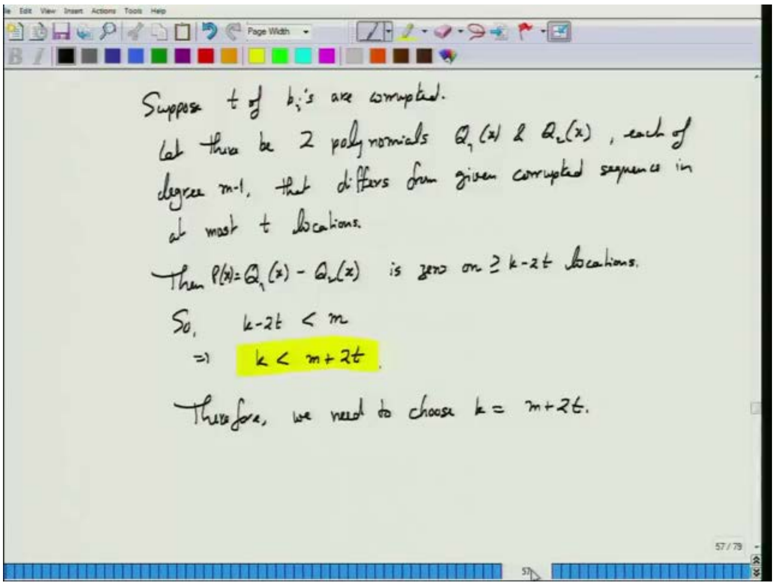
Modern Algebra
Prof. Manindra Agrawal
Department of Computer Science and Engineering
Indian Institute of Technology, Kanpur

Lecture – 18
Application of Fields

We left at this point yesterday.

(Refer Slide Time: 00:17)



Suppose t of b_i's are corrupted.

Let there be 2 polynomials $Q_1(x)$ & $Q_2(x)$, each of degree $m-1$, that differs from given corrupted sequence in at most t locations.

Then $P(x) = Q_1(x) - Q_2(x)$ is zero on $\geq k-2t$ locations.

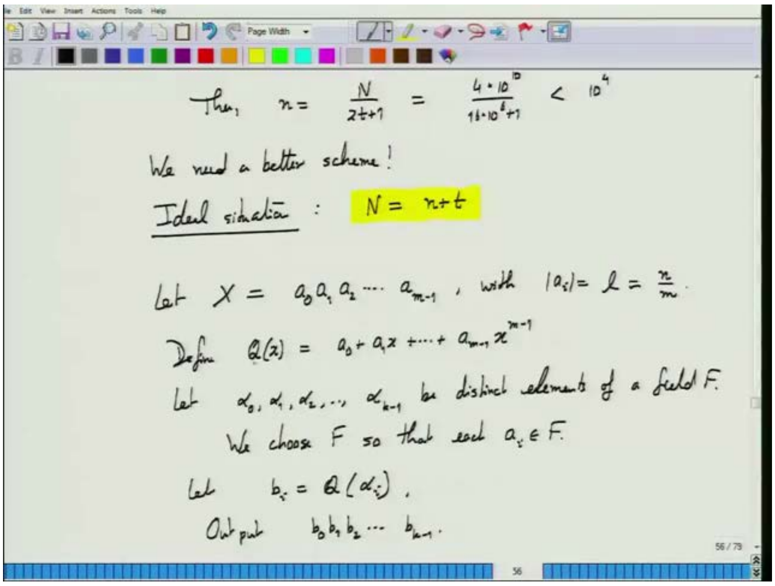
So, $k-2t < m$

$\Rightarrow$ $k < m+2t$.

Therefore, we need to choose $k = m+2t$.

The polynomial p was defined which we know since given Q 2 differ on at most the locations with the fixed string, so this polynomial is 0 on at least k minus 2 t locations. If we choose k equal to m plus 2 t, if we guarantee that we only are a unique polynomial within the errors of (Refer Time: 00:49). So from this now let us try to see, what is the size blow up that happens? We started with n bits, if you remember.

(Refer Slide Time: 01:26)



Then, $n = \frac{N}{2t+1} = \frac{4 \cdot 10^{10}}{16 \cdot 10^{6}+1} < 10^{4}$

We need a better scheme!

Ideal situation : $N = n+t$

Let $X = a_0 a_1 a_2 \cdots a_{m-1}$, with $|a_i| = l = \frac{n}{m}$.

Define $Q(x) = a_0 + a_1 x + \cdots + a_{m-1} x^{m-1}$

Let $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_{k-1}$ be distinct elements of a field F.

