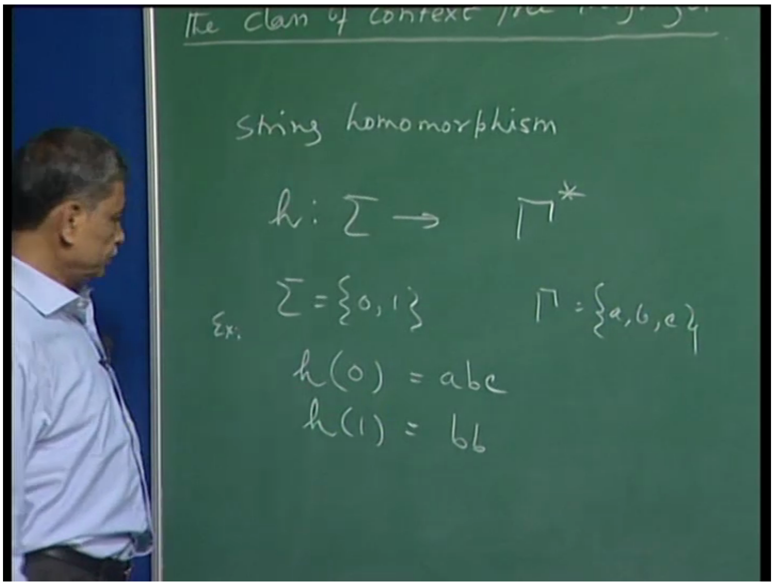
Theory of Computation
Professor Somenath Biswas
Department of Computer Science and Engineering
Indian Institute of Technology Kanpur
Lecture 28
Closure Properties Continued
CFLS not closed Under Complementation

First we see today a few more closure properties of the class of context free languages. One operation which is called the string homomorphism. Now this considers two alphabet let ussay sigma and gamma. And stringhomomorphism is a map from sigma to gamma star. So what it means is say for example your sigma is let us say 01 and gamma this alphabet is a b c, alright? So this is a map which is assigning a stringover gamma to every symbol of sigma. So it could be that h of 0 for example is a b c and h of 1 is let us say b b.

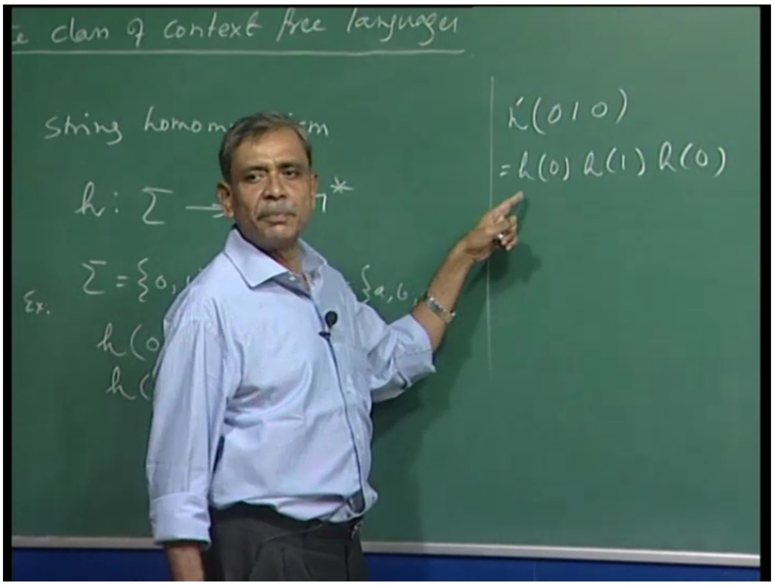
(Refer Slide Time: 01:40)



Now this h clearly extends to strings of sigma star to strings of gamma star. See what we mean is that suppose I have this string 0 1 0 so what we will do is we will apply this map h individually to each of the symbols in the stringto get a new string over gamma. So h of 0, h of 1, h of 0 and these three strings are concatenated.

So in other words by definition what we mean is so this extended thing if I call it h dash which is a map from sigma star to gamma star, in this case is h of 0 concatenated with h of 1 followed by h of 0 because these are the three symbols. Each symbols individually we are applying the map h.

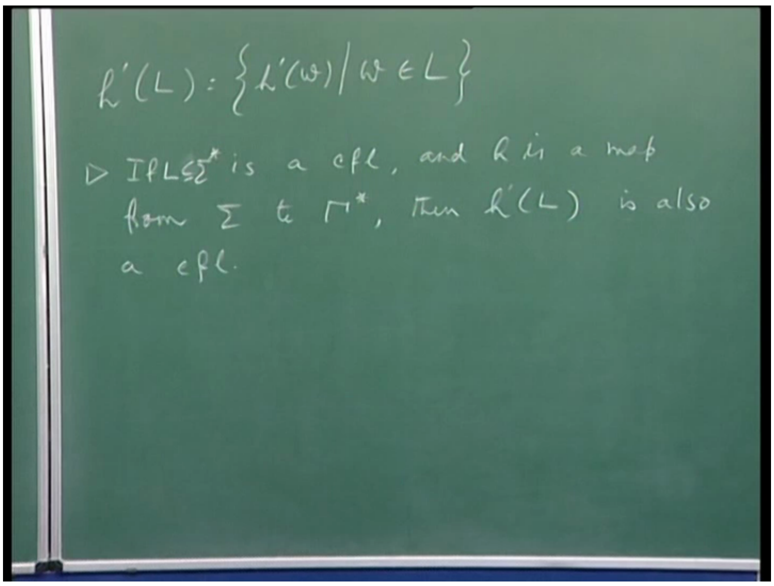
(Refer Slide Time: 02:49)



For this example what we have? H of 0 is a b c, then h of 1 is b b, then again a b c, right? So clearly this map h dashwill map a language over sigma to a language over this other alphabet gamma, right? In a very natural way that is, so in fact we can write thish dash of L is the set of all h dash of w such that w is in L, right? So you can see h dash of L will be alanguage over thealphabet gamma, right?

So theclosure properties is that if Lis a context free language and h is a map from sigma to gamma star. Or I should have saidif Lis a subset of sigma star so that means the alphabet of L is sigma and h is a map from sigma to gamma star. That is h is a string homomorphism then h dash of L, the way we have defined h, is also a CFL, right?

(Refer Slide Time: 05:25)

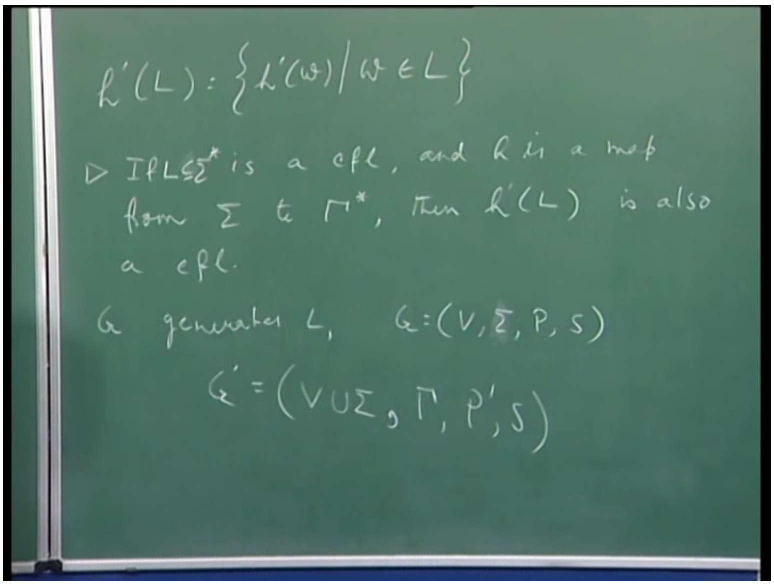


So one way of saying this is that this operation of taking a language, applying h dash to L, so unary operation. So this operation does not take a context free language to a non-context free language therefore we say that any context free languageor the class of context free language is closed under this operation which is known as the homomorphism or string homomorphism.

Now this is actually fairly simple to prove. Why? Because you see L is a CFL, right? So let us say G generates L and G is V, sigma, P, S. And now what we can do is we can consider another grammar G dash which is as follows. We will say that this grammar G will have also these terminal symbols that is the symbols of sigma as nonterminals, right? So what would that mean? That means our nonterminal for G dash is union sigma.

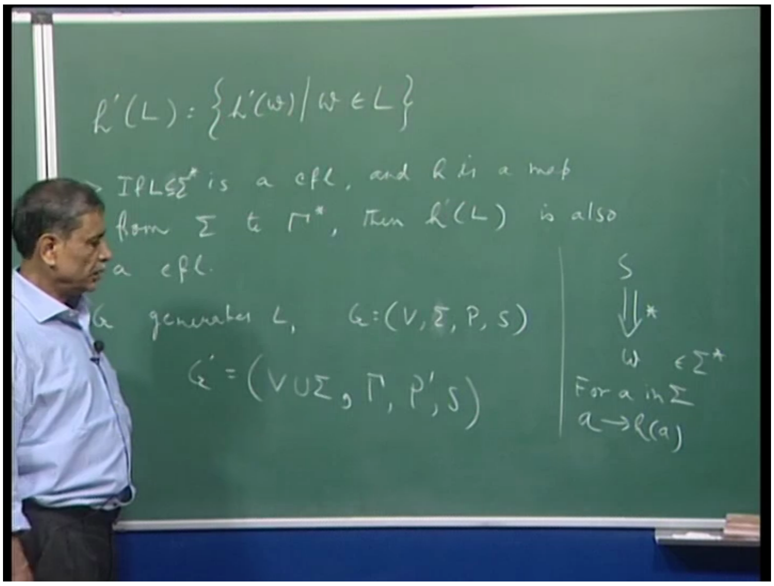
There is no problem in doing this. And my terminal symbols, becausewe are going to give a grammar which will generate h dash L so here of course it will be gamma. The alphabet of this languageh dash L is gamma so therefore the terminal symbol is gamma, P dash and S, okay.

(Refer Slide Time: 07:51)



So P dash is what we are going to do the idea is very simple. First usethese productions P to generate a string from S. That string will be over sigma. So basically generate a string of L, alright? So from S we generate a string let us say win sigma star. But then sigma star is no longer or the alphabet sigmathese are not terminals anymore, these are nonterminals, right? So we just add these productions. Let us say for a in sigma add the production a goes to h of a, right?

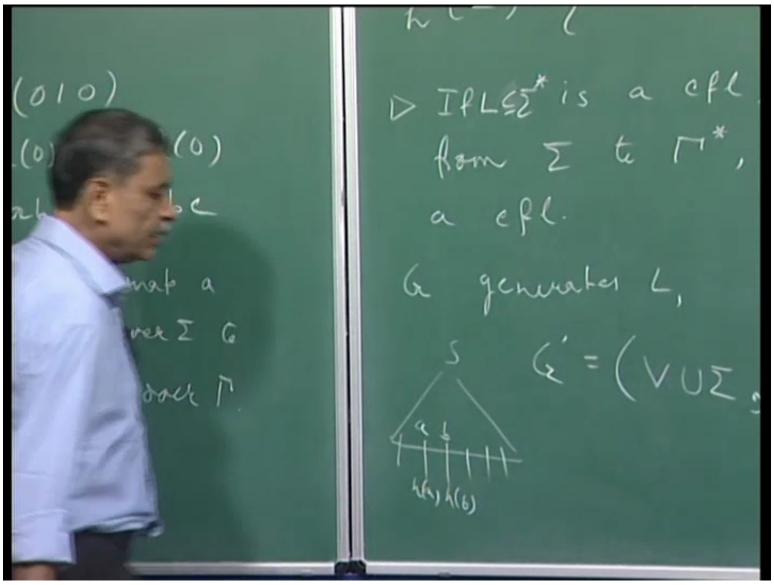
(Refer Slide Time: 08:59)



For each a in sigma we add this productions and so your new P dash, this is fairly now simple to see what we are trying to do. New P dash is old P plushow many productions are there? Asmany as we have symbols in sigma.

