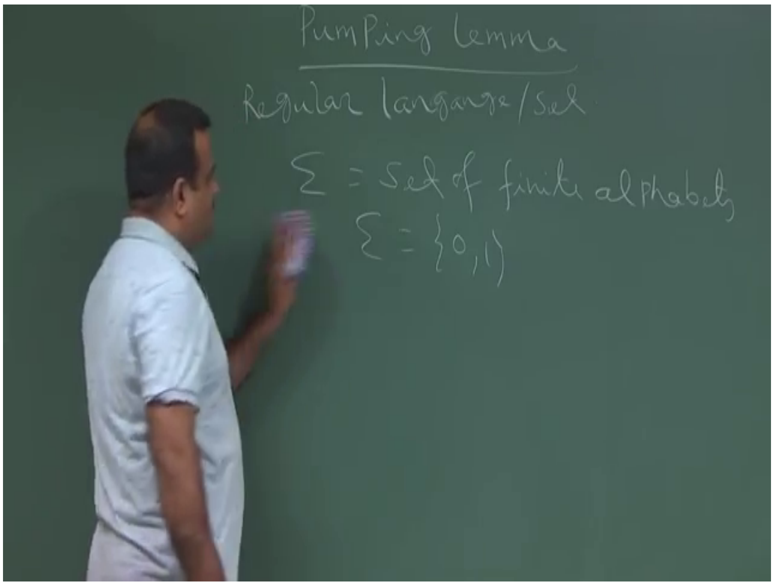
Introduction to Automata, Languages and Computation

Prof. Sourav Mukhopadhyay
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 27
Pumping Lemma

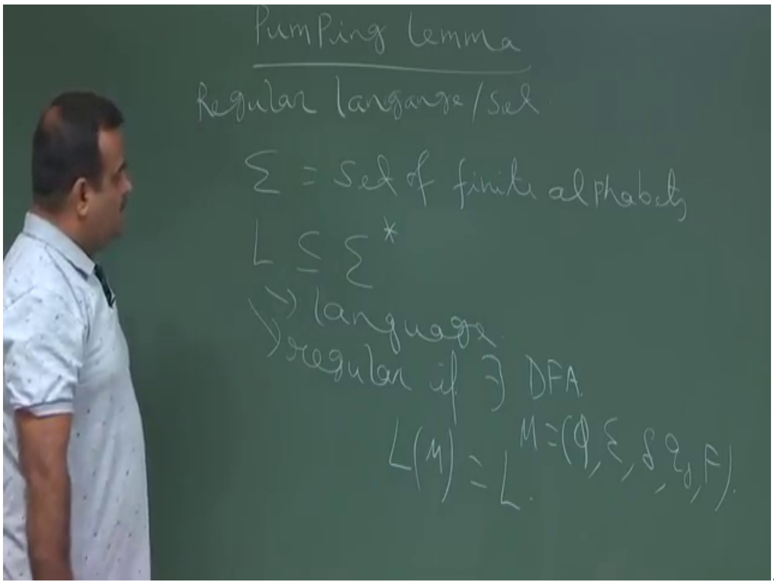
So we are talking about the regular language, regular set. So, today we will discuss the necessary property of a regular language. Suppose, a language is regular, then we have this property, which is called pumping lemma and this is very powerful tool to test whether a language is I mean if it is not regular. I mean if it is failed, the pumping lemma then we can straight away, say the language is not regular, because if it is regular then the pumping lemma will be will hold. So, what is pumping lemma?

