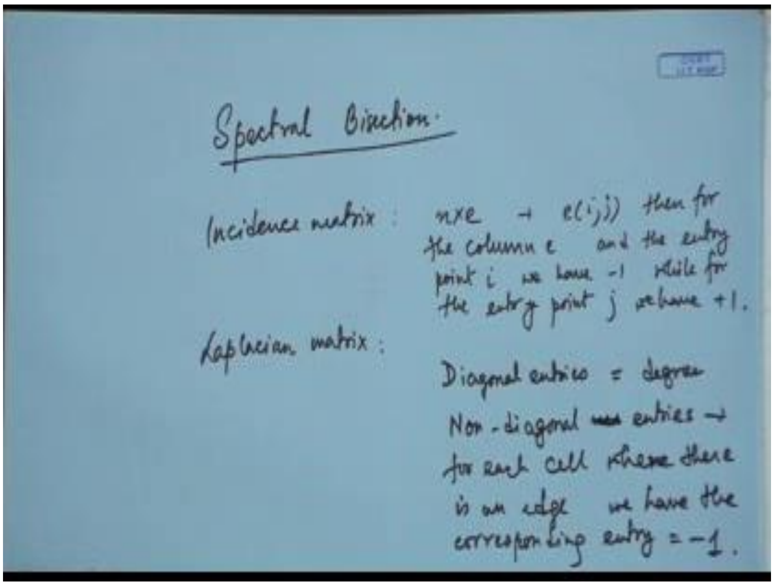
Complex Network: Theory and Application
Prof. Animesh Mukherjee
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

Lecture - 16
Community Analysis – V

In the last lecture, we have already discussed about how to do community analysis, and we have seen quite a few techniques to do community analysis, and we also started with this concept of Spectral Bisection.

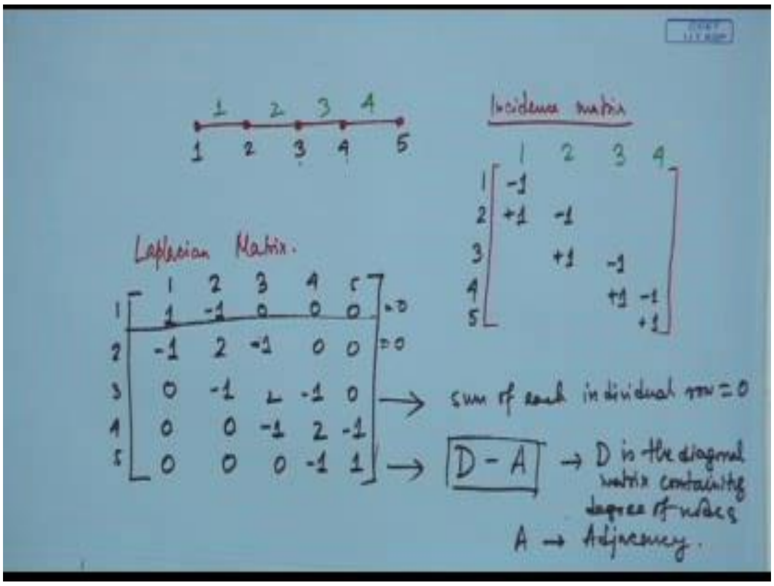
(Refer Slide Time: 00:38)



So, today we will continue with this idea of spectral bisection. And in order to do spectral bisection, last class I already define 2 basic matrices that we will need, for the rest of the analysis; one is what we call the incidence matrix, and the other is the Laplacian matrix. So, the incidence matrix as we said, is a n cross e matrix where if there is a edge e (i, j), then for the column e and the entry point i, we have minus one, while for the entry point j we have plus one. So, we will look at typical examples soon. We also looked into the definition of the Laplacian matrix, which are the diagonal entries at the degrees.

Whereas the non-diagonal entries, for each cell, where there is an edge, we have the corresponding entry equal to minus one. So, that is how we define the 2 basic matrices; the incidence matrix as well as the Laplacian matrix. So, now we will take an example and see how these 2 matrices are constructed given a undirected graph. So, let us take a very simple example and start. Let us say we have this graph.