So for each symbol which is now a nonterminal. Each symbol of sigma is now a nonterminal of this grammar G dash and you can rewrite that with h of a which is a string over gamma. So what is going to happen? The essential idea is that from S you will generate a string over sigma. So let us say this is a, this is b, so here we will use the production h of a and so on, h of b, right?

(Refer Slide Time: 10:12)

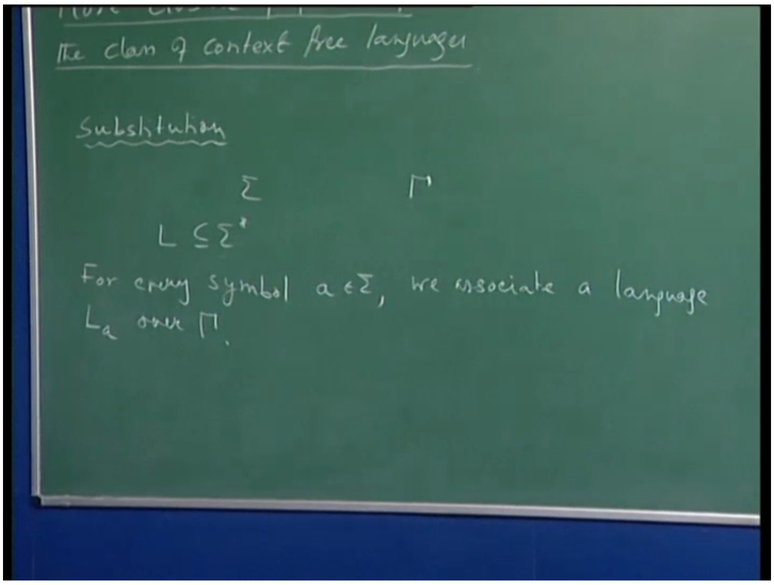


And then you are going to get a string over this alphabet. Now clearly it is very easy to see the grammar that we have described this G dash is going to generate h dash L. And since this is a context free grammar therefore h dash L is also a context free language. So therefore it is very simple to see that if L is a context free language then h dash L is also a context free language which is what we needed to prove.

Using very similar idea actually we can doanother closure property which is somewhat more general and let me explain that property and that property is called or that closure property or that operation is called that of substitution, okay. Now next property that we will look at is substitution. In case of substitution what we have is again two alphabet, one is sigma the other is gamma.

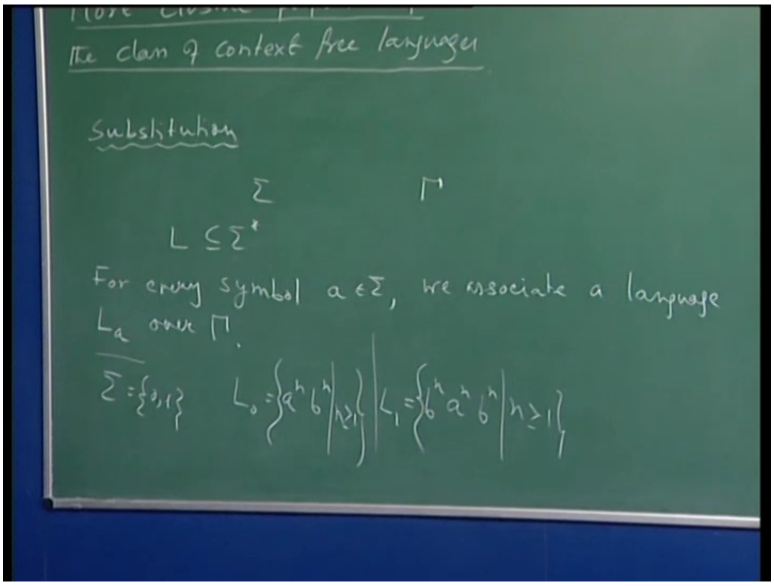
And now let us say we have a language L over sigma, right? Andwhat we do,we associate a language over gamma for every symbol (sig) in sigma. So let me write it down. For every symbol a in sigma we associate a language, let me call it L a over, right?

(Refer Slide Time: 12:34)



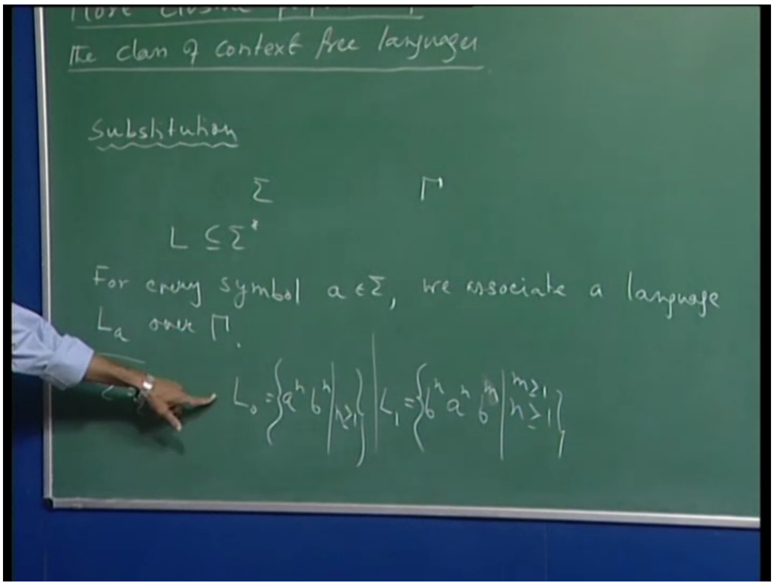
So for every symbol a we associate a language L a over this other alphabet. Now you see in the same manner now we can take a string over sigma. So let me again give an example. So let us say sigma is 0 1 and nowso L 0 this language let me write it as a n b n and L 1 let me write it as b n a n b n, okay. So what I mean is this is n greater than 1. This is one language and this is the other language, okay.

(Refer Slide Time: 13:48)



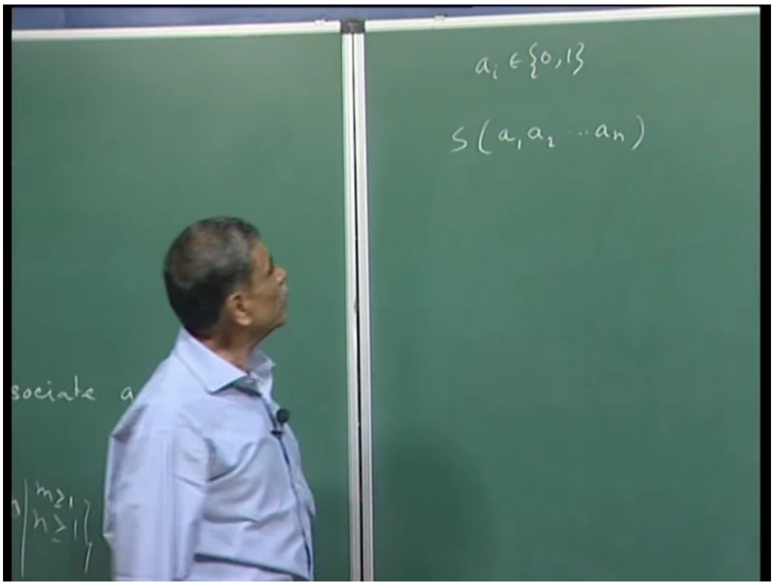
Orpoint I am going to make I will say a little later. So m is greater than 1, n is greater than 1. So we have these two languages L 0 and L 1.

(Refer Slide Time: 14:04)



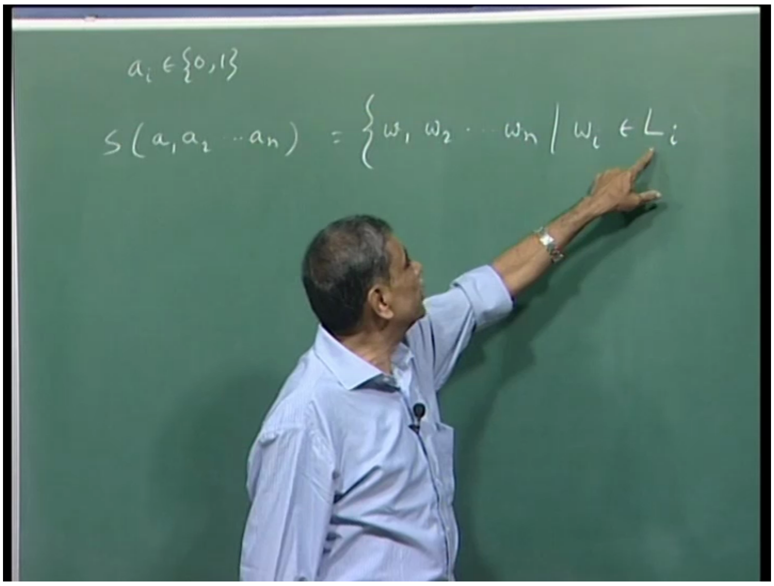
Nowa substitution essentially tells you first of all that this association of every symbol of sigma is associated with a language over gamma, right? Now so let me callthis particular substitution as S. And now you see what we can do is ifyou take a string, let us say a 1 a 2 a n, right, where each a i is an element of sigma. In our case is 0 1, that we said. Thenthis actually defines a language.

(Refer Slide Time: 15:19)



How? Thissetw 1 w 2 w n concatenation, now these w's are stringswhereas these a's were symbols over sigma such that w i is an element of the language associated with the corresponding symbol. So remember the way we associated we saida iis some particular symbol. So thereis some language associated with it and that is L i. That is the thing that we were saying.

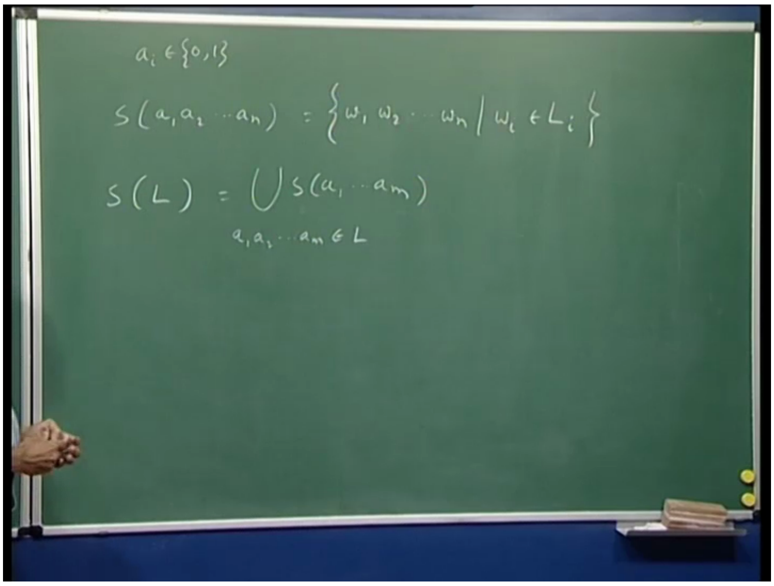
(Refer Slide Time: 16:05)



So now what we have isI get a new string but it is not just a string.So hereI could substitute for w i any string from this language. So that is whythis substitution of a string is now a language. And substitution operation when I do this for a language L,what is it going to be?