(Refer Slide Time: 00:53)



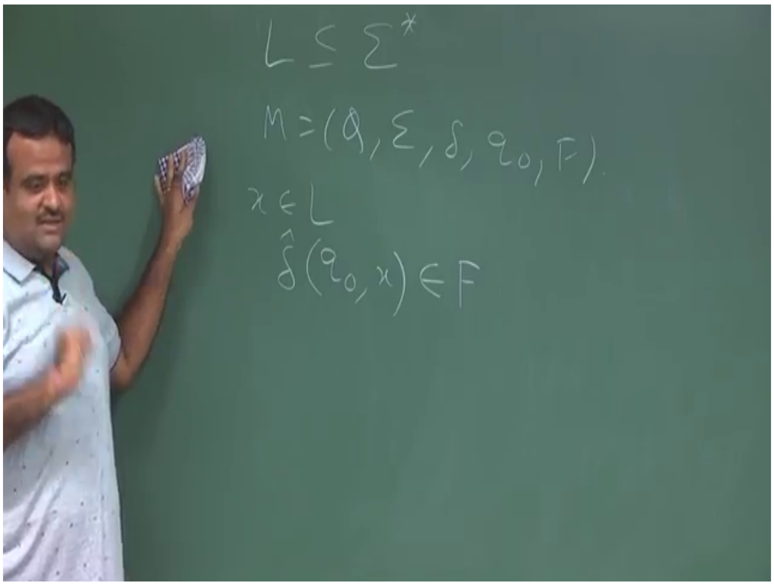
So, this is for a regular set or regular language. So, again what is the regular set? regular set is suppose, we have a, this is the set of finite alphabet, this is alphabet set. It could be 0 1, it could be A B anything. I mean this is the alphabet set and we have a L which is a subset of sigma star.

(Refer Slide Time: 01:32)



Sigma star means we, it includes the epsilon also. Now, this is called language, any subset of sigma star is called language and when we say this language is regular or regular set, if there is a automata which is accepting this. That means, if every string of this language, every member of this every string belongs to this language is accepted by that automata. So, this is a regular set if there exist a D F A, if there exist a D F A M with the same set sigma delta q 0 F such that the language of the D F A is same as language of L; that means, every string of this set must be an accepted string of this D F A ok.

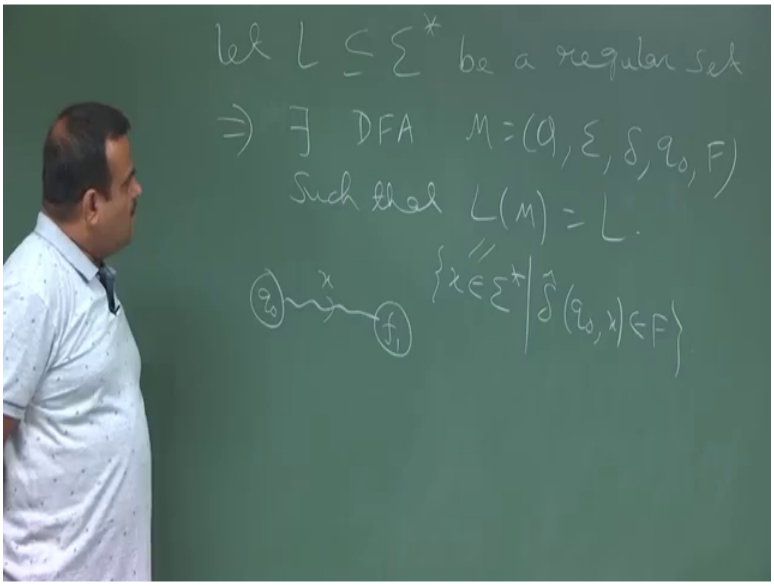
(Refer Slide Time: 02:50).



So; that means, every. So, given this I mean you must have this D F A with the q is the finite number of states and this is the sigma this is the delta transition rule and q 0 is the starting state and M is the final state. So, L will be the regular, if the language of this D F A same as language of L so; that means, if x belongs to L then delta add up q 0 x this must belongs to F.

So, every string of L is accepted by this D F A. So, that is the regular language ok. Suppose, we have a regular language so, then we will have a necessary condition for a regular language that is called pumping lemma. So, suppose L is regular.

