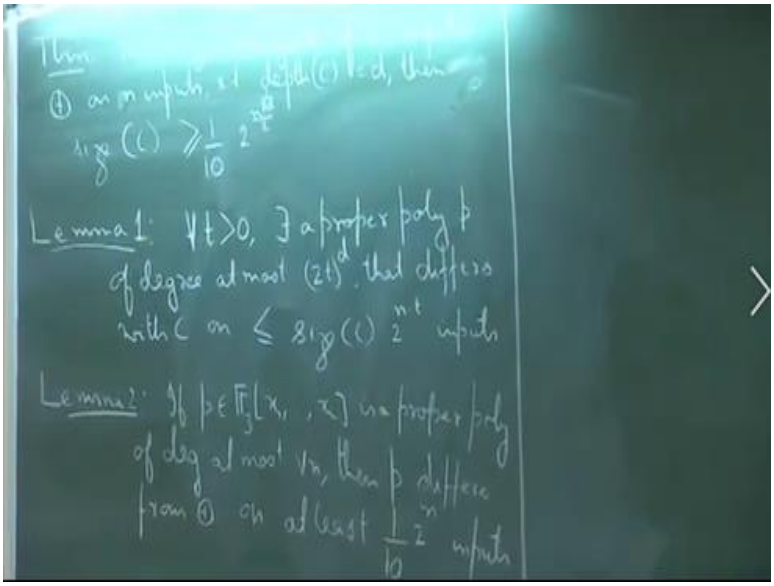
Computational Complexity Theory
Prof. Raghunath Tewari
Department of Computer Science and Engineering
Indian Institute of Technology, Kanpur

Lecture -37
Introduction

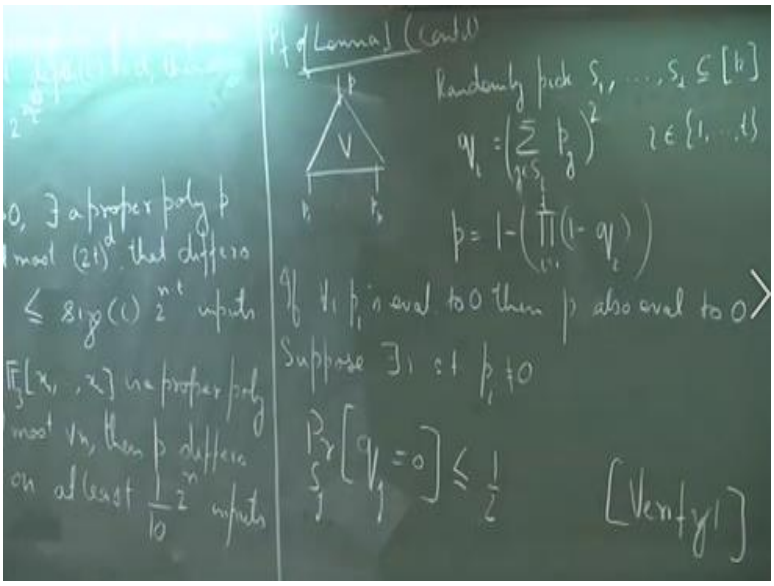
(Refer Slide Time: 00:16)



So, I think this is what we had last time so basically what this lemma says is that if I have an AC 0 circuit then it can be approximated well by a proper polynomial of not too high a degree. So, we will complete the proof of this Lemma today and we will also prove other Lemma (()) (03:43-03:45)) for the sake of completeness. If P belonging to (()) (04:07)) proper polynomial of degrees almost root n, then P differs from parity on at least 1 10 fractions of n.

So, we will see I mean we will see how the combination of these two will give us the theorem and how why each of these two lemmas hold independent.

(Refer Slide Time: 05:34)



So, let us continue the proof of Lemma 1. We saw that we can assume that our AC 0 circuit has only or and not gates and for not gates we can consider the polynomial 1- b if the polynomial at its input edge is b and for or gate the polynomial that we design was the following. So, suppose this is an or gate it has the polynomials p 1 through p k at its input edges. So, we define a polynomial p for the output edge as follows.

