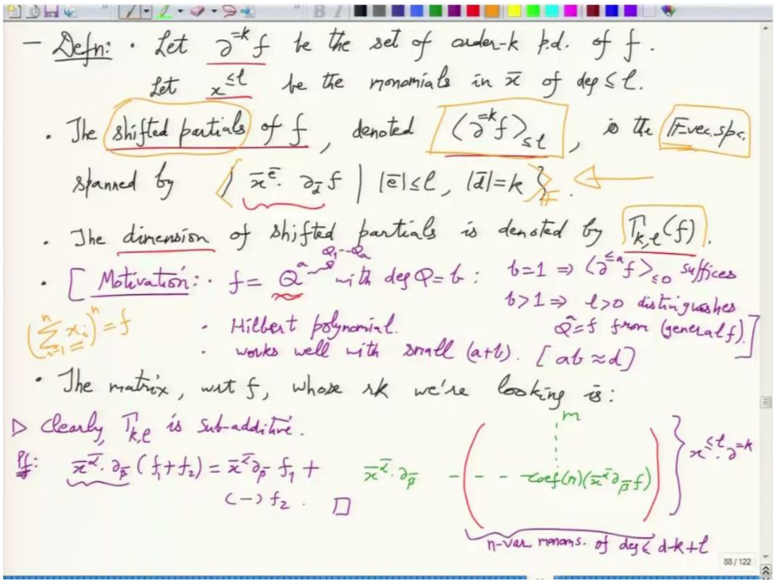
Arithmetic Circuit Complexity
Prof. Nitin Saxena
Department of Computer Science and Engineering
Indian Institute of Technology Kanpur

Lecture-20
Arithmetic Circuit Complexity

(Refer Slide Time: 00:21)



– Defn: · Let $\partial^{=k} f$ be the set of order-k p.d. of f.
Let $x^{\leq \ell}$ be the monomials in $\bar{x}$ of deg $\leq \ell$.

· The shifted partials of f, denoted $\langle \partial^{=k} f \rangle_{\leq \ell}$, is the $\mathbb{F}$-vec.sp. spanned by $\{\bar{x}^{\bar{e}} \cdot \partial_{\bar{d}} f \mid |\bar{e}| \leq \ell, |\bar{d}| = k\}$.

· The dimension of shifted partials is denoted by $\Gamma_{k,\ell}(f)$.

· [Motivation: · $f = Q^{a}$ ($Q_1 \cdots Q_a$) with deg $Q = b$: $b=1 \Rightarrow \langle \partial^{\leq a} f \rangle_{\leq 0}$ suffices
$b>1 \Rightarrow \ell>0$ distinguishes $\hat{Q}=f$ from (general f).]

$(\sum_{i=1}^{n} x_i)^n = f$

· Hilbert polynomial.

· works well with small $(a+b)$. $[ab \approx d]$

· The matrix, wrt f, whose rk we're looking is:

▷ Clearly $\Gamma_{k,\ell}$ is sub-additive.

Pf: $\bar{x}^{\bar{\alpha}} \cdot \partial_{\bar{\beta}}(f_1+f_2) = \bar{x}^{\bar{\alpha}} \partial_{\bar{\beta}} f_1 + (-) f_2$. □

$$\left.\begin{array}{c} \\ \bar{x}^{\bar{\alpha}} \cdot \partial_{\bar{\beta}} \\ \\ \end{array} \underbrace{\begin{pmatrix} & m & \\ - - - & \mathrm{coef}(m)(\bar{x}^{\bar{\alpha}} \partial_{\bar{\beta}} f) & \\ & & \end{pmatrix}}_{n\text{-var. monoms. of deg} \leq d-k+\ell}\right\} x^{\leq \ell} \cdot \partial^{=k}$$

What we did what we started few weeks ago is shifted partials measure. So, what are shifted partials of f. So, this is there are two parameters one is to differentiate. So, that is ∂ so all the derivatives of order k and the second parameter is l which will refer to after you have differentiated you multiply by a monomial of degree up to l. So, this is the space we are looking at. So, you can think of this as a set of derivative time some monomial and all of them all possible such things or you can also look at the vector space.

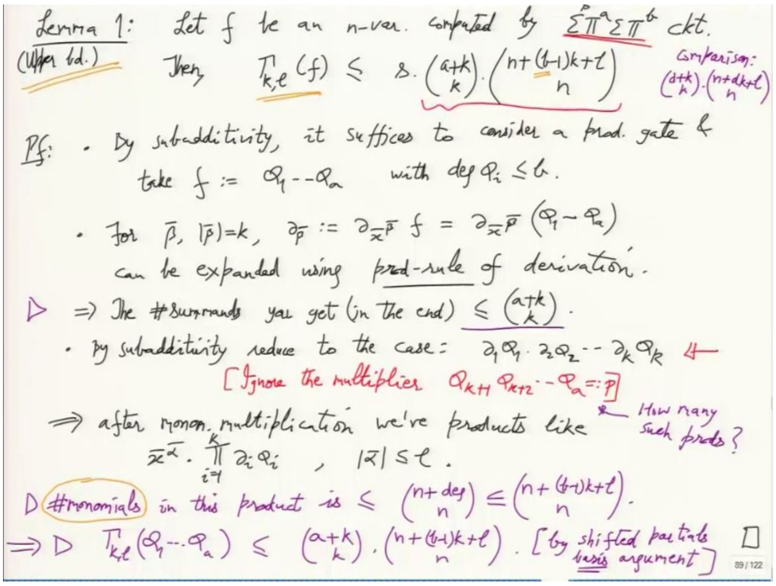
So, this field is F and you are looking at this F, vector space over the field F. So, we can use this notation to generate the vector space over these field of constants. The dimension of this will be denoted by Γ, $\Gamma(f)$ is the dimension. And very quick motivation for this was that if you look at a power of a polynomial then so in the simplest case if you just look at $(x_1 + \cdots x_n)^n$. So, it actually produces 2^n monomials, it produces exponentially many monomials.

But if you look at its derivatives and the rank that is only n so, this $\left(\sum_{i=1}^{n} x_i\right)^n$, if you look at this polynomial then it has a huge number of monomials but, the measure is very small, I mean the measure if you just look at the derivatives then it is very small it is only n in this shifted partials, we are also multiplying by monomials. So, why are we doing that will be clear from the calculations? So, because the thing is that we will not be working only with linear polynomials but quadratic cubic and higher degree in general degree b.