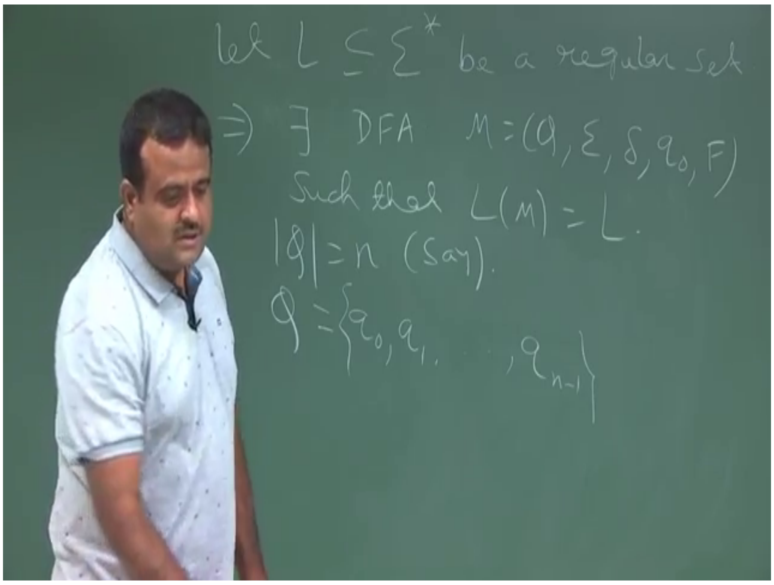
(Refer Slide Time: 03:53)



Let L which is a coming from sigma star be a regular set or regular language, then this imply was, this imply we have a D F A there exist a D F A which is accepting this say Q q 0 f, which accept this such that the language of this D F A is said as language of L.

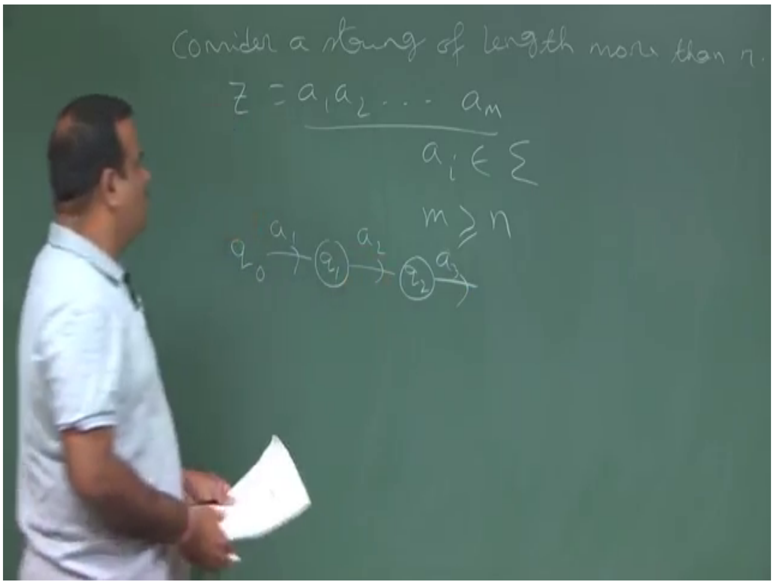
Language of D F A is nothing, but this is the set of all string which is accepted by this D F A, you start with q 0 with this x we must reach to one of the final state. So, we are at q 0 with the x. We must reach to a one of this final state. It may have one or more final states ok. So, this is the now, this is the D F A. So, it has a finite number of states Q is finite.

(Refer Slide Time: 05:09).



So, suppose Q is n, say let n is the total number of states. So, you just denote Q by say q 1 q 2 like this. So, we will just denote this by. So, q 0 q 1 q i these are the, no this is not q i. So, these are the states say q n minus 1 if we have n number of states and some few states are the final state, at least one and this is the starting state ok. Now, we will consider a string of length more than n. We consider a string from this of length more than n ok.