So, randomly pick sets S 1 to S t which are subsets of elements from 1 to k. So, this t basically comes from the parameter of this Lemma, basically this lemma holds for any t. So, whatever is the t that we start off with the subsets that we pick will be equal to that t. So, then q i is defined as sigma i E i t. So, for each subset S i, I define this polynomial q i which is just basically taking the sum of those respective P j's and squaring it up.

And finally this polynomial P will be equal to 1 minus the product of 1 minus q i. So, what can we say? So, suppose all these polynomials on a particular input all these polynomials evaluated to 0. What can we say about p in that case? p will be 0, because no matter what subsets we pick if all the p i's are 0 all these sums will also be 0, the square of the sums will also be 0, so all the q i's will be 0 if all the q i's are 0 or the 1 minus q i's are 1 so this product is equal to 1 so 1 minus 1 is 0.

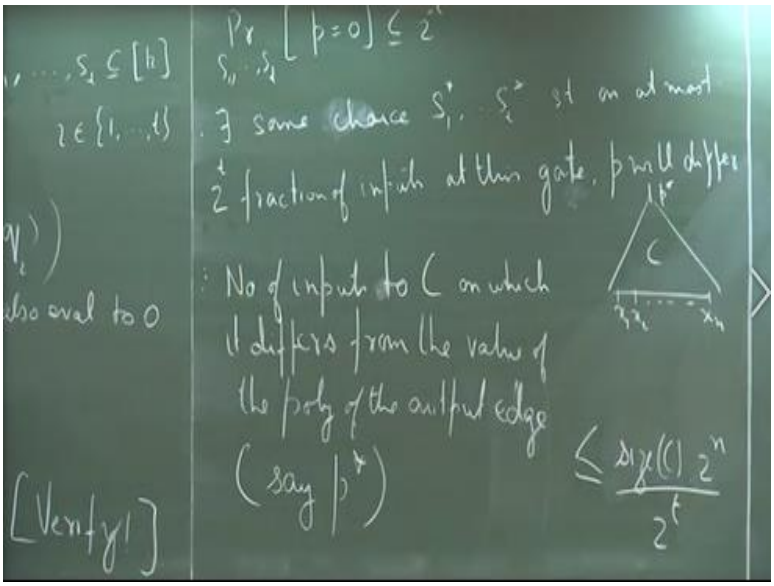
So, it does not matter so if for all i p i's evaluate to 0, then p also evaluates to 0. So, now comes the other case, so suppose on some inputs suppose there exists an i such that p i is not equal to 0 otherwise p i's equal to 1. So, can p be equal to 0, it can still be equal to 0 because basically since we are picking these sets randomly it can happen that we do not, so let us say that this p i' is the only p i that is 1 and in these random sets in none of these random sets we end up picking that particular p i.

So, then again by the same logic this p will turn out to be 0 although that is not what we want. But let us bound that probability, what is the probability that such a thing will happen? So, over the choice of one particular S j let us say what is the probability that this q j is equal to 0 suppose there exist an i such that p i is not zero. What is the probability? That a particular q j becomes 0. So, I claim that so let me make my claim so I claim that this is less than half, why is it so?

There is one very easy argument because you see the following so let, so one way to look at this is that. So, suppose so there can be two cases so either S j contains the ith index or it does not contain the ith index. So, suppose if it contains the ith index then this quantity can sum to 0. So, in other words so this quantity cannot sum to zero for both the cases the first case where S j contains the ith index and in the second case where S j does not contain the ith index.

If I just consider this S j minus this one particular index i, in these two cases it cannot be the thing that this sum becomes 1. Because since p i's not equal to zero at least in one of those cases that thing will not be 0 and therefore the probability that this is equal to 0 therefore is, so just verify this here this is just a half intuition why this is true but convince yourselves.

(Refer Slide Time: 12:33)



So, if this is the case then the probability that over all these sets S 1 to S t the probability that p = 0 is how much? Basically, it is just a product of this because if you note that the way p is computed p is nothing but the power of the q i's. So, even if one of the q i's is equal to 1, p will be equal to one because then this product will be 0 and therefore p becomes 1. So, the only time when p becomes 0 is when each of these q i's are 0 and the probability of that happening is bounded by half so the probability of this is bounded by 2 to the power -t.

