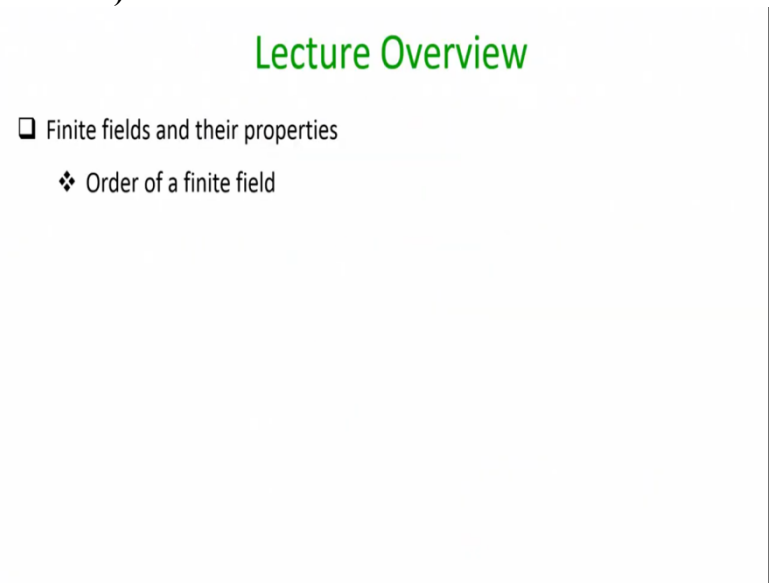
Discrete Mathematics
Prof. Ashish Choudhury
International Institute of Information Technology, Bangalore

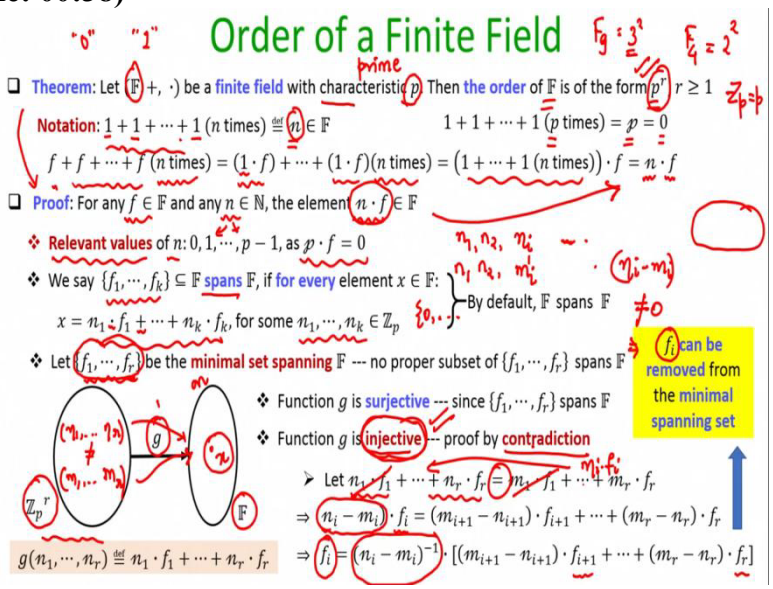
Lecture - 69
Finite Fields and Properties II

(Refer Slide Time: 00:23)



Hello everyone, welcome to this lecture. The plan for this lecture is as follows. In this lecture, we will continue our discussion on finite fields and we will discuss what we call as order of a finite field and what are the properties of an order of a finite field.

(Refer Slide Time: 00:38)



So, basically order of a finite field is the number of elements in your set F if F is the field and we can prove a very strong property; a very strong statement regarding the order of a finite field. So, the statement here is the following: imagine your field F is a finite field and

suppose its characteristic is p. Now, as per the discussion that we had in the last lecture, we know already that this number p is a prime number.

What we can prove actually is that the number of elements in this field is of the form p^r where r greater than and equal to 1. That means you take any finite field, the number of elements in the field will be of the form p^r where p is a prime number, the characteristic of your field. Before going into the proof of this theorem, you can recall easily that the fields that we had discussed in the last lecture.

We saw there a field consisting of 9 polynomials, its cardinality is 3^2, because its characteristic was 3. In the same way, we saw the abstract field consisting of 4 symbols its cardinality is 2^2 because the characteristic of that field was 2, if you take the field $\mathbb{Z}_p$, its cardinality is p^1 because its characteristic is p and so on. So, what we are actually going to prove is that it is not the case that it is happening accidentally only for F_9 F_4 $\mathbb{Z}_p$.