So, they are actually we have to look at a much bigger space the space of derivatives is not enough for us. We have to look at a bigger scale. So, we actually multiply by monomials and although the scale is bigger still when we will compare this dimension with the dimension of determinant, corresponding dimension of determinant there will be a gap. Now, this is currently you just have to take it on faith that derivatives will not work, but shifted derivatives will work.

So, you can look at it as a matrix. So this is the matrix interpretation which is just every row you are applying the operator and then writing the polynomial in full, monomial by monomial the coefficients are written and then look at the rank.

(Refer Slide Time: 04:01)



Lemma 1: Let f be an n-var. computed by $\underline{\Sigma^{s}\Pi^{a}\Sigma\Pi^{b}}$ ckt.
(Upper bd.) Then, $\underline{T_{k,\ell}}(f) \leq s \cdot \underbrace{\binom{a+k}{k} \cdot \binom{n+(b-1)k+\ell}{n}}$ Comparison: $\binom{d+k}{k} \cdot \binom{n+dk+\ell}{n}$

Pf: • By subadditivity, it suffices to consider a prod. gate & take $f := Q_1 \cdots Q_a$ with $\deg Q_i \leq b$.

• For $\bar{\beta}$, $|\bar{\beta}|=k$, $\partial_{\bar{\beta}} := \partial_{\bar{x}^{\bar{\beta}}} f = \partial_{\bar{x}^{\bar{\beta}}}(Q_1 \cdots Q_a)$ can be expanded using prod-rule of derivation.

▷ ⇒ The #summands you get (in the end) $\leq \binom{a+k}{k}$.

• By subadditivity reduce to the case: $\partial_1 Q_1 \cdot \partial_2 Q_2 \cdots \partial_k Q_k$ ⇐
[Ignore the multiplier $Q_{k+1} Q_{k+2} \cdots Q_a =: P$]

⇒ after monom. multiplication we're products like $\bar{x}^{\bar{\alpha}} \cdot \prod_{i=1}^{k} \partial_i Q_i$, $|\bar{\alpha}| \leq \ell$. ← How many such prods?

▷ #monomials in this product is $\leq \binom{n+\deg}{n} \leq \binom{n+(b-1)k+\ell}{n}$.

⇒ ▷ $T_{k\ell}(Q_1 \cdots Q_a) \leq \binom{a+k}{k} \cdot \binom{n+(b-1)k+\ell}{n}$. [by shifted partials basis argument] □

89 / 122

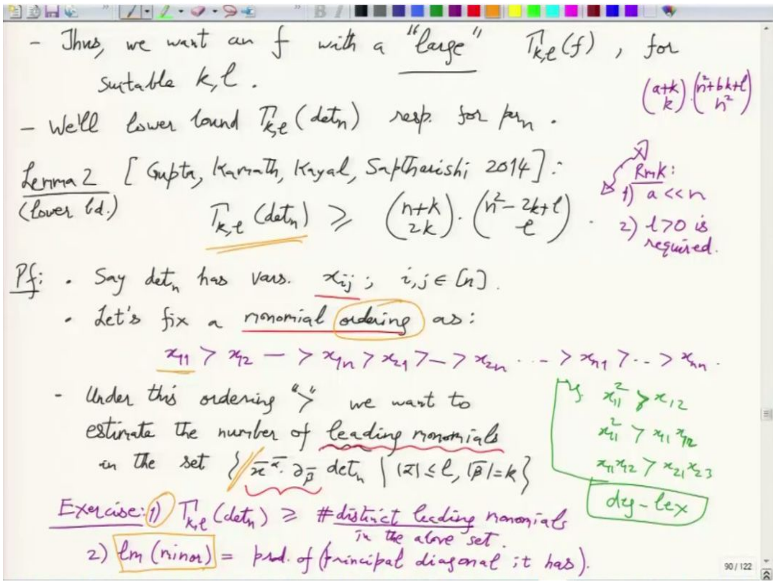
So, the next thing we prove was the upper bound. So this upper bound lemma says that shifted partials of ΣΠΣΠ with which is size s. In fact, the top finding is s and then the top product finding is b a and the bottom product finding is b. So, with these parameters and k, l parameters, the measure is bounded by the product of these 2 binomials. So, this is considered a good upper bound because you are getting $\binom{a+k}{k}$ and not $\binom{d+k}{k}$.

If you look at a general polynomial of degree d, so, you would have gotten $\binom{d+k}{k}$ in our application, we will assume a to be something like $\sqrt{b}$. So this estimate will be better, it is a better approach; b appears in a more complicated way. And so the way we will deal with this second binomial number is we will pick a larger l because l is in our hands, k and l r a lot of our parameters, so we will actually pick k to be small, but l to be large.

And then the two things will still give us a lower bound using determinant. There is no simple answer at this point. I am just jumping the gun and I am telling you about what k and l we will choose the motivation may not be very clear at this point. The only thing is that this l is already coming with the big things these already n + b k sitting there. So, we will actually take l to b comparable to that there is no point taking l to b very small.

Because then l will have no effect then it will be likely only looking at derivatives. So, since l should have some application otherwise we would not even in introduce shifted portions. So, actually l will be taken large k will be taken very small. So, any questions about this proof this upper bound we have already done.

(Refer Slide Time: 06:56)



- Thus, we want an f with a "large" $T_{k,\ell}(f)$, for suitable k, ℓ.

$\binom{a+k}{k} \cdot \binom{n^2+bk+\ell}{n^2}$

- We'll lower bound $T_{k,\ell}(\det_n)$ resp. for per_n.

Lemma 2 (lower bd.) [Gupta, Kamath, Kayal, Saptharishi 2014]:

$$T_{k,\ell}(\det_n) \geqslant \binom{n+k}{2k} \cdot \binom{n^2-2k+\ell}{\ell}.$$

Rmk: 1) $a << n$ 2) $\ell > 0$ is required.

Pf: • Say $\det_n$ has vars. x_{ij}; $i,j \in [n]$.