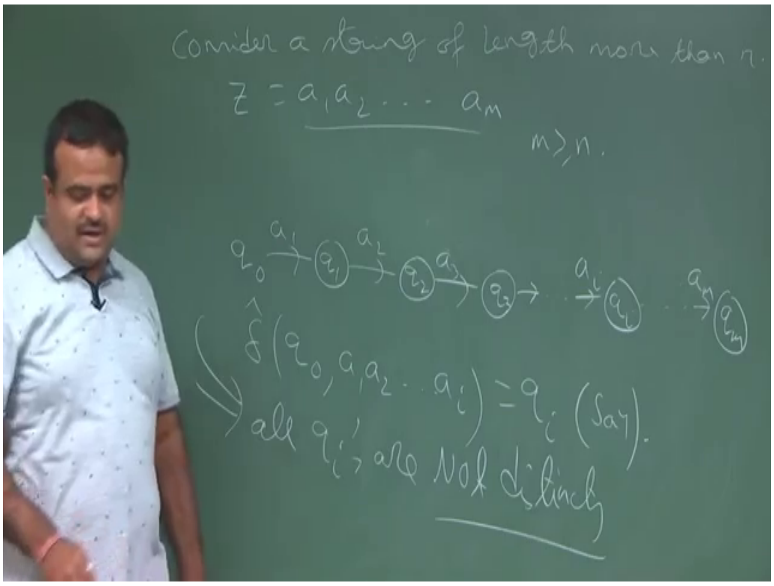
(Refer Slide Time: 06:15)



Say we write this x as a 1 a 2 a n and ais are coming from sigma and m is greater than n or greater than equal to n. So, this we write as Z ok. One of the send note, send in the class note ok. So, let us consider, consider a string of, consider a string of length more than n, n we know, n is the number of states more than n and that is the string, say. If it is strictly greater than; then we have this strictly, otherwise we can have greater than equal to ok.

So, now, we will read the string. So, we start with q 0 proposed with a we go to q 1. So, this is with a 1 and with a 2 we go to again q 2 like this, we do a 3. This is we are not, I mean this is some names ordering of some states ok. So,. So, with q q 2 we go to q q 3 like this dot dot dot, say suppose with a i we go to q 1 ok.

(Refer Slide Time: 07:47)

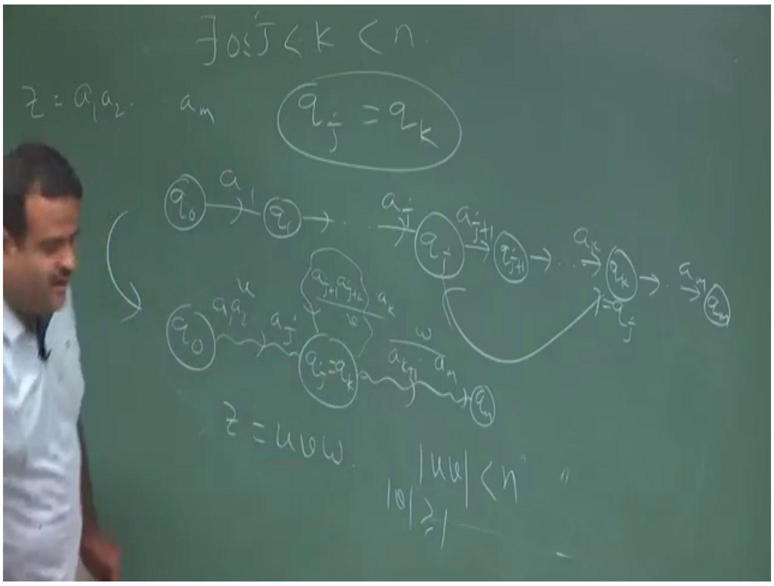


So, ; that means, delta of delta hat of q 0 a 1 a 2 a i. So, this is the substring a 1 a 2 a i is basically q i say. This is a, it has to reach to a state which is q i ok. Now,. So, we are at q i. So, if we continue this so, then up to this we will be at say q m, after this a m ok. Now, we have only now, we have only how many states? We have only n number of states.

