E$  is separable. So, we just need to show that any embedding  $\phi : L \rightarrow \overline{E} = \overline{K}$  has image  $\phi(L) = L$ . But,  $\phi$  being the identity on  $E$  implies that it is the identity on  $K \subseteq E$ . So,  $L$  being Galois over  $K$  implies that  $\phi(L) = L$ .  $\square$

**Lemma 3.4.** *If  $K \subseteq E \subseteq L$  then*

$$|L : K| = |L : E| \cdot |E : K|.$$

*Proof.* Suppose that  $|L : E| = n$ . Then  $L$  has a basis  $x_1, \dots, x_n$  over  $E$ . This means that, for any  $a \in L$ , there are unique elements  $e_i \in E$  so that  $a = \sum e_i x_i$ .

If  $|E : K| = m$ , then  $E$  has a basis  $y_1, \dots, y_m$  over  $K$ . This implies that every element of  $E$ , for example  $e_i$ , is a  $K$ -linear combination of the elements  $y_j$ . So  $e_i = \sum a_{ij} y_j$  for some  $a_{ij} \in K$ .

Claim:  $\{x_i y_j\}$  is a basis for  $L$  over  $K$ .

We already know that this set spans  $L$  since

$$a = \sum e_i x_i = \sum a_{ij} x_i y_j.$$

This set is also linearly independent since

$$\sum a_{ij} x_i y_j = 0 \Rightarrow (\forall j) \sum a_{ij} x_i = 0 \Rightarrow (\forall ij) a_{ij} = 0.$$

So,  $|L : K| = nm = |L : E| \cdot |E : K|$ .  $\square$

**Theorem 3.5** (Galois correspondence). *Suppose  $L$  is a Galois extension of  $K$  with Galois group  $G$ . Then there is a 1-1 correspondence between the intermediate fields  $K \subseteq E \subseteq L$  and subgroups  $H \leq G$ . The correspondence maps  $E$  to  $H = \text{Gal}(L/E)$  and it maps  $H \leq G$  to the fixed field*

$$L^H = \{a \in L \mid \sigma(a) = a \forall \sigma \in H\}.$$

*Proof.* It follows from the definitions (and Theorem 3.3) that  $H = \text{Gal}(L/E)$  is a subgroup of  $G = \text{Gal}(L/K)$ . And it is straightforward to show that  $L^H$  is an intermediate field. The key point is to show that  $E = L^H$ .

Since  $H = \text{Gal}(L/E)$  fixes  $E$  by its definition,  $E \subseteq L^H$ . So, it suffices to show that the degree of this extension is 1, i.e.,  $|L^H : E| = 1$ . But,  $L^H$  is an intermediate field. So,  $L/L^H$  is a Galois extension. The degree of this extension is the size of the Galois group which I claim is

$$\text{Gal}(L/L^H) = H = \text{Gal}(L/E).$$

The reason is that the elements of  $H = \text{Gal}(L/E)$  fix  $L^H$  by definition of  $L^H$ . So,  $H \leq \text{Gal}(L/L^H)$ . And any element of  $\text{Gal}(L/L^H)$  will also fix  $E \subseteq L^H$ . So,  $\text{Gal}(L/L^H) \leq H$ . So, they are equal and their degrees are the same. So, the formula

$$|L : E| = |L : L^H| \cdot |L^H : E|$$

implies that  $|L^H : E| = 1$ . So  $L^H = E$ .

The other part of the proof I did not say in class: We should take an arbitrary subgroup  $H$  of  $G = \text{Gal}(L/K)$  and show that  $\text{Gal}(L/L^H) = H$ . Let  $S = \text{Gal}(L/L^H)$ . Then  $S$  contains  $H$  and  $|S| = |L : L^H| = n$ . So, all we need to do is to show that  $H$  contains at least  $n$  elements, i.e.,  $|L : L^H| \leq |H|$ .

For this I need to use the primitive element theorem which I mentioned earlier. It says that  $L$  is generated by one element:  $L = L^H(a)$ . Let  $C = \{\sigma(a) \mid \sigma \in H\}$ . Then  $C$  is a subset of  $L$  having  $\leq |H|$  number of elements which is invariant under the action of  $H$ . This implies that

$$f(X) = \prod_{c \in C} (X - c)$$

is invariant under the action of  $H$ , i.e.,  $f(X) \in L^H[X]$ . Since  $f(a) = 0$ ,  $f(X)$  is a multiple of  $\text{irr}(a, L^H)$ . So,

$$|L : L^H| = \deg(\text{irr}(a, L^H)) \leq \deg(f) = |C| \leq |H|$$

which is what we needed to prove.  $\square$

**Theorem 3.6.** *If  $L$  is a finite separable extension of  $K$  then  $L = K(a)$  for some  $a \in L$ .*

*Proof.* First we can exclude the case where  $K$  is finite because, in that case, the units of  $L$  form a cyclic group generated by one element. So, assume  $|K| = \infty$ .

Let  $a \in L$  be an element of maximal degree, say  $n$ . Then I claim that  $L = K(a)$ . Otherwise there is an element  $b \in L$ ,  $b \notin K(a)$ . Then

$|K(a, b) : K| = m > n$ . Let  $\phi_1, \phi_2, \dots, \phi_m : K(a, b) \rightarrow \overline{K}$  be the distinct embeddings of  $K(a, b)$  into  $\overline{K}$  over  $K$ . Then I claim that there is an element  $x \in K$  so that

$$\phi_i(a)x + \phi_i(b) = \phi_i(xa + b)$$

are all distinct. The reason is that the polynomial

$$p(X) = \prod_{i < j} ([\phi_i(a)X + \phi_i(b)] - [\phi_j(a)X + \phi_j(b)])$$

is nonzero (each factor being nonzero) and therefore there is an  $x \in K$  so that  $p(x) \neq 0$  (since  $p(X)$  has only a finite number of roots and  $K$  is infinite). But then,  $ax + b \in L$  has degree at least  $m > n$  which is a contradiction.  $\square$

**3.3. quadratic extensions.** Suppose that  $L$  is a degree 2 extension of  $K$ . Then  $L = K(a)$  and the irreducible polynomial of  $a$  is  $f(X) = X^2 + bX + c$ . Suppose that  $\text{char}(K) \neq 2$ . Then the roots of this polynomial are

$$\frac{-b \pm \sqrt{b^2 - 4c}}{2}$$

and  $a$  is one of these roots. Since  $b, 1/2 \in K$ , the extension  $K(a)$  is equal to

$$K(2a + b) = K(\sqrt{b^2 - 4c}) = K(\sqrt{\Delta}).$$

**Definition 3.7.** The *discriminant*  $\Delta$  of a polynomial  $f(X)$  is equal to the sum of squares of differences between the roots:  $\Delta = \delta^2$  where

$$\delta = \prod_{i < j} (\alpha_i - \alpha_j).$$

Note that  $\delta$  is only well defined up to sign.