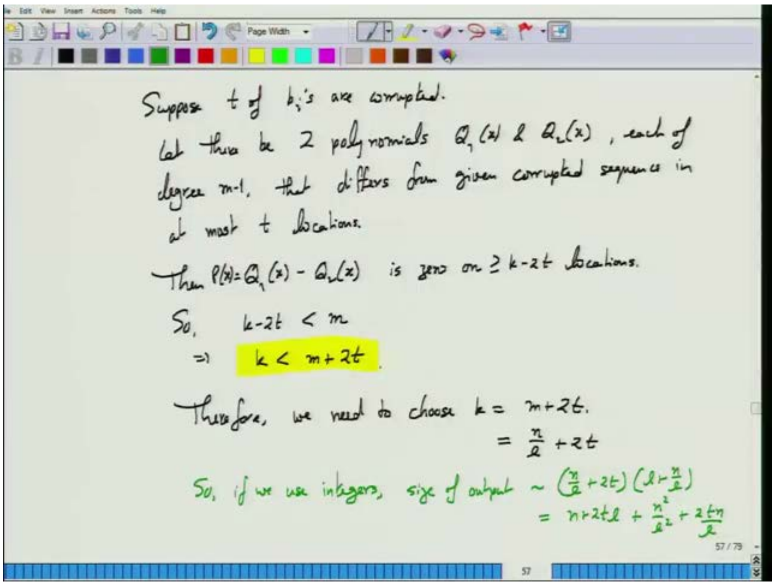
We choose F so that each $a_i \in F$.

Let $b_i = Q(\alpha_i)$.

Output $b_0 b_1 b_2 \cdots b_{k-1}$.

And in this algorithm what we did was that we choose m, such that m is n by l, l is the length of each a j we chop the message into l bit locks. And k was the number of points on which we evaluate the polynomial q which is of degree m minus 1.

(Refer Slide Time: 02:10)



Suppose t of b_i's are corrupted.

Let there be 2 polynomials $Q_1(x)$ & $Q_2(x)$, each of degree m-1, that differs from given corrupted sequence in at most t locations.

Then $P(x) = Q_1(x) - Q_2(x)$ is zero on $\geq k-2t$ locations.

So, $k-2t < m$

$\Rightarrow k < m+2t$.

Therefore, we need to choose $k = m+2t$.

$= \frac{n}{l} + 2t$

So, if we use integers, size of output $\sim \left(\frac{n}{l}+2t\right)\left(l+\frac{n}{l}\right)$

$= n + 2tl + \frac{n^2}{l^2} + \frac{2tn}{l}$

So, let us just put it all together continuing from there m is n by l plus 2 t; t is the number of errors we wish to tolerate, and this tells us the number of points k on which we need to evaluate. Now this is dependent on two parameters; one is l, and one is k. If the number of point on which we evaluate the polynomial is k and we concatenate all the evaluations what is going to be the resulting size of the string we get.

Look at this. This is a polynomial q, we evaluate this polynomial on alpha j's and we get p i which is q of alpha i and then concatenate all the q's. What is the size of this output? It is about k times the size of each b i. Of course, this different b i's may have different size, but if you assuming that they are about approximately the same then approximate estimate would be k times of the size of b i. So, what is the size of b i? Well, depends really on the field on which we are working. Let us say begin with integers, that is the most natural starting point that each a i i treat as the integer between 0 and 2 to the l minus 1 and then evaluate this polynomial and then see what is the result we get.

So, what is going to be the size of the number b i when you evaluate polynomial q on alpha i. It is a degree m minus 1 polynomial.

Student: Maximum possible alpha is (Refer Time: 03:22).