But you take any prime any finite field with characteristic p, the number of elements in the field will be of the form p^r. So, before going into the proof of this theorem, we are going to introduce some notations which we will be using in the proof. So, remember the additive identity of the group is 0 and the multiplicative identity of the group is 1. I will use the notation n with different font here (n) to denote the element, which I will obtain by adding the multiplicative identity n number of times. So, typically in regular integer arithmetic where this element 1 is the numeric 1 if I add 1 to itself n times we get n but, this element 1 need not be the numeric 1 it is the multiplicative identity. If you add this element 1 to itself n number of times and as for the closure property of the field you will obtain an element from the field, that can be any abstract element. But just for the sake of simplicity, I will use this notation n with a different font to represent a result of 1 added to itself and n number of times. In the same way. 1 added to itself p number of times will be represented by this special font p and since the characteristic of the field is p itself this element p with a special font is nothing but the element 0.

That comes from the definition of the characteristic of a field. In the same way if I take any abstract element, an arbitrary element f from the field and add f to itself n number of times, I can obtain the same result by saying that each of these f s can be replaced by the product of 1

and f because 1 is the multiplicative identity and this I am doing n number of times and then I can distribute the + over dot.

And 1 added to itself n number of times as per my notation is n, which is an element from the field and that element multiplied with f will be the overall result. So, now let us go into the proof of this theorem; the proof is very interesting here. So, you consider any element of the field and any multiple n from the set of natural numbers, my claim is that the operation or the result of adding f to itself n number of times, which will give me this element ($n \cdot$f) is also an element of F and that comes from your closure property. So, as per our notation; as per our definition the element n added it to itself f will give me this element ($n \cdot$ f) and as per the closure property this element will be an element of the field itself. Now, what are the relevant values of n; relevant in the sense which will give me nonzero elements.

So, the relevant values are 0 to p - 1, by relevant I mean only those multiples which will give me distinct elements. So, 0 times f will of course, give me the element 0 as per the definition. Now f added to itself, or 1 times f actually, to be more precise is just element f. Now, f + f will give me some element from the field. So, I can call it as two times f; this is not the numeric 2. But some 1 added to itself two number of times that is a representation here and like that if I continue, then I can say that f added to itself (p – 1) number of times will give me (p – 1) ·f and after that if I add f to itself once more I will get the element 0 because p is the characteristic of the field. In that sense the only relevant multiples of f are 0 to p - 1 because after that you take the higher order multiples of the element f you will start getting the same elements which you could have generated by taking the multiples of f in the range 0 to p - 1.

Next let me define what I call as the span of the field. So, a collection of k elements, so, here is your field F which is finite and which has some number of elements. So, if I focus on a collection of values which are called as f_1 f_2 f_i f_k I will call the collection of these elements as the span of the field if the following hold. You take any element x from the field that can be expressed as a linear combination of the elements from your collection f_1 to f_k, where the linear combiners are from set 0 to p – 1. Why I am focusing on the linear combiners which are in the range 0 to p - 1 because as I said here, the relevant multiples of any element from the field are where when you take the multiples to be in the range 0 to p - 1. So, basically

span means that it is actually the subset of those elements from the field in terms of which you can express any element of the field, by taking various linear combinations. And when you say linear combination by that I mean that I am doing the plus operation and the dot operation as per the field. So it is easy to see that a trivial span of the field is the entire field itself. You take any element x from the field that can be always represented as 1 times x + all other elements from the field being multiplied with 0.

So 0 times the first element and so on. That is why the entire field is of course a span of itself. Now let me next define what we call as the minimal spanning set of the field. So the minimal spanning set of the field is the collection of elements from the field which is minimal in the sense that you cannot remove any element from this collection.

If you remove any element from this collection then it is no longer the case that reduced collection still spans the entire field. That means no proper subset of this collection spans the entire field F in that sense it is minimal it is essential collection. And there could be multiple minimal sets spanning your field, it is not the case that it is always unique. It may be possible that a collection of first 3 elements from the field constitutes a minimal set spanning the field or say the last 2 elements from the field they are the essential elements and so on.

Now why I am focusing on value r here because remember our goal is to show that the order of the field is of the form p^r that is why I am taking r here. So I am basically saying that a collection of r elements from your field f will be considered as a minimal set spanning the field if it is the bare minimal collection elements whose presence is required to express every element from your field as a linear combination.