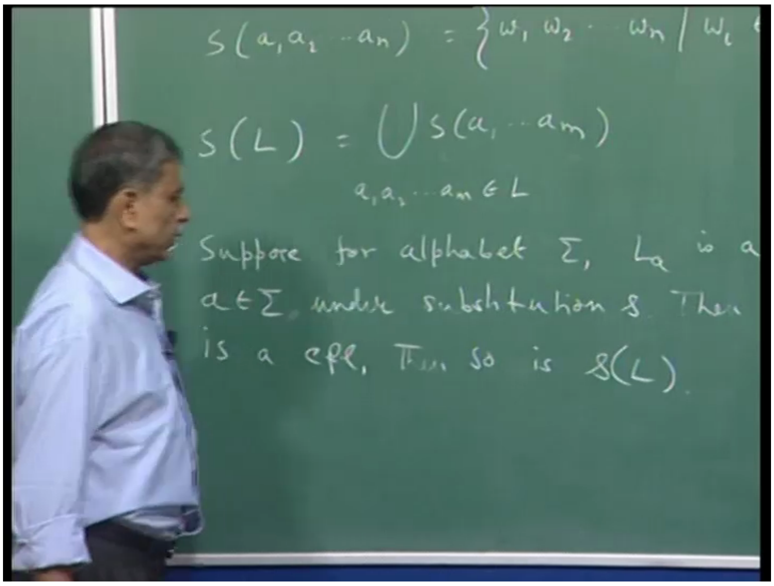
It is going to be the union of you know all these, right? Soremember this is going to give mea language. So it is okay to talk of this and what we were saying is that a 1 a 2 a n this string is an element of the language L, right?

(Refer Slide Time: 17:23)



So essentially what we are doing is that for every string of the language L, right, on substitution I get a language and then we take the union of this language. So now basically our theorem is that suppose for alphabet sigma, L a is a context free language. What is L a? For each a in sigma. In other words what we are saying that suppose for every symbol in sigmathe substitution associates with that symbol is a context free language under substitution S. Then if L over sigma star is a CFL then so is S of L, okay.

(Refer Slide Time: 19:23)

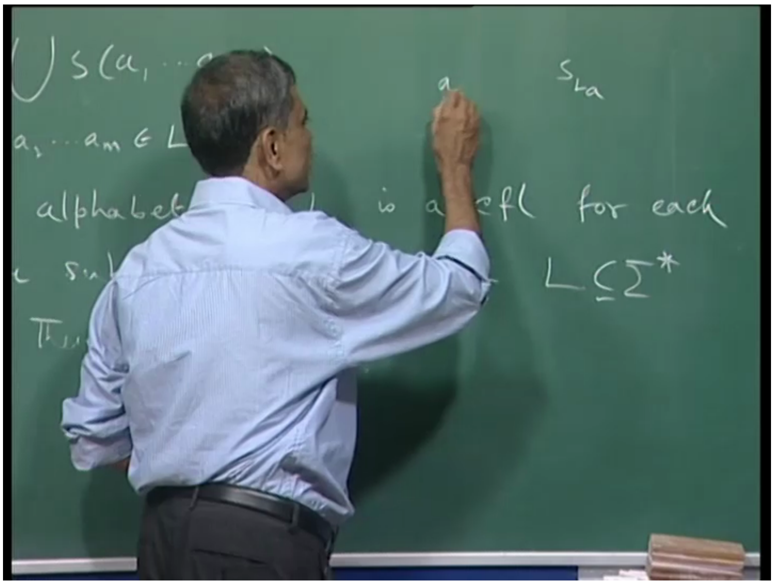


See remember again that like string homomorphism, substitution also is an operator which takes a language andyou know returns another language, right? And what we have said or we were trying to assert in this that if each language that we associate with a symbol under the substitution operator S that you have defined is a CFL and then you apply the substitution to a CFL L then S of L also is going to be a context free language. You knowall this definition took a long time to explain but the proof is again very simple.

And in thisthe manner is very similar to the string homomorphism case. Soyou know previously what we did in case of string homomorphism? That we just considered that elements of sigma to be part of the nonterminal set for thelanguage which we obtained under string homomorphism.

Herefor every such symbolinstead of generating a stringwhat we can do is that we think of generating a stringbutin that string for a we have the start symbol of let us say S of L. So let us say there is a grammarsincethe grammar is CFL for you know L a, L a we said is a CFL, so there is a start symbol for L a. So let me just denote it as start symbol of L a. And this was the association, right?

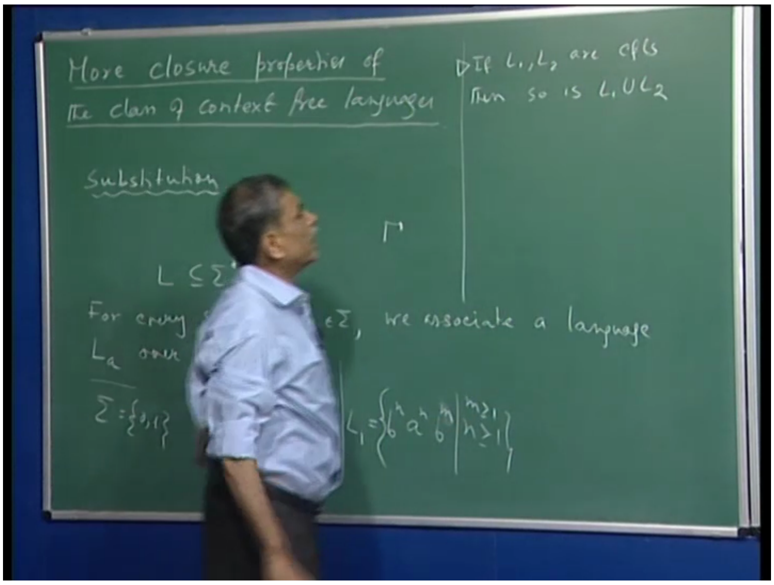
(Refer Slide Time: 21:36)



What we are saying that generate a string which will generate, you know youstart the derivation of L and then when you finish the derivation, instead of getting a string over sigma you get a string over corresponding start symbols. And then use thosestart derivations from those start symbols using the grammars of L a, etc. to generate the other string. This is a very simple idea and this is going to work.

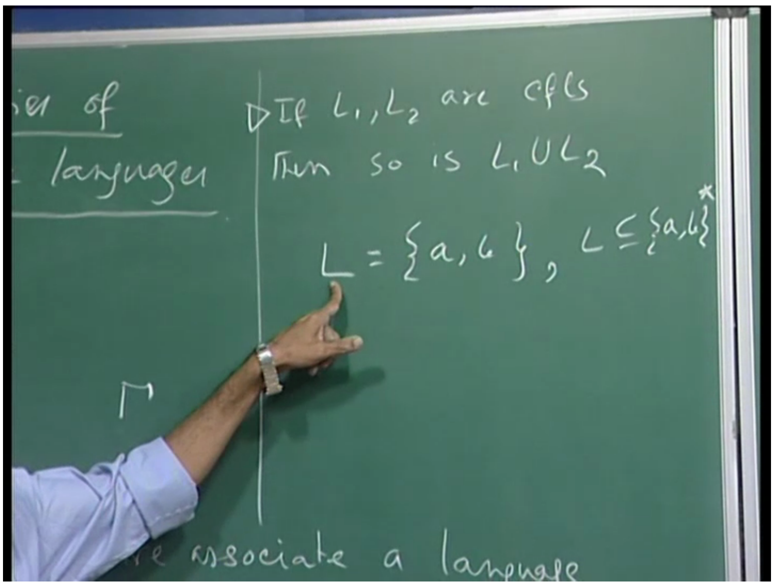
Now what I can show you that once we understandthat substitution preserves context freeness, this operation does not take the class of context free languages to anything else. Then certain things like all those concatenation union which we proved,for those operations the class was closed under the context free languages. For example let us take this case. We proved that if L 1 L 2 are CFL then so is L 1 union L 2, right? This is one result that we had.

(Refer Slide Time: 23:08)



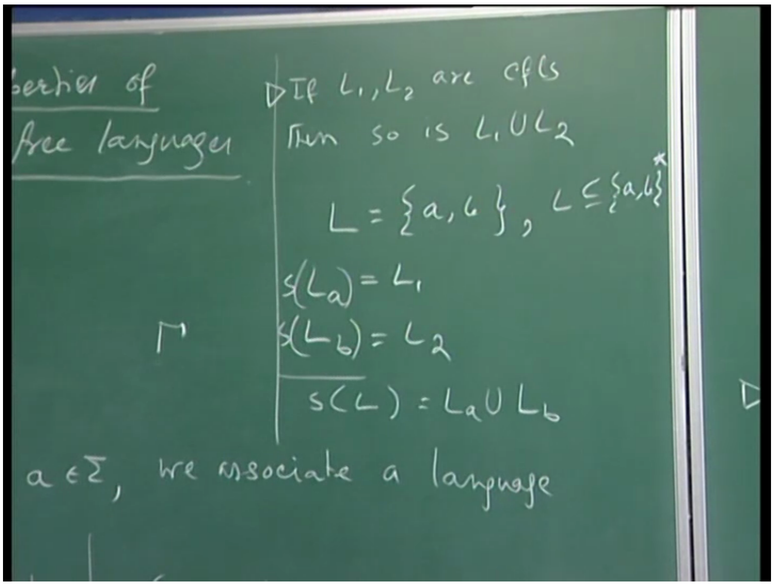
You know this you can prove very simply using this substitution idea. Considera very trivial language whichconsist ofjust these two strings a and b, okay. So therefore this language L clearly is a subset of a b star. That is okay but this is a very trivial language. This language L has just two strings a and b.

(Refer Slide Time: 23:36)



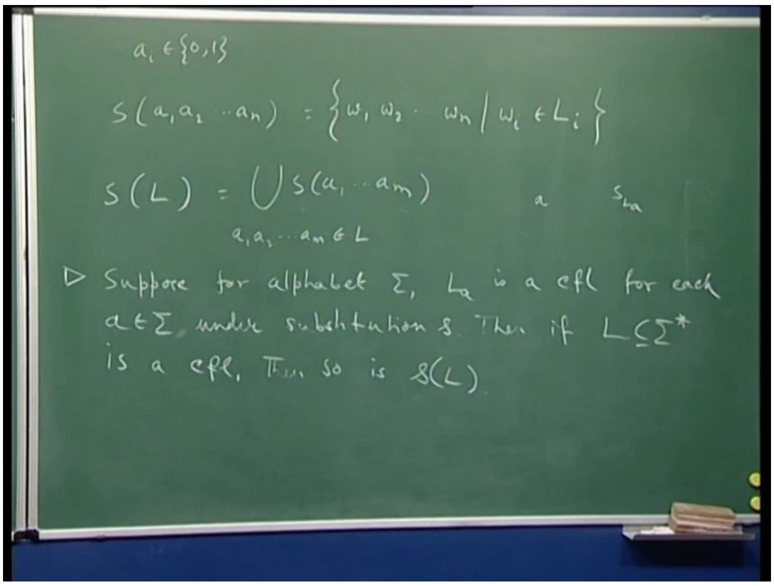
And now what you say L isthis symbol a, the language we associated with L a is let us sayL 1 and L b the language we associate. So let me just say this S of L a is L 1, S of L b is L 2. And now it is easy to see S of this language is nothing but L a union L b which is nothing but L 1 union L 2.