**Theorem 3.8.** *If  $\text{char}(K) \neq 2$ , every quadratic extension of  $K$  has the form  $L = K(\sqrt{D})$  for some  $D \in K$ ,  $D \notin K^2$ .*

The irreducible polynomial of  $\sqrt{D}$  is  $X^2 - D$  which has two distinct roots  $\pm\sqrt{D}$  both of which lie in  $K(\sqrt{D})$ . Therefore,  $K(\sqrt{D})$  is a Galois extension of  $K$  with Galois group  $\mathbb{Z}/2$ .

**3.4. cubic extensions.** Now suppose that  $L$  is a cubic (degree 3) extension of  $K$  and  $\text{char}(K) \neq 2, 3$ . Then  $L = K(\alpha)$ . Suppose that  $f(X) = \text{irr}(\alpha, K)$ . Then there are two possibilities:

- (1)  $L$  is the splitting field of  $f(X)$ .
- (2)  $f(X)$  factors as

$$f(X) = (X - \alpha)g(X)$$

where  $g(X)$  is irreducible over  $L$ .

In the first case,  $L$  is a Galois extension of  $K$  and the Galois group is  $\mathbb{Z}/3$ .

In the second case, the splitting field of  $f(X)$  is  $L(\beta) = K(\alpha, \beta)$  where  $\beta$  is a root of  $g(X)$ . Since  $g(X)$  is quadratic,  $|K(\alpha, \beta) : K(\alpha)| = 2$  and

$$|K(\alpha, \beta) : K| = |K(\alpha, \beta) : K(\alpha)| |K(\alpha) : K| = 2 \cdot 3 = 6.$$

The Galois group  $\text{Gal}(K(\alpha, \beta)/K)$  is the symmetric group on 3 letters. This follows from the following observation.

**Theorem 3.9.** *The Galois group of the splitting field of a polynomial  $f(X)$  is a subgroup of the group of permutations of the roots of  $f(X)$ .*

*Proof.* I said in class that this is “obvious.” That is because the splitting field  $L$  of  $f(X)$  is generated by the roots of  $f(X)$  by definition:

$$L = K(\alpha_1, \dots, \alpha_n)$$

and any automorphism  $\phi$  of  $L$  over  $K$  is determined by its effect on the these generators. Furthermore,  $\phi$  must take roots to roots since it is a homomorphism which fixes the coefficients of  $f(X)$ . So,  $\phi$  permutes the roots of  $f(X)$  and is determined by this permutation.  $\square$

In terms of permutation groups the two cases are

- (1)  $G = A_3 = \langle (123) \rangle$  the alternating group since this is the only subgroup of  $S_3$  with 3 elements.
- (2)  $G = S_3$ .

We discussed the role of the discriminant:  $\Delta = \delta^2$  and the fact that  $\delta$  is *alternating* in the sense that it changes sign if you switch two of the roots.

**Lemma 3.10.**  $\delta = \prod_{i < j} (\alpha_j - \alpha_i)$  has the property that it changes sign if two of the roots are switched. In fact, for any permutation  $\sigma \in S_n$  we have:

$$\prod_{i < j} (\alpha_{\sigma(j)} - \alpha_{\sigma(i)}) = \text{sgn}(\sigma)\delta.$$

*Proof.* It suffices to take the case where  $\sigma$  is a transposition of consecutive integers:  $\sigma = (i, i + 1)$ . In that case, the factor  $\alpha_{i+1} - \alpha_i$  changes sign and all other factors remain the same.  $\square$

**Corollary 3.11.**

$$\delta \in L^{A_n}.$$

**Theorem 3.12.** *Let  $f(X)$  be an irreducible cubic polynomial over  $K$  with discriminant  $\Delta \in K$ . Assume  $\text{char}(K) \neq 2, 3$ . Then the splitting field of  $f(X)$  has degree 3 over  $K$  if and only if  $\delta = \sqrt{\Delta} \in K$ .*

*Proof.* In class we figured out that one direction is clear: If the degree of splitting field is 3 then the Galois group is  $A_3$  and  $\delta \in L^{A_3} = K$ .

As Roger pointed out after class, the converse is also easy: If the splitting field has degree 6 and the Galois group is  $S_3$  then  $\delta$  is not an element of  $K = L^{S_3}$  since it is not fixed by the action of  $S_3$ .  $\square$

**3.5. Vandermonde.** This product  $\delta$  is also the sign of the *Vandermonde* determinant:

$$\delta = \det \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \alpha_3^{n-1} & \cdots & \alpha_n^{n-1} \end{pmatrix}$$

The proof is by induction on  $n$ . Let  $X = \alpha_n$ . Then, the Vandermonde determinant, call it  $V_n$ , is a polynomial in  $X$  of degree  $n - 1$ . If  $\alpha_n = \alpha_i$  for some  $i < n$  then the determinant of the matrix is zero. So,  $\alpha_1, \alpha_2, \dots, \alpha_{n-1}$  are the roots of the polynomial. So:

$$V_n = C \prod_{i=1}^{n-1} (X - \alpha_i)$$

where  $C$  must be the leading coefficient, i.e.,  $C = V_{n-1}$  is the smaller Vandermonde which we know by induction:

$$V_{n-1} = \prod_{n-1 \geq j > i \geq 1} (\alpha_j - \alpha_i).$$

Multiplying these we get

$$V_n = V_{n-1} \prod_{i=1}^{n-1} (X - \alpha_i)$$

which gives the formula we wanted when  $X = \alpha_n$ .