Now, what I am going to define is the following: I am going to now define a mapping g from the $\mathbb{Z}_p{}^r$ to the field F. Now, what is the $\mathbb{Z}_p{}^r$? so as per the definition of Cartesian product, $\mathbb{Z}_p{}^r$ is nothing but the Cartesian product of $\mathbb{Z}_p$ which itself r times. That means if I consider an r tuple present in $\mathbb{Z}_p{}^r$ then by that I mean that I am talking about r elements where each of the elements are from set $\mathbb{Z}_p$.

Now how exactly this mapping g is defined? So if you want to map an r tuple as per the mapping g then what basically you have to do is the following, you have to take a linear

combination of the elements in your minimal spanning set as per the linear combiners in your r tuple. That is the way I have defined my mapping g. Right now I am not making any claim about this mapping g whether it is injective, bijective, surjective.

It is just a function right now, I am just giving you the definition of the function that definition is you give me any r tuple then I will match that r tuple to a finite field where the mapping is obtained or where the image is obtained by taking a linear combination of the elements in the minimal spanning set as per the linear combiners in my r tuple. Now I am going to make certain claims about this function g.

I am going to prove that this function g is a bijection and if it is a bijection then as per the rules of cardinality it shows that the cardinality of $\mathbb{F}$ is same as the cardinality of $\mathbb{Z}_p^r$ and what is the cardinality of $\mathbb{Z}_p^r$? The cardinality of $\mathbb{Z}_p^r$ is nothing but p^r because as I said the definition of $\mathbb{Z}_p^r$ is you take the Cartesian product of $\mathbb{Z}_p$ r times. So there are p^r possible elements or p^r number of r tuples present in the Cartesian product of $\mathbb{Z}_p$ r times.

And assuming that g is a bijection which I am going to show assuming that this statement is true, it shows that the cardinality of $\mathbb{F}$ is same as the cardinality of $\mathbb{Z}_p^r$ and hence it shows that the number of elements in my field $\mathbb{F}$ is some p^r so that is the proof strategy here. Now everything boils down to proving that my mapping g is indeed a bijection and as per the definition of a bijection I have to prove that the mapping g is a surjection and it is an injection.
Well, proving that g is a surjection is trivial. That comes from the definition of your spanning set. Since as per my definition, the collection of f_1 to f_r is a spanning set. That means you give me any element x it will have a pre-image. Why? Because as per the definition of a spanning set this element x can be expressed in terms of these r elements as per a linear combination.

Where the linear combiners will be from $\mathbb{Z}_p$ and how many such linear combiners I will need? I will need r such linear combiners and if each of them is an element of $\mathbb{Z}_p$ basically the collection of the corresponding linear combiners is going to be an r tuple from this $\mathbb{Z}_p^r$. So, that trivially proved that is function g is a surjective function. Now, I want to prove that this function g is also an injective function and that I will prove by contradiction.

So as per the contradiction assume that the mapping g is not injective. That means imagine you have 2 different r tuples so you have an r tuple say n_1, n_2, n_r. So, let me write down this different r tuple here itself because I will need the space. So, imagine you have 2 different r tuple n_1 up to n_r and another r tuple m_1 up to m_r and say both of them gets mapped to the same element x as per the mapping g.

What does that mean? It means that you take the linear combination of the elements of your spanning set as per the combiners n_1 to n_r and if you take the linear combination of the elements of your spanning set as per the linear combiners m_1 to m_r you get the same element same field element that is what it means when I say the mapping g is not injective.

If this is the case, I have to arrive at a contradiction. Basically, I will try to arrive at a contradiction that the collection of r elements, which you assumed to be the minimal spanning set is actually not a minimal spanning set that means there are some unnecessary redundant elements which have been added unnecessarily in this collection which can be simply removed.

I will arrive at that contradiction. How do I arrive at that contradiction? Well what I can; what I know about this r tuples is that they are different. That does not mean that the entire set of the r values in the first r tuple and all the r values in the second r tuple they are different; there might be some of them which are same. So it might be the case that, say, the first r elements in both the r tuples are same.

So you have say n_1 n_1 and n_2 n_2 occuring. But suppose i is the first index where in the first r tuple you have the value n_i and in the second r tuple you have the value m_i where n_i and m_i are different but the first i - 1 components in both r tuples suppose they are the same. So, I am focusing on the first index i where the r tuple n and r tuple m they are different; that index i could be any index in the range 1 to r.

