CS681 **Computational Number Theory**

Lecture 7: Towards Factorization over Finite Fields

Instructor: Piyush P Kurur *Scribe: Ramprasad Saptharishi*

Overview

We shall slowly move into factorization of univariate polynomials (polynomials with just one variable) over finite fields. We are given a finite field K and a polynomial over one variable X whose coefficients are from K. We are to find the factorization of this polynomial into irreducible factors.

Before we get into this question, we need to first understand if it even makes sense. How can we be sure that such a factorization exists? And even if it did, how do we know if it is unique?

We shall first answer a lot of questions in the algebra related to it before going to factorization as such.

1 Rings, Ideals, Factorization etc.

We know that integers can be uniquely factorized into product of prime powers. However, not all rings are as well-behaved as the integers are. We first need to ask if the algebraic structure has this property of unique factorization. Let us look at an example where this fails.

Look at the set of integers modulo 8. This is called $\mathbb{Z}_8$ and we know that this forms a ring. Suppose we look at polynomials over this ring, polynomials of a single variable X whose coefficients come from $\mathbb{Z}_8$, does this ring have the property of unique factorization? Here is a counter example in $\mathbb{Z}_8[X]$.

$$X^2 - 1 = (X-1)(X+1) = (X-3)(X+3)$$

But $\mathbb{Z}_8$ is a bad ring, in the sense that non-zero elements can multiply to give 0 ($2 \times 4 = 0$ here). As for another example, look at the set of all number of the form $a + b\sqrt{-5}$ where $a, b \in \mathbb{Z}$. This forms a ring and over this ring $6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$.

Hence it's not always true that factorization is unique. However, fortunately for us, we have unique factorization over $K[X]$ whenever K is a field.

Definition 1. *A ring is said to be an integral domain if and only if there are no non-trivial zero divisors. That is, if $a, b \in R$ such that $ab = R$, then either $a = 0$ or $b = 0$.*

The ring $\mathbb{Z}$ is an integral domain but the ring of integers modulo 6 is not (since $2 \times 3 = 0$ over the ring).

Inorder to define factorization, we need a notion of primes over arbitrary rings. Let us first look at the definition of primes over integers. Let us first look at the wrong definition.

A number p is said to be prime if for all a that divides p, either $a = 1$ or $a = p$.

This translates to the ring definition of a maximal ideal and not a prime ideal.

This is the common definition in school but generalization based on this is erraneous. Though it happens to correct over the set of integers, it is not true in general. Here is the right definition.

Definition 2. *A number p is said to be a prime if and only if for all a, b such that p divides ab, either p divides a or p divides b.*

Thus this gives the definition of prime ideals in the setting of rings.

Definition 3. *An ideal $\mathfrak{a} \subseteq R$ is said to be prime if and only if for all $a, b \in R$ such that $ab \in \mathfrak{a}$, either $a \in \mathfrak{a}$ or $b \in \mathfrak{a}$.*

Any element $p \in R$ that generates a prime ideal is called a prime element of R.

Definition 4. *An ideal $\mathfrak{a} \subseteq R$ is said to be maximal if and only if for every ideal $\mathfrak{a}' \supseteq \mathfrak{a}$, either $\mathfrak{a}' = 1R = R$ or $\mathfrak{a}' = \mathfrak{a}$.*

This basically means that no proper ideal of R properly contains $\mathfrak{a}$. Note that not all prime ideals are maximal. We were just lucky that this was true on $\mathbb{Z}$ and hence both definitions of prime numbers were equivalent. This is not true over arbitrary rings.

Definition 5. *An ideal $\mathfrak{a} \subseteq R$ is said to be a principle ideal if the ideal is generated by a single element. That is, $\mathfrak{a} = aR$ for some $a \in R$.*

Definition 6. *An integral domain R is said to be a*

- *principle ideal domain (PID) if every ideal in it is principle (every ideal is generated by a single element).*

- *unique factorization domain (UFD) if every element can be uniquely factorized in to product of prime elements of the ring.*

We already saw an example of a ring (and a domain) that was not a UFD. Here is an example of a ring that is not a PID. Consider a field K and look at the ring of polynomials on two variables X, Y over this field. This is denoted by $K[X, Y]$.