Since, we have n number of states and here m is greater than n or greater than equal to n then some of the then all the qis cannot be distinct. If the all the qis are distinct then we have how many states? We have total m plus 1 state which is more than m which is more than m plus 1.

So, that is not possible. So, these all the states over here cannot be distinct. So, that is one observation all qis are not distinct. So; that means, some of the qis has to be repeat, because we do not have that many state we have total n number of states and here we have m plus 1, which is greater than n so; that means, some of the qis has to be repeated so; that means, what?

(Refer Slide Time: 09:51)



That means, we should have a state over here, that means they are exist a i and k which is less than n such that because such that that q i q J will be q k q J will be q k. So; that means, after this say let me draw it again. So, we are going like this q 0 a 1 we go to q 1 dot dot dot say we had q J with a J.

Now, suppose up to q J it is ok, all are distinct now then q J plus 1 then up to q k, this q k must be same as q J then this is a k dot dot dot then this is q m a m. So, this will be the situation. So, these state and these state has to be same, because we have only n number of states. So, they are there will be in the middle. There will be two states, which will be same and that is two state has to be number from 1 to a 0 to n and J must be less than k. So, this is the observation. So, these two states has to be same.

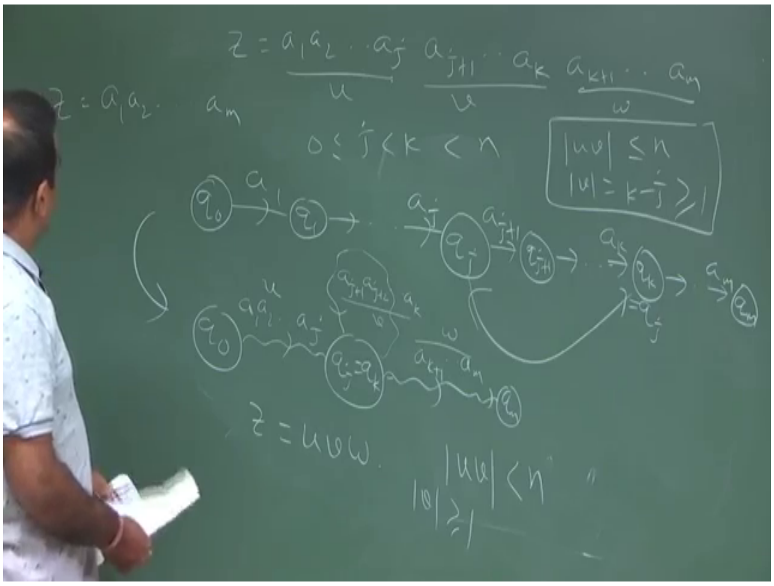
So; that means, after q J it will again come back, by the string of a J plus 1 like this. So, if you write this, this is similar to; so q 1, so we go like this. So, we reach to the state q J, which is same as q k and this is the string of a 1 a 2 up to a J all are over distinct over here, then it has to repeat. So, this repetition this is reading the string of what a J plus 1 a J plus 2 up to a a k.

This is repeating and then we have a k plus 1 to a m and then we are reaching to q m ok. So, this is the situation, this is equivalent to this, this is the starting state. So, if we have a string of length more than n that is say a 1 a 2 a n then we must have this situation, what is the situation then we must have this string I mean this two states in the middle which will be same; that means, in that path we have a repetition.

So, we denote this by let us use some symbol say this is u and this is v and this is w. So, this is our Z. So, Z will be written as u v w, where length of u v is less than n and length of v, this is v, length of v must be greater than 1, because length of v is nothing, but this is our v.

So, let me write it again.

(Refer Slide Time: 13:53).



