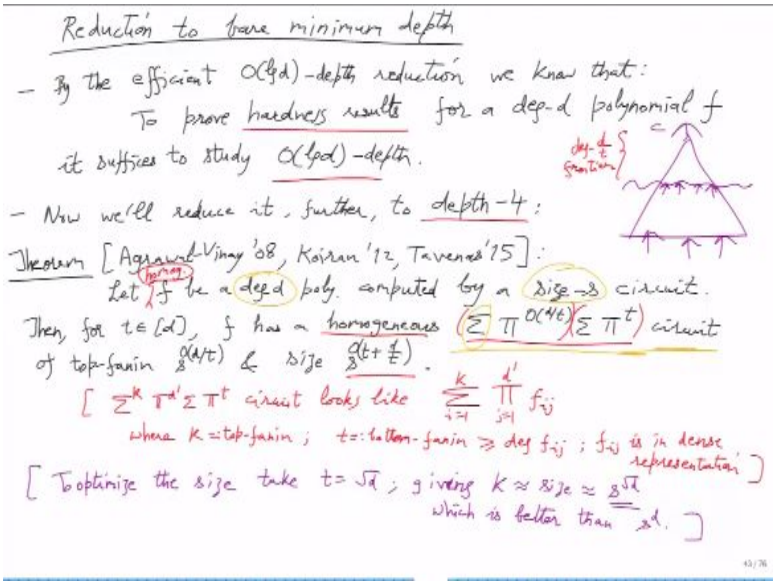
**Arithmetic Circuit Complexity**
**Prof. Nitin Saxena**
**Department of Computer Science and Engineering**
**Indian Institute of Technology-Kanpur**

**Lecture - 10**

So we finished the proof of theorem that gave us a depth 4 reduction.

**(Refer Slide Time: 00:17)**




So any degree d polynomial with size s circuit with no restrictions. The only restriction is that it is a homogenous circuit. Then with that you can bring it down to depth 4 where the two multiplication fanins are the product fanins are very special. So the top multiplication fanin is $d/t$ and the bottom multiplication fanin is only t for any t. With size only $s^{O(t+d/t)}$.

So if you take $t = \sqrt{d}$ then multiplication fanin $\sqrt{d}$ could give you a depth 4 circuit of size $s^{\sqrt{d}}$ instead of trivial $s^d$. We will not discuss that or maybe we will discuss it after the mid sem, some lower bound theorems that will suggest that this reduction is kind of optimal.

So in depth 4 you cannot improve, although $s^{\sqrt{d}}$, but you cannot really improve it. The circuit size may only be s but when you squash it down to depth 4 then the price you have to pay will usually be $s^{\sqrt{d}}$.

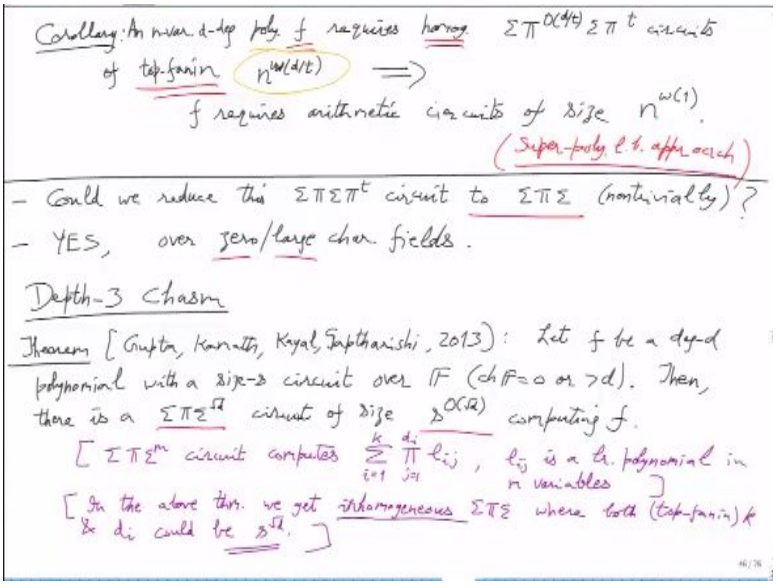
**[student]** We know we can't reduce it, or we know if we reduce it then something bad will happen?

**[Professor]** No, so we will give a, no I think it will be an unconditional result.

So I think we will show that for iterated matrix multiplication for which there is obviously a small arithmetic circuit that can do the multiplication. We will show a lower bound of $s^{\sqrt{d}}$ for homogeneous ΣΠΣΠ representation. So if you want to do iterated matrix multiplication say you want to compute the top left entry, you are say multiplying $n \times n$ matrices d many of them so the degree of the polynomial you are computing is d.