(Refer Slide Time: 03:11)



This is basically align graph of 5 nodes; 1 2 3 4 and 5. Now, the edges can also be numbered; this is edge 1, edge 2, edge 3 and edge 4. So, now, as we have defined the incidence matrix would look like this. So, as we have said, it is the n cross e matrix. So, basically it will be like this, we have the n c r 1 2 3 4 and 5. Whereas, we have the e is on the columns 1 2 3 and 4.

Now, we see that between the node 1 and 2 we have an edge. So, one entry point becomes minus 1, while the other entry point becomes plus 1. Similarly, we see that we have an edge between 2 and 3. So, in the column of 2, which is the column of the edge 2, we have one entry point as minus 1 at 2 and the entry point as plus 1 at 3 in this way we have an edge between 3 and 4. So, in the column of 3, we have one entry point as minus one here, and one entry point as plus 1 here, and similarly for the last case, we have one entry point as minus 1 here, and one entry point as plus 1 here.

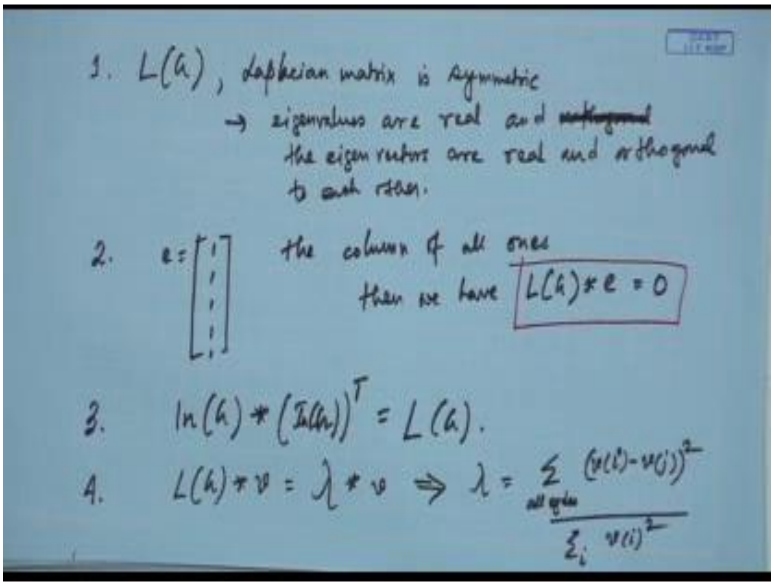
So, in this way we can construct the incidence matrix. Note that each column here represents a particular edge; edge number 1, edge number 2, edge number 3, edge number 4. And if the edge is between points i and j then point i gets minus 1 point j gets plus one, in this way we have constructed the incidence matrix. Similarly, we can also construct the Laplacian matrix. So, as we said that the Laplacian matrix is a n cross n matrix. So, you have 1 1 3 4 5, n 1 2 3 4 5. And now each entry; so as you see the diagonal entry is nothing, but the degree of the node.

So, here the diagonal entry for 1 1 is the degree of the node 1 is 1. The degree of node 2 is 1 plus 1 2. The degree of node 3 is 1 plus 1 again 2. The degree of node 4 is 1 plus 1 2, and the degree of node 5 is 1. So, in this way you can construct the degree of each of the nodes, and in this way the diagonal entries are filled up. Now, look at the other entries. So, there is only 1 edge between 1 and 2. So, 1 and 2 there is an edge. So, you put a minus 1 here, there is no other edge, at least in this particular row. So, the other rows become, the other entries are equal to 0. For 2 1, so, 2 1 you have an edge. So, this is minus 1, and you also have an edge from 2 to 3. So, this is minus 1, rest is 0.

Similarly, for 3 1 you do not have any edge, so it is 0. From 2 3 2 there is an edge, and from 3 4 there is an edge. So, you have minus one entry here, and you have 0. So, one 4 you do not have any edge, 2 4 you do not have any edge, 3 4 you have an edge. So, it is minus 1 and 4 5 you have an edge. So, it is minus 1. So, for 5, all the entries are 0 accept 4 5 which is a minus 1 at the entry 5 4. So, now you see the interesting property of this particular Laplacian matrix, is that some of each individual row, is always equal to 0.