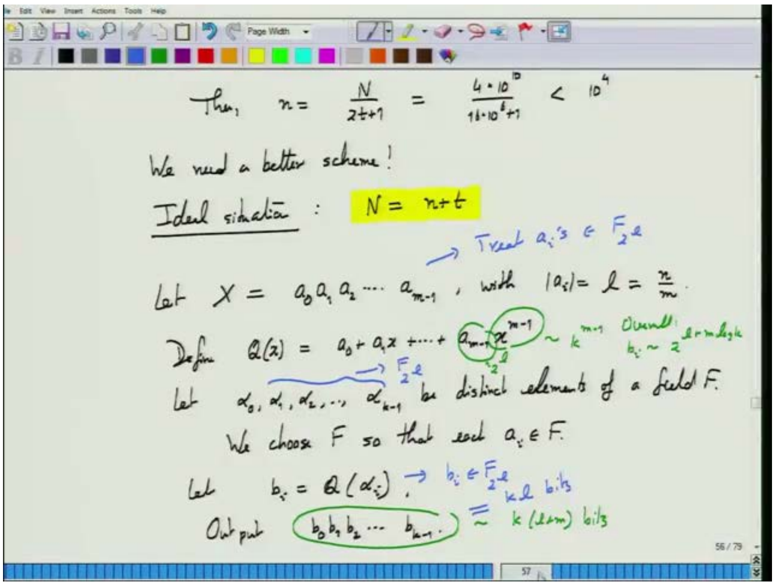
Let us say maximum alpha (Refer Time: 03:24). In fact, since we have to choose k alpha i it could simply be l 1 to k. We could have done with any k alpha i, so maximum is k of alpha i. Then look at this polynomial, so for x will replace at the max k. We will get k to the m minus 1 here multiplied with a m minus 1; this will be the biggest number, all the other numbers will be smaller. So if we just stay with this the magnitude of this number a m minus 1 has magnitude about 2 to the l and this will have magnitude about k to the m minus 1.

So, overall magnitude would be you can say b i is roughly 2 to the l plus m log k. That means, how many bits will each b i require. Helpless I am locked a bit. The total size even if we ignore log k factor, let us say l plus m bits. So total size here would be k times l plus m, that many bits would be needed. In fact, there will be more bits needed, but this is a very conservative estimate of the bits required. And the value of k we derived from here was this.

So, if we use integers size of a output is n by l plus 2 t times l plus n by l, and this is n plus 2 t l plus n square by l square plus 2 t n by l. This is the output size. And if you notice typically we want l to be a small, we do not want to be deal with too big numbers. So, the dominant term in all of this is going to be this; n square by l square, because t is also comparatively much smaller than n, l is comparatively much smaller than n. So, n square by l square is the dominant term. So which tells us that for starting with n bit of data we have to roughly go to about n square bits of data in order to recover from t bits of error.

Well, not depending on what the value of t is this may or may not be improvement on the trivial scheme. There it was about 2 t n here is about the n square, it really depends on the value of t that this may be better this may even be more standard at this in the previous scheme. So, it does not seem to be gaining us much until we realize or notice that all of these calculations we did could have been done over any field. It was of necessary to do it over integers or rationales.

(Refer Time: 07:40)



Then, $n = \frac{N}{2t+1} = \frac{4 \cdot 10^{10}}{16 \cdot 10^{6}+1} < 10^{4}$

We need a better scheme!

<u>Ideal situation</u> : $N = n+t$

Let $X = a_0 a_1 a_2 \cdots a_{m-1}$, with $|a_i| = l = \frac{n}{m}$. → Treat a_i's $\in F_{2^l}$

Define $Q(x) = a_0 + a_1 x + \cdots + \boxed{a_{m-1}}\boxed{x^{m-1}}$ $\sim k^{m-1}$ $\sim 2^l$ Overall: $b_i \sim 2^{l+m\log k}$

Let $\underbrace{\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_{k-1}}_{\to F_{2^l}}$ be distinct elements of a field F.

We choose F so that each $a_i \in F$.

Let $b_i = Q(\alpha_i)$. → $b_i \in F_{2^l}$

Output $\boxed{b_0 b_1 b_2 \cdots b_{k-1}}$. kl bits $\approx k(l+m)$ bits

And the problem with all this raises over integers is that the size of this b’s becomes humongous very large, you started as are l bit long and suddenly we have see b;s become l plus m bit long so there is a whole value m gets added which is a pretty large number.

And b's are i’s or as a result of arithmetic operations over a’s. You know a way of keeping the size of b limited arrays can be choose a field of operations so that the b’s will also has small size

Student: (Refer Time: 08:19).

Yes.

Student: (Refer Time: 08:25).