This obviously you can write down a circuit which will be small, but when you squash it down to depth 4 then the depth 4 representation homogeneous ΣΠΣΠ with $\sqrt{d}$ fanin representation will be this $s^{\Omega(\sqrt{d})}$. So this big O actually becomes a theta in that case then for that specific problem of IMM. So I think we will do that after the mid sem as a lower bound result.

**(Refer Slide Time: 03:46)**




And if you can do something better, so if you can find a polynomial for which this type of a representation requires size $n^{\omega(\sqrt{d})}$. So instead of big omega if you can make it little omega, then actually you have proven a super polynomial circuit lower bound result. So such a polynomial cannot be written in a small sized circuit representation.

So that is really the key motivation for the depth reduction to this extreme extent. And now we will do something even better. So the question we ask is could we reduce this $\Sigma\Pi\Sigma\Pi^t$ to depth 3 in some non-trivial way and so the answer to this will be yes as we will prove. But this will assume high characteristic or zero characteristic.

What comes into the picture is just $1/n!$. If that does not exist then there is a problem. So this was called depth 3 chasm result. And the previous theorem that we showed the Agarwal Vinay result was originally called depth 4 chasm. So chasm means that this is the depth at which point things become suddenly very general. So depth 2 for example, is kind of the trivial case.

And if you go to depth 3 then suddenly you are able to connect with general circuits with unrestricted depths. So this theorem is due to Ankit Gupta, Pritish Kamath, Niraj Kayal, and Ramprasad Saptharishi. It is not too old. Says that same thing with $\Sigma\Pi\Sigma$. So let f be a degree d polynomial with a size s circuit over f and let us think of f as the field of complex rationals. It could be any characteristic zero field.

In fact, you will see in the proof that it suffices if the characteristic is more than d. So maybe I also add that. Just more than the degree. Then there is a $\Sigma\Pi\Sigma^{\sqrt{d}}$ circuit of size $s^{\sqrt{d}}$ exactly computing f. which should be a shocking result. So you were given a general circuit of size s and you are able to squash it and bring it down to depth 3 which is just $\Sigma\Pi\Sigma$.

This is just a representation where you are adding a bunch of products of linear polynomials. So this was not really conjectured before it was shown. It never looked possible. Even if you are given depth 4 chasm result. Even after that this does not look possible because how do you compress two multiplication gates into one that they are doing very different things.

So the proof will be highly counterintuitive, as you will see. So one very counterintuitive thing will be that although f is a degree d polynomial, this multiplication gate in $\Sigma\Pi\Sigma$ will have a huge fanin. Its fanin will be $s^{\sqrt{d}}$. So for these products of linear polynomials they will be computing a product polynomial which has far more degree than the degree of f.

But then these all these product gates will come together and things will cancel out and give you f in some non-trivial way. This by the way did not happen in depth 4. In depth 4 the multiplication fanins were even smaller than d. And together they were giving you only things around degree d. They were not really computing something much bigger than degree d. But here it is different.

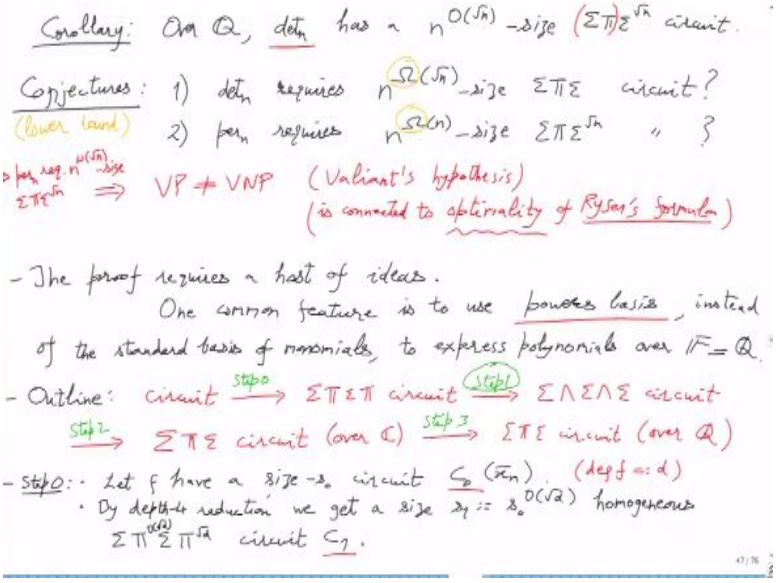
So something non something counterintuitive will happen in the proof. So let us just realize what this model is. So $\Sigma\Pi\Sigma^m$ model computes things like $\sum_{i=1}^{k}\prod_{j=1}^{d_i} l_{ij}$ and $l_{ij}$ is a linear polynomial in how many variables? So this m is the fanin of $\Sigma$. So in m variables. m may be much smaller than n which will be kind of the case we are in.