(Refer Slide Time: 24:22)



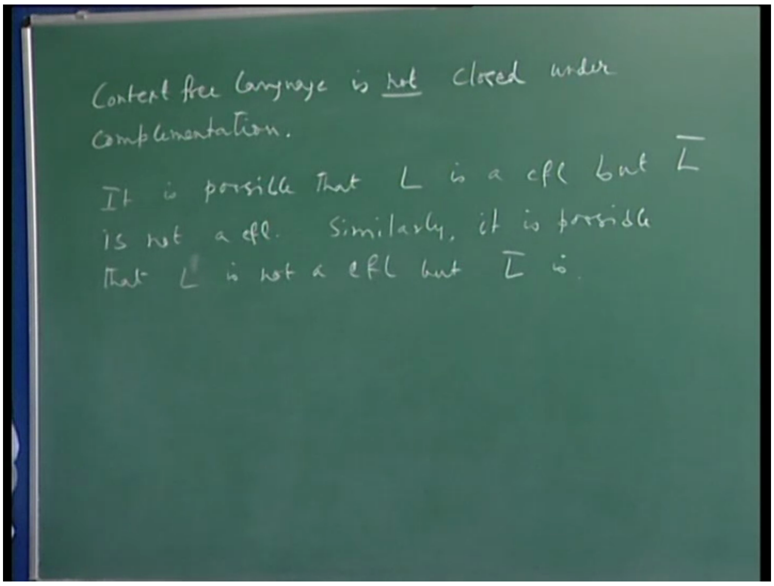
So in the same manner we can show just using substitution that context free languages are closed under concatenation for example or even Kleene star. So all those things reallycan follow from, though we have proved them individually, we can also see them as corollaries of this particulartheorem thatthe class of context free languages is closed under substitution.

(Refer Slide Time: 24:49)



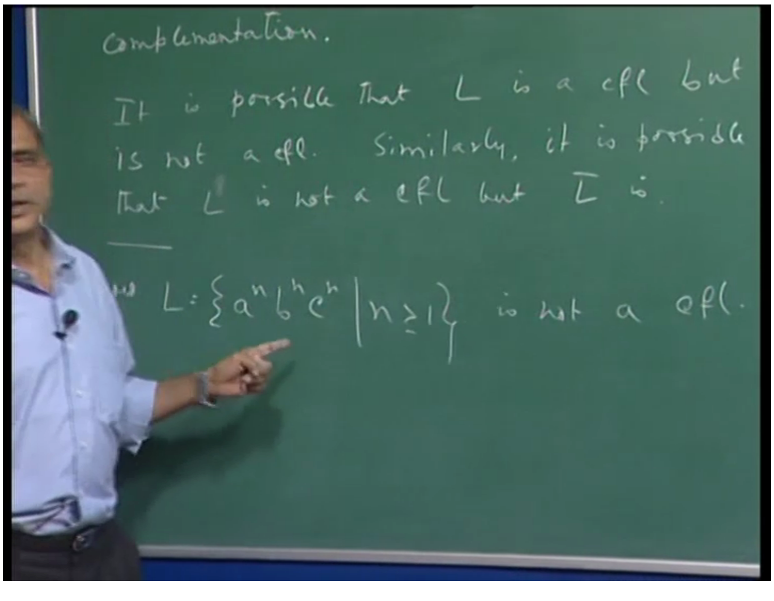
We had proved that (con) context free language class is not closed under complementation. This means that it is possible that L is a CFL but L complement is not a CFL. And also the other way similarly it is possible thatL is not a CFL but L complement is.

(Refer Slide Time: 26:40)



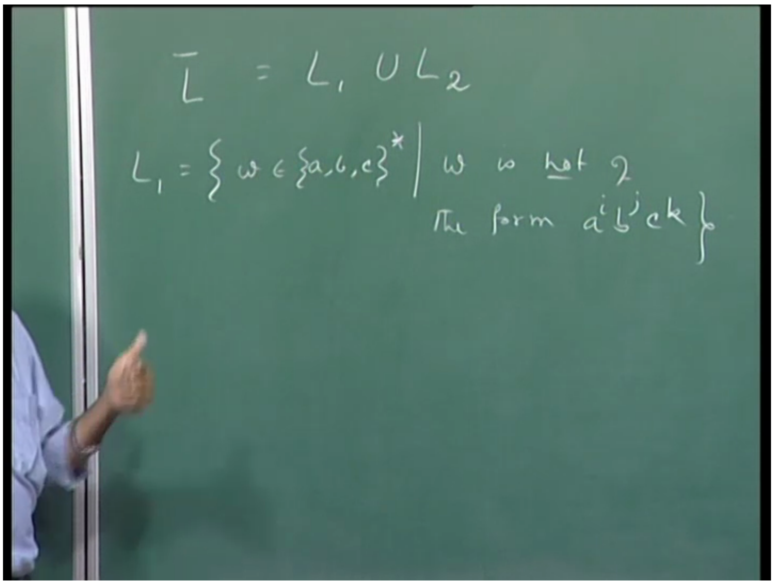
And it will be nice to get examples of such pairs of languages that one is of that pair that pair is a language and its complement, one is a CFL, the other is not a CFL. And in fact the very first language that we had taken as the example of a language which is not a CFL, what was that? We said this language a n b n c n, n greater than equal to 1. We knowthis L is not a CFL, right? We had proved this using quite simply our pumping lemma. That we have proved and we know that this language is not a CFL.

(Refer Slide Time: 27:34)



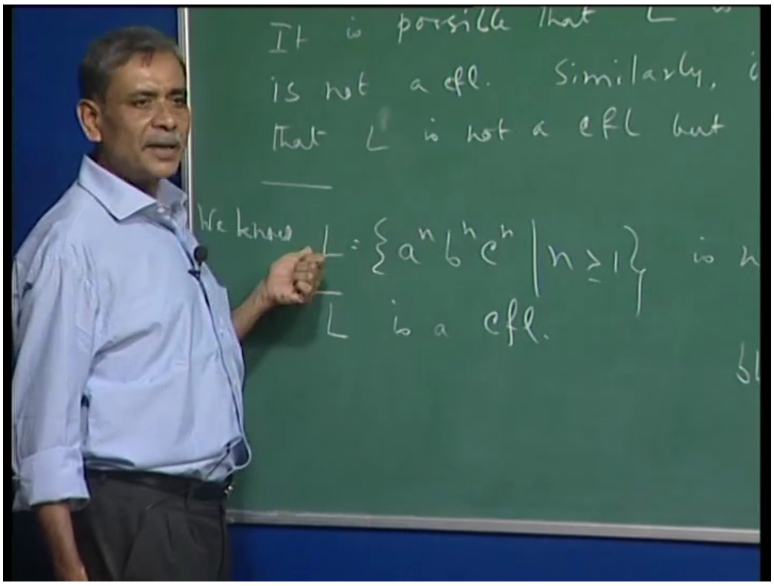
But what we can show that L complement however is a CFL. Now what is L complement? If you see that L complement is really the union of two sets. One is you can say L compliment is L 1 union L 2 where L 1 is all those w's over this alphabet a b c such that w is not of the form a i b j c k, okay. So here we are not insisting that the i j k could be anything. But essentially this is not of that form, right?

(Refer Slide Time: 29:19)



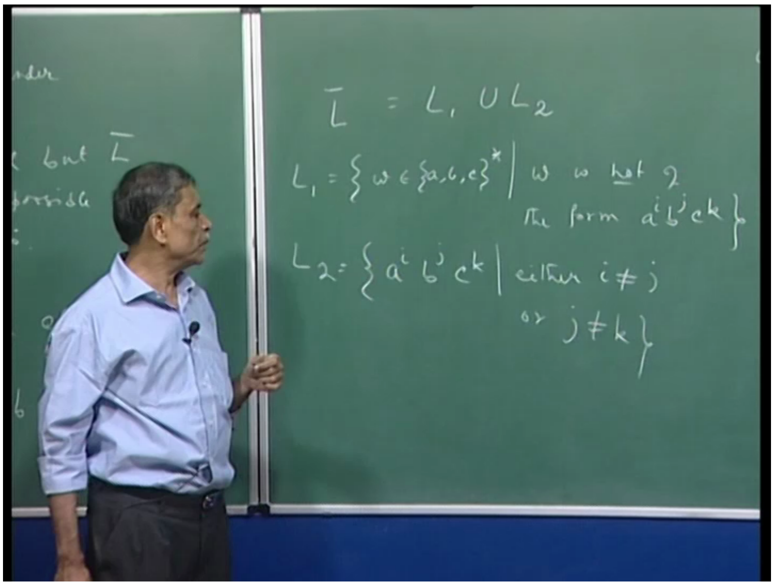
And what is L 2?L 2 is a i b j c k where either i is not equal to j or j is not equal to k. This is not too difficult to see. So you see something may not belong to this language because for a reason that it is something like let us say you know this string b b a a a b, you know something like this. So in this kinds of strings all a's will precede all b's and they will precede all c's, right? They come in a's, then b's, then c's will come. But if that order is destroyed then you get some string. So clearlythat cannot be the language L.

(Refer Slide Time: 30:35)



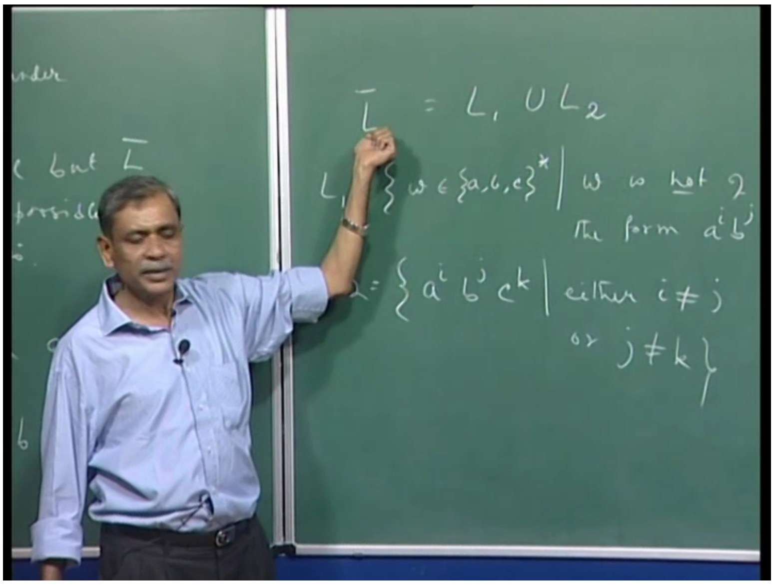
Soyou know these are the strings which form the language L. Now can you see that L 1 is a regular language? Claim L 1 is regular. Why? Becausenow L 1 complement is what? W is of the form a i b j c k but that can be checked by a finite automata that first all a's, then all b's, then all c's and so complement of L 1 is regular therefore L 1 is regular because regular languages are closed under complementation.

