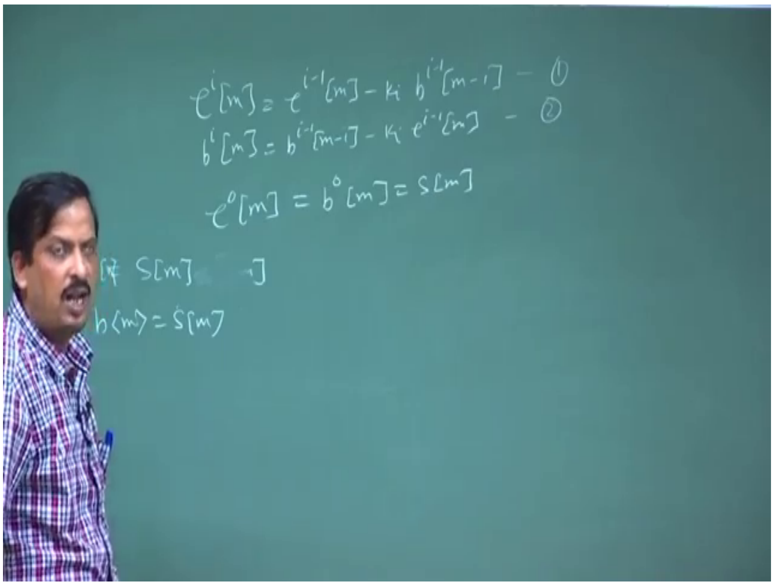
Digital Speech Processing
Prof. S. K. Das Mandal
Centre for Educational Technology
Indian Institute of Technology, Kharagpur

Lecture - 25
Lattice Formulations Of Linear Prediction (Contd.)

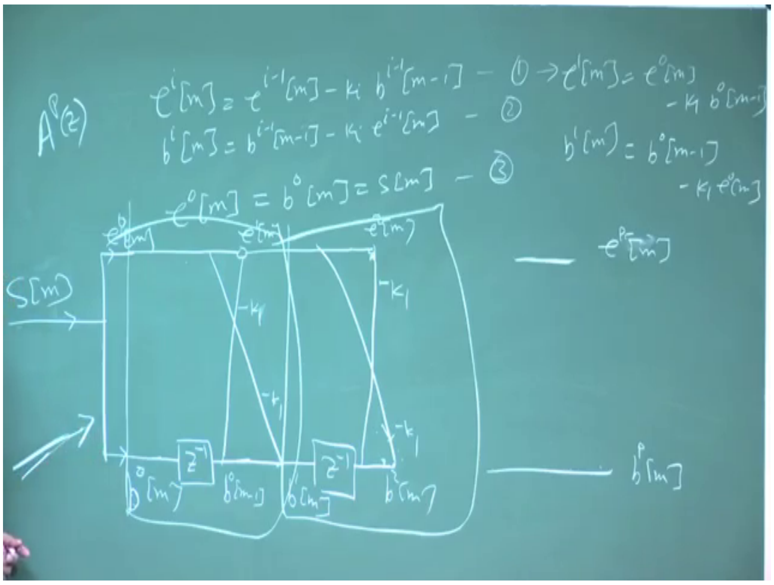
(Refer Slide Time: 00:18)



So we have derived these 2 equation. So, forward error ith order forward error is nothing, but a previous 1 minus previous backward error multiply by the k i, so this 2 equation; we have derived. Now, if I am not predict anything; what is e 0 m; what is e 0 m? e 0 m means that what is error? Error is that sample value minus estimated value. So, if order is 0; that means, I am not estimating anything. So, error is nothing, but a S m. So, similarly backward error is also is nothing, but a S m. So, I can say e 0 m is equal to b 0 m is equal to S m.

Now, try to draw the signal flow diagram or I try to draw the equation lattice filter structures using these 3 equations. So, let this is the equation number 3.