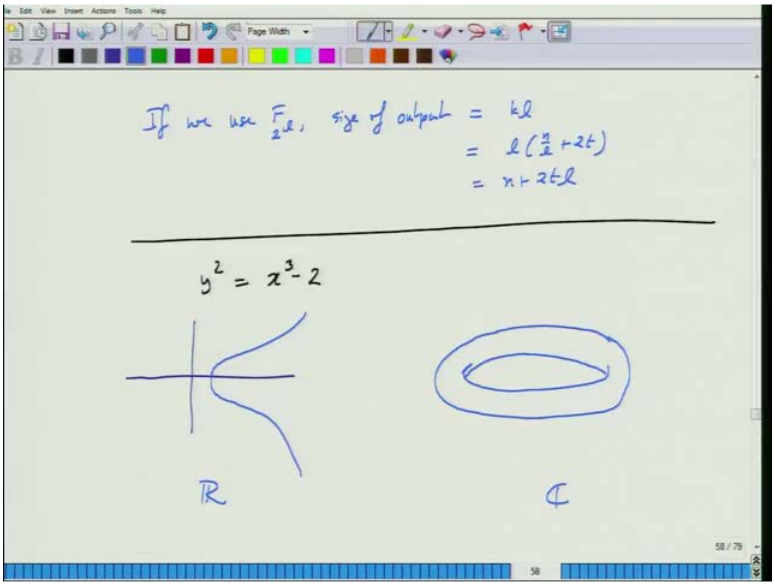
Different values it should be larger, but that is certainly true. My question is can we reduce the size of b a’s so that the output raise is smaller. When we choose everything our integers or rational (Refer Time: 08:53) about to feel this size is growing up.

Student: (Refer Time: 08:59).

Exactly, do it over finite field because no matter how much the arithmetic you do at f b you stay with an f b and the size of an element of f b is only log p bits log. In fact, going back to the particular finite I earlier talked about it very naturally fits in here. We choose if each a i is l bit long. So we treat a i’s in f 2 to the l. Now these will also be f in f 2 to the l, but here the important point as that we need k distinct alphas so 2 to the l must be at least k that is important. So, i cannot be chosen very small. We have to choose l slightly large enough so that (Refer Time: 09:59).

Now, what is the size of b when we do this evaluation? Well, b i’s is also an f 2 to the l, so in it is just an l bit lon number. So, instead of this we now get this to be exactly equal to k times l bits, is b i l bit lon k of them k times l bits. So now, let us go to this.

(Refer Slide Time: 10:36)



If we use F_{2^l}, size of output $= kl$

$$= l\left(\frac{n}{l}+2t\right)$$

$$= n+2tl$$

$y^2 = x^3 - 2$

We just start from here again and do a calculation size of output is k times l and there is l times n by l plus 2 t which is n plus 2 t l. Now we get something very significant. This is telling us n is the size of input, size of output is n plus 2 t l just thing additive part. And notice that n is going to be very large on the order of giga bytes when you talk about storing in a DVD. T is, I said may be 1 make to i to this to example may be 1 MB 40, if you remember 1 MB right I picked number of (Refer Time: 11:39) 1 MB. So about 1 mega bits here and l is choice avail is r. So we choose it to be small enough to taking k values, so it would not be very large.

So, overall if you see the addition in there it is going to be much smaller compared to n, and that is what gives us the real error correction with very little overhead; I tell you the overhead was again I mentioned here n plus t, start with n and n plus t is the increase in size. Here I am getting n plus 2 t times l so is not quite the best, but it comes very close to being best specially for CD’s DVD because if you the kind of error that occur there is scratch it would not erase a bit. A scratch is if you amplify it is many bits long in width. So, we will basically remove large number of contiguous bits. Now if you store each element here in l contiguous bits. So when we remove a bunch of contiguous bits and say t bits contiguous in the o x stream case it basically means we are removing t by l elements. And so this l factor of l will go away.

So, the glass is simply n plus 2 t in that case is a rough that I am giving, but it has come down to very small. And that is why the code that I described is variation of it is the one that is used to store the data on CD's and DVD's because it gives excellent performance with very little additional space required for correcting this. How did this become possible? In finite fields we were not there it would just not be possible. In fact, all communication this is the code that I describe is used for storage in DVD's CD's, but for communication over microwave radio, waves variation of this algorithm is used but that also critically uses finite fields. In fact, just over any error correction algorithm uses finite fields and for the same reason. Otherwise, the size after doing arithmetic keeps on doing out.

Although, seemingly in awkward and also seemingly very abstract entities finite fields are at the heart of model life I would say, simply model life will not be possible no none of phones or because if you do not have error correction if you call somebody you would not be able to hear what the other side is saying. No data storage may be possible proper on that very easily there will be errors on the data. Communication, the internet would not exist nothing will exist. And we arrived at finite field just by this pure abstract algebra.