(Refer Slide Time: 31:30)



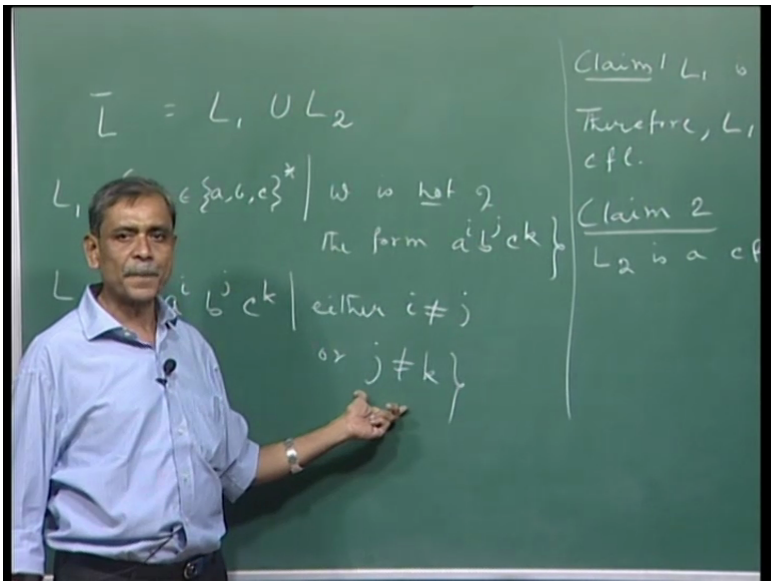
So therefore L 1 is a CFL, claim 1. Now claim 2 and this is where the main part of our work is there is that L 2 is a CFL, right? Now since L complement is the union of L 1 and L 2 and oncewe have these two claims, claim 1 and claim 2, both L 1 and L 2 are context free languages. So remember that context free languages, this class is closed under union so therefore if both of them are context free languages then it has to be also context free language.

(Refer Slide Time: 32:33)



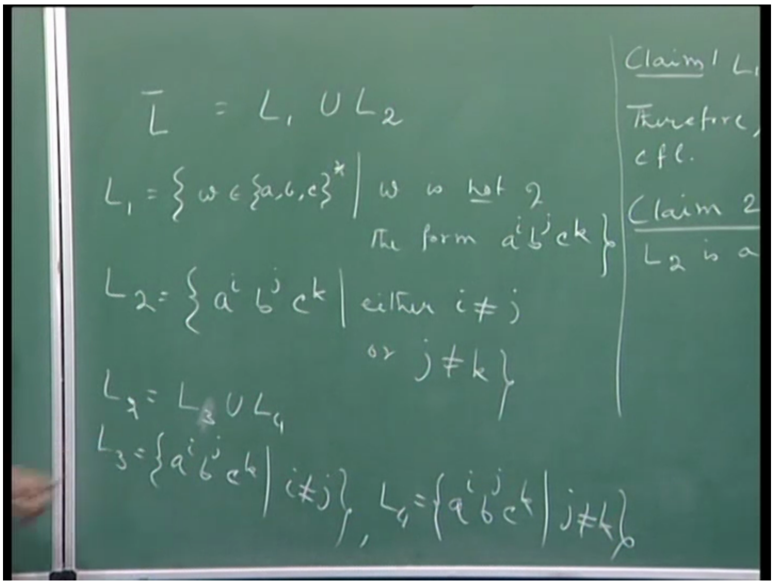
So that is the idea behind this group and I kind of gave you the reason why claim 1 is true which you can convince yourself of. So let us try to prove that claim 2 which is that this language L 2 which is of the form a i b j c k where either i is not equal to j or j is not equal to k. This language is CFL, right?

(Refer Slide Time: 33:01)



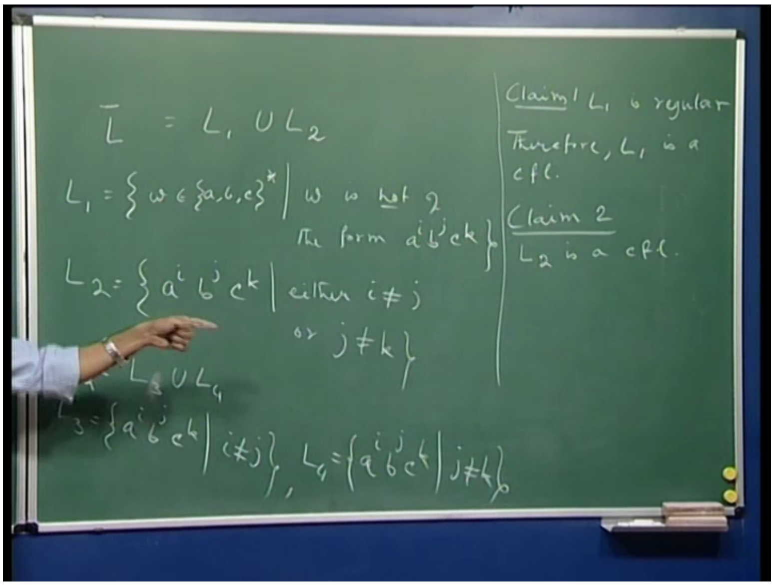
Further what is L 2? I can take L 2 as,again you know I can say that L 2 is union of two languageslet me say L 3 union L 4. What is L 3? L 3 is a i b j c k such that i is not equal to j and L 4 is a i b j c k such that j is not equal to k, okay. So you can easily see that L 2 is the union of these two languages.

(Refer Slide Time: 34:11)



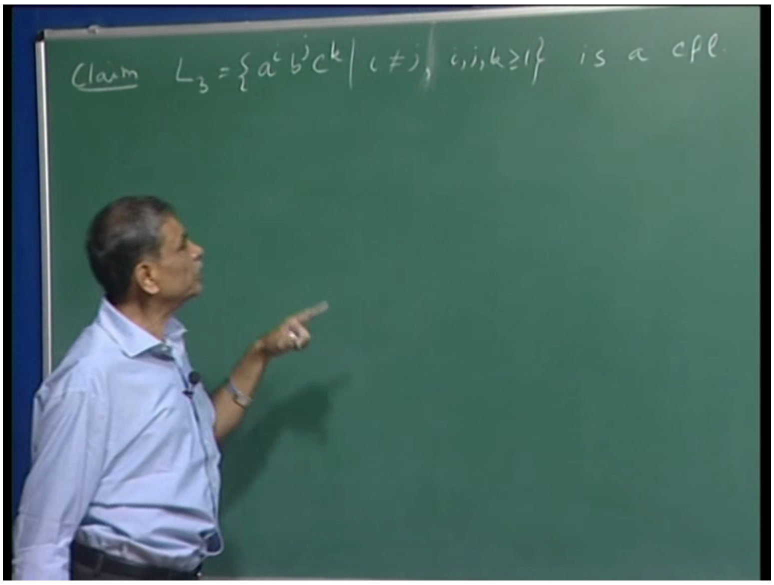
So because every string in L 2 satisfies either this property or this property. So separately we are considering these two properties and then taking their reunion these languages.

(Refer Slide Time: 34:21)



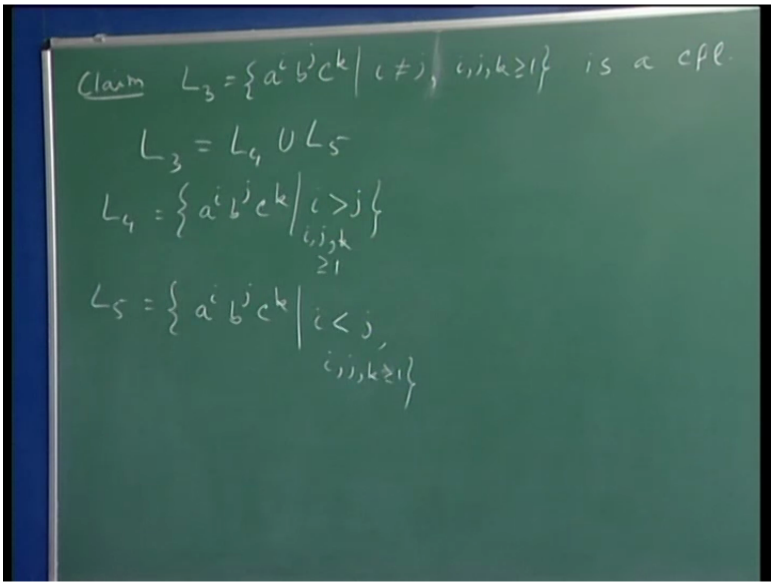
Therefore I am getting the language L 2. What wewill show individually that both L 3 and L 4,each of them is context free language. And you can see that if I show for example L 3 is context free, in the same way we would be able to show L 4 is also context free, right? So let us try to prove that L 3 is a context free language. We were trying to prove this claim that this language L 3 a i b j c k where i is not equal to j. That you recall that we want i j k to be greater than equal to 1. So I am adding that also. This language is a CFL.

(Refer Slide Time: 35:11)



Now again you see L 3 now is the union of two other languages L 4 union L 5 where L 4 is a i b j c k such that i is strictly greater than j and i j k greater than equal to 1 and L 5 a i b j c k, i is strictly less than j and again I j k is greater than equal to 1.

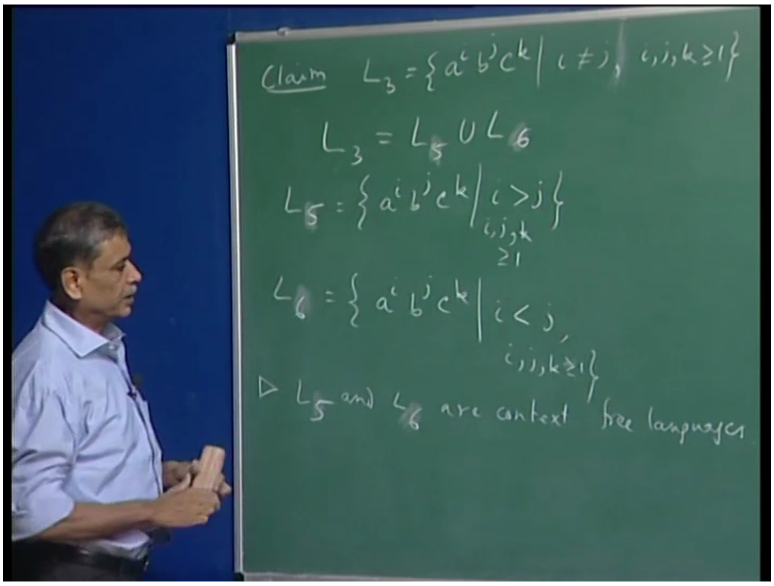
(Refer Slide Time: 36:19)



Now furtherwhat we can claim is thatand we can prove this that L 4 and L 5 are context free languages. You see how we are making use of the closure property? Suppose I managed to show L 4 and L 5 our context free languages that is we prove this claim, right, then what do I have?

That L 3 is also a context free language because it is the union of two context free languages. In a similar manner I can show that L I am sorry I mean I used up L 4 here. Maybe then I should use L 5 L 6, right? So this is L 5 and this is L 6 and then L 5 and L 6, right?