You see this row, take this row the sum is 1 plus minus 1 plus 0 plus 0 plus 0 that is 0. So, this is minus 1 plus 2 minus 1 plus 0 plus 0 that is 0. So, in this way sum of all the rows is equal to 0 in this particular Laplacian matrix. And these Laplacian matrix can also be written as D minus a, where D is the diagonal matrix, containing degree of nodes and a is nothing, but the adjacency matrix. So, basically if you implement D minus a, you will recover the Laplacian matrix. So, now, with these 2 notions of the incidence matrix and the Laplacian matrix, we will introduce certain properties, which will finally allow us to actually do a partitioning of the graph, based on spectral ideas.

(Refer Slide Time: 09:09)



So, some of the properties that these particular matrices follow are that I write here; number one is that L g which is the Laplacian matrix is symmetric. Therefore, its Eigen values are real, and the Eigen vectors are real and orthogonal to each other. So, basically, since it is a symmetric matrix, since L(G) is a symmetric matrix, because the graph g is an undirected graph. So, that is what our assumption to start off with.

So, since we assume that graph g is an undirected graph. So, we have a L(G) matrix which is the symmetric matrix in this case, and therefore, all the Eigen values are real, and the Eigen vectors are real and orthogonal; that is perpendicular to each other the vectors are perpendicular to each other in the vector space. So, that is the first property. The second property is that, if we have a vector; say e which looks like this 1 1 1.

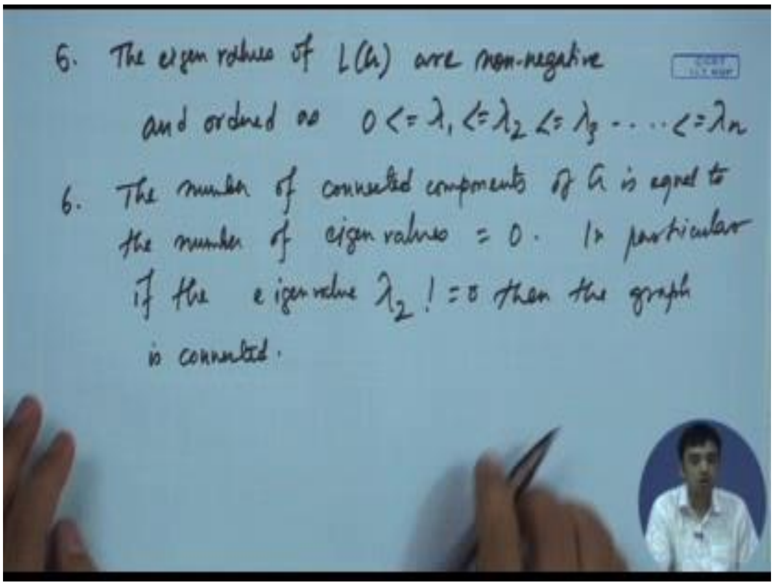
So, this is a vector of all ones, the column vector of all ones. So, then we have L(G) into e is equal to 0. So, why this is true? So, for instance if you look at the previous example that we have seen. So, take this example, here as we have seen that, sum of all the individual rows is equal to 0.

Now, if you put a column vector of all ones what will happen. The first entry will do nothing, but sum all the entries of this particular row. The first value the product with this one will do nothing, but sum all the entries of the first particular row. Similarly, the multiplication with this one will sum all the entries of this particular row. So, the sum of this first entry is equal to 0 as you know, the second entry is equal to 0 and so on and so forth. Therefore, from this observation we have this particular interesting property, that L(G) into a vector of all ones is equal to 0.

Now, we have 2 more important properties which we will prove in a while. Let us first write the properties; the incidence matrix multiplied by the transpose of the incidence matrix, these actually returns the Laplacian matrix. So, the product of the incidence matrix and it is transpose together, actually this product the incidence matrix and the transpose of the incidence matrix, that actually returns you back the Laplacian matrix. So, this is our very nice property which we will be using in a while.

And, the forth property is that, the Laplacian multiplied by an Eigen vector v is equal to sum constant lambda into the Eigen vector v; so where lambda is the Eigen value. So, these actually tells you that lambda is equal to nothing, but sum over all edges v(i) minus v(j) whole square by sum over all nodes v(i) whole square. So, basically what this says is that lambda. So, if you look at this formula the expression at the numerator is a sum of the squares where as expression at the denominator is also a sum of the square. So, both of them have to be positive and therefore, this ration as to be positive. So, the lambda values are positive; that is what we will prove in a while.