We said, let us abstract our right in the beginning the heart of or the key operations or concepts of arithmetic and see whether we can generalize it, we generalize it and generalize it and we say here are rings, question rings, fields and then we found a there are whole bunch of kind of fields some of them are extremely used (Refer Time: 16:11). So, I hope this convinces you that the algebra that is we is studied is really fundamental and with required this process of abstraction can lead to unexpected benefits.

So, this ends one example, I will end with very quickly of describing another example which again something I promise probably in the first class, geometry. So far most of this course is has been spent on looking at arithmetic numbers, see in what way we can study number by generalizing it and abstracting it out. The other fundamental objects of study in mathematics are; geometric objects, curves surfaces, and you would like to study them also. Studying them is more difficult than studying numbers. The numbers they have nice structure we have already identified of excellent structure and you can use this lot of tools have been developed through analyze them. If you some how simple ones I gave you glimpse of this.

On the other hand geometric objects have a problem that you first have to visualize them. When you want to talk about a circle you have this picture in mind that is a circle and then you have a tangent and fine do something. For with simple geometry objects circle it is fine, but any time you go let me more complex of times to say high order got. Or even a circle over complex numbers, can you visualize it. X square plus y square equals 1, where x and y are complex numbers. You cannot, because it is no longer a two dimensional object. In fact, if you look at this curve this also came up somewhere in one of the analysis, minus 2. And in order to solve for this we had to go up to not just integer, but we have to extend the ring of integers to this algebraic integers and I would give out of that.

There is a complex number also that came in there right, some part of i it is y the square root of minus 2. (Refer Time: 19:20) square root of minus 2 rings. Basically, there is some complex numbers also come in, so we are actually looking at this curve this will over complex numbers. But now, we dint use geometry at all, we just simply use the basic numbers or arithmetic to derive the solution that you want; but that were only we found only one solution.

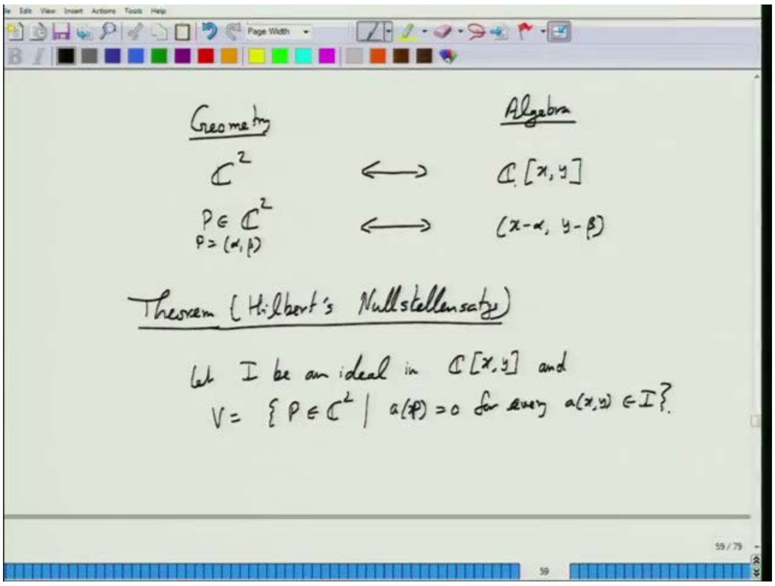
Student: We (Refer Time: 19:50).

Right we gave one solution and we can say that other solutions cannot exist. But curves like this for this specific curve we can do it, but in general of this curves of this kind if you want to analyze we are not always able to do or solve it as simply as we can do it for this curve. So, there it becomes important to analyze the geometry of this curve. But how do you analyze geometry of this curve when it is all complex numbers. Or yes how does this curve look like. When x cube is between 2 or negative, in fact x less than cube root of 2 it does not have any solution and x above cube root of 2 it will have solutions, so it will look something like this I think something like symmetric around x axis whether the y square root of (Refer Time: 21:01) will look something like this, that is all real’s.

How does it look like where complex numbers are to even imagine. Now with some more advance analysis we can figure out what it looks like what complex numbers. It is nothing that you would imagine. This looks like a donut or tyer. It is not a two dimensional object any anymore it is a three dimensional object, because now you have complex numbers so there are four independent variables in the complex number, because their y and x both are complex so shape of this curve is like donut shape.