(Refer Slide Time: 01:23)



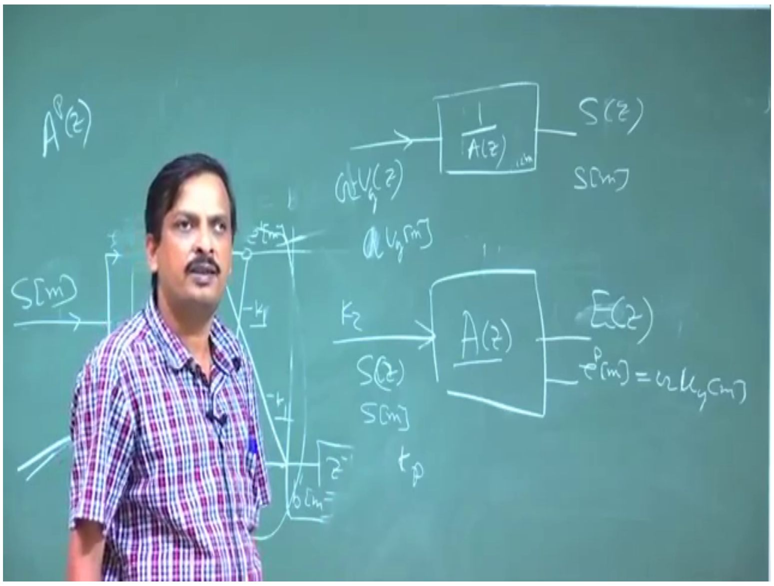
So, what I say if I no S m I know S m. So, which give me the e 0 m and b 0 m; now let us I want to find out the e 1 m; I want to find out e 1 m. So, what is e 1 m? From this equation if I say e 1 m is nothing, but a e 0 m minus k i; K 1 into b 0; m minus 1. So, what is b 0? This is b 0 m. So, if it is b 0 m; I want to produce b 0 m minus 1. So, I can say Z to the power minus 1.

So, S m is feeded here; speech signal is feeded here. So, e 0 m; b 0 m is equal to S m; Z to the power minus 1 delay if I applied, I get b 0 m minus 1. Now, e 0 m; e 0 1 is nothing, but a e 0 m minus K i into K 1 into b 0 m minus 1. Similarly, suppose I want to predict b 1 m. So, b 1 m backward prediction error; b 1 m again I put the; i equal to 1 in here. So, I can say b 0 m minus 1 minus K 1; e 0 m, so I can say this one is nothing, but a this is e 0 m, this will be minus K 1.

So, if I say that I want to predict e 2 m and I want to predict b 2 m. So, again I have to delay; one delay Z to the power minus 1 and I can say this signal should be added here with minus K 1 and this signal should be added here to get the btn minus K 1. So, if you see the same structure is repeated here. So, this is called lattice structure; single lattice, second lattice, third lattice and that way I can get the value of epm; if the pth order prediction.

So, I get epm and from here I get b P m; so, that is called lattice structure. So, what is epm? epm is nothing, but a error. So, if I say I want to implement A P Z; this is the implementation of A P Z; why? If you remember in my first class, what I am saying that if the whole speech production system is linearly model; LTI system or linearly model then I can say; if I apply if this is my H Z is my vocal track transfer function; H Z in here.

(Refer Slide Time: 04:51)



So, what I said; if we apply impulse A U, G Z; glottal impulse with amplitude, then I get speech signal; which is S Z or I can say in time domain I can write A U g m; this is and I get S m.

What is H Z? H Z is nothing, but a gain either A or write G; gain; A or G; I can write the gain as G also. So, I write the gain as A, G. So, G by 1 minus K equal to 1 to P alpha K Z to the power minus k. So, this is nothing, but a prediction error; which is H Z. So, I can say H is equal to G by A Z; G is a constant.

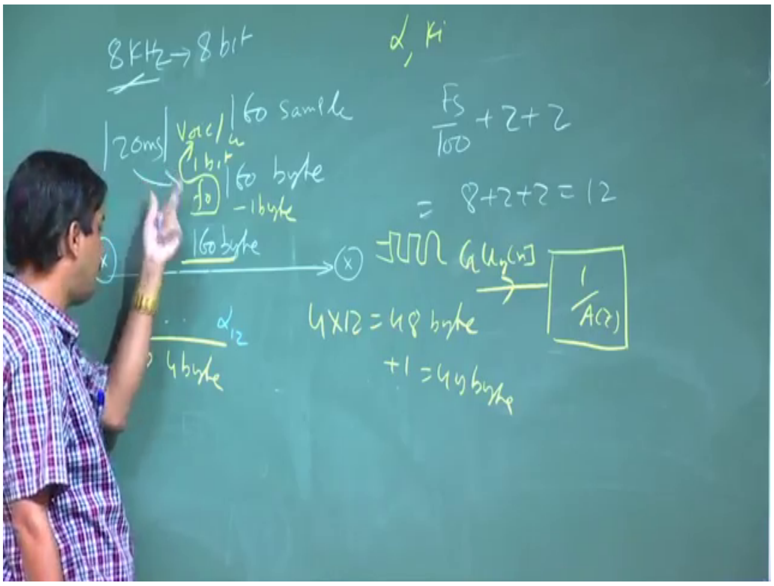
So, I can replace that A Z by 1 by A Z or I can say; if I apply A Z, if I implement A Z and apply S Z; I can get error E Z; which is nothing, but a I can get E Z or forward prediction error or backward prediction error; at the pth state both will be same E Z or I can say, if I apply S m; time domain signal and if I implement this A Z, I will get E Z; epm or which is nothing, but a G into U g m.