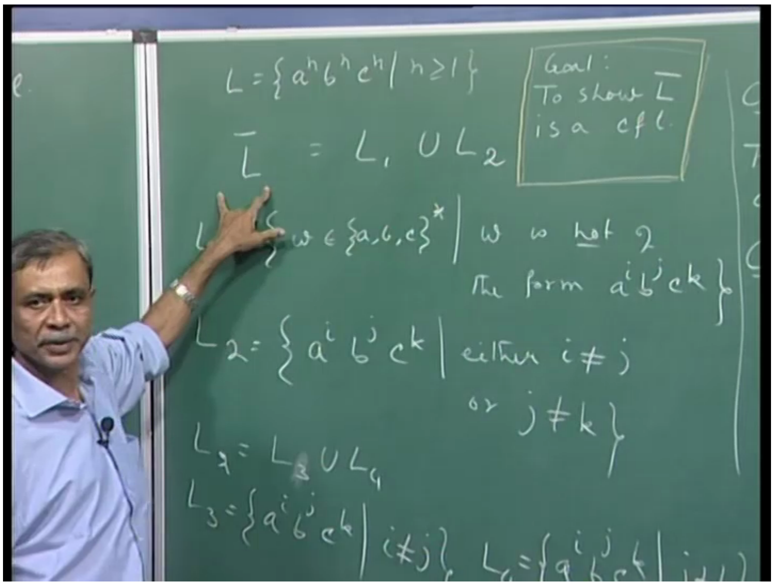
(Refer Slide Time: 37:36)



So anyway the idea is very simple that L 5 and L 6 are context free languages. Weneed to prove this but suppose we have done it already. In that case L 3 will become a context free language because this is the union of two context free languages. So L 3 is a context free language. In the same manner very similarly you see L 4 is very similar to the L 3 and L 2 is the union of two context free languages.

After we manage to show L 3 is context free as well as L 4 is context free then L 2 is a context free language. But what is L 1? We said L 1 is regular therefore L 1 is a context free language, right? So we have proved L 1 and L 2, they are context free languages. Their union is L complement.

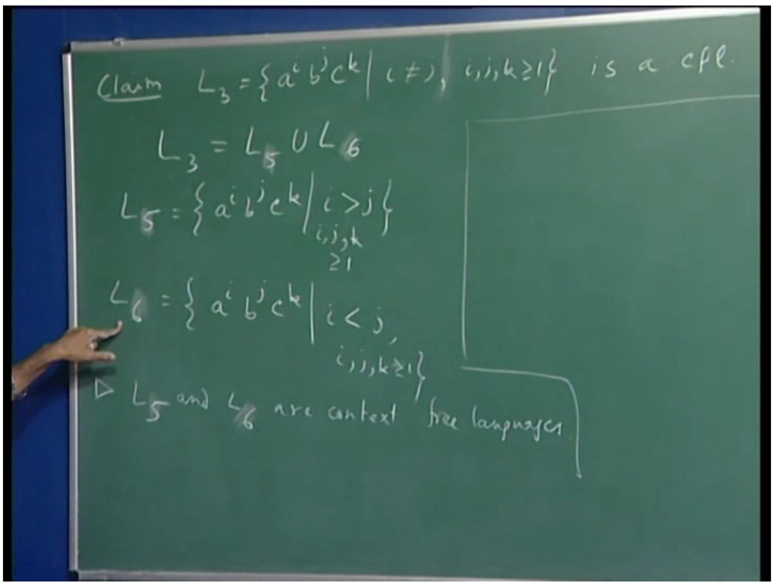
(Refer Slide Time: 38:36)



So therefore L complement is also a context free language and that settles our goal to show that L complement is acontext free language. You see so this is one way of making use of closure properties of languages. Here we are using them to show that something is context free language and similarly you can sometimes use closure properties to show that something is not a context free language.

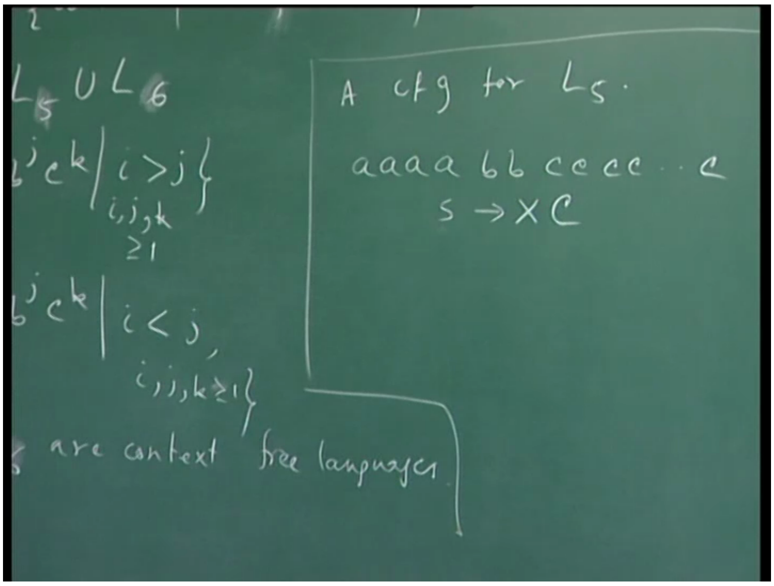
But you know that example we have not seen yet but let usjust work on this. You see you will agree that you know if I can show you that howwe can generate L 5by a context free grammar thereby proving that L 5 is a context free languageand then very similarly you can show that there is a context free grammar for L 6 also.

(Refer Slide Time: 39:39)



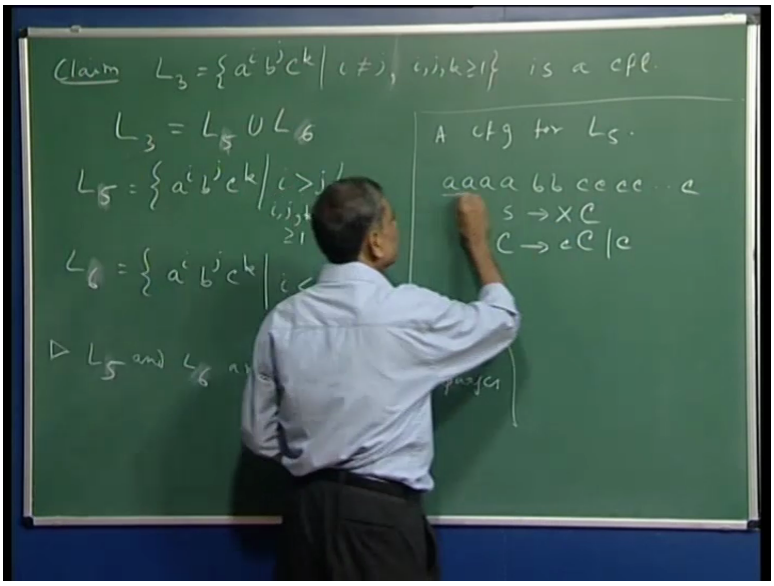
Sothen of course we will be able to prove that L 3 is a context free language. So let us see a CFG for L 5. So what is L 5? You know think of a typical string of L 5. So you know it is number of a's is strictly more than number of b's. So something let us say a a a a b b and thenyou have some c's. You do not care actually about the number of c's so long it is greater than or equal to 1. So you see suppose what I have? I have astart symbol S for the grammar for L 5 and then let me use let us say x and then capital C.

(Refer Slide Time: 40:48)



See the role of this capital C is to generate a string ofone or more c's. So that is very easy to do, right? So C is either small c capital C or simply small c. So you knowthis is going to generate any number of c's, at least one or more number of c's. So that takes care of this. With X I would like to generate this part of the string.

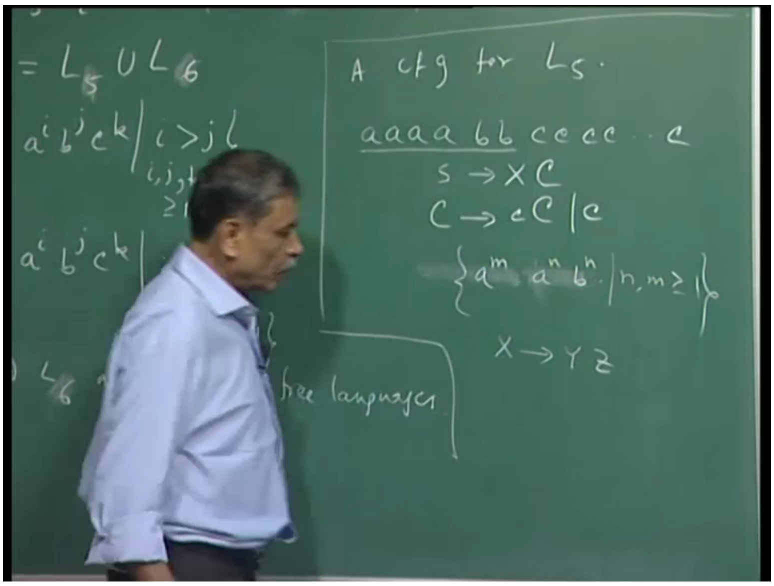
(Refer Slide Time: 41:19)



What is that part of the string? That is it has number of a’s followed by number of b’s but the number of a’s is strictly more than number of b’s. So like in this case. But here what you can see is that you see there will be some number of a’s then equal number of a n b n and this is some other a m wheren m are greater than 1.

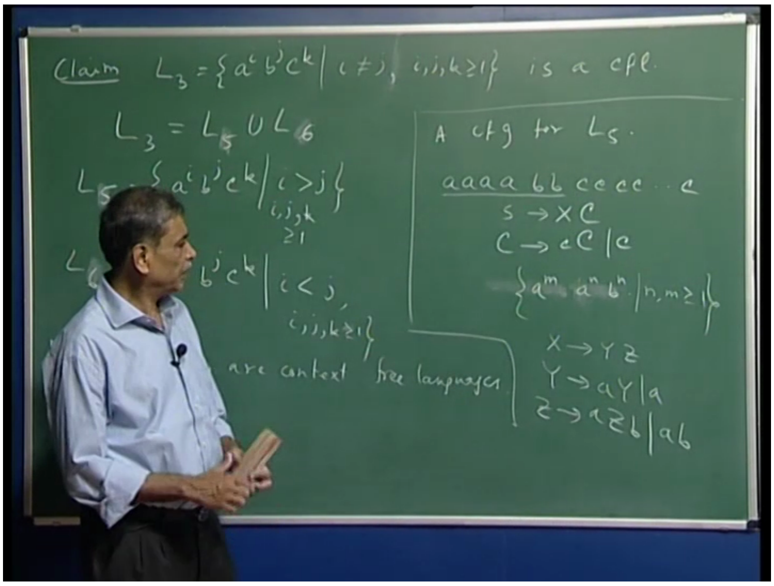
This language captures strings of this kind where the number of a’s will strictly more than the number of b’s and this is fairly easy to generate, right? Weknow sobasically X should generate strings over this language. Sowhat we can do I can write X as Y followed by Z.

(Refer Slide Time: 42:31)



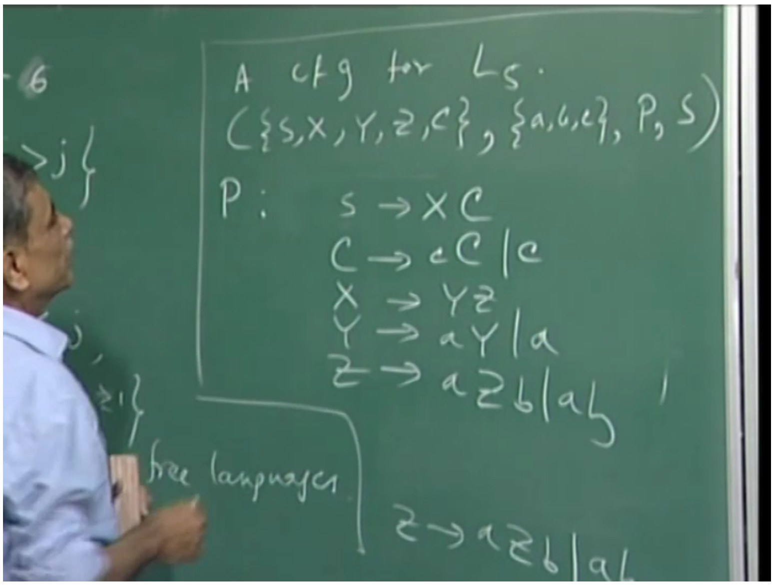
What is the role of Y? To generate this initial part of a’s. So Y is again very simple,a Y followed by a. So it will generate one a or any number of a’s. And what is Z? Z we should generatenumber of a’s followed by equal number of b’s. Again we know how to do that. So we can say a Z b, a b, right?