(Refer Slide Time: 14:22)

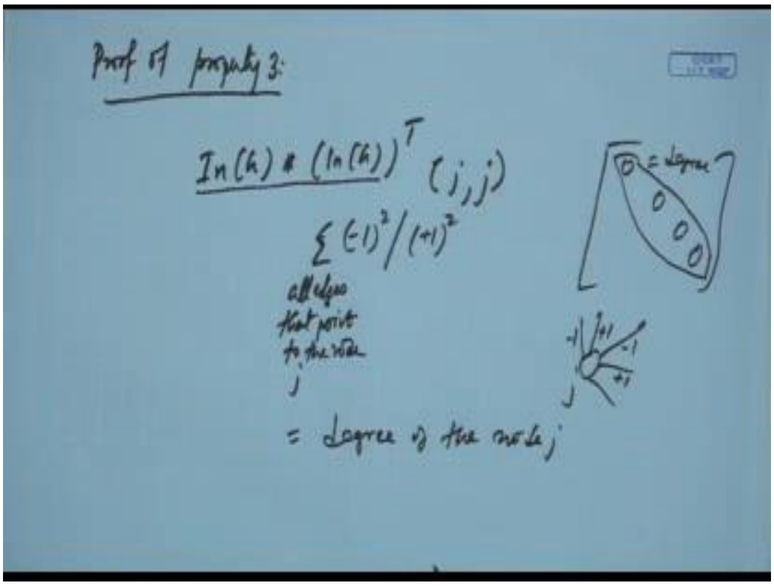


Then there is this fifth property which says that the Eigen values of L(G) are as I said non-negative, and ordered as 0 less than equal to lambda 1, less than equal to lambda 2, less than equal to lambda 3, and upon to less than equal to lambda n. So, in this way you can order the different Eigen values of the Laplacian matrix. So, and then you have the last property which tells you that the number of connected components of g is equal to the number of Eigen values equal to 0.

In particular if the Eigen value lambda 2 is non 0, then the graph is connected. So, basically this is why as I was telling you last day, the second Eigen value is very important. So, there is one point to consider here, since we are assuming the Laplacian matrix which is D minus the adjacency matrix, we are interested in the second smallest Eigen value. If on the other hand we deal with the adjacency matrix directly, and look at it is Eigen vectors and Eigen values, then we will have to look into the second largest Eigen value.

Since, you are looking at a matrix which is D minus a, so you have to look into the second smallest Eigen value, and here the second smallest Eigen value as we have noted is lambda 2, and especially if you have observed that lambda 2 is not equal to 0; that means, there are more than one connected component in the graph; that means, the graph is disconnected. So, if this is non 0, then only you can make the observation that the graph is connected. So, now, one by one we will prove properties 3 4 5 and 6. So, proof of property three. So, this is a very easy proof to see.

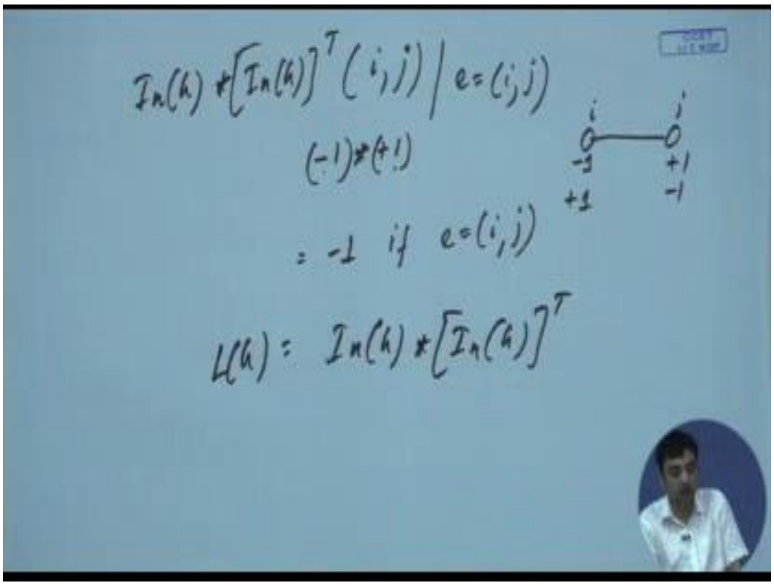
(Refer Slide Time: 17:16)



So, now, assume the product of ING and IN (G) transpose. So, this is actually if you are considering, the cell j j; that is the cell where it is a diagonal. So, imagine the matrix. So, you are considering the diagonal values. Now the diagonal values in this particular matrix and this particular product will be nothing, but actually minus 1 square or plus 1 square. So, depending on whether j was assigned a minus 1 or a plus 1. So, it is either minus 1 square or plus one square, and this will sum over all edges; that point to the node j.