So, therefore what this implies is there exists some choice of subsets S 1 let us call them S 1 star to S t star such that so there exists some choice of subsets, such that on at most 2 to the power t fraction of inputs at this gate P will differ. So, suppose now if I again look at this gate and if I look at all the inputs to this gate by this argument what we can say is that on at most so, much fraction of the inputs the output of this gate will differ from whatever the value that p is computing.

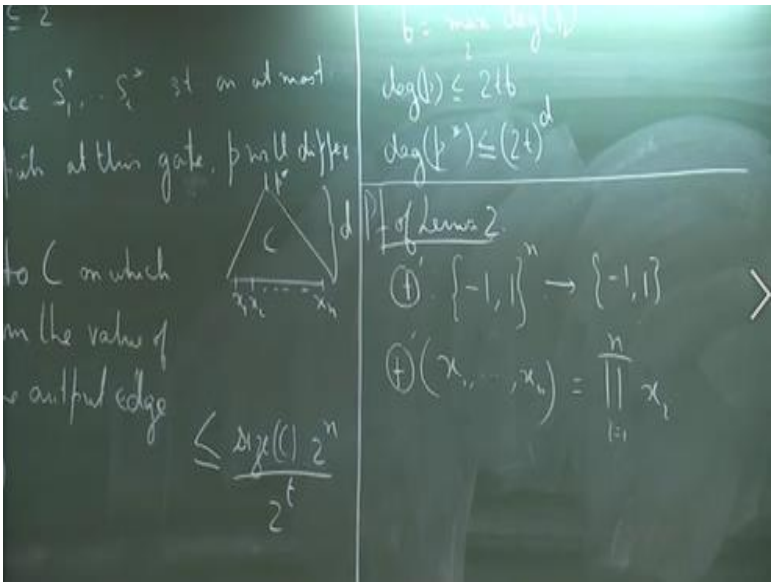
So, now I can just apply a summation pump. So, how many gates do we have in the circuit? So, the total number of gates is size of C so therefore the number of inputs to see on which it differs from the value of the polynomial of the output edge of that circuit. Denote this polynomial by p star. So, basically what I am saying here is that suppose now this is our entire circuit C it has inputs x 1 x 2 up to x n and this is the output edge.

So, by this construction we have associated a polynomial with each edge of this circuit and let us say that p star is the polynomial associated with the output. So, on how many inputs will the value of this polynomial differ from the value computed by the circuit? So, basically, I can just take a very rough upper bound so the total number of gates is size of C on each gate I mean I do not know I do not even care how many NOT gates or how many OR gates are there if it is an OR gate, I am lucky because there will not be any error.

But let us assume that all of them AND OR gates and on all the gates I make an error of this much amount. So, the total number of inputs on which it will differ is size of C times the total number of possible inputs which is 2 to the power n divided by the amount of error and that is what we want. So, the last thing that needs to be checked is how much is the degree of p star? So, the degree of p star so I briefly mentioned this towards the end of last lecture.

So, if you look at a NOT gate, the NOT gate does not increase the degree because if the input to a NOT gate is p the output is just 1 - p, the degree remains the same. At and or gate how much is the increase in the degree?

(Refer Slide Time: 18:46)



So, suppose for and OR gate over all its input polynomials let p be the maximum degree, borrowing that same rotation. Then how much is the degree of p? How do you get then? So, if the maximum degree of these p j's is p then this the maximum degree of this q i is 2 b and by

that product it becomes 2tb. So, now if you just go on level by level I mean if you just do a simple you can write down a simple recursive formula also and what you get is that degree of p star will be 2t d place to the parity.

Because at the base level at this level all the degrees are 1 then they go up to maybe 2t and then 2t whole square 2t whole cube and since the total depth is d, the total degree is bounded by 2t to the power d. So, that completes the first Lemma. So, why is it a proper polynomial why is p star a proper polynomial? So, what happens at a not k? If the input is a prop, so again from the base case so all the x i's are proper polynomial.