(Refer Slide Time: 43:09)



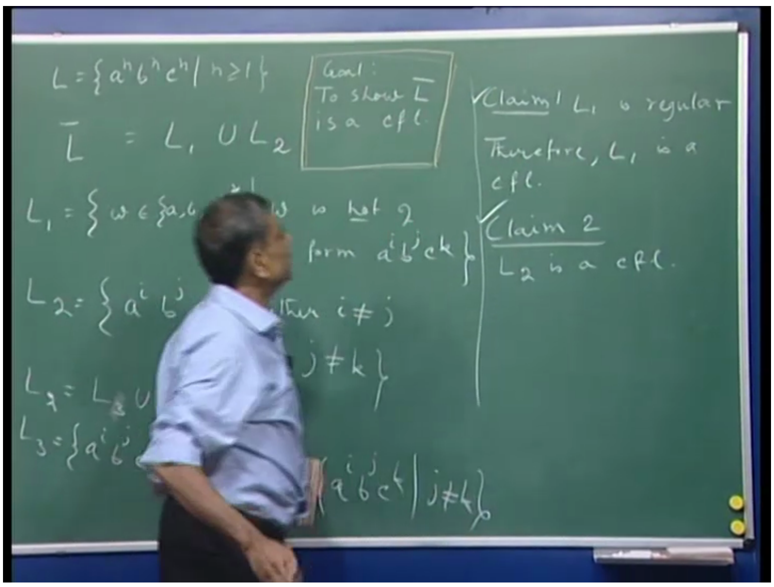
So I can claim that this grammar which consists of start symbol and these nonterminals X C, Y Z and the terminals are a’s, b’s and c’s, right? So in fact let me clean it up and clean this what I am making. So let me say this is my set of productions and let me just rewrite it here so that they are all in one place. This is P and your grammar is thereforeS, X, Y, Z, C. These are your nonterminals. Your terminals is of course a b c.this is the set of terminals, this is your set of productions P and of course the start symbol is S, this grammar.

(Refer Slide Time: 44:40)



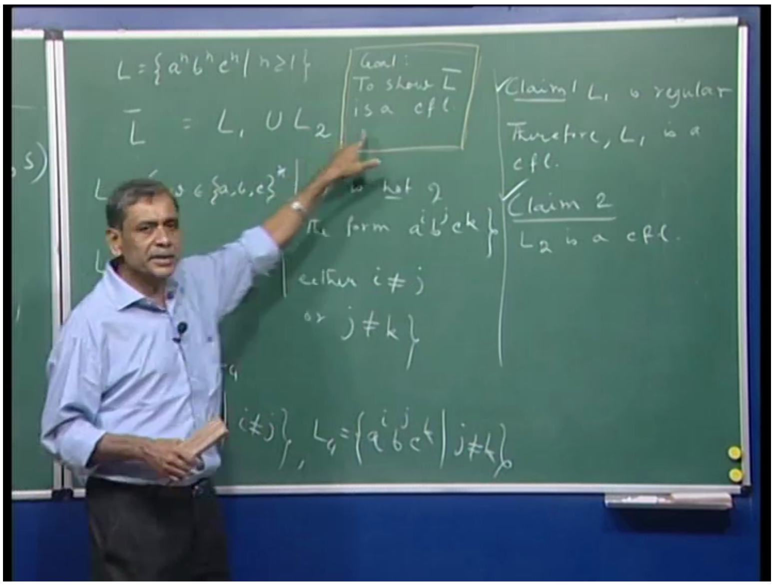
This grammar generates L 5 and if you are convinced that this grammar generates L 5 this then on your own it would be very easy to write a grammar for L 6. So that proves this claim that L 5 and L 6 are context free languages. Since L 5 and L 6, their union is L 3 and each is a context free language. So therefore L 3 is a context free language. And in a similar manner we can prove that L 4 is also context free language, right? Now L 2 was L 3 union L 4, so L 2 is a context free language. So this claim is also there. Claim 1 was already you know we argued.

(Refer Slide Time: 45:34)



So L compliment is L 1 union L 2. Each one is a context free language so L complement is a context free language. So what we have finally managed to show that L complement is a (con) context free language. So what is the point? That I have a language L which is a n b n c n, right? N is greater than equal to 1. This set of strings that was not a context free language butthat languages complement is L complement which is we managed to show that it is a CFL.

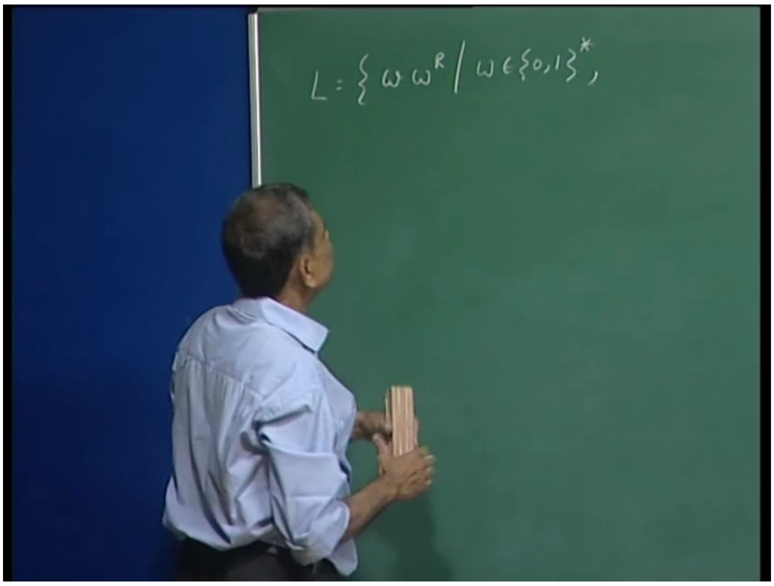
(Refer Slide Time: 46:18)



So here is a pair of languages. One is a CFL, its complement is not a CFL. In this manner you know this is something very different. This behavior is very different from regular languages because regular languages are closed under complementation. If some language L is regular then complement is also regular. But here sincethis class of context free languages, this class is not closed under complementation we have such a possibility having a pair of languages, one is a complement of the other, one is a CFL the other is not.

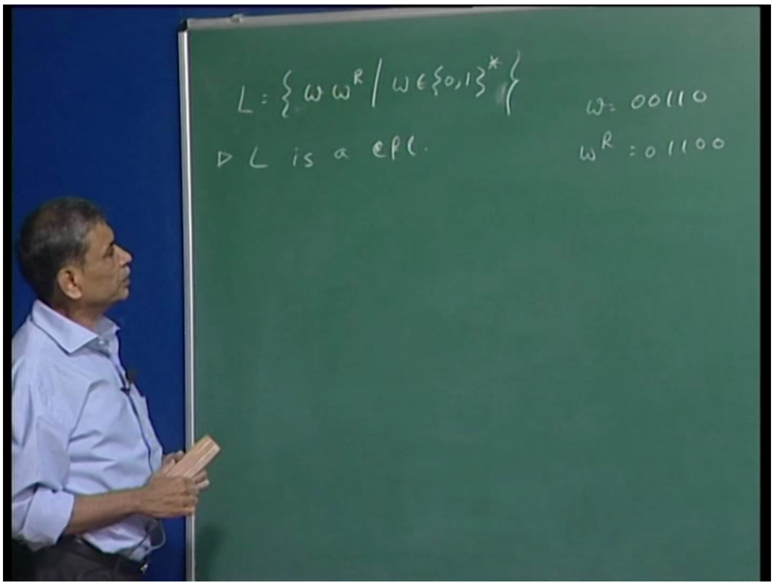
You knownot in the last class maybe the class before that we tried to informally saywhat kinds of languages are context free what are not by looking at some examples after we had proved the pumping lemma. And for example I will elaborate that a little bit that point. Consider this language L w w r where w is a string over 0 1.

(Refer Slide Time: 47:50)



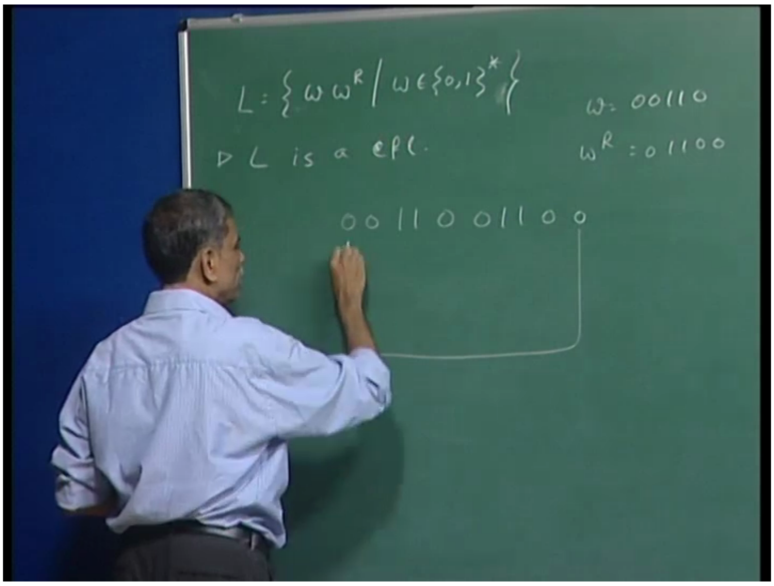
And by w r what does it means? Thesame string reversed, right? Sosuppose w is 0 0 1 1 0, then w r is reverse, right? So you start from here, 0 1 1 0 0, right?What we said or what we are trying to saythe language which consists of two strings such that one is the reversal of the other then this language is a CFL. So L is a CFL and you can prove it easily. I think we even wrote a grammar for it.

(Refer Slide Time: 48:39)



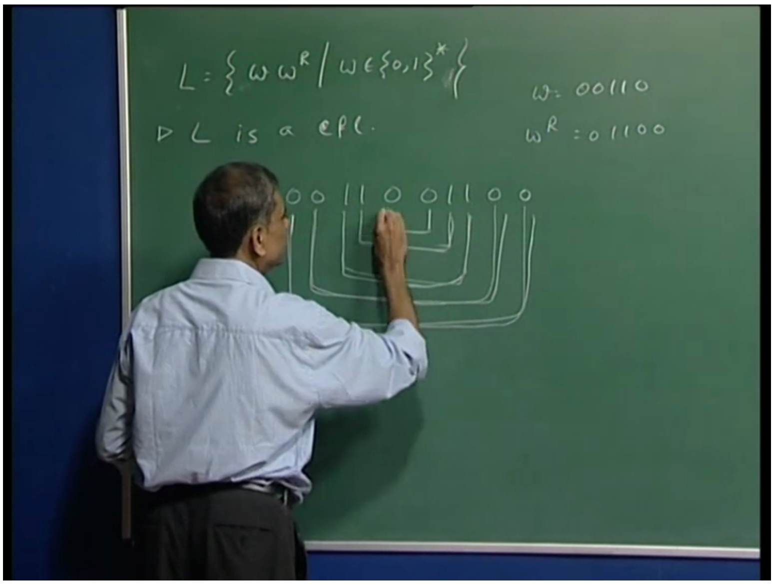
But you see what is happening? Solet us just take this. This is w, this is w r. So this is w and this is w r, 0 1 1 0 0. Now there is a correspondence. What is the correspondence between I mean pairs of symbols? This corresponds to, so this and this has to be the same, is not it?