$\Sigma$ fanin we will have actually only $\sqrt{d}$ the bottom sigma fanin. So it might be actually adding just few variables taking a linear combination and k is the top fanin. So that is one thing about the model. The second thing is that this model is inherently inhomogeneous. We get inhomogeneity, that has to be the case where both top fanin k and the $d_i$'s could be very high $s^{\sqrt{d}}$.

So it is inhomogeneous and these other things which are unspecified k and $d_i$'s, these can be as high as the size. This inhomogeneity has to be there because well your product gate is multiplying a lot more things than you needed. So if this model was homogenous, then you will never be able to cut back f. So it is actually the inhomogeneity that is the beauty that can cancel things out.

So it produces not only high degree monomials but also all these lower degree monomials and then the higher ones cancel and the lower ones remain to give you the sum as f. So let us go to one interesting consequence, which you can see as a new classical result.

**(Refer Slide Time: 13:53)**




Something is telling you about the determinant which was never known and not even believed. So determinant has a non-trivial depth 3 representation. So by definition determinant is defined as a sum of product of variables right. But then these products are how many? For $n \times n$ matrix it is $n!$ many, so $n^n$ many. So that $n^n$ can be reduced to $n^{\sqrt{n}}$.

So there is an efficient, well not really efficient but a non-trivial depth 3 representation. So over Q or over complex whatever; $n \times n$ determinant has a $n^{\sqrt{n}}$ size $\Sigma\Pi\Sigma^{\sqrt{n}}$ circuit. So remember that determinant is defined in this $\Sigma\Pi$ form. And then its size is $n^n$. But if you look at it $\Sigma\Pi\Sigma$ with this bottom $\Sigma$ fanin $\sqrt{n}$ representation then it can be done compressed much more all the way to $n^{\sqrt{n}}$.

So from the definition of determinant could you ever deduce this? Right you cannot even conjecture this. How do you show this now using the previous theorem? So you have to observe that determinant $n \times n$ has $n^2$ variables but degree is only n and you have seen in multiple ways you have seen circuits of size poly n for determinant.

So you get $s^{\sqrt{d}}$ which is $n^{\sqrt{n}}$. This actually sheds light on something very basic. You learn something new about determinant and that you have been doing every other week to learn new things about determinant. So what we do not know is these two conjectures. So first we conjecture that it cannot be done any better.

So determinant requires $n^{\Omega(\sqrt{n})}$ size $\Sigma\Pi\Sigma^{\sqrt{n}}$ circuit. So this $\Sigma\Pi\Sigma$ representation is optimal. That is in fact, any depth 3 representation. So depth 3 representation for determinant is optimal. This is the optimal, this is the only way in a sense to express determinant in depth 3. And the second conjecture is for permanent.

So what do you want to say about permanent in depth 3 representation? So what do you know about deputy 3 representation of permanent? $n^n$ is by definition, but there is something better you have seen. The Ryser's formula gave you $2^n$. But in that $2^n$ if you recall the formula, the bottom fanin was not $\sqrt{n}$. In the bottom actually it was an inner product of all the variables with fully supported vector.

So bottom fanin was actually n. Intuitively, it seems that if you want to reduce that to square root and bottom fanin then what should be the size? Just the definition of permanent which is $n^n$, so $n^{\Omega(n)}$. At the level of depth 3 already, these are reasonable conjectures. And they suggest that permanent and determinant are very different. In the exponent there should be a gap of square root or in the exponent there is some squaring happening

So remember that these conjectures are on $\Omega$. These are lower bound conjectures. So these are both are big open questions if you prove both these conjectures what do you get? It will imply that VP $\neq$ VNP. You do not even need to show 1, you just show 2. If you show that permanent requires $n^{\omega(\sqrt{n})}$ size $\Sigma\Pi\Sigma$ $\sqrt{n}$ circuit, then there cannot be a $poly(n)$ sized circuit for permanent because if there was then the from the previous theorem you would have also obtained $n^{O(\sqrt{n})}$ depth 3 representation of this special type in fact.

So just 2 here actually implies something basic conjecture we started with, Valiant's hypothesis. But we are conjecturing more than that. So here actually we are conjecturing that permanent and determinant they are very different. VP, VNP is not really the same. It does not talk about permanent versus determinant. It is just permanent versus circuits.

And we do not know whether the determinant is complete for circuits. Circuit for all we know maybe slightly more general than the determinant model. So determinant is equivalent to what? Yeah determinant is complete for ABP so which we define the class VBP. But VBP for all we know could be strictly smaller than VP. So this is what we have now. And this gets connected, so Valiant's hypothesis is now connected to the optimality of Ryser's formula right and who knew?

This was not clear before. The VP VNP question is just about whether this Ryser's formula is optimal. It gives you $2^n$ and if you can show that that is optimal you cannot improve that depth 3 representation. Then again by the sequence of connections you will prove VP different from VNP. Then we will dive into the proof and there will be a lot to dive in.