• Let's fix a monomial ordering as:

$$x_{11} > x_{12} - > x_{1n} > x_{21} > - > x_{2n} \cdots - > x_{n1} > \cdots > x_{nn}.$$

- Under this ordering "$>$" we want to estimate the number of leading monomials in the set $\{\bar{x}^{\bar{\alpha}} \cdot \partial_{\bar{\beta}} \det_n \mid |\bar{\alpha}| \leq \ell, |\bar{\beta}| = k\}$

$x_{11}^2 > x_{12}$

$x_{11}^2 > x_{11}x_{12}$

$x_{11}x_{12} > x_{21}x_{23}$

deg-lex

Exercise: 1) $T_{k,\ell}(\det_n) \geqslant$ #distinct leading monomials in the above set.

2) $\boxed{\mathrm{lm}(\text{minor})}$ = prod. of (principal diagonal it has).

So, the thing we could not finish and which is more complicated is the lower bound on determinant. So, now we will complete this. Lower bound for the determinant will be the proof will be based on what is called monomial ordering and the reason is that when you differentiate determinant you get a minor and these minors are complete polynomials. So, just a priori it is not even clear whether the binomial whether the monomials are linearly independent.

Even to prove that you have to use monomial ordering that is the simplest proof. So, we will continue with that technique. So, now what we are doing is we are differentiating determinants so we get either monomial minor or 0 so, when we get a minor we actually will multiply with monomials because this is what shifted partials do they differentiate and then multiply by a monomial.

Then we will extract the leading monomial and if we can show that we get a large number of distinct monomials leading monomials in this way, then that will lower bound the rank because if in a set of polynomials if the leading monomials are distinct then that is a lower bound on the rank I mean the proof is simply by contradiction you take a linear combination of these polynomials from which you got distinct leading monomials and suppose that combination vanishes equal to 0.

So, now leading monomial on the left hand side and leading monomial on the right hand side will give you something nonzero equal to 0. So, that is a contradiction. It is just a contradiction you see that if the leading monomials are distinct then there is no chance of d getting cancelled and hence the polynomials. So, in a direct way if you start with 2 distinct monomials and just add arbitrary monomials of lower order you cannot get dependent belongs x we are just using that fact.

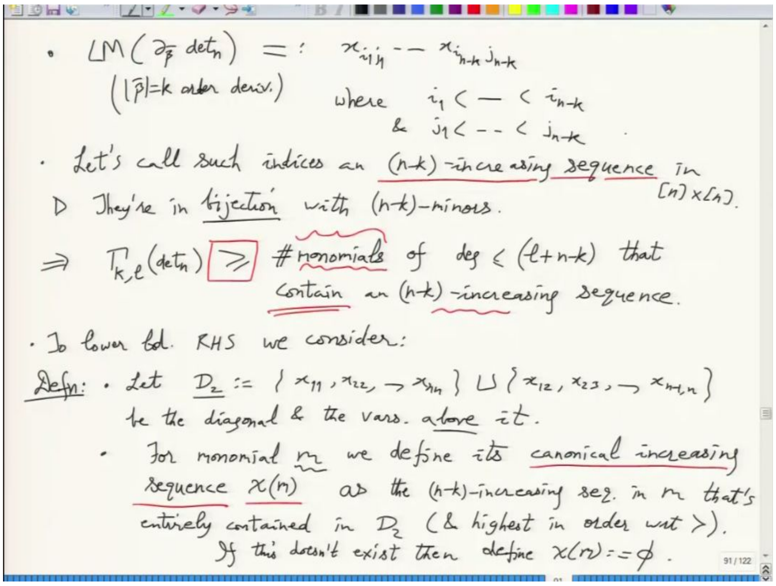
Now this may not be a very good estimate but it will work for us. This is a lazy this may be a lazy estimate, but it will still be high enough. So, we will see that the monomial ordering is just this you start with the top left corner that is the most important location and then decrease along the first row and successively decrease in the next rows. This is our ordering. It will change if you change the ordering drastically. So if you after x_{11}, you jump to and I am just saying that x_{11} is greater than x_{12} will receive a column.

But that is simply it is an isomorphic ordering because there is a permutation which can give you the other thing. But that is well not just permutation. There is a even if you take a random permutation, arbitrary permutation on the variables you might have to consider some other permutation but that will be determined by your permutation yes actually by isomorphism argument anything will work.

So, all those proofs will go through this intuition which we will now see that the focus will always be on the principal diagonal of the matrix and hence the minors. So, under this ordering when you look at the leading monomial of a minor that will be equal to the monomial that is produced by the diagonal that is one observation; the first thing we have already talked about that our measure on determinant will be at least the distinct leading monomials we can produce.

So, now we will study the leading monomial operator on this. So, we differentiate multiply by a monomial and look at the leading monomials give a lower bound. So, any questions about ordering and leading monomials and these 2 properties, so, this must be did last time, I mean the only ways constructive will actually identify these monomials which are distinct.

(Refer Slide Time: 12:44)



- $LM(\partial_{\bar{\beta}} \det_n) =: x_{i_1 j_1} \cdots x_{i_{n-k} j_{n-k}}$ where $i_1 < \cdots < i_{n-k}$ & $j_1 < \cdots < j_{n-k}$.
  ($|\bar{\beta}| = k$ order deriv.)
- Let's call such indices an (n-k)-increasing sequence in $[n] \times [n]$.

▷ They're in bijection with (n-k)-minors.

$\Rightarrow$ $T_{k,\ell}(\det_n) \boxed{\geqslant}$ # monomials of deg $\leqslant (\ell+n-k)$ that contain an (n-k)-increasing sequence.

- To lower bd. RHS we consider:

Defn: • Let $D_2 := \{x_{11}, x_{22}, \to x_{nn}\} \sqcup \{x_{12}, x_{23}, \to x_{n-1,n}\}$ be the diagonal & the vars. above it.

- For monomial m we define its canonical increasing sequence $x(m)$ as the (n-k)-increasing seq. in m that's entirely contained in D_2 (& highest in order wrt $>$). If this doesn't exist then define $x(m) := \phi$.

91 / 122

So, the leading monomial of derivative of the determinant is what is the $\overline{\beta}$ that we are using, that is given by this, $\overline{\beta}$ is I mean the derivative is basically order k. So, we have n - k by n - k minors and that is the number of principal diagonal entries. So, let us define them i_1 j_1 dot dot so on $i_{n-k}j_{n-k}$,when you differentiate by order, when you differentiate by k variables, you will be left with n - k variables.