So, from output speech; speech is nothing, but a excitation into the transfer function. So, this is A Z. So, if I this if I subtract this one only remaining is the glottal excitation. So, that is gain into glottal excitation; so, error function is nothing, but a gain into glottal excitation.

So, suppose for a given speech signal; find out the prediction error or prediction error signal which is nothing, but a gain into glottal excitation signal. So, I can apply the speech signal and implement A Z; I can implement easily A Z; if I know K 1; value of K 1; K 2; in second stage; it will be K 2. So, third stage it will be K p; so, if I know the value of K 1, K 2, K P and if I implement this diagram; this thing in digital computer and then if I apply speech signal, I get the excitation signal.

So, this is called prediction lattice filter for A Z or error filter implementation lattice filter for A Z or error filter implementation. Now, suppose in speech synthesis; so, let us I told you a one thing; suppose I have a speech signal which has let us 8 kilohertz sampled signal.

(Refer Slide Time: 08:25)



And let us I take a window of 20 millisecond and it is encoded yet let us 8 bit 8 kilohertz 8 bit. So, in 20 millisecond; how many samples will be there? 160 samples will be there. Now, each sample is 1 byte; so, I can say this 20 millisecond window can be stored or can be required 160 byte memory to store.

Now, suppose the problem is that I want to transmit this 20 millisecond speech from this point to this point. So, my bit rate is 160 byte I have to transmit it; I have to transmit it

160 byte from this point to this point. Can I reduce it using this LPC method? So, instead of transmitting 20 millisecond, so since it is 8 kilohertz; what should be the order of the prediction? So, for F S is equal to 8 kilohertz; so I can say F S by 1000 plus 2 0for radiation and 2 for glottis. So, I can say it is nothing, but a 8 plus 2 plus 2; twelfth order prediction is required; 12 to 14 order prediction is required.

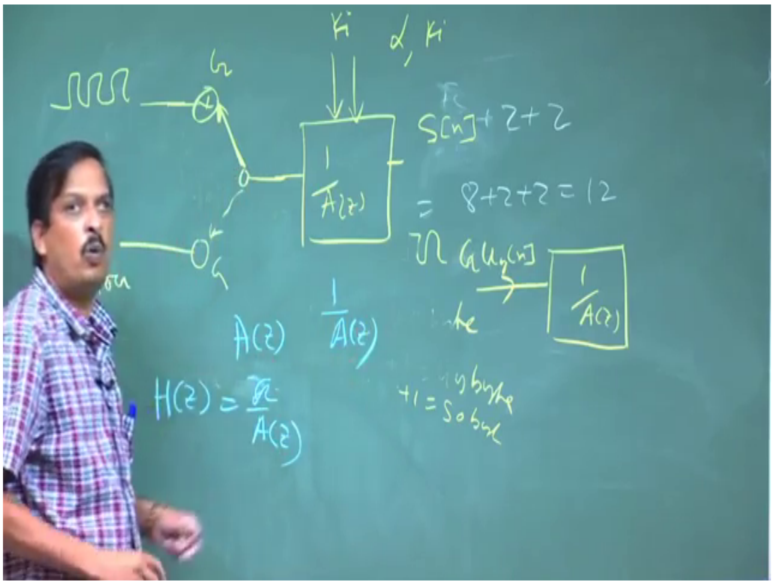
What is the meaning is that; that means, that I required alpha 1 to alpha 12. So, if it is alpha 1 to alpha 12 and each one of the alpha is encoded with let us encoded with 4 byte let us take 4 byte. So, 4 into 12; 48 byte; I require to encoded the alpha value; now it is said that if I know the alpha value or K value K i value let us say K i value. So, if I know the K i value; then at the receiving side I can say that I can implement the filter which is 1 by A Z and I can implement 1 by A Z.

And if I apply the gain and error signal; what is error signal nothing, but a excitation. So, excitation if it is voice; then it excitation is nothing, but a impulse if it is unvoiced excitation nothing, but a noise. So, what information extra things I require? Whether of this segment is voiced or unvoiced that require 1 byte; 1 bit, if it is one then it is voice; if it is 0 then it is unvoiced.