So the proof requires a host of ideas. And initially you will not understand why we are doing, what we are doing in the proof okay. So it will be, for a while it will be magical. So just to guide you through that process. So one common feature is change of basis to express polynomials.

So one common feature is to use powers basis let us say or basis of powers of polynomials instead of the standard basis which is the basis of monomials to express polynomials over this for now characteristic 0 field F. Already this idea is very suspect right? To go from depth 4 to depth 3, why are we looking at powers, sum of powers, right? That already I really have no answer for that. So let us just continue.

So the outline is that from a general circuit we will go to depth 4 representation which is $\Sigma\Pi\Sigma\Pi$ circuit. This you have already seen. Then we will, from here we will go to something you can call it depth 5. So $\Sigma \wedge \Sigma \wedge \Sigma$. So wedge ($\wedge$) is like a product gate, but all the inputs are the same. So since the inputs are the same, your multiplication is just actually giving you a power of this one single input of the first input.

It is a special multiplication gate and then it is depth 5 now. So we actually increase the depth. We will go the opposite way, instead of reducing the depth. And from there we will suddenly jump for some reason to $\Sigma\Pi\Sigma$. So from 4 we go to 5 and from 5 we go to 3. This we will do so okay let us specialize this to, the field is the field of rationals.

Because given a rational polynomial, this, $\Sigma\Pi\Sigma$ will actually be complex, will have complex coefficients. But that is also so you are going to have constants above your field. So then you have to in one step, you will have to come down to rationals, $\Sigma\Pi\Sigma$ circuit over rationals. This will be the end. So we will convert our polynomial f from general circuit to $\Sigma\Pi\Sigma$ circuit over Q.

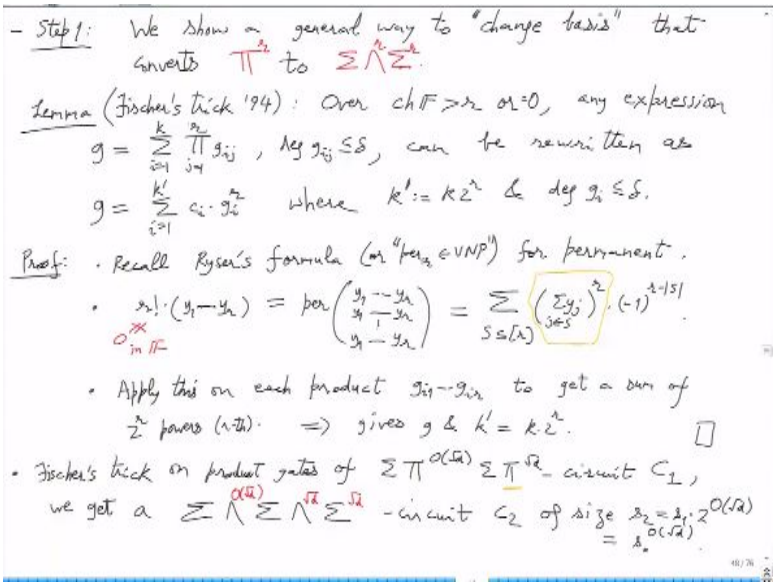
And we have to keep analyzing the size. So this is already done, this is step 0. This is step 1, step 2 and step 3. These will be the 3 new steps that we have to describe. Is the outline clear? **"Professor - student conversation starts"** Is it because of the depth 5 circuit we are getting the characteristic of 0. So step 1 requires characteristic 0. Only step 1. **"Professor - student conversation ends".**

Other steps are characteristic independent. Step 0 definitely was. And step 2 and step 3 will also be characteristic independent. Well, step 3 you have to define what does it even mean? So step 3 is actually the weakest step. It is not of great interest. Step 1 and step 2 are the most interesting ones. But step 1 is the reason why we need characteristic 0, or bigger than the degree d. Step 2 is also very interesting.

But fortunately, it works for any field. Does not matter. I mean, for any characteristic as long as the field is big, which we can always assume. Just to recall step 0 is done. So what is what do you mean by done? So let f have a circuit of size $s_0$. Let us call it $C_o(\overline{x_n})$ of n variables. So by this depth 4 chasm or depth 4 reduction we get a size $s_1 = s_0^{O(\sqrt{d})}$.

Homogenous $\Sigma\Pi\Sigma\Pi$ circuit, let us call it $C_1$. So $C_0$ you start with. $C_1$ is your depth 4 with this these special product fanins both square root d and the size is $s_0$ to the $\sqrt{d}$. Degree of f is d. So now we want to convert this into $\Sigma\wedge\Sigma$. So basically I want to replace these general multiplication gates by very special ones. So that will be step 1.