And there definitely is one such index i because as per my assumption the entire n tuple and the entire m tuple they are different. So, if they are completely different definitely there must be some component, some index i where the component in the n tuple component and the component in the m tuple, they are different. I do not know what exactly is that index, but

that index i definitely exists. So, I am focusing on that index i and assuming that the first i - 1 components they are same in both the n tuple and m tuple.

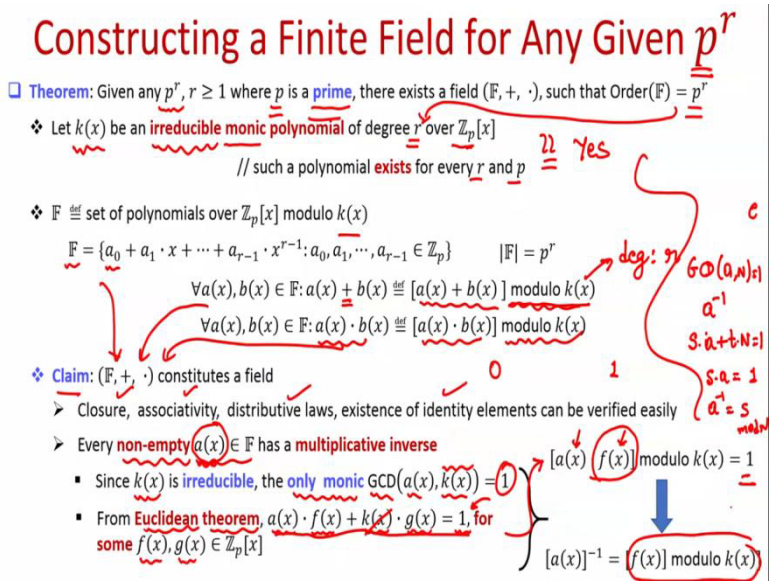
Then I can cancel them out both from the LHS and RHS because if I have n_1 times f_1 occurring in the LHS and imagine m_1 is same as n_1 then I can cancel out n_1 f_1 from both sides. In the same way if n_2 is same as m_2 I can cancel out into n_2 times f_2 both from LHS and RHS and so on. But then when I come to the ith term what I have done here is I take the term m_i times f_i here to the LHS here.

And whatever is the remaining part of the expression in the LHS part I took it and bring it into RHS. I have simply arranged the terms here. Now, if this is the case, if I get this equation what I can say is the following: if I multiply both sides of the equation by the multiplicative inverse of this element and the multiplicative inverse of the element $n_i - m_i$ exist because as per the definition and n_i is not equal m_i, that means $n_i - m_i$ is not 0.

And if it is not 0 then as per the definition of a field, I do have a multiplicative inverse of this element that means, I do have an element which I can denote by this notation $(n_i - m_i)^{-1}$, which when multiplied with the difference of n_i and m_i will give me the multiplicative identity namely 1. So, if I multiply with the multiplicative inverse on both the sides, I basically get the fact that f_i can be expressed in terms of f_{i+1}, f_{i+2} ... f_r.

That means, it shows that I can remove f_i safely from my supposedly minimal spanning set of f, it is not necessary to keep f_i in this collection, because I can express f_i in terms of the remaining elements in this spanning set which spans the finite field F and that goes against my assumption that this collection of r elements f_1 to f_r was the minimal spanning set of the finite field and why I came to this contradiction. Because I assumed that my function g is not an injective mapping. So, that means, whatever I assumed about g is incorrect and that shows that indeed my mapping g is an injective mapping and that shows that my function g is a bijection that means the cardinality of $\mathbb{Z}_p^r$ and the finite field are same and that proves this theorem.

(Refer Slide Time: 22:30)



So, we have proved that you give me any finite field with characteristic p it will have p^r number of elements. Now, let us see how exactly we can construct finite fields for any given p^r where p is a prime number and this is very interesting because it says the following you give me any prime number p, I will show the existence of a finite field whose characteristic will be that prime number p.

And the number of elements in the field will be p^r and how exactly we construct such a field. So, for constructing such a field we will take the help of some irreducible monic polynomial where the coefficients are over $\mathbb{Z}_p$ and the degree of the polynomial will be r. Why r? Because r is also given as part of your input. So you are given a prime number p and value r, my goal is to show the existence of a finite field with characteristic p and with p^r number of elements.