Next if it is voiced then I have to generate the impulse who supply the impulse do you know? That is also transmitted. So, I can say again the value of f 0 will be also transmitted. So, if f 0 value is transmitted using 1 byte; let us f 0 value and this voice unvoiced information is transmitted using 1 byte. So, 48 byte plus 1 byte let us 49 byte is required.

So, if I transmit instead of 160 byte; I can transmit forty nine byte and require the signal same segment at the receiver end this is called LPC encoding. So, in LPC encoding; I only transmit this alpha value; why that is a voiced and unvoiced or f 0 value and then receiver side I connect. So, I can say if I want to draw the decoder. So, I can say if it is voiced let us this is a impulse generator.

(Refer Slide Time: 12:34)

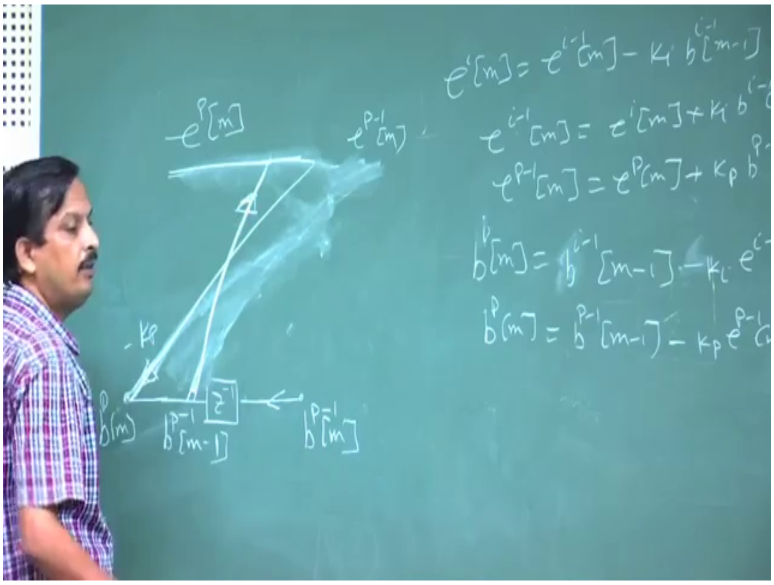


Then there is a gain let us gain also transmitted G 1 byte. So, gain is transmitted using 1 byte; so, it is 50 byte. So, this is gain this is multiply with gain; this is gain, this is noise source. There is a switch; if it is voiced, the switch will be connected to here; if it is unvoiced it will be connected to here. And then I implement 1 by A Z which is supply from K i value; then I get the speech signal back.

So, how do you implement A Z? I know how to implement A Z? Now, I know that if it is A Z and H Z is nothing, but a G by A Z. So, I want to implement 1 by A Z; I have to implement 1 by A Z how do I implement 1 by A Z? A Z implementation is already explained. So, I have to again I have to implement 1 by A Z; which is nothing, but a H Z. This is required; this is called a lattice formation of vocal track transfer function, which if I implement that using K i value; from this information, I can generate the speech signal.

So, how do you implement 1 by A Z filter? Let us try to 1 implement 1 by A Z; A Z I can implement from that equation. So, 1 by A Z; so, what I want? I want to supply error signal; epm I know inverse; I want to inverse it.

(Refer Slide Time: 14:21)



So, that why I have applied S m and I get epm; now I say I want to apply epm and reverse the signal and get S m back that is 1 by A Z. So, at the output; what I said? I said epm is the error signal. So, I know epm; let us say epm I know is the error signal I know.

So, if you remember the equation eim is nothing, but a ei minus 1 m minus K i into bi minus 1; m minus 1. So, if I know epm; what I have to estimate? I have to estimate ep minus 1 m. So, if I say e i minus 1 m; is equal to eim plus K i into b i minus 1; m minus 1. So, I can say ep minus. So, instead of i; I put P; e P minus 1 m is equal to epm which I know; plus I have to know this; I have to know b P minus 1 m; m minus 1; b P minus 1 m minus 1 or not.

So, if I say this point is my b P m; so, I have to apply a Z to the power this way; coming this way I can apply Z to the power minus 1 to get b P m minus 1 or not. So, if it is that b P minus; P minus; so, if it is b let us this 1 is b P minus 1 m let us this 1 is b P minus 1; this is b P minus 1 m minus 1.