3.6.  $\alpha = \sqrt{2} + \sqrt{3}$ . We computed the irreducible polynomial of this element by squaring it twice:

$$\begin{aligned}\alpha^2 &= 2 + 3 + 2\sqrt{6} \\ x(\alpha^2 - 5)^2 &= \alpha^4 - 10\alpha^2 + 25 = 4 \cdot 6 = 24\end{aligned}$$

So,

$$f(X) = \text{irr}(\alpha, \mathbb{Q}) = X^4 - 10X^2 + 1.$$

This polynomial is irreducible over  $\mathbb{Q}$  because it is irreducible over  $\mathbb{Z}$ . This in turn follows from that observation that, if  $f(X)$  factors over  $\mathbb{Z}$ , it must be as a product of two quadratic terms. This would imply that two of the roots

$$\pm\sqrt{2} \pm \sqrt{3}$$

have both product and sum equal to an integer. But that is not so.

I pointed out that sums of these roots give:

$$\begin{aligned}(\sqrt{2} + \sqrt{3}) + (\sqrt{2} - \sqrt{3}) &= 2\sqrt{2} \\ (\sqrt{2} + \sqrt{3}) + (-\sqrt{2} + \sqrt{3}) &= 2\sqrt{3}\end{aligned}$$

This means that the splitting field of  $f(X)$  contains  $\sqrt{2}$  and  $\sqrt{3}$ . And conversely,

$$\pm\sqrt{2} \pm \sqrt{3} \in L = \mathbb{Q}(\sqrt{2}, \sqrt{3}).$$

So,  $L$  must be the splitting field of  $f(X)$ .

Questions

- (1) What is the Galois group  $\text{Gal}(L/\mathbb{Q})$ ?
- (2) Find all subgroups of  $G$ .
- (3) What are the corresponding subfields of  $L$ ?

(1)  $\text{Gal}(L/\mathbb{Q})$

We know that the Galois group has 4 elements. So, it must be abelian (any group of order  $p^2$  is abelian). So, it is either  $\mathbb{Z}/4$  or  $\mathbb{Z}/2 \oplus \mathbb{Z}/2$ . The group  $\mathbb{Z}/4$  has only one nontrivial proper subgroup,  $2\mathbb{Z}/4$ . By the Galois correspondence this would mean there is only one intermediate field. But we have at least three intermediate fields:

$$\mathbb{Q}(\sqrt{2}), \mathbb{Q}(\sqrt{3}), \mathbb{Q}(\sqrt{6}).$$

So,  $\text{Gal}(L/\mathbb{Q}) \cong \mathbb{Z}/2 \oplus \mathbb{Z}/2$ .

(2) Subgroups of  $\text{Gal}(L/\mathbb{Q})$ .

This group has three subgroups of order 2. These must correspond to the three obvious intermediate fields list above. The question is: What is the correspondence?

The elements of the Galois group are:  $\sigma, \tau, \sigma\tau, 1$  where  $\sigma, \tau$  are defined by:

$$\begin{aligned}\sigma(\sqrt{2}) &= -\sqrt{2}, & \sigma(\sqrt{3}) &= \sqrt{3}, \\ \tau(\sqrt{3}) &= -\sqrt{3}, & \tau(\sqrt{2}) &= \sqrt{2}, \\ \sigma\tau(\sqrt{2}) &= -\sqrt{2}, & \sigma\tau(\sqrt{3}) &= -\sqrt{3}.\end{aligned}$$

To prove this, use the fact that the Galois group acts transitively on the set of roots of  $f(X)$ . Then  $\sigma, \tau, \sigma\tau$  are the elements  $Gal(\mathbb{Q}(\alpha)/\mathbb{Q})$  which send  $\alpha = \sqrt{2} + \sqrt{3}$  to

$$\sigma(\alpha) = -\sqrt{2} + \sqrt{3}, \quad \tau(\alpha) = \sqrt{2} - \sqrt{3}, \quad \sigma\tau(\alpha) = -\alpha.$$

(3) Corresponding intermediate fields.

Since  $\sigma, \tau, \sigma\tau$  fix  $\sqrt{3}, \sqrt{2}, \sqrt{6}$ , resp., we have:

$$\begin{aligned}Gal(\mathbb{Q}(\alpha)/\mathbb{Q}(\sqrt{2})) &= \langle \tau \rangle \\ Gal(\mathbb{Q}(\alpha)/\mathbb{Q}(\sqrt{3})) &= \langle \sigma \rangle \\ Gal(\mathbb{Q}(\alpha)/\mathbb{Q}(\sqrt{6})) &= \langle \sigma\tau \rangle.\end{aligned}$$

## 4. REED-SOLOMON CODE

The Reed-Solomon code is a simple algebraic code using polynomials over finite fields. It is used on all CD's and the Berlekamp-Massey decoding algorithm allows the CD player to correct 16 mistakes out of 255 bytes.

4.1. **the code.** The encoding formula is extremely simple. Start with:

- (1)  $F$  a finite field (usually  $F = \mathbb{F}(2^8)$ , the field with  $q = 2^8$  elements).
- (2)  $k$  is a positive integer less than  $q$  elements (usually  $k = 223$ ).
- (3)  $n$  is a positive integer

$$k < n < q = |F|.$$

(usually  $n = 255$ )

- (4)  $r = n - k$  is a positive even integer (usually  $r = 255 - 223 = 32$ )
- (5)  $m = r/2$  is usually 16.

You start with data:

$$a_0, a_1, \dots, a_{k-1} \in F.$$

You take this block of data and make it into a monic polynomial of degree  $k$ :

$$f(X) = X^k + a_{k-1}X^{k-1} + \dots + a_0 \in F[X].$$

You take  $n$  fixed nonzero elements  $\alpha_1, \dots, \alpha_n \in F$ . Then the *Reed-Solomon code* is the  $n$ -tuple of elements of  $F$  given by:

$$(f(\alpha_1), f(\alpha_2), \dots, f(\alpha_n)) \in F^n.$$

Usually you let  $n = q - 1$  and you take all of the nonzero elements of the field.

The theorem of Berlekamp and Massey is that, assuming that at most  $r/2$  of the number  $f(\alpha_i)$  are read incorrectly, we can efficiently determine the original polynomial  $f(X)$ .

4.2. **decoding with no errors.** There is a simple formula to recover  $f(X)$  if we know  $f(\alpha_i)$  for more than  $k$  choices of  $\alpha_i$ .