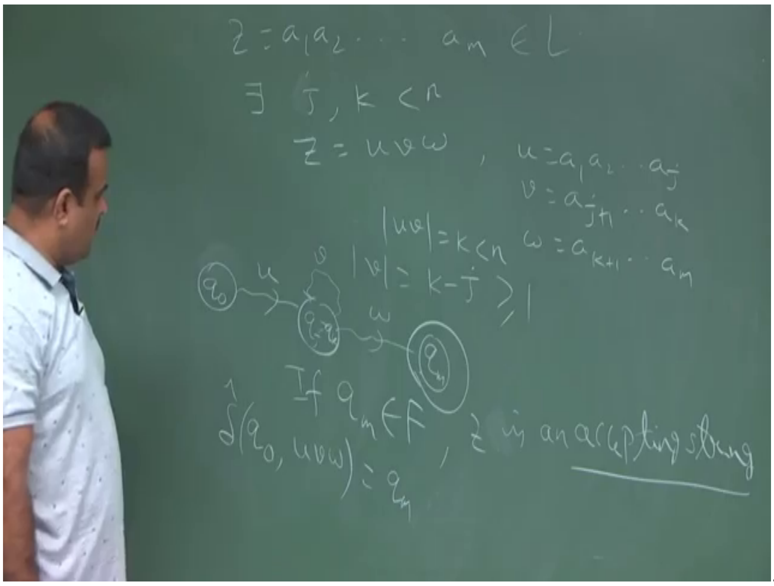
So, Z will be written as a 1 a 2 up to a J. So, that is this u and then a J plus 1 to a k. This length is k minus 1, this is our u v and a k plus 1 to a m and this is our w ok. So, this is length of so now, length of this, because i J are coming i J are less than k and i is less than sorry, i J k sorry J k J and k are less than n and J is less than k.

So, from here we can write u v, u v length of u v is k which is less than n or less than equal to n, a length of v is k minus J which is greater than equal to 1. So, this is the thing and with u, we are going to this state q i q J, which is same as q k and there we need a string, again we come back to here.

So, there are some repetition of the state, there may be some other states which is also repeating which may be distinct. So, this will again come back here, but at least this is the first state which is equal to the which is, which will be repeating, because it has to be there, because otherwise we have only n number of states and if we need to have all the string states then we need total m number of states and sorry m plus 1, which is greater than m strictly greater than n, because n is greater than equal to n.

So, m plus 1 is greater than equal to n plus 1. So, that is not possible. So, that is why it is there ok. So, this is the observation. So, let me refresh this. So, this we will write in a theorem form. This is basically called pumping lemma, I mean you have to write in a theorem form. So, what we have?

(Refer Slide Time: 16:25)



.

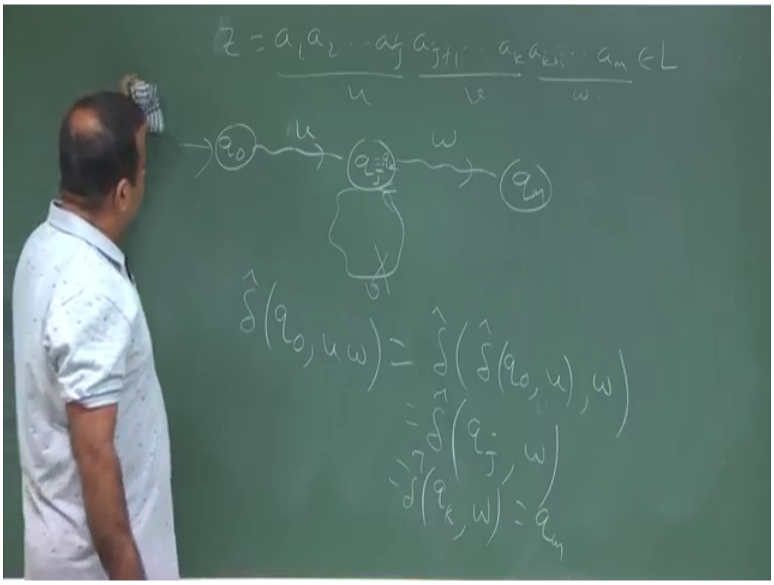
So, we have given a string which is more than this, then there exist a J and k which is less than n such that Z will be written as u v w, where u is a 1 a 2 up to a J, v is a J plus 1 to a k and w is a k plus 1 to a m u v w ok.