So, those are the ones says i_1j_1 to $i_{n-k}j_{n-k}$ are the n - k locations that remain and produce the principal diagonal where, so keep them ordered. So $i_1 < \cdots < i_{n-k}$ and $j_1 < \cdots < j_{n-k}$. So i_1j_1 is the most important and so on. Note that this is strict inequality because it is a principal diagnosis. So we will call such sequences. Such indices we will call an (n - k)-increasing sequence in the universe and n cross n.

So in this Cartesian product n^2, if you will call n - k elements in order to be an increasing sequence if this happens, basically, it is a principal diagonal of some minors. So, these are obviously in bijection with n - k minors. So, given an n - k increasing sequence, you have an associated $(n-k) \times (n-k)$ minors and converse is also true. So, it is in bijection that is trivial so, which gives us a lower bound indication on determinant. So, the $\Gamma_{k,l}(det_n)$ is at least the number of monomials.

So, remember that now, there is also this l so we have this n - k minors or increasing sequence and then we are multiplying by n up to l degree monomials. So, we will get, we will put here the number of monomials of degree l + n - k that contain n - k increasing sequence. So, the proof of this is just a simple observation that when you look at a monomial of degree up to l + n - k that contains an n - k increasing sequence.

So, you can look at the sequence times this extra monomial that is multiplying it and from the sequence you identify the unique minor $(n-k) \times (n-k)$ minors. So, that gives you the operator also the derivative operator which gives you this minors. So, you got the shifted partial that corresponds to this production. So, and these are linearly independent, because when you apply the leading monomial operator you will get exactly this, monomial.

So, since they are distinct, the corresponding operators, operator values are linearly independent and hence, the $\Gamma_{k,l}$ is at least that big here. I want to prove this, that is all. So for this I just have to show operator values that are linearly independent and this is the procedure. So what you are seeing is that there are several operator values which are dependent but that I do not care, I just want to construct enough linearly independent values and that you can construct.

Which are less so, they are linearly dependent as they are lower bound on the rank. So, now our construction will be these monomials. We now want to construct them, there might be many ways to construct them, but the way the authors did it was in a very special way, which is just focusing on the principal diagonal n entries and the entries above this. So, which is n + n - 1 variables. So, they use only this variables, in this band to come up with all these monomials it will be even more special.

The thing is it will be easier to analyze and count. So to lower bound the right hand side we consider the following band. So $D_2 = \{x_{11}, x_{22}, \cdots, x_{nn}\} \cup \{x_{12}, x_{23}, \cdots, x_{n-1,n}\}$ so these are 2 n - 1 variables. This is the diagonal and the variables above. So, basically our (n - k)-increasing sequences will all be contained in D_2 and we to produce the monomials on the RHS we will multiply that subset by some extra monomials.

So, for monomial m we define its canonical increasing sequence χ_m, as what? So for any monomial what should be χ_m? So, this canonical, I mean this increasing sequence which you want to identify with a minor, it should be completely contained in D_2. That is what we want to define. So as the (n - k)-increasing sequence in m. That is entirely contained in D_2 and it should be the highest in order well but so are geographically our biases towards the center.

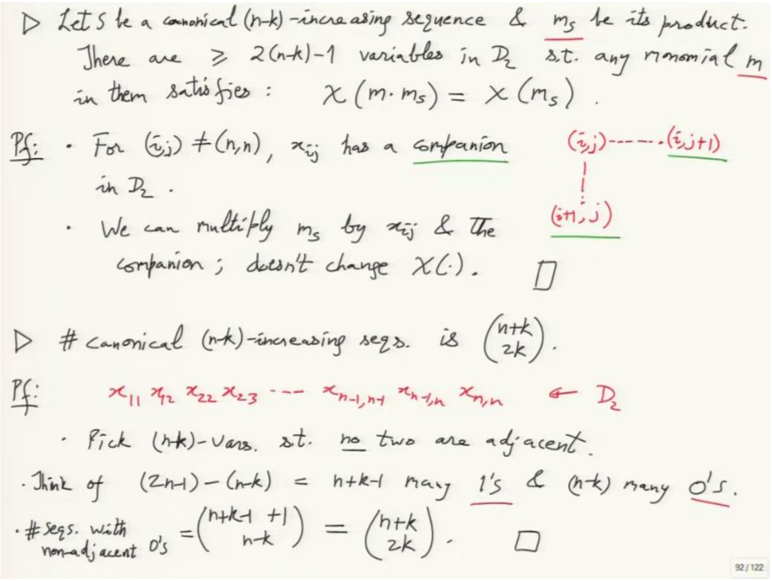
So the minor that we look at should be around the principal diagnosis and maybe slightly allowed to be slightly shifted up but not down it but there are minors which are not of this type so there are minors which are extreme left or extreme right. So they so for those minors what is χ_m? Well, you define it to be empty. So if the latter does not exist, if this does not exist then define, so that is also possible.

But that will be 'don't care' cases for us. So we will be interested in and we will just count these m$_s$ with χ_m 's χ_m exists, and χ_m is distinct. So, once we have defined the map χ, what we are interested in is the size of the pre image of χ so, how many monomials are there for which χ maps to a single quantity? So, how big is that size and then we will multiply it by the number of χ_m ? So, that will be our bound.

So, that will be our, in the lemma statement we are product of 2 binomials, so that product will now get. As I said there are these extreme minors also. We are counting here only the ones that are central in the matrix because it is easier to count them and they intuitively dominate. They are the ones that are the most. To come up to a large no of course, you will not leave before seeing a proof that is a promise yes, and l will be quite large.

So, basically this shifted partials will increase the canvas. So it goes to nearly exponential or even exponential rank, which a priori may seem useless because determinant also has exponential rank. Then it will quantify the exponential. So, here this exponential will be smaller than the determinant exponential and that the ratio will give us the lower bound which will excel the exponential, that will be possible by l. But you do not see that unless you do the analysis or if you are ultra-genius or no belief always works.

(Refer Slide Time: 26:26)



▷ Let S be a canonical (n-k)-increasing sequence & m_S be its product. There are $\geq 2(n-k)-1$ variables in D_2 s.t. any monomial m in them satisfies : $\chi(m \cdot m_S) = \chi(m_S)$.

Pf: • For $(i,j) \neq (n,n)$, x_{ij} has a companion in D_2 .

• We can multiply m_S by x_{ij} & the companion ; doesn't change $\chi(\cdot)$. □

(i,j) - - - - $(i,j+1)$

$(i+1,j)$

▷ # canonical (n-k)-increasing seqs. is $\binom{n+k}{2k}$.