**Definition 4.1.** Let  $L(X) \in F[X]$  be the monic degree  $n$  polynomial given by

$$L(X) = \prod_{j=1}^n (X - \alpha_j).$$

And, for each  $i$ , let

$$L_i(X) = \frac{L(X)}{X - \alpha_i} = \prod_{j \neq i} (X - \alpha_j).$$

If  $n = q - 1$  and  $\{\alpha_i\} = F^\times$  then

$$L(X) = X^n - 1$$

and

$$L_i(X) = \frac{X^n - 1}{X - \alpha_i}.$$

Since  $L_i(\alpha_j) = 0$  for  $i \neq j$  we have the following formula.

$$(4.1) \quad \frac{L_i(\alpha_j)}{L_i(\alpha_i)} = \delta_{ij}.$$

**Theorem 4.2.** *If  $f(X)$  is a polynomial of degree less than  $n$  then*

$$f(X) = \sum_{i=1}^n \frac{L_i(X)}{L_i(\alpha_i)} f(\alpha_i).$$

*Proof.* The equation holds for  $X = \alpha_i$  for all  $i$  by (4.1). This means the difference between the two polynomials is a polynomial of degree  $< n$  with  $n$  roots. So, this difference must be zero.  $\square$

Since  $f(X)$  has degree  $k$ , we can multiply  $f(X)$  by another polynomial of degree  $< n - k$  and still have the same equation:

**Corollary 4.3** (orthogonality condition). *For  $0 \leq s < r = n - k$  we have:*

$$f(X)X^s = \sum_{i=1}^n \frac{L_i(X)}{L_i(\alpha_i)} f(\alpha_i) \alpha_i^s.$$

*Comparing coefficients we get:*

$$\sum_{i=1}^n \frac{f(\alpha_i) \alpha_i^s}{L_i(\alpha_i)} = \delta_{s, r-1} = \begin{cases} 1 & \text{if } s = r - 1 \\ 0 & \text{if } s \leq r - 2 \end{cases}$$

**4.3. errors.** Suppose that there are errors in the transmission or reading of the code. Let  $A$  be the set of indices  $i$  for which the value  $f(\alpha_i)$  is misread and let  $B = \{1, 2, \dots, n\} - A$  be the complement. Thus we have the correct values of  $f(\alpha_j)$  for  $j \in B$  but we don't know what  $f(\alpha_i)$  is for any  $i \in A$ . Let  $L_A(X) = L^B(X)$  be defined by

$$L_A(X) = L^B(X) = \prod_{j \notin A} (X - \alpha_j) = \prod_{j \in B} (X - \alpha_j)$$

This is a polynomial of degree  $n - |A| = |B|$ . Let

$$L_i^B(X) = \prod_{j \in B, j \neq i} (X - \alpha_j) = \frac{L_A(X)}{X - \alpha_i}$$

for all  $i \in B$ . Then, as a special case of Theorem 4.2, we have:

**Corollary 4.4.** *For any polynomial  $f(X) \in F[X]$  of degree  $< |B|$  (i.e., if  $|A| < r = n - k$ ) we have*

$$f(X) = \sum_{i \in B} \frac{L_{i,A}(X)}{L_{i,A}(\alpha_i)} f(\alpha_i).$$

The conclusion is: We need to know the set  $A$  of indices for which  $f(\alpha_i)$  has been misread and we need  $|A| < r$ .

**4.4. finding the errors.** We assume that the error set  $A$  has at most  $r/2$  elements. The correct code is  $f(\alpha_i)$ . But, with errors we will read this as

$$c_i = f(\alpha_i) + \epsilon_i$$

where  $\epsilon_i$  is the *error*. Thus  $\epsilon_i \neq 0$  only for  $i \in A$ . The orthogonality condition Corollary 4.3 implies that

$$\sum_{i=1}^n \frac{c_i \alpha_i^s}{L_i(\alpha_i)} = \sum_{i=1}^n \frac{f(\alpha_i) \alpha_i^s}{L_i(\alpha_i)} + \sum_{i=1}^n \frac{\epsilon_i \alpha_i^s}{L_i(\alpha_i)} = \delta_{s,r-1} + \sum_{i=1}^n \frac{\epsilon_i \alpha_i^s}{L_i(\alpha_i)}.$$

Thus we can compute the sequence of  $r$  numbers:

$$d_s := \sum_{i \in A} \frac{\epsilon_i \alpha_i^s}{L_i(\alpha_i)} = \sum_{i=1}^n \frac{c_i \alpha_i^s}{L_i(\alpha_i)} - \delta_{s,r-1}$$

for  $0 \leq s < r$ . The key point is:

**Theorem 4.5.** *The numbers  $d_s$  satisfy a homogeneous linear recurrence of degree  $|A|$  and the roots of the minimal polynomial  $p(X)$  of this recurrence are  $\alpha_i$  for  $i \in A$ . Furthermore, if  $|A| \leq r/2$ , this linear recurrence and the polynomial  $p(X)$  are uniquely determined.*

To see this we need to review the solution of homogeneous linear recurrences.

#### 4.5. homogeneous linear recurrence.

**Definition 4.6.** We say that a sequence of elements  $d_0, d_1, \dots \in F$  satisfies a *homogeneous linear recurrence* of order  $m$  if there are fixed elements  $a_0, a_1, \dots, a_{m-1} \in F$  so that

$$d_s + a_{m-1}d_{s-1} + a_{m-2}d_{s-2} + \dots + a_0d_{s-m} = 0$$

for all  $s \geq m$ . If  $m$  is minimal, the degree  $m$  monic polynomial

$$p(X) = X^m + a_{m-1}X^{m-1} + \dots + a_0 \in F[X]$$

is called the *minimal polynomial* of the recurrence.

In class I used the *Fibonacci sequence*

$$1, 1, 2, 3, 5, 8, \dots$$

which satisfies the degree 2 homogeneous recurrence

$$d_s - d_{s-1} - d_{s-2} = 0$$

with minimal polynomial  $p(X) = X^2 - X - 1$ . To solve the recurrence we set

$$d_s = C^s.$$

Then the equation is

$$C^s - C^{s-1} - C^{s-2} = 0.$$

Assuming  $C \neq 0$  we get

$$C^2 - C - 1 = 0.$$

I.e.,  $C$  is a root of the polynomial  $p(X)$ . So,

$$C = C_{\pm} = \frac{1 \pm \sqrt{5}}{2}.$$

This means that the general solution of the recurrence is

$$d_s = e_0 C_+^s + e_1 C_-^s$$

where  $e_0, e_1$  are obtained from the initial conditions  $d_0 = 1, d_1 = 1$ .