Now, such that this v, u v which is of length k which is less than n, because k has to be less than n and the v, the length of v which is nothing, but k minus J which is also greater than 1 greater than equal to 1. Now, if q m. So, it is ultimately going to the q m, we are starting with q 0, we are reading this u and we are going to q J, which is same as q k and then with this v we are again come back to this and then with w we are going to q m.

Now, if Z is a accepting string; that means, this if q m belongs to F if q m is a final state of the automata then Z is an accepting string, accepting string ok. So, if you take Z from the, L which is a regular language then every every string of L has to be accepted by this automata so; that means, if Z is like this then q m has to be n final state.

Now, from here we will draw some observation. So, q q m is final state means. So, this is delta Z of q 0 with Z, Z is u v w, this is nothing, but q m ok. Now, from here we can have some observation. So, can you right this?

(Refer Slide Time: 19:36)

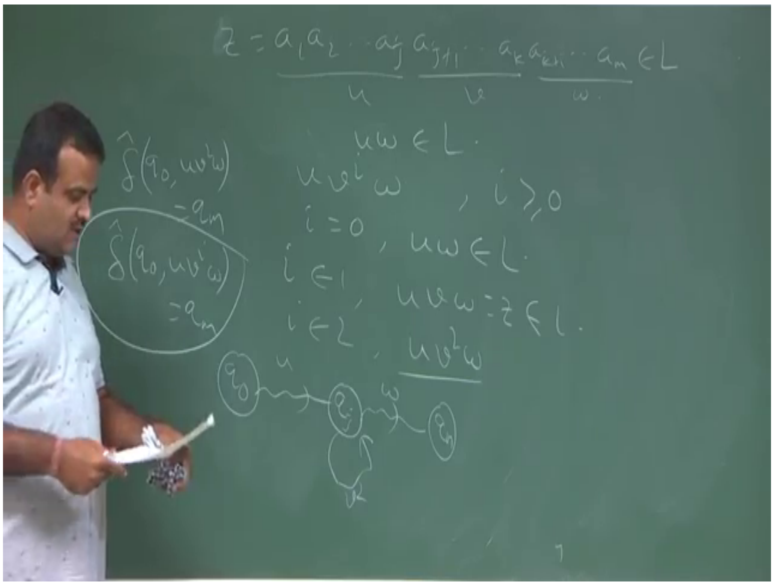


So, we know Z is a k plus 1 a m. Z is a acceptance string and which is written as this is u, this is v, this is w and we know that from q 0 with u we are going to some state q J and with v we are coming to same state that is q k, which is same as q J. So, this is with v we are coming to same state and then with w we are again going to q m.

Now, if we discard this string then also you are reaching to the final state with the starting state, because if we just have a string that u and v u and w sorry, then also delta hat of what is delta hat of q 0 u and w, it is also q m, because we are by u we are, because this is same as delta hat of delta hat of q 0 u then we can read w.

This is the way, this is easily can be easily proved, this is the recursive definition of delta hat, this is the way how we read delta hat. So, this is nothing, but delta hat of, so, where is q 0 u? We are going to, which is same as q k w. Now, from q k w we are reaching to q m with the by reading the string w. So, we can just omit this v, still we are reaching to the final string. This is one observation even we can repeat that v. So, how to do that?

(Refer Slide Time: 22:06)



So, if we if this is L then we have observe u v u w is also belongs to L not only that u v to the power i w v to the power i means, i is greater than equal to 0. If i is 0 then v is not there if i is 0 then it is u v and just now, we have seen u v is belongs to L.