**(Refer Slide Time: 30:59)**




How do you change the product gates? So we show a general way to change the basis that converts $\Pi$ to $\Sigma$ wedge. So change of basis will basically be this changing product gate to wedge. But then this as you can immediately realize is impossible, but it somehow will become possible if you take a sum of powers. So we are basically writing a product as a sum of powers.

This product is your elementary symmetric function and previously you were only interested in expressing it as a function of sum of powers. But here, it is even stronger, we want it to be a sum of powers, not a function. There should be no extra

function. This will just be a syntactic translation. This does not need any assumptions on what the product is.

So this is actually given by Fischer's lemma or Fischer's trick. So it is a simple statement if the characteristic of the field is bigger than r or 0, any expression g where you have sum of products of polynomials in the product you are multiplying r things. So you have this kind of depth 4 representation where in every product gate there are only r inputs. Characteristic is bigger than r.

This you can rewrite as a sum of powers. So we are replacing each of these products by a sum of powers and then we are taking the sum. So ultimately g can be expressed as a sum of powers. So what is $k$` ? It is at least k, but how much will it blow up? We will show that it blows up by $2^r$ exactly.

Every product we will express as a sum of $2^r$ many r$^{th}$ powers. And the degree of $g_i$ does not blow up. That bound remains the same. Essentially, we will just take linear combinations of $g_{ij}'s$. So whatever was there degree bound remains the same remains for $g_i$. This proof you have seen in Ryser's formula. The proof you have already seen actually.

So recall Ryser's formula proof which was basically this why is permanent in VNP, that proof. So $per_n$ is in VNP. So that there you have seen the ideas. So this $y_1 \cdots y_r$ this monomial. Right, this is essentially permanent of a matrix. What is the matrix? Same rows. So this equals rows, permanent of this is just $r!$ times the product of these row entries.

And in that proof you were able to express this permanent as a sum of products of linear polynomials where these linear polynomials are just inner products. So you can just directly write it as this for all subsets. So well the only new thing is this, this part. Now this is a perfect power. And why is that happening? It is an r$^{th}$ power because in

that old formula, these linear polynomials you got in the product gate, they were the same inner product.

So now since all the rows are the same, you get the same factor. In other words you get r[th] power. So $y_1 \cdots y_r$ this product can be written as a sum of powers with very simple coefficients, but the problem is this $r!$. So $(r!)^{-1}$ should exist. Otherwise even if it vanishes, then you do not get anything about $y_1 \cdots y_r$. So this should not vanish, that is all.

So as long as it does not vanish, we have a representation for the product for general multiplication gate. And then we can apply this on each product. **"Professor - student conversation starts"** Are we converting the product gate to be sum of powers of sum, right. Sum of powers of linear polynomials. Yeah, so that should be like $\Sigma \wedge \Sigma$ right? Oh, in the idea you are saying.

Yeah, sure. That is true, yes. **"Professor - student conversation ends".** I did not mean the sum of monomial powers. That would be impossible again. So apply this on the i[th] product gate, just computing $g_{i1} \cdots g_{ir}$. So you will replace these $y_j's$ by $g_{ij}'s$. It is just a simple combination of the $g_{ij}'s$. So the degree is obviously bounded and what is $k$` ? So how many summands, how many powers are there?

The number of subsets, right so that is $2^r$ to get a sum of $2^r$ powers r[th] powers which implies the lemma statement. The cost is that fanin increases exponentially in r. But it is a simple exponential function. Can you do this thing over other characteristic? So will this Fischer's trick exist for other characteristic?

So that is in the assignment, solve the assignment. There is a question that it is impossible. So this business is actually I mean this $\Pi$ to $\Sigma \wedge \Sigma$ representation, this basis change, requires characteristic to be large. Otherwise it does not exist. It is provably impossible. Once we have this, we will use it on $\Sigma\Pi\Sigma\Pi$ representation. So on product gates with fanin $\sqrt{d}$ on circuit $C_1$ we get.

So yeah this has to be thought about carefully. So you have these two layers right on the bottom $\Pi$ you will apply this transformation. **"Professor - student conversation starts"** So take the polynomial for characteristic p, take the monomial $x_1 x_2 x_p$. There are only 3 variables. Yeah, so let us say characteristic is 2 and you are looking at x 1, x 2. Yeah.

So now if I can express it as $c_i g_i$ to the sure yeah that would mean that by binomial I can that means $x_1 \cdots x_p$ is like power t which is not true. This is a multinomial. Yeah, so you have to solve it in the assignment. **"Professor - student conversation ends".** We cannot discuss all the details here. There are some details. You have to also remember that this r is merely an upper bound, this when you are multiplying r things, many of these things may just be 1.