Pf: $x_{11}\ x_{12}\ x_{22}\ x_{23}\ ---\ x_{n-1,n-1}\ x_{n-1,n}\ x_{n,n} \leftarrow D_2$

• Pick (n-k)-vars. s.t. no two are adjacent.

• Think of $(2n-1)-(n-k) = n+k-1$ many 1's & (n-k) many 0's.

• # seqs. with non-adjacent 0's $= \binom{n+k-1\ +1}{n-k} = \binom{n+k}{2k}$. □

I mean, what monomials to multiply with, let us first analyze that, so let S be an (n - k)-increasing sequence, entirely contained. Well, so we call it canonical. It is the canonical increasing sequence and m_s be its product m_s is the monomial, m_s is essentially a canonical increasing sequence and so there are at least twice many variables in D_2 such that any monomial in them, any monomial m in them satisfies

$$\chi(m \cdot m_s) = \chi(m_s).$$

So, basically if you multiply m_s with m it will not hurt you so, the canonical sequence will not change.

So, you should read this as there are a large number of variables in D_2 which any monomial in them you can multiply and χ will give you the same value. So, this is a way to blow up the rank. So, how do you prove this? Well, the easiest proof is by this picture. So, if x_{ij} is a variable in S. So, and S has n - k variables. For each of these i, j look at the, so remember that i,j is either in the diagonal or above the diagonal because it is a canonical increasing sequence.

So, it has two move associated locations either on the right side which will be same i or below which is same j. It has these two locations, I mean one of the two actually usually both of them will be available to you know if, so, we have to count how will we get doubles the doubles should not be important. I think order n - k no. So, you can also use i j, x_{ij} itself. So, that is why it will be double. So, use x_{ij} which was in s and one of the companions.

Which is available so, if x_{ij} is in the diagonal then you can use something you cannot go down because then you will go outside D_2, but you can go right. If you are above the diagonal x_{ij} is above the diagonal, then you can go down to the diagonal. So, you will have one extra possibility together with x_{ij} and when you multiply with any of these variables to m_s, χ does not change well both of them are smaller, both the companions are smaller, which so whatever is the companion it is smaller.

And if you multiply by x_{ij} making it x_{ij}^2 for example, it does not change the canonical sequence and does not change the value of that. So, then we have one minors here, basically double. So for i j except this corner one, the bottom corner except that, x_{ij} has a companion in D_2. So that is in the picture. And so we can multiply m_s by x_{ij} and the companion. So this will not change χ so that is the proof. So here we use quite strongly we use the monomial ordering the way we have defined.

So, going right or going down decreases the reference in the ordering. That is essentially defines the order. So this will now give us a blow up. Because for 1 canonical increasing sequence, you have all these associated free images, which also work well, so first we count how many canonical sequences there are. So we count them, it is not a question of whether they are distinct. It is a question of counting them. Number of canonical (n - k) - increasing sequence. Yes, so what should be the number so here again, we will use the way we have defined D_2.

So let me just write the expression and the proof we will be using that. So, let us write this right D_2 suggestively in this way so x_{11}, x_{12} then x_{22},x_{23} and so on $x_{n-1\,n}$. So, n - 1 you will get n - 1 n - 1 and n - 1 n and then you will get n, n. So, that is your D_2 in this order and so, canonical increasing sequence will be you pick n - k variables from this such that no 2 should be adjacent because if it is adjacent then see that a substance subscript gets repeated.

So, that violates the definition of increasing sequence. So, it is simply a question of picking n - k things. So, pick n - k variables such that, no two are adjacent in the above order, no in the order the way I have written it in read from this no two should be adjacent you pick n - k out of this. You can observe that if you pick non adjacent variables you and will x_{ij} and $x_{i'j'}$ then they can be extended to an increasing sequence, pair you can extend to an increasing sequence.

So, like x_{11} and $x_{2\,2}$ for example; this is fine. And then x_{12} and x_{23} is also fine. Basically, you want the first indices to be distinct and the first subscript to be distinct and the second subscript also. Exactly so, that fits well with this ordering and so, what is the count now from these 2 n - 1 elements, you want to pick n - x that are that no two are adjacent. That is a class 11th question of counting. So, you have 2 n - 1 things.

And you want to pick n - k. So, remove, those first remove these n - k places and so you will get what? That is n + k - 1. So, basically think of the rest as all ones you have n + k - 1 and the 1’s you want to pick call them 0 between every two 0 you want at least 1 ? So that count will come out to be $\binom{n+k-1+1}{n-k}$. So think of these many 1’s and n - k many 0’s. So just this abstraction will work.

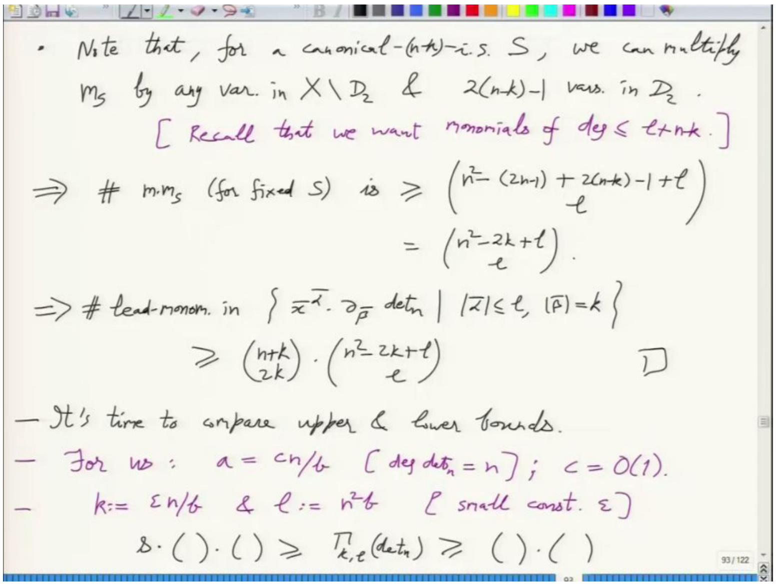
You want to look, you want to count the number of 0 1 strings using these many 0 1 such that 0 should not be closed. So the 0’s do not go along. So they should be seated apart. That is the, so this is the count. So, that is n + k chooses 2 k. So, that is the number of canonical increasing sequences or increasing sequences based in D_2 contained in D_2. Now, we use the above property to look at a look at the size of the preimage of some S. How many m’s are there?