In this field, look at the ideal generated by X and Y. That is, the set of polynomials of the form $Xf(X, Y) + Yg(X, Y)$, those polynomials that do not have a constant term. This is clearly an ideal but this isn't principle.

A similar example is over $\mathbb{Z}[X]$ and the ideal being (p, X) where p is any prime number.

Fact 1. *For any field K, $K[X]$ is a PID.*

Fact 2. *Any PID is also a UFD*

The two facts together tell us that we can indeed talk of factorization of polynomials in $K[X]$. Another useful fact is the following, and this helps us see that factorization makes sense even on multivariate polynomials.

Fact 3. *If R is a UFD, so is $R[X]$.*

The following theorems are very useful.

Theorem 1. *A ring R is a field if and only if the only ideals of R are the 0 ideal and the whole ring R.*

Proof. First we shall show that a field has no non-trivial ideals. Suppose The field had some ideal I that contained some element $x \neq 0$. Since it is a field, the inverse of x exists. Since I is an ideal and $x \in I$ would mean that $xa \in I$ for all $a \in R$ and in particular $xx^{-1} = 1 \in I$. But if $1 \in I$, then for every element a in the field, $1a \in I$ which would then force I to be the entire field.

As for the other direction, suppose the ring R was not a field. We want to show that there exists some non-trivial ideal in this ring. Since we assumed that it isn't a field, there must be some non-zero element a whose inverse does not exist. Look at the ideal generated by it, aR. This ideal certainly contains a and it cannot 1 since if it did, it would mean that a is invertible. And hence this is an ideal that is non-zero and also not the whole of R; a non-trivial ideal. □

Theorem 2. *For any ring R*

1. *if an ideal $\mathfrak{a}$ is prime, then $R/\mathfrak{a}$ is an integral domain.*
2. *if an ideal $\mathfrak{a}$ is maximal, then $R/\mathfrak{a}$ is a field.*

Proof. We have to show that if $\mathfrak{a}$ is prime, then $R/\mathfrak{a}$ is an integral domain. Suppose not, then there exists two non-zero elements a, b such that $ab = 0$ in $R/\mathfrak{a}$. This means that $a \bmod \mathfrak{a} \neq 0$ and $a \bmod \mathfrak{a} \neq 0$ but $ab \bmod \mathfrak{a} = 0$ or in other words $ab \in \mathfrak{a}$ but neither a nor b belongs to $\mathfrak{a}$. This contradicts the assumption that $\mathfrak{a}$ and hence R has to be an integral domain.

As for the case when $\mathfrak{a}$ is maximal, assume that $R/\mathfrak{a}$ is not a field. Then there exists some non-zero element that is not invertible. Look at the ideal generated by this element. As in the earlier theorem, this is a non-trivial ideal (neither 0 nor the entire ring). But in the map from R to $R/\mathfrak{a}$, ideals of $R/\mathfrak{a}$ corresponds to ideals in R that contain $\mathfrak{a}$. Since we just found a non-trivial ideal in $R/\mathfrak{a}$, this would translate to a non-trivial ideal in R that properly contains $\mathfrak{a}$ thus contradicting the maximality of $\mathfrak{a}$. Thus R has to be a field. □

1.1 Some Insights

This is not completely a part of the course but it would be useful to know this to understand factorization. In any ring, we can talk of a tower of prime ideals. What this means is a series of the form $0 \subseteq I_1 \subseteq I_2 \subseteq \cdots \subseteq I_n \subseteq R$ such that each ideal I_j is a prime ideal. The number n is called the Krull Dimension of the ring R.