So when you are looking at $x_1 x_2$, it does not mean that you have to express it as a sum of squares. Maybe you can express it as a sum of cubes or sum of force powers or some of hundred powers. **"Professor - student conversation starts"** So in the g equal to $c_i g_i$ to the r that is like $r$` where $r$`, no not $r$`. It is an upper bound on the multiplicative product fanin. **"Professor - student conversation ends".**

But many of these things, all of these things maybe one. It is merely an upper bound. So when we say that the basis change is impossible, then you have to show that no powers whatever be the exponent can produce your polynomial. Yeah, but do not get distracted here, can do it at home. So what you have to think here is what happens when you replace this pi by Fischer's trick.

So this is not a single $\Pi$, obviously it is in the layer, bottom layer. So there are many multiplication gates. Each of them you will replace by a $\Sigma \wedge \Sigma$ and that already has increased the depth by one. This is now depth 5 and what happens to the top layer of multiplication gates? So when you replace them by $\Sigma \wedge \Sigma$ since it is already sandwiched between sigmas this will not increase the depth.

Okay, so you get depth 5. So this gives you a $\Sigma\wedge\Sigma\wedge\Sigma$ circuit $C_2$ of size $s_2$ which we have to calculate. But first let us look at the fanins here. So what is the fanin of the bottom wedge? So since it was $\Pi^{\sqrt{d}}$, so if this is $\Pi^r$ then $\Pi^r$ is going to $\wedge^r$. And not only that, this bottom $\Sigma$ is also just r. This is the change of basis we have deduced from Fischer's trick.

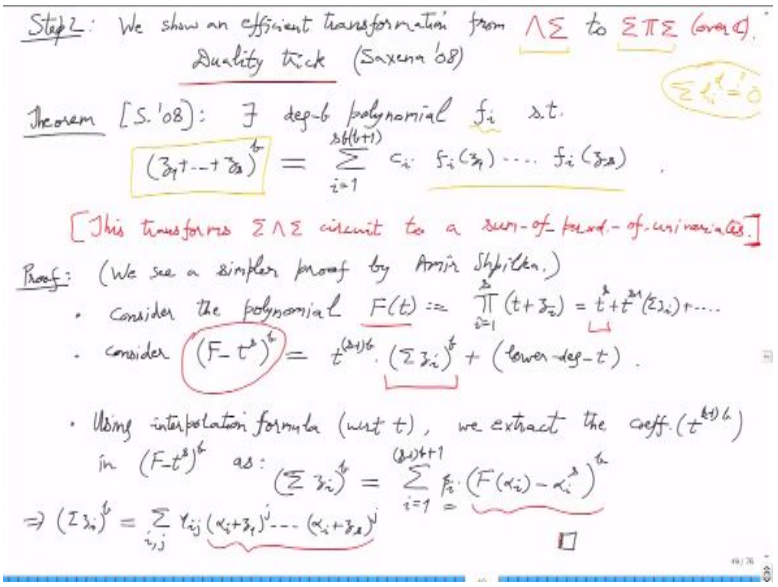
So using that you get these r $\sqrt{d}$, both of them and what else do you know? The top $\wedge$ that is also $O(\sqrt{d})$. Others are general. So they can be very large. So these 3 fanins you have guaranteed to be small and what happens to the size. So $C_1$ was of size $s_1$ let us say. So on top of that how much multiplicative growth are you doing? So each transformation will cost you by $2^{\sqrt{d}}$.

So this is $2^{O(\sqrt{d})}$. That is it. You do it $s_1$ times, which is ${s_0}^{O(\sqrt{d})}$. So that does not change. So we actually without paying any price, you have gone from this depth 4 chasm to a very special depth 5, which is

$$\Sigma\wedge^{O(\sqrt{d})}\Sigma\wedge^{\sqrt{d}}\Sigma^{\sqrt{d}}.$$

Well, but what have we gained? This is no gain, right? This is just increasing the depth. So now what do you want to do?

**(Refer Slide Time: 48:06)**



So that will be our step 2. So step 2, now we want to do the opposite which is convert $\wedge$ to $\Pi$ gate. So we show an efficient transformation from $\wedge\Sigma$ to $\Sigma\Pi\Sigma$ over complex. Wait what did I define step 2 as? Step 2 here is, so what I am saying is this $\wedge\Sigma$, I will convert this into $\Sigma\Pi\Sigma$. And the other one also I will convert to $\Sigma\Pi\Sigma$. And how will that help?

You get back to depth 4, right? That you do not want to do. So why are we doing all this? Yeah, so you will see. While we do this conversion, you will see that you will actually in the end when you do these things simultaneously, you will get to depth 3 and not back to depth 4. You will see why from the trick. So this thing worked well so I christened it, duality trick in 2008. So I will show that.

So let me state that. So the duality trick is the following theorem. So there exists degree b polynomial. Maybe I should mention why I call it a duality trick. So it is duality because in some sense this product addition gate is being swapped. So powering is being written as a sum of products. It is in some sense I mean obviously it is a transformation, but in some sense it is also a dual

So quantitatively what happens is any degree b polynomial sorry any sum of variables let us say s variables raised to b. So $b^{th}$ powers of linear form can be expressed as a sum of products of univariate polynomials. So let me write that down. So there exists degree b polynomial $f_i$ such that the following identity holds.