So, we count them so, how many variables we can use to generate m? So we can use variables in D_2 which are these many? Twice and - k - 1 many. And what about the variables outside D_2? Can we use them? Well outside D_2 we can use any variable to multiply because that would not change chi. They are all smaller than what is the argument for? If you pick a variable outside D_2 and multiply with m_s, what is χ?

So, basically, you have to see as an exercise that this extra variable that you picked outside D_2 this cannot give a; cannot contribute to canonical increasing sequence. Actually we are defined canonical to be in D_2. By definition, it cannot contribute here. It is not just an increasing sequence, it is canonical increasing sequence. That is our definition of χ so you just use that definition. It should be entirely contained in it.

So otherwise, we will just call it empty. If you cannot find any such thing. Anything outside D_2 you can actually multiply to m_s and that cannot change χ if previously there was a value of χ, then that will continue to be, if previously there was no value. Then now also it cannot appear just go through that argument.

(Refer Slide Time: 42:04)



- Note that, for a canonical-$(n-k)$-i.s. S, we can multiply m_S by any var. in $X \setminus D_2$ & $2(n-k)-1$ vars. in D_2.

[Recall that we want monomials of deg $\leq l+n-k$.]

$$\Rightarrow \# \; m \cdot m_S \text{ (for fixed } S) \text{ is } \geq \binom{n^2-(2n-1)+2(n-k)-1+l}{l} = \binom{n^2-2k+l}{l}.$$

$$\Rightarrow \# \text{ lead-monom. in } \{ \bar{x}^{\bar{\alpha}} \cdot \partial_{\bar{\beta}} \det_n \mid |\bar{\alpha}| \leq l, |\bar{\beta}| = k \} \geq \binom{n+k}{2k} \cdot \binom{n^2-2k+l}{l} \qquad \Box$$

— It's time to compare upper & lower bounds.

— For us: $a = cn/b$ [$\deg \det_n = n$]; $c = O(1)$.

— $k := \varepsilon n/b$ & $l := n^2 b$ [small const. ε]

$$s \cdot (\;) \cdot (\;) \geq \Pi_{k,l}(\det_n) \geq (\;) \cdot (\;)$$

93 / 122

And so now we know what variables to use. So note that canonical (n - k)-increasing sequence S, we can multiply m s by any variable outside D_2 which is x \ D_2 and 2(n - k) - 1 variables in D_2. So, what does that give us? And also remember that we want monomials or leading monomials of degree upto l + n - k. But because you differentiate by k and multiply up to l degree monomials . So what is this count? So number of m times m_s for fixed S. This is? So well the m_s already has n - k. So, you have a possibility of l degree monomial.

And for that l, how many variables are there? So,

$$\binom{n^2-(2n-1)+2(n-k)-1+l}{l} = \binom{n^2-2k+l}{l}.$$

And hence, we saw now go over all the S and all the pre images for every S. So, that will give you so, the number of leading monomials in the set determinant, operator on the determinant. That is at least $\binom{n+k}{2k} \cdot \binom{n^2-2k+l}{l}$ as promised.

Will you believe this? Have we counted them differently? That is the question. So nobody's claiming that for different α or β or these things are different. We have just said that these many are at least appearing and they are different. We are not counting $\bar{\alpha}$ or $\bar{\beta}$. We are counting the action of lm on this set. The action of lm on this set is what we are counting because we are first coming up with the leading monomial or normal monomial. And then we are claiming that this appears as a leading monomial in the set.

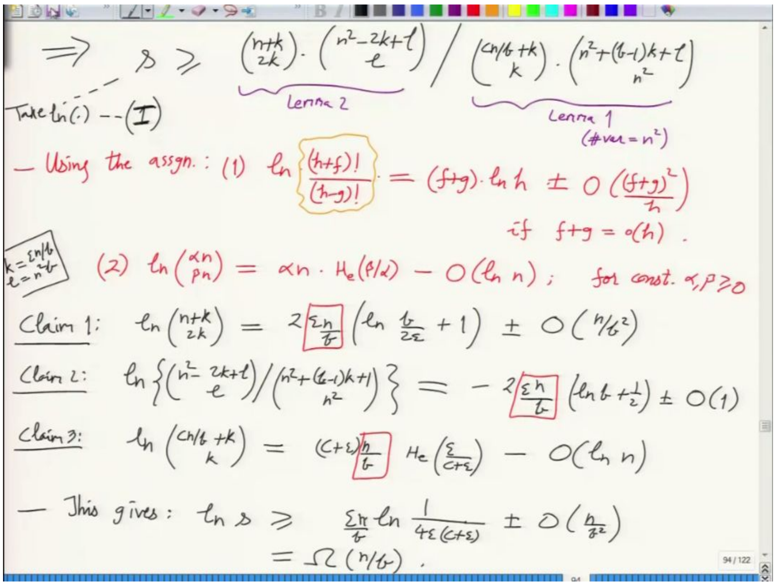
We do not care how many times it appears, that is not our concern. The more it appears, the more $\overline{\alpha}$ or $\overline{\beta}$ there will be, but that is not what we are interested in. So, we have now both upper and lower bounds. So, if you want to take the easy route it will be just to stop here and go home and do the calculations. But let us anyways go through the calculations here. So, some painful calculations follow when you compare the two. So, it is time to compare the two, to compare upper and lower bounds.

What will help is the is the assignment that we just solved. So, in that assignment the binomial estimates were given. We will now invoke them. For us, in the applications, we will take a and b to be when it was $\Sigma\Pi^{a}\Sigma\Pi^{b}$, computing say a polynomial of around degree n so, we want a times b to be n this is the setting we want in the application. So, we will take a to be c n / b where c some constant.

So, think these are as a b = n and n is the degree and it is also the number of well, in the case of determinant it is the degree and the number of variables is n square. So, that intuition is right. So, degree of determinant n equal to n. So, that is the justification for taking a = c n / b c is a constant that will fix in the end. We will take k l parameters as follows. So, k will be around this a because we got a + k choose k. So, we will take it around a, $k = \varepsilon n/b$ and l we will take quite large, comparable to $n^2 b$, ε will be small enough that also we will fix later.