This elementary argument works in general. Given any root  $\alpha$  of the polynomial  $p(X)$ , the sequence  $d_s = \alpha^s$  is a solution of the homogenous linear recurrence with minimal polynomial  $p(X)$ . If  $p(X)$  has  $m$  distinct roots  $\alpha_1, \dots, \alpha_m$  then the general solution of the recurrence is

$$d_s = \sum_{i=1}^m e_i \alpha_i^s$$

where  $e_i$  are determined by the initial values  $d_0, \dots, d_{m-1}$ . If we insert

$$e_i = \frac{\epsilon_i}{L_i(\alpha_i)}$$

we get the numbers from the Reed-Solomon code (assuming  $A = \{1, 2, \dots, m\}$ ). Since we are assuming that  $m \leq r/2$ , the problem is now reduced to the question:

Given the numbers  $d_0, \dots, d_{2m-1}$  can we find the degree  $m$  linear recurrence satisfied by these numbers?

The answer is given by the Euclidean division algorithm.

**4.6. Euclidean algorithm.** Suppose we have polynomials  $f(X), g(X) \in F[X]$  and we want to find the greatest common divisor  $h(X)$ . Since  $h(X)$  is the generator of the ideal  $(f, g)$  generated by  $f$  and  $g$ , there are polynomials  $a(X), b(X)$  so that

$$(4.2) \quad h(X) = a(X)f(X) + b(X)g(X).$$

The Euclidean algorithm will find all three:  $h(X), a(X), b(X)$ .

4.6.1. *the usual algorithm.* Note that the equation (4.2) is a matrix product:

$$h(X) = (a, b) \begin{pmatrix} f \\ g \end{pmatrix}.$$

The Euclidean algorithm starts with the two solutions of this equation:

$$\begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix}.$$

These are the first two rows of an array which will have the solution at the bottom. The third line is obtained by multiplying the second line by  $Q_1(X)$  and subtracting from the first line. Here  $Q_1(X)$  is the quotient  $f/g$ :

$$f(X) = Q_1(X)g(X) + R_1(X)$$

where  $\deg R(X) < \deg g(X)$ . This produces:

	$h$	$a$	$b$
$f(X)$	1	0	
$g(X)$	0	1	
$f - Q_1g = R_1(X)$	1	$-Q_1(X)$	

Next, divide  $R_1(X)$  into  $g(X)$ :

$$g(X) = Q_2(X)R_1(X) + R_2(X)$$

to get:

	$h$	$a$	$b$
$f(X)$	1	0	
$g(X)$	0	1	
$f - Q_1g = R_1(X)$	1	$-Q_1(X)$	
$g - Q_2R_1 = R_2(X)$	$-Q_2(X)$	$1 + Q_1Q_2$	

In each row, the term on the left has strictly smaller degree and Equation (4.2) holds. The last nonzero remainder is the greatest common divisor.

4.6.2. *finding the recurrence polynomial.* Suppose we have a recurrence  $d_0, d_1, \dots, d_{2m-1}$ . Then the procedure is to take  $f(X) = X^{2m}$ ,

$$g(X) = d_{2m-1}X^{2m-1} + \dots + d_0$$

and perform the Euclidean division algorithm only half way. You stop as soon as the remainder has degree less than  $m$ .

Here is an example. Take the sequence: 1,1,2,3. Then  $f(X) = X^4$  and  $g(X) = 1 + X + 2X^2 + 3X^3$ . The algorithm gives:

$Q$	$h$	$a$	$b$
	$X^4$	1	0
	$3X^3 + 2X^2 + X + 1$	0	1
$\frac{1}{9}(3X - 2)$	$\frac{1}{9}(X^2 - X + 2)$	1	$-\frac{1}{9}(3X - 2)$
$9(3X + 5)$	-9	$-9(3X + 5)$	$9(X^2 + X - 1)$

*Remark 4.7.* Since the  $Q$ 's are quotients of successive terms in the  $h$  column, their degrees add up to the difference in degrees between the first term  $X^{2m}$  and the second to last term in the  $h$  column which has degree  $\geq m$  (otherwise we would have stopped). The last  $b$  has degree equal to the product of the  $Q$ 's (by induction). So,  $\deg b \leq m$ .

**Theorem 4.8.** *The polynomial of the recurrence is given by*

$$p(X) = \frac{X^{\deg b}}{b(0)}b(1/X).$$

The formula  $X^{\deg b}b(1/X)$  reverses the coefficients of the polynomial  $b(X)$ . You need to divide by  $b(0)$  to make  $p(X)$  into a monic polynomial. In the example we get:

$$p(X) = \frac{X^2}{-9}9 \left( 1 + \frac{1}{X} - \frac{1}{X^2} \right) = X^2 - X - 1.$$

*Proof.* To see why this is the polynomial of the recurrence, note that the last row of our chart gives:

$$b(X)g(X) = h(X) - a(X)X^{2m}$$

Since  $h(X)$  has degree  $< m$ , there are no terms of degree  $s$  for  $s = m, m+1, m+2, \dots, 2m-1$ . This means that

$$b_0d_s + b_1d_{s-1} + \dots + b_md_{s-m} = 0$$

for all  $m \leq s < 2m$  where  $b(X) = b_0 + b_1X + \dots + b_mX^m$ . This is the linear recurrence with coefficients indexed backwards. So, the reverse polynomial  $p(X)$  is the polynomial of the the recurrence.  $\square$

**4.7. proof of uniqueness.** Today I explained why the algorithm always gives the correct answer. Namely, the minimal polynomial is uniquely determined and the algorithm will always give it to you (assuming the number of errors is  $\leq m = r/2$ ).