So, basically this will sum over all edges that point to note j. So, what are the all edges that point to node j? If you have this node j and. So, each will have come with minus one or plus one. Now, you are taking a square of this. So, basically that gives you nothing, but the degree of the node j. So, basically you have shown half of it. So, you have shown that the diagonal entries of this product matrix will be equal to the degree. So, as we have seen in the Laplacian matrix also, the diagonal entries are nothing, but the degree of the nodes. So, that part we have already shown. Now you are left to show the, think for the other part.

(Refer Slide Time: 19:23)

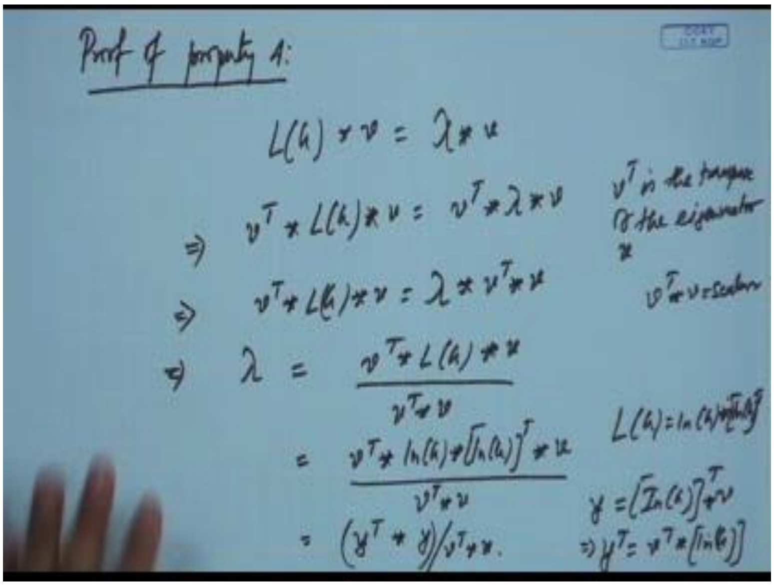


So, now again consider ING times IN(G) transpose, but now i and j are such that, there is an edge between i and j. If there is an edge between i and j then what you will have, either you will have minus 1 into plus 1 or plus 1 into minus 1. This is because one side is i and the other side is j and there is an edge connected between them. If this is minus 1 then this is plus 1. If this is plus 1 then this is minus 1, as per the definition of the incidence matrix.

Now, if you take a product it will be either minus 1 into plus 1 or plus 1 into minus 1. So, these leads to minus 1 if there is an edge existing between e(i,j) and that is what is exactly the definition of a Laplacian matrix. So, we have shown that the product of the incidence matrix and it is transpose will give you in the diagonal entries the degree values, and in the non diagonal entries, if there is an edge between the 2 nodes corresponding to that cell then it will be a minus 1.

So that means this act is exactly equal to the definition of the Laplacian matrix. So, that is why, we have proved the property three, which says that L(G) is equal to ING into IN(G) transpose. So, this is how we can easily prove that there is interesting relationship between the incidence matrix and the Laplacian matrix. Next, we will have to look into the proof of property 4.

(Refer Slide Time: 21:08)