Suppose if I have a NOT gate the polynomial associated with the NOT gate is 1- t so if p is proper then 1- p 0 but what about for or gate? If all these p j's are proper then note that q i is a proper polynomial because the way we define this. And if all the q i's are proper then this again takes value in 0 and 1 therefore 1 minus this plugin take value is 0 and 1. So, any other doubts? So, now let us move to lemma 2.

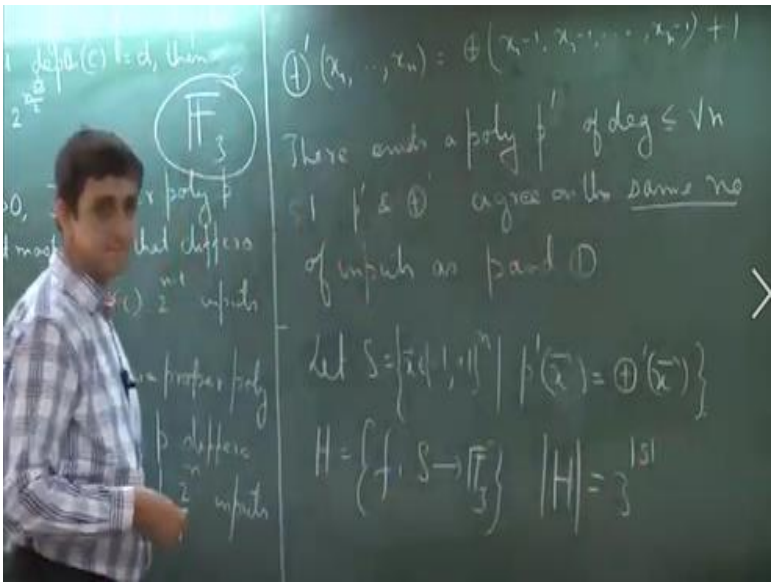
So, let us look at lemma 2, so lemma 2 says that parity cannot be well approximated by a small degree proper polynomial. So, here the trick is what we will do for Lemma 2 is that we will actually translate the problem. So, instead of looking at this parity problem we will look at another problem which is very similar to parity in the sense that it basically can be translated from parity by a very easy conversion and we will see that on how many inputs this other function equals the proper polynomial that we have?

And from that we will basically find out the number of inputs on which parity will differ from this probability. So, let me write it down so the polynomial the I mean the function that we will consider so let us just denote it by parity prime. So, this is a function from the domain -1, 1 to the power n so it takes n bits from -1, 1 and it outputs a value in -1,1 and it is defined as nothing but just a product of its input. So, what was parity?

Parity was the sum of its input bits modulo 2 and that was over 0, 1 here I am looking at another function which is just a product of its input bits over -1 and 1. And we see that why this function

works for us and what is the relation between this function and parity? So, first let us look at this second part, what is the relation between parity prime and time?

(Refer Slide Time: 24:26)



So, what can I say about this parity prime of x 1 to x n, if I want to write parity prime as a function of parity what can I do that? What is the exact translation? Parity is from 0,1 to the far end absolutely so firstly I have to translate these the domain of my x i's. So, here one looks like exactly so here one is behaving like 0 because 1 does not change whatever the product value is but minus 1 basically switches it between plus and minus 1.(())(25:26-25:45) Many ways in which you can do this change but the easiest is just to x 1- 1 x 2 - 1 and so until x n - 1.

So, what we get here is so 1 gets mapped to 0 and - 1 gets mapped to 1, so now when I compute the parity of this (()) (26:18-26:43) it is so this entire computation that we are doing is in this field F 3. So, this is the reason why we have chosen it because it is F 3 which is the smallest feed in which both these operations can be defined together with this translation and we have been working in F 2 then this kind of a translation would not have been possible.

So, -2 is the same as 1 so basically now this is parity of x 1-1 up to x n -1+ 1. So, there exists a polynomial P prime of degree at most root n, such that P prime and parity prime agreed on the same number of inputs as p and parity. So, what does the hypothesis of Lemma to give us? The

hypothesis of Lemma 2 gives us a proper polynomial of degree at most root n basically that should agree with parity on a certain number of it.