(Refer Slide Time: 49:07)



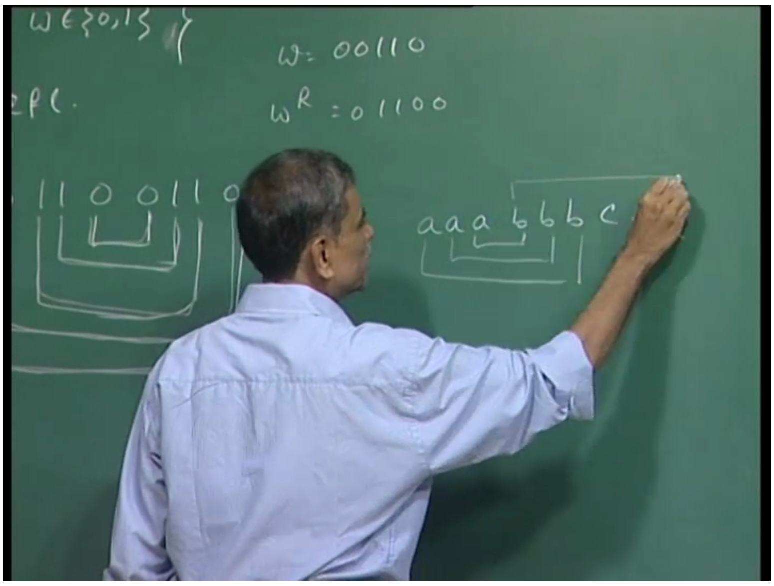
Because you know w you have reversed it so the first symbol of w becomes the last symbol of w r. Similarly second symbol of w is the second symbol of w r and so on. So you see these symbols kind of corresponds, okay. And in this correspond mean here we are saying that they are same. It could have been different also. That is not the point but the point is you see these correspondences, thepairs of symbol which correspond they nest. This is called nesting, right? That this encloses this, this encloses this.

(Refer Slide Time: 50:03)



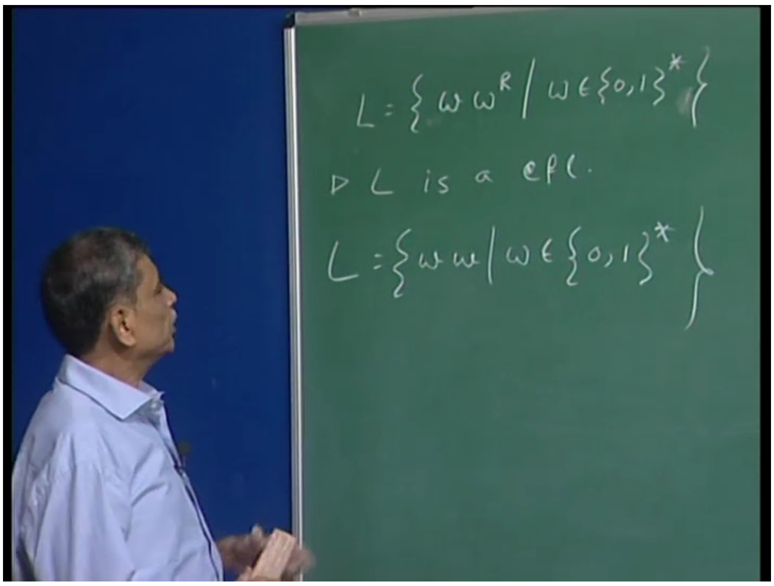
You know I am saying this informally but you understand nesting. Might have come across in some other context. So context free languages are very good in taking care of nesting, okay. Correspondences which are basically nesting that can be taken care of. Now you see informally why a n b n c n is not a CFL? So we can say this corresponds to this. Of course this corresponds to this, this corresponds to this. We also wanted to say things likelet us say that this corresponds to let us say suppose we say this.

(Refer Slide Time: 50:56)



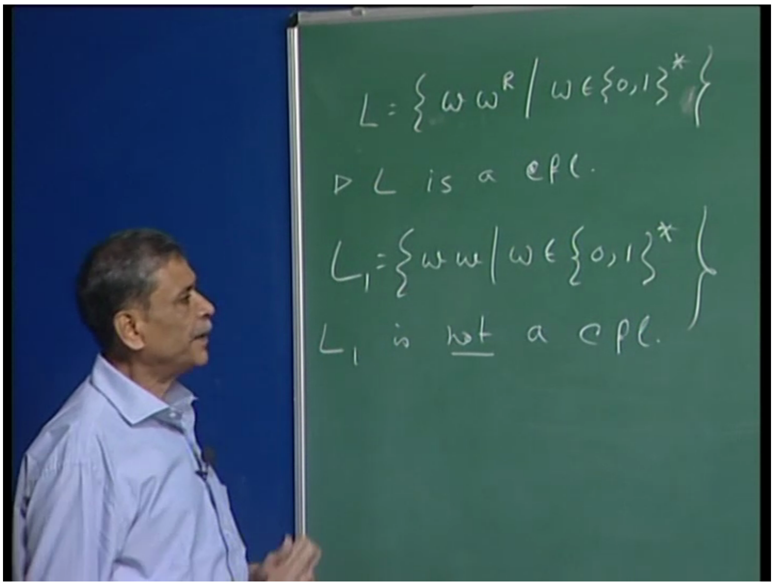
You know we are trying to nest. So but you see these two nestingare kind of clashing, right? This is something we cannot take care of in context free languages. So on the other hand if I had just one set of nesting which we can take care of. So there is something looks very similar but yet it is not a context free language which is the following language, okay.

(Refer Slide Time: 51:45)



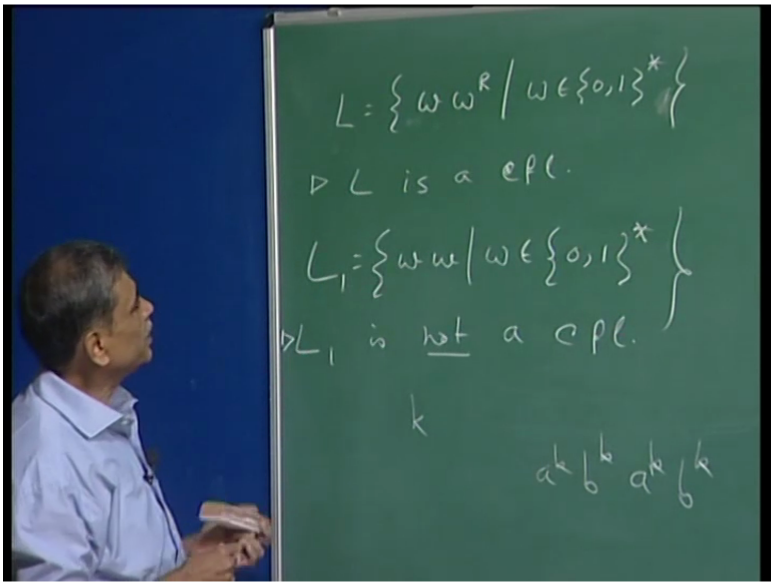
So what we are saying that this language L consists of two strings, right? So basically one string w followed by the same string w, right? You can take any binary string, write it twice, you will get a member of the language L. Now let me call it L 1. L 1 is not a CFL, okay.

(Refer Slide Time: 52:18)



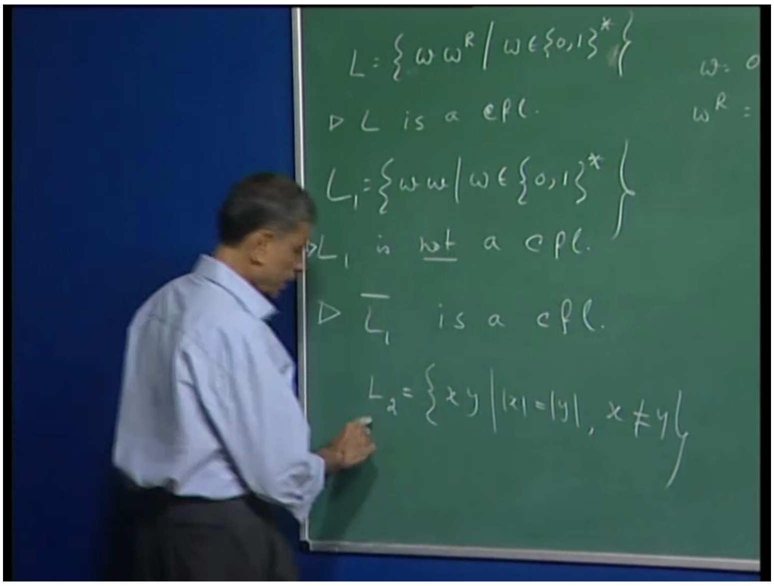
Now we can prove this very simply by using pumping lemma that suppose your pumping lemma constant for this language was k then consider the string a k b k a k b k and you will get a contradiction, right? So it is easy to show this L 1 is not a CFL. Of course we know Lis a CFL.

(Refer Slide Time: 52:47)



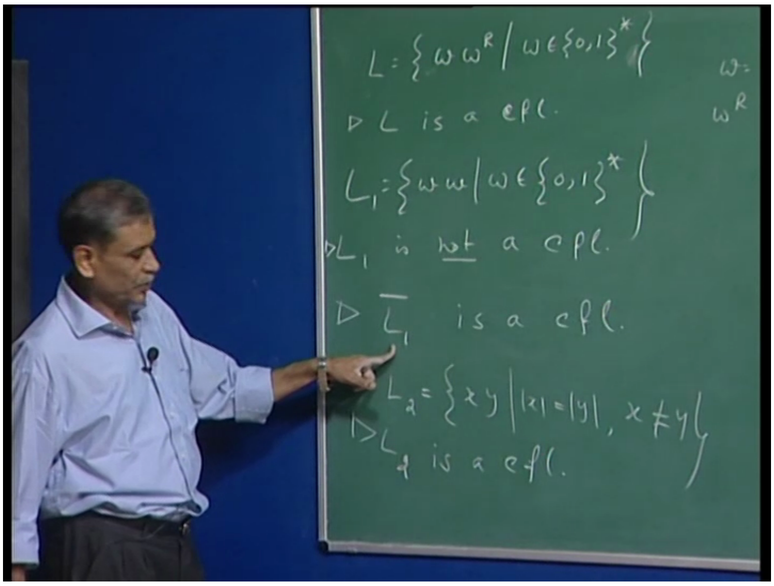
And now in fact this L 1 gives another example that L 1 is not a CFL however L 1 complement is a CFL. Andthis proof is quite interesting. That is why I thought proof that L 1 complement is not a CFL, we would like to see this. What I am going to show which is slightly different from this language but you will be able to show thatyou can take care of L 1 complement once we take care of that language. So let me just say L 2 is this language x y such that length of x is equal to length of y and x is not equal to y, okay.

(Refer Slide Time: 54:00)



So essentially L 2 is a set of strings. They have length given such that the first half is different from the second half. Today we are running out of timeso we will just claim or state what the proof that we are going to show in the beginning of the next class that L 2 is a CFL. And you see my point is because L 2 is a CFL then you will be able to prove that L 1 complement is also a CFL.

(Refer Slide Time: 54:48)



However this proving that L 2 is a CFL is quite interesting which is what we are going to do in the next class.