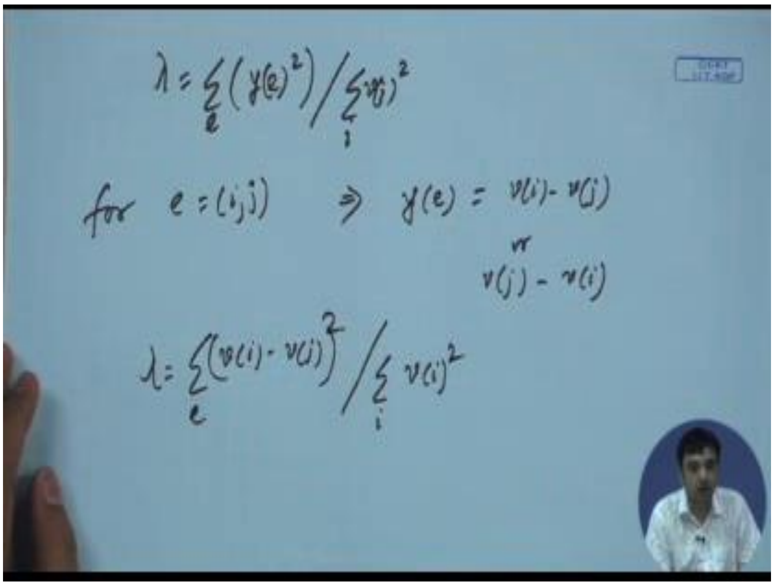


Now, what have we seen in the property 4 that L(G) multiplied by an Eigen vector, is nothing, but lambda multiplied by an Eigen vector. So, now we can write something like this; v transpose L(G) star v is nothing, but v transpose star lambda star v, where v transpose is the transpose of the Eigen vector v. So, now, since v transpose is the transpose of the Eigen vector v. So, if you take a product of v transpose v, this will be nothing, but a scalar value. So, that means, we can write v transpose star L(G) star v is equal to lambda we can take as out as constant v transpose star v.

Therefore, we can write an expression for a lambda as follows; this is v transpose star L(G) star v divided by v transpose star v. And we can do this division since we transpose v is nothing, but a scalar, because you are multiplying the Eigen vector which is transpose so, it will translate into a scalar value, it is no longer a vector.

So, now this can be further simplified. So, v translates into. So, now, L(G) you know that L(G) is equal to ING into IN(G) transpose. So, you can put this in the formula and rewrite the expression as ING multiplied by IN(G) transpose times v by v transpose times v. Now, we will call y is equal to IN(G) transpose v; that will imply y transpose is equal to v transpose ING. So, replacing this expression by y, you have v transpose into ING, is nothing, but y transpose into y divided by v transpose into v. So, we have nicely simplified the expression to y transpose multiplied by y by divided by v transpose multiplied by v. So, now what is v transpose multiplied by v?

(Refer Slide Time: 24:42)



So, now we can further rewrite lambda as. So, the denominator v transpose multiplied by v is nothing, but square of v i’s or sum over all i is basically. So, the v transpose v is nothing, but, you are taking each entry of v multiplying with it with v transpose as entry, that will give you v i square for each entry point i for each node i. Whereas, the numerator is nothing, but sum over all edges y e whole square, because you are now looking at the incidence matrix.

So, the incidence matrix is defined on the edges. So, one side of the edge is i and the other side of the edge is j. 2 nodes are there the node i and the node j. and now if you multiply IN(G) transpose into v into v transpose into ING. This will give you nothing, but the square of these values. So, basically if there is an edge e(i,j), for the edge e(i,j) we can easily see that ye is nothing, but. So, you have one entry for v(i) and the other entry for v(j). So, you have one entry for the node i and the other entry for node v(j). So, what you are doing, you are basically multiplying the v vector with the incidence matrix.

So, the v vector is nothing, but v 1 v 2 v 3 up to v n; all the individual entries for each individual node, and the incidence matrix is nothing, but plus one for one end of the node and the minus one for the other end of the node. So, you have v(i) into a plus one then v(j) the other end of the node will have a minus one. So, you will either have v(i) minus v(j) or v(j) minus v(i). So, it is either v(i) minus v(j) or the opposite v(j) minus v(i). This comes from the fact that you are multiplying the vector of entries in v with the incidence matrix.

So, the vector of entry is, if at one point the vector is v(i) and at the other point the there is a vector v(j) and there is an edge between these 2 vectors, then you will have for v(i) will multiply with plus one where as for v(j) we will multiply with minus one, since there is an edge between them and then when you sum you will have v(i) minus v(j). So, you will get this for each individual edge. So, therefore, you can write lambda is equal to v(i) minus v(j) whole square sum over e divided by sum over i v(i) whole square.

So, in this way you have come to the conclusion that, the numerator as well as the denominator of this expression, is a sum of squares. Therefore, each of them is positive and therefore, this whole expression is positive.

In the next lecture we will continue proving the next 2 properties.