4.7.1. *a, b are relatively prime.* The algorithm starts with  $f(X) = X^{2m}$  and

$$g(X) = d_0 + d_1X + \cdots + d_{2m-1}X^{2m-1}$$

where the coefficients satisfy some homogeneous linear recurrence of degree  $\leq m$ . The algorithm produces polynomials  $h(X), a(X), b(X)$  so that  $\deg h < m, \deg b \leq m$  and

$$h(X) = a(X)X^{2m} + b(X)g(X).$$

**Lemma 4.9.** *a(X), b(X) are relatively prime.*

*Proof.* The reason is that the  $2 \times 2$  matrix formed by the last two entries in the  $a$  and  $b$  columns has determinant  $\pm 1$ . This is by induction. We start with the  $2 \times 2$  matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

which has determinant 1. Then, to get the third line, we subtract  $Q_1(X)$  times the second row from the first row. This is an elementary row operation which does not change the determinant of the matrix. However, the new row goes to the bottom so the second matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & -Q_1(X) \end{pmatrix}$$

has determinant  $-1$ . Then we subtract  $Q_2(X)$  times the third line from the second line. These are rows 2 and 1 in this second matrix. So, again, the determinant remains 1 in absolute value and changes sign. At the end we have a matrix of determinant  $\pm 1$  whose last row is  $(a, b)$ . So, the greatest common divisor of  $a, b$  is 1.  $\square$

4.7.2.  $b(0) \neq 0$ . The next point I made was that  $b(0) \neq 0$ . This is the same as saying that  $X$  does not divide  $b(X)$ . This is important for two reasons. First, we use the equation:

$$p(X) = \frac{X^{\deg b} b(1/X)}{b(0)}$$

to obtain the polynomial of the recurrence. The second reason is that, this will imply that  $h(X), b(X)$  are relatively prime. By the following lemma, this will imply that  $a, b, h$  are unique up to a scalar multiple.

**Lemma 4.10.** *Suppose that  $a_0(X), b_0(X), h_0(X) \in F[X]$  and  $a_1, b_1, h_1 \in F[X]$  are two triples of polynomials so that  $\deg h_i < m$ ,  $\deg b_i \leq m$  and*

$$h_i(X) = a_i(X)X^{2m} + b_i(X)g(X).$$

*Suppose also that  $(b_0, h_0) = 1$ . Then*

$$a_1 = a_0q, \quad b_1 = b_0q, \quad h_1 = h_0q$$

*for some  $q \in F[X]$ .*

*Proof.* The point is that  $h_i \equiv b_i g$  modulo  $X^{2m}$ . This implies that

$$h_0 b_1 \equiv b_0 b_1 g \equiv h_1 b_0.$$

But  $h_i b_j$  has degree less than  $2m$ . So, we must have

$$h_0(X)b_1(X) = h_1(X)b_0(X).$$

Since  $b_0, h_0$  are relatively prime,  $b_0|b_1$  and  $h_0|h_1$  with the same quotient  $q$ . This also implies that  $a_1 = qa_0$ .  $\square$

**Lemma 4.11.** *If the coefficients of  $g(X)$  satisfy a homogeneous linear recurrence of degree  $\leq m$  then, the polynomial  $b(X)$  obtained by the division algorithm has nonzero constant term.*

*Proof.* I will go over the same example that I did in class, then point out how the argument can be generalized. To make the general argument rigorous, we need to work a little harder.

Suppose that the minimal polynomial is  $p(X) = X - 1$ . Then the recurrence relation is  $d_n = d_{n-1}$  which means that all of the coefficients of  $g(X)$  are equal. The reversal of  $p(X)$  is the same:  $b(X) = X - 1$ . Suppose that the algorithm gives a different polynomial:  $b(X) = X(X - 1) = X^2 - X$ . Since  $(a, b) = 1$ , we know that  $a(0) \neq 0$ . But then, the equation

$$h(X) = a(X)X^{2m} + b(X)g(X)$$

tells us that the coefficients of  $X^{2m-1}, X^{2m-2}$  in  $g(X)$  are not equal, giving a contradiction:

$$\begin{aligned} a(X)X^{2m} + b(X)g(X) &= (X^2 - X)(d_{2m-1}X^{2m-1} + d_{2m-2}X^{2m-2} + \dots) \\ &= (\text{terms of deg } \geq 2m+1) + (a(0) + d_{2m-2} - d_{2m-1})X^{2m} + (\text{lower terms}) \end{aligned}$$

Since  $\deg h(X) < m$  we get

$$d_{2m-2} - d_{2m-1} = -a(0) \neq 0$$

which is a contradiction.

In general what happens is that, if  $b(0) = 0$ , we must have  $a(0) \neq 0$  and  $h(0) = 0$ . If we divide by  $X$  we then get:

$$\frac{h(X)}{X} = a(X)X^{2m-1} + \frac{b(X)}{X}g(X)$$

which implies (by induction on  $m$ ) that the coefficients of  $g(X)$  up to degree  $2m-2$  satisfy a uniquely determined recurrence of order  $\leq m-1$  and furthermore  $a(0) \neq 0$  implies that the coefficient of  $X^{2m-1}$  does not satisfy this recurrence. Since the recurrence is unique, the coefficients of  $g(X)$  do not satisfy any recurrence of degree  $< m$ . The argument below implies that  $g(X)$  will not satisfy a recurrence of degree exactly equal to  $m$ . So we get a contradiction.

Claim: Under the conditions above,  $g(X)$  does not satisfy any recurrence of degree exactly equal to  $m$ , i.e., with minimal polynomial  $p(X)$  of degree  $m$ .

Suppose it did. Then, letting  $b_0(X)$  be the reverse polynomial

$$b_0(X) = \frac{X^m p(1/X)}{p(0)}.$$

(From the definition of the minimal recurrence polynomial we get  $p(0) \neq 0$ . Also the equation for  $b_0$  implies  $b_0(0) \neq 0$ .) Then  $b_0(X)g(X)$  has no terms of degree  $m, m+1, \dots, 2m-1$ . So, there is a polynomial  $a_0(X)$  so that

$$h_0(X) = a_0(X)X^{2m} + b_0(X)g(X)$$

has degree  $< m$ . But then  $b_0, h_0$  must be relatively prime (otherwise, we could divide the entire equation by the common factor to obtain a recurrence relation on  $g$  of order less than  $m$ .) Therefore, by the previous lemma (4.10),  $b_0$  must divide  $b$ . But this is a contradiction since  $b$  is  $X$  times a polynomial of degree  $< m$ .  $\square$

Putting these two lemmas together we get the following.

**Theorem 4.12.** *Assuming that the coefficients of  $g(X)$  satisfy a homogeneous linear recurrence of order  $\leq m$ , the division algorithm gives the minimal polynomial  $p_0(X)$  of this recurrence and any other recurrence of degree  $\leq m$  has polynomial a multiple of  $p_0(X)$ .*

4.7.3. *proof of uniqueness.* Let me go back to the beginning. The polynomial  $g(X) = \sum_{s=0}^{2m-1} d_s X^s$  has coefficients

$$d_s = \sum_{i \in A} c_i \alpha_i^s, \quad c_i = \frac{\epsilon_i}{L_i(\alpha_i)}.$$

The scalars  $c_i$  are nonzero for all  $i \in A$  by definition of the error set  $A$ . Therefore, the numbers  $d_s$  satisfy a recurrence with polynomial

$$p(X) = \prod_{i \in A} (X - \alpha_i).$$

The algorithm gives the minimal polynomial  $p_0(X)$ . So, we still need to show that  $p(X)$  is minimal. In class I just said it is because the number  $\epsilon_i$  are nonzero. But here is a more detailed proof.