Now, if i is 1 then this is u v w this is nothing, but Z this is belongs to L then if i is 2 then u v square w. Now, how this is be in this will be again reach to the q m, because so to read this what we do? We start with q 0 with u we go to q J.

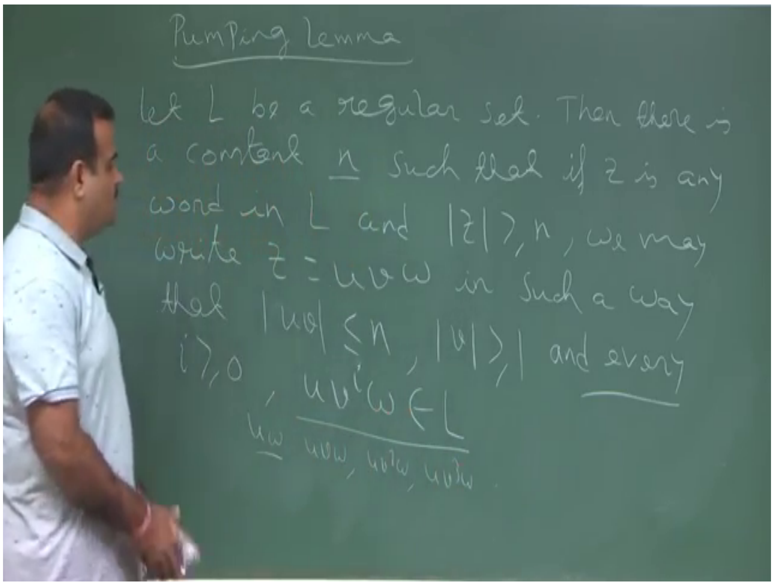
Now, with v we go here again with same v, we will off here pumping we are pumping this state. So, 2 times we will do that and we will then after that we will read w, this is 2 times, this is v square, we reach to the q m.

So; that means, delta hat of q 0 u v square w is also q m. So, not only v square any i delta hat of q 0 u v to the power i w is q m. This is nothing, but called pumping lemma, we can pump. There is a there is a substring v, which we can pump there, we can pump there.

So, let us write it in a formal way, but this is nothing, but pumping and this is the necessary condition. If our language is regular then it should have this property.

If our language is regular then we should have a n, n is nothing, but the number of states for that automata and if we take a string more than n then in that string there is a substring, which we can reuse it we can pump there to get a bigger string and still we that that should also belongs to L. So, this is the necessary condition let us write this lemma.

(Refer Slide Time: 24:28)



This is the formal way let L be a regular set, then there is a constant n such that this constant is nothing, but the number of state for that D F A which we have seen just now. Such that if Z is any string any string or word in L and if length of this is more than n, we may write ; so, there has to be some u v u v w in such a way such a way that u and v length is less than n, a length of v must be greater than 1.

There has to be such string and for every, for every i greater than equal to 0, this u v to the power i w must belongs to L and this must be true for all i. So, this is called pumping lemma ok. This is the statement of the pumping lemma.

So, what it is telling? It is telling that if we have a regular set L then we must have a constant n, that is nothing, but the number of states in the D F A. Anyway, we must have a constant n such that if you take a string from this L whose length is more than n, then we must able to write this in terms of u v w. These are all substring such that the length of this substring u v is less than n and this substring has to be greater than 1 and we can just, this is for pumping we take u, because if i is 0 then this is nothing, but this if i is 1. This is u v w, if i is 2 u v square w, if i is 3 u v q w, like this.

So, all must belongs to n this is the necessary condition, if this is regular then we have this, but this is not sufficient. If this condition satisfied we cannot say that a language is regular. So, if a language is regular then it must have this property. So, to check your language is non regular this property is very much useful, this lemma. So, that we will discuss in the next class, we will take some example and there will prove that this language are non regular which is, because of failing this pumping lemma. We will talk about in the next class.

Thank you.