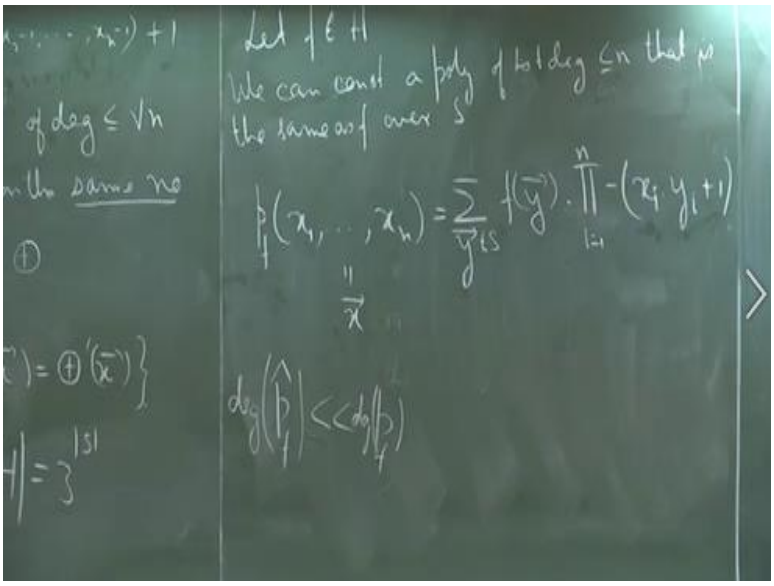
So, if I have a proper polynomial of degree at most root n, I can easily get another polynomial p prime which has the same degree and it agrees with parity prime on the same cardinality of inputs as p agrees with parity, because if I just apply this simple conversion you see that the degree will not increase. Because if I just look at p of x 1 to x n and I define p prime of x 1 to x n by just a simple substitution there will not be any increase in degree;

And because of this equality the cardinality of points on which these two polynomials are these two think of them as functions these two functions agree will be same as the cardinality of points on which these two functions. But there is a crucial difference I mean it is the cardinality on which I mean it is a cardinality which is equal it is not the set of points which are equal because for this pair of functions I am considering 0, 1 as my domain but here I am considering -1 and 1. So, very important that we; are looking at the number and not the exact set.

But the inputs can be mapped, but it can be mapped to each other because both these domains have the same cardinality so ultimately what we want is just the numbers so we can actually look at this entire thing from the perspective of these two functions. And that is what we do. So, any questions so far? So, let us see why this is true? So, let us S b the set of inputs on which these two conquer so the p prime of using this arrow notation to denote the fact that this is a vector.

So, p prime of x squared is the same as parity prime of x and let H is a set of functions f from S to this F 3, what is the cardinality of H? How many functions do we have from S to F 3? So, any point here can go to one of these three points. So, each point has three choices so the total number of functions is just 3 to the power S. In fact, if you have any domain a and b the number of functions from a to b is the cardinality of b raised to the carnality of b. So, let so basically what we will show is that now we can associate a polynomial with every such function.

(Refer Slide Time: 33:12)



So, let f be one such function in H. We can construct a polynomial of total degree at most n that is basically the same as f. So, we call this polynomial p f in fact we can explicitly define this so the way this polynomial f is defined is we have inputs x 1 to x n I just denote this by the vector x is equal to sum over all vectors y in s of, f evaluated at that point times the product of minus x i dot y i + 1 so i going from 1.

So, basically what I am saying is that you take any function from this set you can always construct a polynomial which have whose total degree is at most I mean these are the n variables and no monomial has a total degree more than n, some monomial is basically just a term in this entire polynomial such that p of f will equal f for all possible inputs from this set from this like over S.

So, why is this I mean why does this polynomial work it is very simple, so suppose you pick and y which is not the same as x, if y is not the same as the given input x then there is at least one index in which y and x will differ, so if there is one index in which y and x will differ so let us say that index is j. What will be the value of this product? So, their product will be minus one, because f over S, one is one and other is minus one.