So, remember the lower bound it was s times something, times something greater than equal to the measure of determinant and measure of determinant is at least the lower bound which is something times something. So, if determinant has a $\Sigma\Pi\Sigma\Pi$ expression right then you get the upper bound, which is the first inequality and the lower bound which is the second inequality. So, you will get that s is greater than equal to product by product. So, let us look at that product now.

(Refer Slide Time: 50:50)



$$\Longrightarrow \quad s \geq \underbrace{\binom{n+k}{2k}\cdot\binom{n^2-2k+l}{l}}_{\text{Lemma 2}} \Big/ \underbrace{\binom{cn/b+k}{k}\cdot\binom{n^2+(b-1)k+l}{n^2}}_{\substack{\text{Lemma 1}\\(\#\text{var}=n^2)}}$$

Take $\ln(\cdot)$ -- (I)

— Using the assgn.: (1) $\ln \boxed{\frac{(h+f)!}{(h-g)!}} = (f+g)\cdot \ln h \pm O\left(\frac{(f+g)^2}{h}\right)$ if $f+g = o(h)$.

$\boxed{\begin{matrix}k=\varepsilon n/b\\ l=n^2 b\end{matrix}}$ (2) $\ln\binom{\alpha n}{\beta n} = \alpha n \cdot H_e(\beta/\alpha) - O(\ln n)$; for const. $\alpha,\beta \geq 0$

Claim 1: $\ln\binom{n+k}{2k} = 2\boxed{\frac{\varepsilon n}{b}}\left(\ln\frac{b}{2\varepsilon}+1\right) \pm O(n/b^2)$

Claim 2: $\ln\left\{\binom{n^2-2k+l}{l}\Big/\binom{n^2+(b-1)k+l}{n^2}\right\} = -2\boxed{\frac{\varepsilon n}{b}}\left(\ln b+\frac{1}{2}\right) \pm O(1)$

Claim 3: $\ln\binom{cn/b+k}{k} = (c+\varepsilon)\boxed{\frac{n}{b}}\, H_e\left(\frac{\varepsilon}{c+\varepsilon}\right) - O(\ln n)$

— This gives: $\ln s \geq \frac{\varepsilon n}{b}\ln\frac{1}{4\varepsilon(c+\varepsilon)} \pm O\left(\frac{n}{b^2}\right)$
$= \Omega(n/b)$.

94 / 122

So, this will give you s greater than equal to is this consistent with? So that is the; that is our lemma 2 expression. This we must have got a lemma 1 that was $\binom{a+k}{k}$ times. So did we get that long time back? Yes, we actually got this. In lemma 1 it says $\binom{n+bk+l}{n}$. But now the number of variables is n^2. Our notation has changed, so it is consistent. It is consistent with that. We have this. Lemma 2 note that we actually got n + k where n is the degree did not get n^2.

But in lemma 2 second expression, we got n^2. Now we will put k and l. But still looking at this will not help because I mean everything looks. So the $\binom{n+k}{2k}$ and $\binom{cn/b+k}{k}$. They look quite comparable. And similarly, the second expression also looks comparable. It is not really clear whether we are getting any lower bound. So, using the assignment what we know are these two estimates that $(h+f)!\,/\,(h-g)!$.

This is? So, if you think about it this ratio I mean without the log should be something like a $h^{(f+g)}$ so, in log it should be f + g ln h but this will not be enough for this calculation. So, we actually also need the error terms. There is a mild assumption needed that f + g are not too big, it is $o(h)$. So, just think of this as an estimate on this saying that this is like $h^{(f+g)}$. But, the important thing is this error term also, for which you really need a calculation. And second is this we have used already in previous lectures. What is $\binom{\alpha n}{\beta n}$?

So, log of this is, nearly $2^{\alpha n}$, it could not have been more that this is the maximum it could be, but how close it will reach at $2^{\alpha n}$ or the log to be αn is dependent on the entropy using natural log. β/α entropy of this error term is not a big problem here this is quite small, exponentially smaller than the main term. So, these are the 2 and here we are assuming again that α, β are constants.

So, the O depends on α, β. So, these two follow from Sterling's approximation and especially the second one is used almost everywhere in computer science including even the elementary estimates on prime number theorem and all. Using this we will now write a sequence of claims. So, first claim is that $\ln \binom{n+k}{2k}$ and basically we have taken ln both sides in the above equation; take ln and call this equation 1. If you take ln both sides, so, we want to now calculate the ln of the sum minus ln of the of another sum.

So, we have 4 ln's. So, we let us first analyze $\ln \binom{n+k}{2k}$. So, $\log \binom{n+k}{2k}$ that is and k remember is what did we choose epsilon? $\varepsilon n/b$ and l is $n^2 b$. So, I think property 2 will be give; the property 2 or let's write it down. So, you want to say this is coming from where? Error term n/b^2, I think is probably need property 1 for this, you actually you cannot apply property 2 because alpha beta are; when you apply property 2 in alpha beta are constants here you have this b hanging around.

So, ε will be a constant small constant, but b is not a constant. So, you actually have to use property 1 in the calculation. So, use the definition of $\binom{n+k}{2k}$! and then do the calculation and this will be 2 f + g = o(h) is true because b is a growing function. So, n / b is actually quite small compared to n. So, you can apply property 1 and you will get this. Next is the ratio of these 2 big terms. So ln of this ratio of the big binomials.

That actually will be negative. So meaning that this denominator is quite large, which it also looks? This will be negative of this so this potentially could hurt the lower bound, because it is a negative term. Again, you have to use property 1 for this, actually, maybe both, both property 1 and 2, because the first binomial is $n^2 + 1$ and l is $n^2 b$. So you could think everything follows from property 1 anyways, completely skipping these calculations.

Third is the term in the denominator which is $\binom{cn/b+k}{k}$. Here definitely will use only property 2. So we will get $(c+\varepsilon)n/b\, H_e(\varepsilon/c+\varepsilon) - O(\ln\ n)$. So this just uses property 2 because k and n / b are comparable. So, for this use 2 to get this entropy expression. Now, all you have to do is claim 1 + claim 2 - claim 3. So, what is that? The focus should be on epsilon n / b, that is the main term. So, notice that it is coming with ln b over epsilon, so, it is more than b and from that you are subtracting ln b.

So, this is fine, still positive. And what is the third thing this also you will subtract, but entropy definition is minus ε minus this fraction times log of the fraction. So, the fraction multiplied by c + ε will give you ε. So, this is minus ε n/b and log of this, but ε and c are constants. So, they cannot be $\ln b$ and there will be another term which is 1 minus the fraction in ε but that again is constant. So, hopefully things will work out.