The Krull Dimension is actually a local property but for it is well defined for rings like $K[X_1, X_2, \cdots, X_n]$ (where K is a field) and $\mathbb{Z}[X]$.

If we were to look at $K[X, Y]$, we have the tower $0 \subseteq (X) \subseteq (X, Y) \subseteq K[X, Y]$. The krull dimension of this ring is 2. Similarly the ring of polynomials on n variables over a field K will have a krull dimension of n.

And the ring $\mathbb{Z}[X]$ has the tower $0 \leq (p) \leq (p, X) \leq \mathbb{Z}[X]$ and hence has krull dimension 2. We shall see soon that factorization of polynomials in $\mathbb{Z}[X]$ is so similar to factorization of polynomials in $K[X, Y]$.

We need to understand the concept of finite fields, extensions, etc before we get into factorization. We shall first spend some time on this.

2 Finite Fields

We shall be studying properties of fields that have finite number of elements in them. A few things to keep in mind, we shall prove them soon, is that any finite field has its cardinality to be a power of prime. There cannot exist a finite field whose cardinality is divisible by two distinct primes. And infact, for any prime p and α, there is exactly one field of size p^α. (and note that this isn't true on the infinite setting. $\mathbb{R}$ and $\mathbb{C}$ both have infinite number of elements but are clearly different)

Definition 7. *A field E is called an extension of a field K if E is a field that contains K. This (also) is denoted by E/K.*

There is a notion of a degree of a field extension but one needs to be familiar with vector spaces to completely understand this. We shall dwell a little on it.

2.1 Vector Spaces

Definition 8. *A vector space V over a field K, with an additive structure and multiplication by elements of K (scalars), satisfies the following conditions:*

- *$(V, +)$ is an additive abelian (commutative) group (additive closure, inverse, identity)*
- *For any vector $v \in V$ and scalar $\alpha \in K$, the element αv is also a vector.*
- *For any vectors $u, v \in V$ and scalar $\alpha \in K$, we have $\alpha(u + v) = \alpha u + \alpha v$.*
- *For any vector u and scalars $\alpha, \beta \in K$, we have $(\alpha + \beta)u = \alpha u + \beta u$ and $\alpha(\beta u) = (\alpha\beta)u$.*

Let us look at a few examples to get ourself familiar with this notion. $\mathbb{C}$ forms a vector space over $\mathbb{R}$. Clearly the above properties are satisfied.

Another example is the plane $\mathbb{R}^2$, set of point (x, y) where both coordinates are from the reals. Scalar multiplication is defined as $\alpha(x, y) = (\alpha x, \alpha y)$.

Another example is the ring of polynomials $K[X]$ over K where K is a field. Scalar multiplication is just multiplying every coefficient by the scalar.

Next we need a notion of linear independance.

Definition 9. *A set $\{v_1, v_2, \cdots, v_k\}$ is said to be linearly independant if the only way*

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0$$

can happen for scalars c_i is when all the c_i's are zero themselves. That is, no non-trival linearl combination of these vectors is zero.

For example, let us look at each of our examples stated and find a linearly independant set. In $\mathbb{C}$ over $\mathbb{R}$, look at the set $\{3, 2+i\}$. Suppose $c_1(3) + c_2(2+i) = 0$, then $(3c_1 + 2c_2) + c_2 i = 0$ and this is possible only when both c_1 and c_2 are zero. Hence the set is linearly independant.

And again, look at the set $\{(1,0),(0,1)\}$ in $\mathbb{R}^2$. This again is linearly independant since the only way $c_1(1,0)+c_2(0,1) = (c_1, c_2) = (0,0)$ is when both c_1 and c_2 are zero.

In the third example, look at the set $\left\{1, X, X^2\right\}$. $c_1 + c_2X + c_3X^2$ can be the zero polynomial if and only if all the c_i's are zero.

This is the notion of linear independance. With a little bit of thought, any vector that can be represented as a linear sum from such a set is infact uniquely represented so.