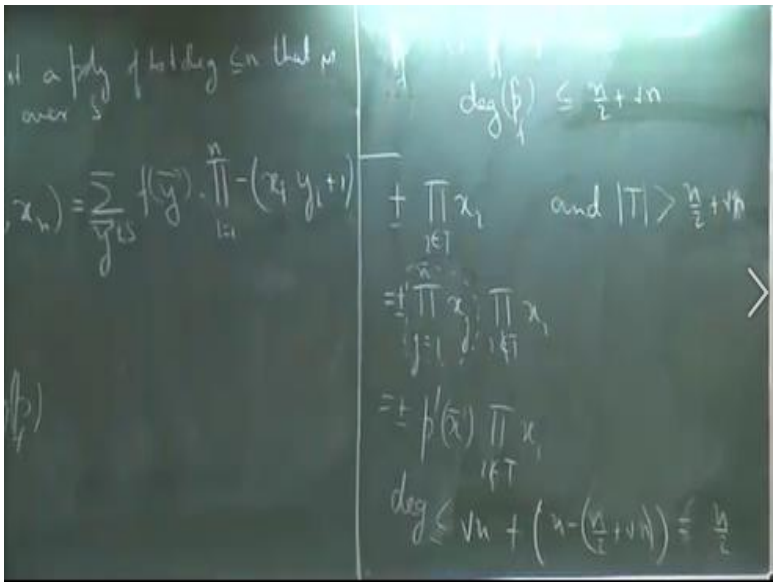
So, there product will be minus one. So, then this sum will be equal to zero and therefore this entire product will be equal to zero, so the amount and therefore this entire thing will be equal to

zero, so the amount that that particular y contributes to this entire sum will be 0. But now let us suppose that y is exactly equal to x, then what happens to this product? So, this is 1 but 1 + 1 is how much?

It is minus one so I have a minus sitting before it so that is plus one then this product becomes equal to one and it is just f of x. So, by definition I mean it is just a contract way of defining this polynomial f p of f but p of f by definition will be equal to where for all possible inputs over S. Actually, there are many ways in which you can find this this is just one way of doing it. But what we have right now is that p of f can have total degree as large as n.

So, our goal is to decrease that degree. So, can I come up with an alternative polynomial let us say some polynomial p f whose degree is asymptotically smaller than the degree of p of x and then the main crux of this proof is that I mean very non-intuitive and it is a very brilliant idea is that you can actually do this. So, let us see how we can do this?

(Refer Slide Time: 38:28)



So, let us pick some monomial so actually what we will do is, we will show that another polynomial we have p f hat exist such that p f hat is equal to f over S and degree of p f hat is at most n by 2 + root n, from n we have come down to n by 2 + 2. So, this is what we intend to do. So, let us pick a monomial so basically this polynomial p f hat will be the same as p f for all monomials which have degree smaller than these quantities will not do any change.

But suppose we have a monomial product of some x i where i comes from some set t and what we have is that T has size more than n by 2 + root n. So, let us take a moment here so what we are doing here is that look at this polynomial p f, all the monomials whose total degree is less than this amount I just leave them as it is, I do not fiddle with them. But all those monomers whose total degree is more than this amount and do some change to them.

So, I just consider one particular monomial which is product of some x i's with a plus or minus one sitting before it and this is the theme that we will do. So, this can be written as product of i going from 1 to n x i times product of i not belonging to p x n. So, if you are more comfortable, I can just use a different index here really matter let me keep this as i I will use. So, all we are doing is that since again we are working over this field F 3 and minus 1 times minus 1 is 1.

And so is 1 times 1 is 1, I can just substitute this monomial this glass of minus is recorded with this particular monomial. So, is this clear to everybody why the substitution is correct simple algebra, but what can we say about this what is this particular term? It is parity prime, and what do we know about parity prime? So, what we know about parity prime is that there exists a polynomial p prime of degree at most root n which can approximate parity prime over this set.