So, to do that I am basically taking a monic irreducible polynomial with coefficients over $\mathbb{Z}_p$ whose degree is r, if you are wondering whether indeed such polynomials always exist for any given r and p, the answer is yes. Such polynomial always exists for every r and p and there are some standard methods for doing that; getting such polynomials but for some well known values of p and r such polynomials are publicly available.

Now, my goal is to construct a field F, so, my set F will be the set of all polynomials with coefficients over $\mathbb{Z}_p$ modulo k(x). In other words, basically the set F is the collection of all polynomials of degree 0, degree 1, degree 2, degree 3 and up to degree r - 1 where the coefficients of the polynomial are from $\mathbb{Z}_p$. So in general, I can say that F is the collection of all polynomials of degree at most r - 1.

So this means degree is at most r - 1 where the coefficients are allowed from the set $\mathbb{Z}_p$, why I am saying it is at most r - 1, because since each of the coefficients are from the set $\mathbb{Z}_p$ and my $\mathbb{Z}_p$ have the elements from 0 to p – 1 that means I can have a polynomial where all the coefficients are 0 that means I can also have a polynomial which is the 0 polynomial. So it is not necessarily the case that a_{r-1}, namely the coefficient of the r - 1 th power of x is always supposed to be there, it can also be 0,

So, it turns out that how many elements I can have; how many such polynomials I can have in my collection $\mathbb{F}$. Since I can have each of the coefficients taken from the set $\mathbb{Z}_p$. Namely each of the coefficients can take p possible values and each of them are picked independently that means it is not the case that the coefficient a_1 depends on the coefficient is a_0, it is not the case that the coefficient a_2 depends on the coefficient is a_0 and a_1 they are picked independently.

So, I can say that from the product rule of counting there are p^r number of possible polynomials in my collection $\mathbb{F}$. So I have defined my collection $\mathbb{F}$. Now I have to give the definition of the abstract plus operation and abstract dot operation. So, my plus operation here is defined to be the addition of polynomials where the coefficients are added as per $\mathbb{Z}_p$ namely addition modulo p and then I take the resultant polynomial modulo the irreducible polynomial.

So that will ensure that my resultant polynomial will have coefficients over $\mathbb{Z}_p$ and its degree will be at most r - 1, because the degree of k(x) is r. To begin with my a(x) and b(x) polynomials both those polynomials will have degree r - 1 and if I add any 2 polynomials of degree r - 1, at most I will still obtain a polynomial of degree at most r - 1.

So in fact, I do not need to take a modulo k(x), because in the sense, the effect of modulo k(x) would not take place. And my multiplication operation is defined to be the product of 2 polynomials, the corresponding 2 polynomials, where the coefficients are multiplied with respect to $\mathbb{Z}_p$ and if the degree becomes more than r, I take modulo k(x). That is my definition of the abstract plus operation and abstract dot operation.

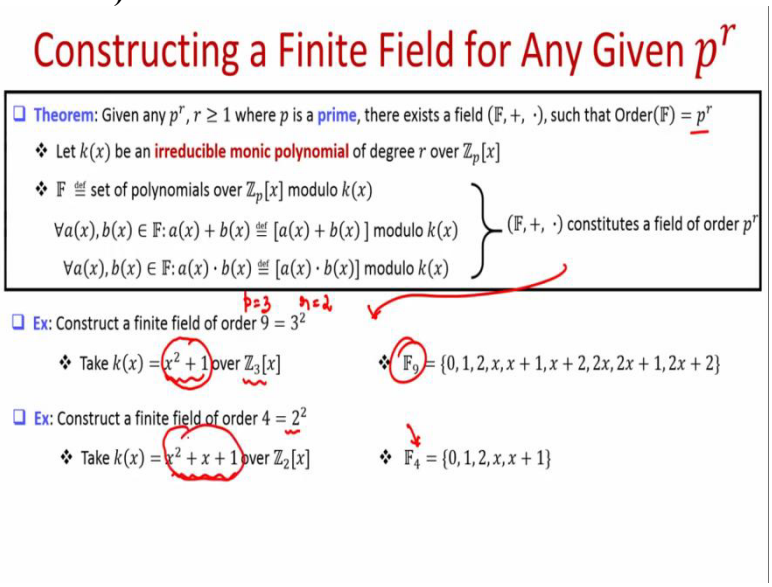
And my claim is that the way I have constructed my F and the way I have defined my plus operation and dot operation they satisfy the properties or they satisfy the field axioms, it can be verified easily. To check, specifically I want to show you that any non-zero polynomial here will have a corresponding multiplicative inverse. Otherwise, remaining properties are easy to verify: the closure, associative, distributive law, existence of identity elements and so on.