Suppose that  $p(X)$  is not the minimal polynomial. Then Theorem 4.12 (in fact Lemma 4.10) implies that  $p_0(X)$  divides  $p(X)$ . Thus

$$p_0(X) = \prod_{i \in A'} (X - \alpha_i)$$

for some proper subset  $A' \subset A$ . But this implies that

$$d_s = \sum_{i \in A'} a_i \alpha_i^s.$$

Subtracting these (letting  $a_i = 0$  for  $i \in A - A'$ ) we get

$$\sum_{i \in A} (c_i - a_i) \alpha_i^s = 0$$

for  $0 \leq s < 2m$ . But this is impossible. This sum represents a linear combination of  $|A| \leq m$  columns of the  $2m \times 2m$  Vandermonde matrix

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_{2m} \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 & \cdots & \alpha_{2m}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{2m-1} & \alpha_2^{2m-1} & \alpha_3^{2m-1} & \cdots & \alpha_{2m}^{2m-1} \end{pmatrix}$$

Since  $c_i - a_i \neq 0$  for  $i \in A - A'$  we have a nontrivial linear relation among the columns of this matrix which is impossible since the Vandermonde determinant is nonzero.

## 5. FINITE FIELDS

On the last day I asked several questions about  $\mathbb{F}_{2^8}$  and we tried to answer them.

**5.1. Galois group.** The first question was: What is the Galois group of  $\mathbb{F}_{2^8}/\mathbb{F}_2$ ? Since  $\mathbb{F}_{2^8}$  is a vector space of dimension 8 over  $\mathbb{F}_2$ , its degree is 8. So,  $Gal(\mathbb{F}_{2^8}/\mathbb{F}_2)$  has 8 elements. Which group is it?

To answer this we looked for intermediate fields. If  $\mathbb{F}_q$  is contained in  $\mathbb{F}_{2^8}$  then  $2^8 = q^n$  for some  $n \geq 1$ . So,  $q = 1, 2, 2^2, 2^4, 2^8$ . So, there are only two intermediate fields:  $\mathbb{F}_4, \mathbb{F}_{16}$ . This means the Galois group has exactly 2 nontrivial proper subgroups. So, it must be cyclic.

$$Gal(\mathbb{F}_{2^8}/\mathbb{F}_2) \cong \mathbb{Z}/8.$$

In fact the Galois group of any finite field is cyclic.

**Theorem 5.1.** *The Galois group of  $\mathbb{F}_{p^n}/\mathbb{F}_p$  is a cyclic group of order  $n$  generated by the Frobenius*

$$\phi(x) = x^p.$$

*Proof.* The Frobenius is a homomorphism  $\phi : K \rightarrow K$  for any field  $K$  of characteristic  $p$ . Its kernel is trivial since  $x^p = 0$  implies  $x = 0$ . Therefore,  $\phi$  is an automorphism for any finite field. So, it is an element of the Galois group. The fixed field of this element is the set of all roots of

$$X^p - X$$

But the  $p$  elements of the prime field  $\mathbb{F}_p$  are roots of this polynomial. So

$$\mathbb{F}_p = \mathbb{F}_p^{\langle \phi \rangle}.$$

By the Galois correspondence, this implies that  $\langle \phi \rangle = Gal(\mathbb{F}_{p^n}/\mathbb{F}_p)$ .  $\square$

**Corollary 5.2.**  *$Gal(\mathbb{F}_{p^{nm}}/\mathbb{F}_{p^n})$  is the cyclic group generated by  $\phi^n$ .*

**5.2. the field  $\mathbb{F}_4$ .** has only 4 elements:  $\mathbb{F}_4 = \{0, 1, \alpha, \alpha + 1\}$ .

**Proposition 5.3.** *The irreducible polynomial of  $\alpha$  is*

$$X^2 + X + 1.$$

*Proof.* 0, 1 are not roots of this polynomial. So, the polynomial is irreducible and defines a degree 2 extension of  $\mathbb{F}_2$ .

In class I gave another proof. There are only 4 polynomials of degree 2: the one above and

$$X^2 + X, \quad X^2 + 1, \quad X^2.$$

But these polynomials are not irreducible:

$$X^2 + X = X(X + 1), \quad X^2 + 1 = (X + 1)^2$$

So,  $X^2 + X + 1$  is the unique degree 2 irreducible polynomial over  $\mathbb{F}_2$ . So, it must be the irreducible polynomial of  $\alpha$ .  $\square$

If we were to represent elements of  $\mathbb{F}_4$  in binary notation we would write the elements as: (0, 0), (0, 1), (1, 0), (1, 1) where

$$(0, 0) = 0, \quad (0, 1) = 1, \quad (1, 0) = \alpha, \quad (1, 1) = \alpha + 1.$$

Addition is coordinate-wise and multiplication is given by the irreducible polynomial.

**5.3. the field  $\mathbb{F}_{16}$ .** We know that  $\mathbb{F}_{16} = \mathbb{F}_4(\beta)$  for any element  $\beta$  of  $\mathbb{F}_{16}$  which is not in  $\mathbb{F}_4$ . We want the irreducible polynomial of  $\beta$  over  $\mathbb{F}_2$  since this will let us write elements of  $\mathbb{F}_{16}$  as strings of four 0's and 1's and tell us how to multiply them.

The irreducible polynomial of  $\beta$  over  $\mathbb{F}_4$  is quadratic. One possibility is

$$f(X) = X^2 + X + \alpha.$$

The four elements of  $\mathbb{F}_4$  are not roots of this polynomial. So, it is irreducible. If  $\beta$  is a root then the irreducible polynomial of  $\beta$  over  $\mathbb{F}_2$  is

$$f\bar{f} = (X^2 + X + \alpha)(X^2 + X + \bar{\alpha})$$

where conjugation  $\bar{\alpha}$  is given by the element of the Galois group,  $\phi$ . So,  $\bar{\alpha} = \phi(\alpha) = \alpha^2 = \alpha + 1$  and

$$\begin{aligned} \text{irr}(\beta, \mathbb{F}_2) &= (X^2 + X + \alpha)(X^2 + X + \alpha + 1) = (X^2 + X + \alpha)^2 + (X^2 + X + \alpha) \\ &= X^4 + X^2 + \alpha + 1 + X^2 + X + \alpha = X^4 + X + 1. \end{aligned}$$

This polynomial has 4 roots. Since  $\mathbb{F}_{16}$  has  $16 - 4 = 12$  generators and each degree 4 irreducible polynomial has 4 roots, there must be two more irreducible polynomials.