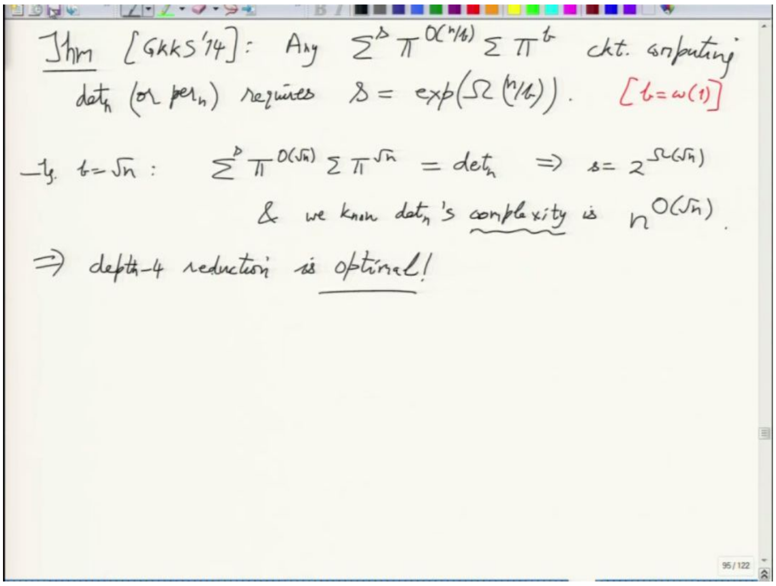
With that hope, let me just say that this gives log of s greater than equal to, so, 1 + 2 - 3 will give you

$$\frac{\varepsilon n}{b} \ln \frac{1}{4\varepsilon(c+\varepsilon)} \pm O(n/b^2).$$

That is what we will get in the end. So, this ε n / b will survive with some constant positive constant factors basically. So, for this ε has to be a fraction as we close it has it should be constant close to 0.

Then this will be positive and also error term is n/b^2 so, that is fine b is growing so, it is smaller item. We can safely see that this is $\Omega(n/b)$, that is the log. So, s is $2^{n/b}$. So, we will write the theorem.

(Refer Slide Time: 01:04:23)



<u>Thm</u> [GKKS'14]: Any $\Sigma^s \Pi^{O(n/b)} \Sigma \Pi^b$ ckt. computing det_n (or per_n) requires $s = exp(\Omega(n/b))$. [$b = \omega(1)$]

-Eg. $b = \sqrt{n}$: $\Sigma^s \Pi^{O(\sqrt{n})} \Sigma \Pi^{\sqrt{n}} = det_n \Rightarrow s = 2^{\Omega(\sqrt{n})}$

& we know det_n's complexity is $n^{O(\sqrt{n})}$.

$\Rightarrow$ depth-4 reduction is optimal!

Let me first write the theorem. So this is GKKS 14 so, Gupta Kamath, Kayal and Saptharishi that any $\Sigma^s \Pi^{O(n/b)} \Sigma \Pi^b$ circuit computing determinant $n \times n$ or permanent. So note that the way we did this determines permanent or Immanant everything's the same; lower bound you get is the same. So, this really this method cannot distinguish them from each other. Requires s to be $exp(\Omega(n/b))$. What is the question now?

I mean you cannot take b to be constant otherwise you do not get anything. So, this does not give you 2 rates to n cannot take b to be 1000 for example, because that in many places, we assume that b is growing, it is a growing function in the analysis, so this requires b to be $\omega(1)$. But it is a function in n it cannot be a constant. Constant would have meant. It was n/b versus n/b^2. So just take b to be growing. That is all extra thing is 1/b.

Now so if b was constant then will there be some inconsistency in the statement. I am not thinking about the proof. So, in the theorem statement, if you just take b to be 1, then will there be some will be a strange statement or I mean, if you take b to be 1 then this will be depth 3 with the product fan in order n. This model is inhomogeneous. Maybe you can also take b to be constant. I do not see any problem, maybe we can give an alternate argument.

But anyways, in the proof technique, we have used b to be growing and or maybe we say b sufficiently large constant. Maybe we do not really need to be growing. Just take b to be much more than the 2, you have a main term you have an error term. Just take b to be large enough 1 billion but that will require looking at these O's which also I do not think is too hard can be done. But then I think you could have directly worked with model and do something. Anyways, I do not want to go into that.

So, this exactly is the theorem we have finished the proof of. So, example is $b = \sqrt{n}$. So, in the $\sqrt{n}$ you are looking at the model $\Sigma^s \Pi^{O(\sqrt{n})} \Sigma \Pi^{\sqrt{n}}$ and $n \times n$ determinant if you try to compute using this model then you get $2^{\sqrt{n}}$. Now, what is the best that you know for determinant in this model? By Agarwal Vinay depth reduction you know $n^{\sqrt{n}}$, this equal to determinant implies that the top fan in s is $2^{\sqrt{n}}$. And we know that determinant complexity is $n^{O(\sqrt{n})}$.

So, this complexity with respect to this model. So, we know that $n^{\sqrt{n}}$ is possible and we know at least $2^{\sqrt{n}}$ is required, so, it is nearly a match. So, this match means that depth reduction cannot be improved, well maybe it can be improved slightly, but it cannot be improved significantly. So, you cannot, for example, get $n^{n^{1/3}}$ or even anything smaller than $\sqrt{n}$ you cannot get except maybe shaving off a log n factor, if not more.

So depth 4 reduction is optimal, that is the consequence of this. Nearly optimal. So we will meet again tomorrow. And on Saturday can we start at 11:40? Extra class. We finished before one. So this Saturday we have an extra class and subsequent Saturday Abhibhav will teach us something related to the limitations of these measures that these measures cannot take you too far. Any questions? That is the feeling, I think in Ramprasad's surveys.

There are calculations that you cannot really get much better than what we have gotten. If l was not there, then you are just looking at derivatives, l = 0 is space of derivatives. But it was not clear a priori, the l we are taking is very large. It is near; it is even more than the number of variables. It is $n^2 b$. So for example $b = \sqrt{n}$, we are taking l to be $n^{2.5}$. So that large l so, so in claim 2 something is happening by picking a larger l. If you take l to be 0 then it will be a big negative. You take l to be a large it will be small negative.