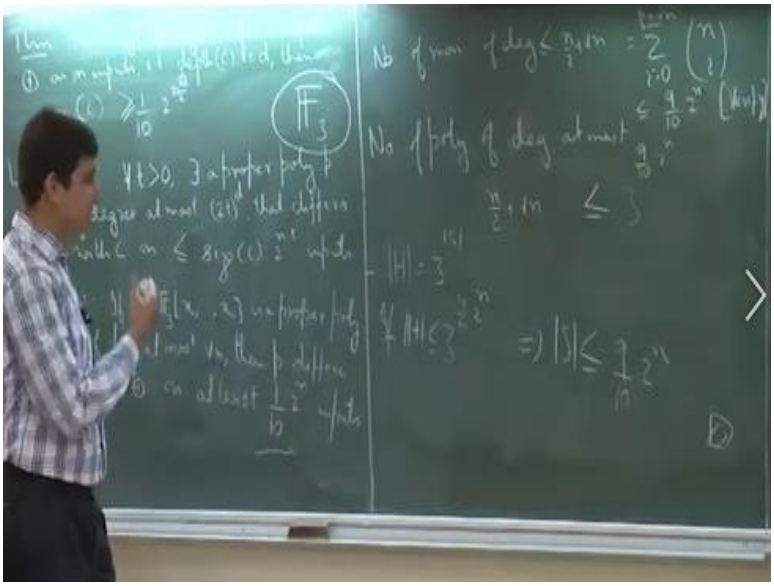
So, this is by definition this is the set of all those points on which p prime approximates parity prime exactly and since we are working only over this set, I can just substitute this monomial with this polynomial p prime of whatever the input exist, that is the clever part that is the trick which they apply. So, now what is the degree of this final polynomial? What is the degree of p prime?

p prime is root n, how much will be the degree of this monomial? T was bounded by I mean T was greater than n by 2 plus root n, so the complement of T is less than n - n by 2 + root n. So, this is n by 2 - root n so root n root n cancels out, so this is less than or this is equal to in fact n by 2. So, what essentially, we did was so again this has the same plus or minus (()) (44:37). So, essentially what we did was we took a monomial which had large degree and we replaced it with another polynomial having small t;

Therefore the claim that we can construct a polynomial e f hat such that which it agrees with f over all inputs from s and the total degrees bounded by this is true and it is constructively true. So, any questions? Because asymptotically because as n grows degree of P f is n but degree of P f hat is n by 2 + root n which is asymptotically smaller than n. So, we will see I mean the reason so this was a use statement to make;

Because exactly what I mean by this; but in the next part we will see that why that is important. But this is clear I mean p f hat will have degree that is less than that how do we make use of this fact.

(Refer Slide Time: 45:56)



So, the number of monomials of degree less than or equal to n by 2 + root n is how much? In fact it is equal to actually, this is equal to sigma i going from 0 to n over 2 + root n and choose i because for each i if I have a monomial which has degree i this is the total number of possible monomials and I am just summing this over i is equal to 0 to the maximum possible degree. So, the number of polynomials of degree at most this quantity is how much?

So, what can be the coefficient of a monomial? So, the coefficient that any monomial takes basically has three choices it can be 1,-1 or 0. So, this is the total number of monomials. The total number of polynomials possible is 3 to the power whatever these quantity is and so I missed

one thing here so this quantity is actually less than so suppose this were to be equal to n then, what is the value of this summation?

It is 2 to the power n. But since this is asymptotically smaller than n what we can do is that for any constant actually we can bound this thing in other words we can say that this is less than 10 by 10 times 2 to the power n. This is again something that you need to verify. Therefore, the total number of possible polynomials is 3 to the power 10 by 10 times 2 to the power n, but now what was the cardinality of H?

So, the cardinality of H was equal to 3 to the power S and any function in H is basically being approximated by a polynomial of so much degree. The total number of possible polynomials is bounded by this so the total number of functions image is also has to be bounded by this, so if H is bounded by this quantity, H is bounded by 3 to the power 9 by 10 to the power times 2 to the power n and H is equal to 3 to the power S, what can I say about S?

S is also bounded by 9 by 10 times 2 to the power n. And S was the set of points on which they were same so the set of points on which they differ is basically at least 1 over ten times 2 to the power n which is what we want. And although this entire thing was with respect to the parity prime function and p prime because of our initial observation that the cardinality does not matter I mean whether we are working with T prime and parity prime or T and parity the cardinality is the same therefore Lemma 2 fold.

So, I will just stop here and why the theorem follows from these two Lemmas is just plugging in the parameters it will not take too much time we will discuss that on Friday.