For example, let us assume that $\{v_1, v_2, \cdots, v_k\}$ was a linearly independant set. Let $v = c_1v_1 + c_2v_2 + \cdots + c_kv_k$. Suppose this could be represented as a linear sum in a different way, we shall obtain a contradiction.

$$\begin{aligned} v &= c_1v_1 + c_2v_2 + \cdots + c_kv_k \\ &= c_1'v_1 + c_2'v_2 + \cdots + c_k'v_k \\ \implies 0 &= (c_1 - c_1')v_1 + \cdots + (c_k - c_k')v_k \end{aligned}$$

And if the two representations were indeed different, there is atleast one i such that $c_i \neq c_i' \implies (c_i - c_i') \neq 0$ but this would give a non-trivial linaer combination of the v_i's to become zero. This contradicts our assumption that they were linearly independant. Hence such linear representations are unique.

An example is that every point (x, y) can be represented uniquely as a linear sum of $(1,0)$ and $(0,1)$ (it is just $x(1,0) + y(0,1)$). The students are encouraged to also check it for $\mathbb{C}$ with our linearly independant set being $\{3, 2+i\}$.

Let us look at our example of $K[X]$. We saw that $\left\{1, X, X^2\right\}$ was a linearly independant subset but the term X^5 can never be written as a linear sum of $1, X, X^2$. Thus, the set $\left\{1, X, X^2\right\}$ doesn't cover or *span* X^5. Since X^5

is not spanned by the set $\{1, X, X^2\}$, we can add it to the set and it would still be linearly independant.

We can keep adding elements to our linearly independant set in this way by picking up some vector that is not spanned by it and adding it. This process can go on indefinitely as well. For the moment let us look at the case where this process stops after finite number of steps. Now we have a set that is linearly independant and it also spans the entire space.

An example would be to look at $\mathbb{C}$ over $\mathbb{R}$. Start with 3. The linear span of this is just elements of the form $3c$ where c is a real number. Hence it does not span elements like $2+i$. Hence we can add $2+i$ to this set and still have a linearly independant set. Now this set $\{3, 2+i\}$ is linearly independant and also spans the entire space. Any complex number $a + ib$ is equal to $b(2+i) + \frac{a-2b}{3}3$.

Such a set that spans the space and also is linearly independant is called a *basis* of the vector space V over K. And every vector in the vector space can be expressed as a linear combination of the basis elements, and uniquely so.

The number of basis elements is called the dimension of the vector space. But wait, how do we know that every basis will have the same number of elements? Is it possible that I can find three complex numbers that are linearly independant over $\mathbb{R}$ and span $\mathbb{C}$? The answer is no. It is not so hard to see that all basis must have the same number of elements. Thus the dimension of the vector space is independant of the choice of basis is hence well-defined.

The vector space $K[X]$ over K has infinite dimension and its basis could be chosen as $\{1, X, X^2, \cdots\}$. And a polynomial, say $80 + 2X + 0X^3 + 3X^4$ can be represented by the tuple $(80, 2, 0, 3, 0, 0, 0, \cdots)$ and every polynomial has a corresponding tuple.

Suppose we choose $\{1, i\}$ as a basis for $\mathbb{C}$ over $\mathbb{R}$, then every element $a + bi$ can be expressed as $a(1) + b(i)$. Now we can represent the number $a + bi$ as the tuple (a, b).

So essentially, a vector space V over K is one where each element in it can be represented as a tuple, whose entries come from K. The arity of the tuple is the dimension of the vector space.

Thus, in the finite setting, if V is a finite dimensional (say d-dimensional) vector space over a finite field K, then the number of elements of V is $|K|^d$. This is clear since it is just the number of d-tuples whose entries come from K.

2.2 Field Extensions

Let E/K be a field extension. This just means that both E and K are fields and that E contains K.