**Theorem 5.4.** *There are exactly three irreducible degree 4 polynomials over  $\mathbb{F}_2$ :*

- (1)  $f_1(X) = X^4 + X + 1$
- (2)  $f_2(X) = X^4 + X^3 + 1$
- (3)  $f_3(X) = X^4 + X^3 + X^2 + X + 1$ .

*Proof.* There are only 8 polynomial of degree 4 with nonzero constant term. The other 5 are reducible since:

$$X^4 + 1, X^4 + X^2 + X + 1, X^4 + X^3 + X + 1, X^4 + X^3 + X^2 + 1$$

have an even number of terms and thus have  $X = 1$  as a root and

$$X^4 + X^2 + 1 = (X^2 + X + 1)^2.$$

I also pointed out that  $f_2$  is the reverse of  $f_1$  and is thus irreducible (the roots of  $f_2$  are the inverses of the roots of  $f_1$ ). The roots of  $f_3$  are 5th roots of unity, so they are not elements of  $\mathbb{F}_4$ .  $\square$

Here is an example of how multiplication is done in  $\mathbb{F}_{16}$  using the irreducible polynomial  $X^4 + X + 1$ .

$$(1101)(0101) = 1101 + 11, 0100 = 11, 1001 = 1, 1111 = 1100$$

where the last two reductions use the irreducible polynomial which is 1,0011:

$$11, 1001 = 11, 1001 + 10, 0110 = 1, 1111 = 1, 1111 + 1, 0011 = 1100.$$

The commas are just to make it easier to read the numbers.

5.4. **the field  $\mathbb{F}_{256}$ .** We didn't get very far with this. The usual irreducible polynomial is

$$g(X) = X^8 + X^7 + X^2 + X + 1 = 1, 1000, 0111$$

5.4.1. *number of irreducible polynomials.* The number of irreducible polynomials of degree 8 is

$$\frac{256 - 16}{8} = \frac{240}{8} = 30.$$

The number of polynomials of degree 8 with nonzero constant term is  $2^7 = 128$ . Half of these have an even number of terms making  $X = 1$  a root. This leaves 64. There are  $2^5 = 32$  polynomials which have  $\alpha$  as a root and half of them have an odd number of terms. This leaves  $64 - 16 = 48$ . We can multiply any two of the three irreducible degree 4 polynomials. There are 6 ways to do that. This leaves 42 left. There are two irreducible polynomials of degree 3, namely,

$$X^3 + X + 1, \quad X^3 + X^2 + 1.$$

And there are 6 irreducible polynomials of degree 5. (Three are 8 polynomials of degree 5 with nonzero constant term and an odd number of terms. Two of them factor as  $X^2 + X + 1$  times one of the two degree 3 irreducibles.) This makes  $2 \cdot 6 = 12$  products leaving  $42 - 12 = 30$  irreducible polynomials of degree 8.

5.4.2. *irreducible polynomial over intermediate fields.* Let  $\gamma$  be a root of the polynomial  $g(X)$ . Then what is the irreducible polynomial of  $\gamma$  over  $\mathbb{F}_4$ ? over  $\mathbb{F}_{16}$ ?

Since  $\phi^4$  generates the Galois group of  $\mathbb{F}_{2^8}/\mathbb{F}_{16}$ , the polynomial of  $\gamma$  over  $\mathbb{F}_{16}$  is

$$(X - \gamma)(X - \phi^4(\gamma)) = (X - \gamma)(X - \gamma^{16}) = X^2 + (\gamma + \gamma^{16})X + \gamma^{17}.$$

Using a computer, I calculated this in binary notation:

$$\gamma^{17} = 1101, 1110, \quad \gamma^{16} = 0110, 1111.$$

So,

$$\text{irr}(\gamma, \mathbb{F}_{16}) = X^2 + 0110, 1101X + 1101, 1110.$$

Some further computer calculations show that

$$\alpha = 1010, 1010, \quad \beta = 1101, 1110$$

satisfy the equations:

$$\alpha^2 + \alpha = 1, \quad \beta^2 + \beta = \alpha, \quad \beta^4 + \beta = 1.$$

So, we can identify them with the generators of  $\mathbb{F}_4$  and  $\mathbb{F}_{16}$  that we chose earlier. In this notation we have:

$$\text{irr}(\gamma, \mathbb{F}_{16}) = X^2 + (\alpha + 1)(\beta + 1)X + \beta.$$

Applying the generator  $\phi^2$  of  $\text{Gal}(\mathbb{F}_{16}/\mathbb{F}_4)$  we get

$$\text{irr}(\gamma^4, \mathbb{F}_{16}) = X^2 + (\alpha + 1)\beta X + \beta + 1$$

The product of these is

$$\text{irr}(\gamma, \mathbb{F}_4) = X^4 + \bar{\alpha}X^3 + \alpha X^2 + \bar{\alpha}X + \alpha.$$

If we multiply this with the conjugate:

$$\text{irr}(\gamma^2, \mathbb{F}_4) = X^4 + \alpha X^3 + \bar{\alpha}X^2 + \alpha X + \bar{\alpha}$$

we get back the original irreducible polynomial

$$\text{irr}(\gamma, \mathbb{F}_2) = X^8 + X^7 + X^2 + X + 1.$$

5.4.3. *the order of  $\gamma$ .* One last point: The calculation

$$\gamma^{17} = \beta$$

implies that  $\gamma$  is a generator of the cyclic group

$$\mathbb{F}_{256}^\times \cong \mathbb{Z}/255 \cong \mathbb{Z}/17 \oplus \mathbb{Z}/5 \oplus \mathbb{Z}/3$$

since  $\beta$  has order 15. (The elements of order 3 lie in  $\mathbb{F}_4$  and the elements of order 5 are roots of the irreducible polynomial  $f_3(X)$ .  $\beta$  is a root of  $f_1(X)$ .)