The additive identity element will be the 0 polynomial, the multiplicative identity element will be the constant polynomial 1 and so on. Let us see the existence of multiplicative inverse. So, imagine you are given a non empty, non empty means non-zero, a non-zero polynomial. I want to show it has a multiplicative inverse and the multiplicative inverse is guaranteed because of the following: since I am taking k(x) to be an irreducible polynomial, so till now you must have been wondering that why I am taking k(x) to be irreducible why cannot I take k(x) to be any polynomial of degree r, there is a reason. If I take k(x) to be irreducible then I know that the only monic GCD of a(x) and k(x) will be the constant polynomial 1. Why so because since k(x) is irreducible, I cannot factorise out k(x). That means I do not have non constant factors of k(x) and hence the only possible monic GCD, the common divisor of a(x) and b(x) could be a constant polynomial 1. That means I can say that I can now apply the Euclidean GCD theorem and as per the Euclidean GCD theorem, the GCD can be expressed in terms of the individual polynomials itself.

So my individual polynomials are a(x) and k(x), then as per the Euclidean theorem I can find out "linear combiners", they are actually not linear combiners they are some polynomials when multiplied with a(x) and k(x) respectively and added will give me the GCD where the GCD in this case is 1 and what can I say about this multiplier polynomials f(x) and g(x)? Each of them are actually polynomials over the fields $\mathbb{Z}_p$ of some degree, need not be of degree at most r - 1. Now, if this is the case, if this equation holds then if I take modulo k(x) on both LHS and RHS, then in my RHS 1 modulo k(x) will give me the polynomial 1 itself. Whereas in my LHS if I divide the LHS by k(x) the effect of k(x) and k(x) cancels out k(x) is completely divisible by k(x). So, overall I get that the product of a(x) f(x) modulo k(x) will give me the constant polynomial 1 and in other words, I have found here 2 polynomials which when multiplied modulo k(x) will give me 1.

So, I can say that f(x) can be treated as the multiplicative inverse of a(x). Now, if f(x) has degree up to r – 1, at most r – 1, well and good. But if that is not the case I can reduce f(x) modulo k(x) and that will give me a polynomial of degree at most r - 1 which when multiplied with a(x) will give me the identity element 1. So, if you see here closely the way I have argued about the existence of multiplicative inverse for a(x) polynomial is precisely the same in which where actually when I showed that if GCD of 2 numbers a and n is 1, then we can find out multiplicative inverse of a and there we argued that as per the Bezout's theorem I can express the GCD of a and n in terms of linear their combination. So, say s and t are Bezout coefficients and then I do modulo n on both sides and I get s times a modulo n is 1 and then I say that a inverse is actually s modulo n. That is what we did in our number theory. The same thing we have generalised in the context of polynomials.

(Refer Slide Time: 32:39)



So, that is a general template for constructing a finite field for any given p and r where the order is p^r. Now let us see how exactly this framework can be applied to construct fields of various fields of order p^r for some given values of p and r. So, imagine I take p to be 3 and r to be 2. So, I need a irreducible polynomial of degree 2.

So, this is an irreducible polynomial ($x^2 + 1$) and my collection F_9 will have all the polynomials over $\mathbb{Z}_3$ namely the coefficients are from $\mathbb{Z}_3$ and the degree of the polynomials can be 0 or 1. So, I get total 9 such polynomials: {0, 1, 2, x, x + 1, x + 2, 2x, 2x + 1, 2x + 2} and my plus operation and multiplication operation will be defined modulo $x^2 + 1$. If I want to construct a finite field of order 4 so characteristic should be 2.

So I will take a irreducible polynomial of degree 2, so if you take this irreducible polynomial ($x^2 + x + 1$) and this to be your F_4 set {0, 1, x, x+1} and my plus and dot operation will be defined modulo this irreducible polynomial and so. So with that I conclude today's lecture. Just to summarise in this lecture we continued our discussion regarding the properties of finite fields and we proved a very nice property about order of a finite field.

We showed that the order of a finite field is always of the form some prime number raised to the power r where the prime number is actually the characteristic of that field. Thank you.