The left hand side is power of a linear form and the right hand side is basically you are evaluating $f_i$ at these different variables of the linear form, taking the product and then taking a sum. Okay, so the sum of products of these univariates will give you the power and this will be an efficient transformation. So there exist these polynomials $f_i's$.

Now obviously LHS you can just expand fully, that also will give you a sum of monomials and obviously a monomial is of this type. It is just multiplying univariates

in fact distinct variables. But what is the problem? The problem is that they are too many. So that is not an efficient transformation. So that is also a dual but it is an inefficient dual. This on the other hand is an efficient dual because you are using only $(sd)^2$ many products.

So that is the point of this over the trivial representation, it is a non-trivial expansion you can think of it like that. This transforms $\Sigma \wedge \Sigma$ circuit to sum of product of univariates. For duality trick, well, motivation is to solve the identity testing problem for this $\Sigma \wedge \Sigma$ model, which is the sum of powers of linear forms. So $\sum (l_i)^d$ model you want to test whether it is 0.

So that is a natural model, but then there is as always, there is an exponential blow up happening in just ${l_1}^d$. So how do you test whether these things sum up to 0 efficiently? So for that I convert ${l_1}^d$ to this dual and then I can do identity testing for this case using linear algebra. That I think we will see if we get time towards the end of the course. I will not give my proof because that is complicated.

It uses exponential functions and what not. So I will give a proof that is much clearer by Amir Shpilka. So this is just by considering auxiliary polynomials and then doing interpolation. So the auxiliary polynomial he considers is, so remember that we are working with only $(z_1 + \cdots + z_s)^b$. So we have reduced all our model difficulties into just this case. We are only focusing on the sum of distinct variables raised to b and we want to express it in an alternate yet efficient form.

So can you guess what this auxiliary polynomial will be using these $z_1$ to $z_s$. So a polynomial whose roots are $z_1$ to $z_s$ right, kind of. So let us consider this. So let us consider this auxiliary polynomial where the roots are essentially $z_i's$ or minus of that and you see a coefficient that computes the sum of $z_i's$. So this is equal to

$$F(t) = \prod_{i=1}^{s} (t + z_i) = t^s + t^{s-1} \left( \sum z_i \right) + \cdots$$

and then more complicated things.

But we will not care about them. So we will bring this $t^s$ on the LHS so that $\Sigma z_i$ becomes the leading term, okay. So consider

$$(F - t^s)^b = t^{(s-1)b}\left(\sum z_i\right)^b + (\textit{lower degree terms in } t).$$

So now what do we want to do? We want to extract this, right. So think of it as given that you want to extract the highest degree t monomial.

So how do you extract, how do you extract coefficients from a polynomial? Here you cannot quotient because the remaining are lower degree. If you quotient then this leading will be killed. So better is to use interpolation. So you actually evaluate this LHS, this function in T for various values of t and then take a suitable linear combination so that all these lower degree t terms get canceled.

These coefficients do not contribute. And the only thing that remains is this leading coefficient $(\sum z_i)^b$. So using interpolation formula with respect to the formal variable t, so we want to eliminate t basically. There are many variables, but our interpretation is that t is our only variable. Everything else is kind of a constant. And I want to extract this particular monomial in t. I want to eliminate t.

So there is an interpolation formula that also I wanted to do, but I think I will skip it because it is for this crowd it is standard. You all know what interpolation is. So let me skip the Lagrange interpolation formula and just say that we can extract using it the coefficient of $t^{(s-1)b}$ in $(F - t^s)^b$. So $(\sum z_i)^b$ will be the value that you will get and this you will get by taking a linear combination of evaluations of that function at $\alpha_i$ and linear combination with $\beta_i$.

And how many $\alpha_i's$ will you have to try? How many evaluations do you have to look at? That is basically the degree with respect to t plus one. So that is $(s-1)b+1$. So those many coefficients are there and if you want to focus on if you want to extract

one of them, still you have to evaluate the polynomial at maximum number of points and then there will be these betas which will be special.

They will be functions of $\alpha_i$ such that when you take the linear combination everything cancels out and you only get the coefficient that you wanted, which is $(\sum z_i)^b$. So this implies that $(\sum z_i)^b$ is so you can expand it. This is just using binomial expansion. Treat F as a single variable so this is just you just have a difference of two things raised to b.