It is a three dimensional because, one dimension is canceled by the equation. You see this is two dimension equations, but say there are equations so it eventually becomes a single dimensional object. So, when you have four terms and there are there is equation setups and there are two equation setup; real part and complex part. So that reduces at the dimension. It becomes a two dimensional object, but it is rendered in a three dimension space just like this is a single dimensional object but rendered in a two dimensional space, that is totally unexpected. There is really no way can make (Refer Time: 22:34) that it would look like this. So the question refer is how do we study the geometry of such curves, imagination just gives up not possible. And there again the algebra that we have developed comes to our (Refer Time: 22:57).

(Refer Slide Time: 23:06)



Geometry $\qquad$ Algebra

$\mathbb{C}^2 \longleftrightarrow \mathbb{C}[x,y]$

$P \in \mathbb{C}^2 \longleftrightarrow (x-\alpha, y-\beta)$

$P = (\alpha, \beta)$

Theorem (Hilbert's Nullstellensatz)

Let I be an ideal in $\mathbb{C}[x,y]$ and

$V = \{ P \in \mathbb{C}^2 \mid a(P) = 0$ for every $a(x,y) \in I \}$.

So, let us look at a simple curve which I will just use an example, may be the simply possible; a circle. Say the nice over reals we can visualize it, but over complex number like I said heart to visualize. What I am going to do is, abstract out this all the geometric properties in an algebraic object. So let us start building up this, we will just call a dictionary between geometry and algebra. So, what are geometric (Refer Time: 23:49) things, curves, points, tangents these are geometric entities and will build a corresponding concepts on the algebra sight.

First, it is a simple two dimensional space that we are working on. So, let us first build that the base and it will be useful to start with an algebraic close fields so we will just stay with complex numbers. The two dimensional space over complex numbers it is where this curves exist. This in not too difficult, this is something we already have been doing. We have been studying geometric objects like circles with algebraic equations like this, and this curve is simply a polynomial which lies in this ring. This is all polynomials in x and y which correspond to all curves over c square two dimension space. So, that is a simple correspondence.

And in the school this is where we adopt and a little bit more, but let us goes beyond points. A point in c square what does it look like. Valuation no will just stay in this ring; just like this is the space encompassing everything all objects we are interested in. Similarly this is now the ring where everything that we want should be there. So, this contains point.

Student: Constant polynomial.

Constant polynomial; what constant polynomial? So it should correspond to let us say p is alpha beta, what object we associate here. We think in terms of curves you will not be able to really (Refer Time: 26:06) correspondence, think in terms of ideals.

Student: (Refer Time: 26:16).

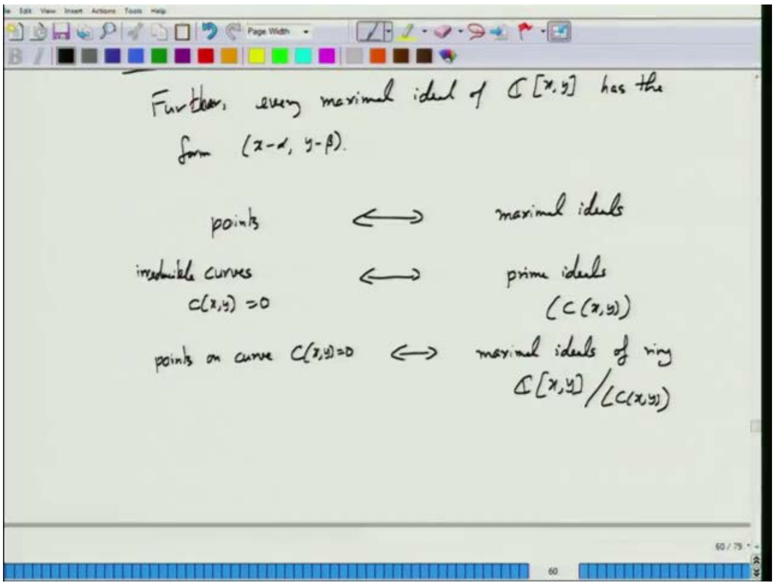
You just look at this principle ideal; x minus alpha y minus beta. And I say that these correspond to each other and this is not a very easy fact. One flight is clear that if you look at the point p and look at this ideal every curve in this ideal vanishes or should say every polynomial in this ideal is 0 at p, because this every polynomial this ideal is so it can be written as something times x minus alpha plus something else times y minus beta and at alpha beta it is 0. So, one way is clear. How about the other way that is, in order to establish this exact correspondence what between this points and ideals. We say here is this point here is an ideal which corresponds to this point I should also say in order to make this correspondence unique and there is no other ideal that corresponds to this point.

In other words every polynomial that is 0 on this point lies in this ideal. So, this was if you think about it is not at all simple to conclude.