Now observe that for all $\alpha, \beta \in K$ and $u, v \in E$, $\alpha(u+v) = \alpha u + \alpha v$ and $(\alpha+\beta)u = \alpha u + \beta u$ etc. Thus all the conditions to call this a vector space hold. Thus, we can think of E as a vector space over K.

An example of this we have seen already. $\mathbb{C}$ is a field that contains $\mathbb{R}$. And $\mathbb{C}$ actually is a vector space over $\mathbb{R}$. Another example would be to look at

$$\mathbb{Q}[\sqrt{2}] = \left\{ a + b\sqrt{2} \ : \ a, b \in \mathbb{Q} \right\}$$

It is easy to check that this is a field and this clearly contains $\mathbb{Q}$. And this also naturally forms a vector space over $\mathbb{Q}$.

Definition 10. *The dimension of E as a vector space over K is called the degree of the extension E/K. This is denoted by $[E : K]$.*

$\mathbb{C}$ over $\mathbb{R}$ is a 2-dimensional extension. $\mathbb{R}$ over $\mathbb{Q}$ is an infinite dimensional extension. $\mathbb{Q}[\sqrt{2}]$ over $\mathbb{Q}$ is a 2 dimensional extension.

Adjoining Elements: An informal discussion

The field $\mathbb{C}$ is just taking $\mathbb{R}$ and adding the element i to it. Once we add i to $\mathbb{R}$, we just take all possible linear combinations, products, inverses to make it a field. We let the set $\mathbb{R} \cup i$ grow into the smallest field containing $\mathbb{R}$ and i. This is formally referred to as $\mathbb{R}(i)$, the field got by adjoining i to $\mathbb{R}$.

It is easy to check that $\mathbb{Q}(\sqrt{2})$ is infact $\mathbb{Q}[\sqrt{2}] = \left\{ a + b\sqrt{2} \ : \ a, b \in Q \right\}$. And similarly one can also check that $\mathbb{Q}(\sqrt[3]{2}) = \left\{ a + b\sqrt[3]{2} + c\sqrt[3]{2^2} \ : \ a, b, c \in \mathbb{Q} \right\}$. From this is it easily seen that $\mathbb{Q}(\sqrt[3]{2})$ is a degree 3 extension over $\mathbb{Q}$.

Given such an adjointed field extension, is it easy to find out the degree? The answer is yes. All we need to do is choose an easy basis for the vector space. For example, let us look again at $\mathbb{Q}(\sqrt[3]{2})$. Let $\alpha = \sqrt[3]{2}$. We want the degree of the extension $\mathbb{Q}(\alpha)/\mathbb{Q}$. Now consider the set $\left\{1, \alpha, \alpha^2, \alpha^3, \cdots ...\right\}$. When does this fail to be a linearly independant subset? We know that $\alpha^3 - 2 = 0$ and hence it loses its linear independance after α^3. This is because α was a root of $X^3 - 2$, a degree 3 polynomial over the $\mathbb{Q}$.

Instead if we were to look any α, any equation of linear dependance would look like $a_0 + a_1\alpha + a_2\alpha^2 + \cdots a_k\alpha^k = 0$ and this would just mean that α is a root of the polynomial $a_0 + a_1X + a_2X^2 + \cdots a_kX^k = 0$. Thus,

the degree of such an extension $\mathbb{Q}(\alpha)/\mathbb{Q}$ is just the degree of the smallest degree polynomial of which α is a root.

$\mathbb{C} = \mathbb{R}(i)$ and i has $X^2 + 1$ as its minimum polynomial and thus $[\mathbb{C} : \mathbb{R}] = 2$. If we were to look at $\mathbb{Q}(\pi)$, π is not a root of any polynomial with coefficients in $\mathbb{Q}$ (this is also referred as 'π is transcendental'). Thus the set $\left\{1, \pi, \pi^2, \cdots\right\}$ would be an infinite linearly independant subset. And hence the extension $\mathbb{Q}(\pi)$ over $\mathbb{Q}$ is of infinite degree.

We shall look at more properties of finite fields and extensions next time.