So you can expand it out and F then is just a product. So you get a sum of product representation. So ultimately you get something like

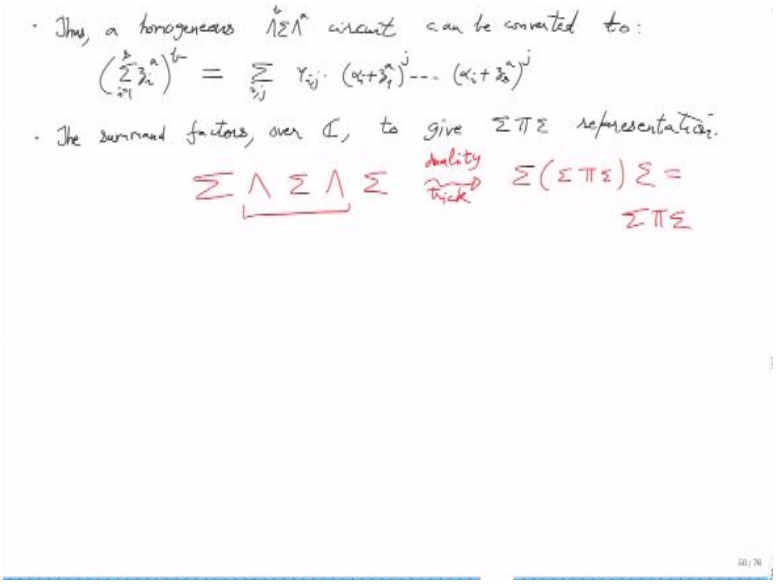
$$\left(\sum z_i\right)^b = \sum_{i,j} \gamma_{ij} (\alpha_i + z_1)^j \cdots (\alpha_i + z_s)^j .$$

So these j's are all the, these exponents are all the same here. In this part the exponents are all the same, because this big F is a single thing and then this is being raised to j. So we actually are getting big F raised to j evaluated at some alpha i. So that is this product.

Many such products for all these $\alpha_i's$ and then linear combination. $\gamma$ depends on $\alpha$ here. This is the $f_i$ that we were talking about. This $f_i$ is very simple, it has a very simple form. So that finishes the proof of duality trick. And it is characteristic independent also. Now what? Now we see once you have seen this now we have to apply this. This was a detour.

So let us go back to where we were stuck. So we were stuck here, right. Now this $\wedge\Sigma$ and the bottom $\wedge\Sigma$ we will use duality trick, let us say we will use it first on the bottom $\wedge\Sigma$. So that will give us a $\Sigma\wedge\Sigma$ representation. So the top $\wedge$ is actually acting on? Okay let us write that.

**(Refer Slide Time: 1:05:14)**




So thus a homogeneous $\wedge\Sigma\wedge$ circuit yeah so maybe I should have said that in the beginning that we will now do it like this. We will focus on this $\Sigma\wedge\Sigma\wedge$. So this $\wedge\Sigma\wedge$ is on which we will apply the duality trick. That will give me $\Sigma\Pi\Sigma$. So let us see why or how. So this $\wedge\Sigma\wedge$ can be converted.

So you are taking power of sum of powers. So in general this is the format $\left(\sum_{i=1}^{s} z_i^{a}\right)^{b}$. So the bottom fanin is a and the top fanin is b. So this is the general format and this now by the duality trick becomes what? So the duality trick actually gave you a very special form. So in that special form when you substitute, when you replace $z_i$ by $z_i^{a}$ it still remains rather special.

It becomes

$$\sum_{i,j} \gamma_{ij}(\alpha_i + z_1^{a})^{j} \cdots (\alpha_i + z_s^{a})^{j}.$$

And I am saying that this is $\Sigma\Pi\Sigma$ representation, is it? So it is a sum of products, but inside the product, this is not a linear polynomial in $z_1$. It is a general polynomial in $z_1$. So why am I calling it $\Sigma$? Technically it is depth 4. Well you factor it. So it is a univariate $(\alpha_i + z_1^{a})$. So over complex, it can always be factored completely. Factors will be linear polynomials.

So it is actually a product of linear polynomials. So the summand actually factors over complex to give $\Sigma\Pi\Sigma$ representation. Yes, so I think I will stop now at this point. We just have to give the quantitative statements now. So qualitatively what we have done is we had this $\Sigma \wedge \Sigma \wedge \Sigma$ and this middle part we will convert it into $\Sigma\Pi\Sigma$ by the duality trick.

Okay, so that is the duality trick with factorization. Duality trick works for any field at least Amir Shpilka's proof and factorization you can do over the algebraic closure. And so together you will have a $\Sigma\Pi\Sigma$ but since it is sandwiched between sigmas it is equal to $\Sigma\Pi\Sigma$. So that is what you get. So that is the moral reason why depth 4 can be reduced to depth 3 for special fields.

We will look at the parameters and especially the fanin, what is the bottom fanin. So because it is actually not just a run of the mill depth 3 representation. It is a very special depth 3 representation. So any general circuit you can actually reduce to a special depth 3 representation. So if you can prove lower bounds against this representation, then you have proven lower bounds against general circuits.

Okay, that is the moral of the story.