Student: (Refer Time: 27:57).

Well, it is not quite like that because here see, what we are shown is one way that every polynomial here vanishes at point p, but can there be any other polynomial not in this ideal which vanishes at point p. This requires theorem to prove and it is a very famous theorem. We prove by Hilbert very famous mathematician is called Hilbert's Nullstellensatz, This theorem was the beginning of algebraic geometry, because this is what allowed mathematicians to establish this dictionary completely. The theorem says away general statement and a more specific statement can be given in terms of this but let us say. So, let I be an ideal in this ring and if corresponding to ideally set it define a set of points p so that all polynomials a x y in ideally I vanish at p to collect all such points. And let b x y be in c x y such that take any polynomial b which is also 0 on all points in v.

(Refer Slide Time: 30:39)



Further, every maximal ideal of $\mathbb{C}[x,y]$ has the form $(x-\alpha, y-\beta)$.

points $\longleftrightarrow$ maximal ideals

irreducible curves $c(x,y)=0$ $\longleftrightarrow$ prime ideals $(c(x,y))$

points on curve $c(x,y)=0$ $\longleftrightarrow$ maximal ideals of ring $\mathbb{C}[x,y]/\langle c(x,y))$

Then there is a power of b that belongs to i. This is one connecting, so this rules out what I just asked earlier. That if you start with an ideal which vanishes on a set of points take any polynomial that also vanishes on the set of points that polynomial may not necessarily blind that ideal but some power of that polynomial lies in that ideal, and this is the best you can say anyway So, this establishes correspondence between ideals and set of points.

And now going back using Nullstellensatz we can say that any polynomial that is 0 on point p, some power of that polynomial must belong to this ideal. And say, this is a maximal ideal in the ring, this ring. I think we have shown this there is a maximal ideal in the ring, I have not shown it is easy to show. This is a maximal ideal in the ring. Since the maximal ideal if power of a polynomial lies in the ideal then that polynomial or that power of an element lies a maximal ideal then their element also lies in the maximal ideal, that is again using maximality we can prove it.

So, which now putting everything together we get that this is precisely the set of ideal of all polynomials that are 0 on this point. So, this ideal uniquely identifies this point and vice versa. So points correspond to maximal ideals. In fact, we can go further and show every maximal ideal. This also requires some bit of work. So now, we have a very nice correspondence that points correspond to maximal ideals. How about curves? With curves we have to be little careful. A curve can simply be union of two lines for example, that is a reducible curve that is you can factor that curve into two lines.

So, when I talk about curves I will talk about irreducible curves, which cannot be for like; circle, l f, parabola, they cannot be factored for the n 2 curve. And again using Nullstellensatz we can prove that this correspond to prime ideals. To prime ideal corresponding to a curve if the curve is let us say c x y it can be shown that the curve is irreducible you can only with which ideal is prime. But this is very preliminary, because really what we want to study is curves in points on that curve, the property the points and various other things on a particular curve, right now we have all points to correspond to maximal ideal, curves corresponding to prime ideal.

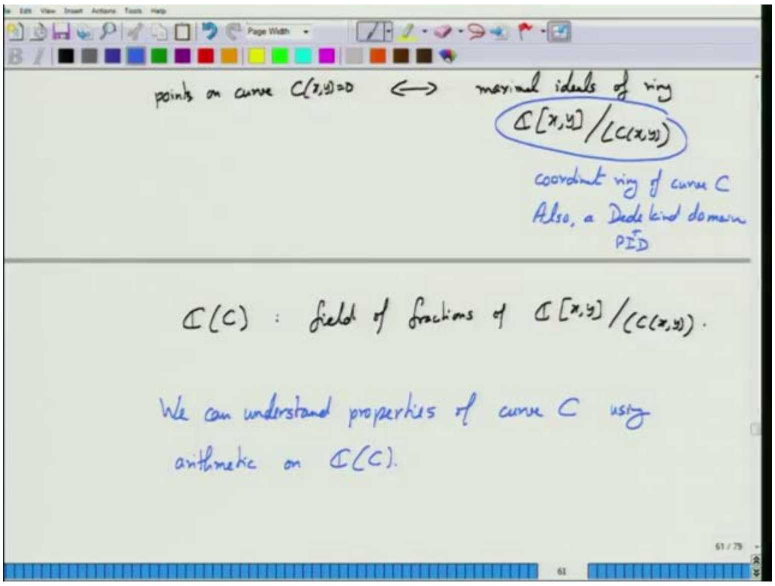
So let us go further; points on curve c x y equal to 0 they correspond to take the quotient ring quotient with ideal generated by c x y. Now maximal ideal of this quotient ring correspond to points on this curve c x y equal to 0.

Student: when we say curve then how (Refer Time: 35:27).

We can say since we already have we can always write a curve in algebra (Refer Time: 35:35) that is y cube equals x square minus 2 or whatever that equation is. So that much often correspondence we already had, but going beyond that points on that particular curve correspond to maximal ideal of this string. If you want to study properties of points of a curve will just need to study the maximal ideal of this ring.

Now, something quite magical happens once we are up to this. This is actually called the coordinate ring of curve. This since sees prime, this quotient ring is in integral domain you have seen this. Since it is an integral domain we can form the field of fractions, of any integral domain you can have field of fractions.

(Refer Slide Time: 36:57)



points on curve $C(x,y)=0$ $\longleftrightarrow$ maximal ideals of ring $\mathbb{C}[x,y]/(C(x,y))$

coordinate ring of curve C
Also, a Dedekind domain
PID

$\mathbb{C}(C)$: field of fractions of $\mathbb{C}[x,y]/(C(x,y))$.

We can understand properties of curve C using arithmetic on $\mathbb{C}(C)$.

So the field of fractions of this curve is called it is represented as c of c. Now, with a little bit more work, not very difficult but it requires some bit of algebra. Of course, these fields of fractions also have the all the maximal ideals there it is now without field of fractions we have to be little careful in defining what that maximal ideal looks like here. Because everything the coordinate ring is subset of field of fractions, so the maximal ideal also are here and this also happens to be a dedicated domain. In fact, this is also a dedicated domain. So, which means that every ideal here can be uniquely factored as product of prime ideals, this is the property of the dedicated domain.

Now there is one more property satisfies that it is also a principle ideal domain, which means that every ideal in this ring is principle, which intern means that every prime ideal is maximal. So, every ideal therefore in this ring can be written as a product of maximal ideals, maximal ideals correspond to prime points. If you think of this as a ring of numbers, again with lot of nice properties you can view them as a ring of numbers. So, the prime numbers in this ring are points to the curve. And since this is a ring we can do full algebra on the ideals here that intern corresponds to doing set some kind lot quite algebra there, but certain very interesting operations on points of the curve. And then you have this field of fractions associated with it this is the full arithmetic you can do in this field of fractions.

So, in this sense we have associated with a curve a field which is just a collection of numbers of certain kind where full arithmetic is possible, and the arithmetic on this field has some very tight connection with the operation to the curve. And since arithmetic we understand and we can do like I just mentioned earlier that we can do arithmetic in a much more nicer way, we have many more tools to do arithmetic, so we can understand that curve and for it is properties by arithmetic in this field and that is the key take away.

And this I indeed resulted in a lots of properties of course which otherwise you would not have been able to prove. Some simple example been say, two curves of degree one curve of degree 2 another curve of degree 3, you understand what degree of curve is the high and degree of the highest degree term in the curve. And how many points can they inter sector matters 6, why? See a line and a degree two curve can intersect in at most two points, in fact a line and degree d curve can intersect in at most d points that is simple.

Student: Higher (Refer Time: 42:20).

But higher degree let us say degree d 1 curve and degree d 2 curve what is your maximum number of points they can inter sector.

Student: (Refer Time: 42:30) d 1 and d 2.

Maximum on d 1 d 2, no; it is d 1 times d 2. Take a circle and an ellipse, you can the max they can intersect at 4 points so that is 2 degree 2 degree 2 that is 4. That is just an example, but in general any degree d 1 curve and any degree d 2 curve, as long as they do not have a factor in term I mean both the curves have a line in common then it is a trivial thing. But otherwise the maximum number of points they can intersect n is d 1 times 0.

Student: (Refer Time: 43:16).

You cannot eliminate a variable; see if you can eliminate a variable yes. But, the curve could be any after may not allow you to eliminate variables. And that is proven using algebraic geometry that is take out the curve look at this coordinate rings and then about argue on this. That is a very simple fact there are many interesting facts you can prove. They are also very useful in finding out integer solutions of equations like, one example was y square equals x cube minus 2 you are looking an integer solutions of that and we saw that you have to go to higher algebraic number can find a solution, but that was not very adequate. We can study solutions on families of curves of certain kind by again looking at the geometry through the coordinator.

So, I will stop here.