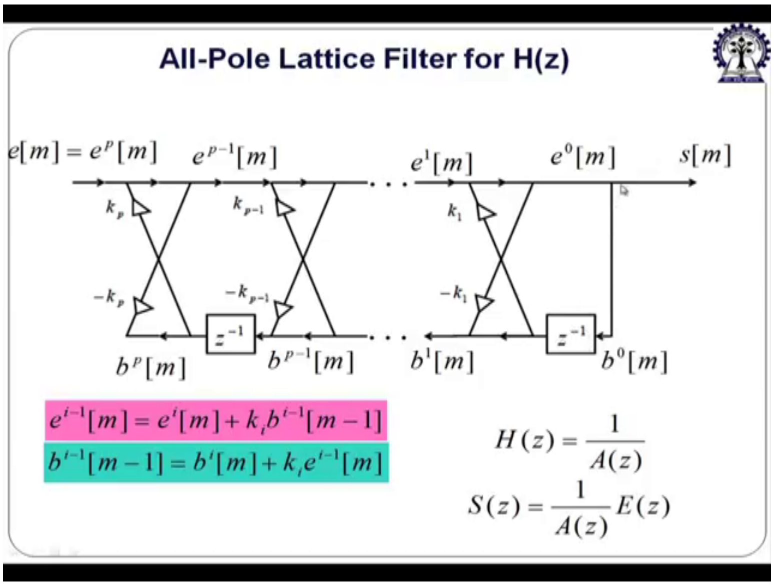
So, I can say aim b P minus 1 m has to be added here; we multiplication factor K i K 1 which is K 1 is K P; i is equal to P. So, I can say e P minus 1 m is equal to epm plus K P into b P minus 1 m minus one. So, if it is b P minus 1 m minus 1; I can get it, now what is b P m; b P m is nothing, but a b P m is nothing, but a what is b P m? b P I can say b i m is nothing, but a b i minus 1; m minus 1 plus minus K i; e i minus 1 m.

Now, I can say b P m is equal to b P minus 1 m minus 1 minus K P e P minus 1 m. So, I can say; if this is my less this is my b P m; this is my b P m. So, if it is my b P m this is my b P m let us this point is my b P m. So, b P m is nothing, but a bp minus 1 m minus 1 is minus k i into; so this has to be; so, this is here. So, this I can say this as to be again come epm.

So, it is b e what is b P m? e P minus 1 m ep minus 1 m has to be come again in here 2 and multiply by minus K P get the b P m. So, if I say I just want to draw the drawing very neatly then I can say this is my figure. So, I can let us this one and I can say this is this one; this direction multiplying factor and these direction multiplying factor. Similarly I can calculate P I minus 2; I can calculate b i minus 2 and draw the figure. So, you can draw the figure in home; you see the slides is there. So, you can see the slides also if you want help; ultimately if you see once I go to the 0; what is e 0 m? So, you can say e 0 m; e 0 m is nothing, but a e 1 m plus K 1 into b 0 m minus 1.

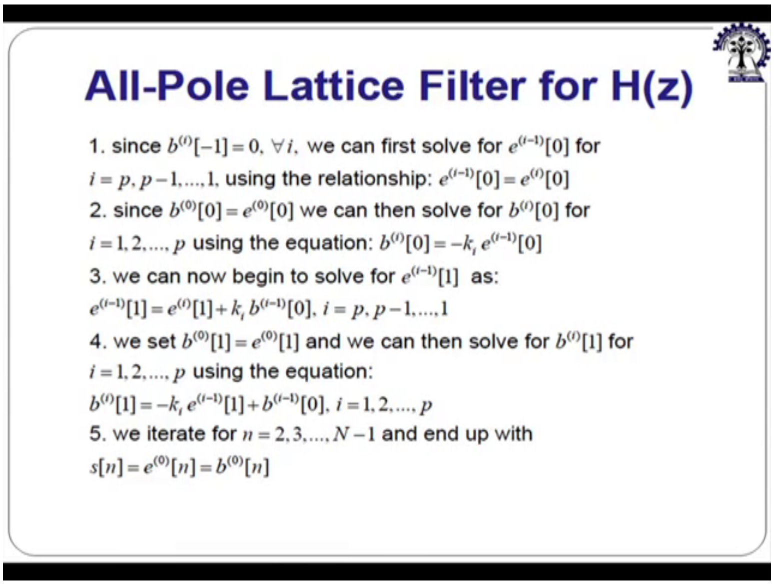
Now, e 0 m is nothing, but a my S m; e 0 m is equal to S m; so, this is the picture.

(Refer Slide Time: 20:48)



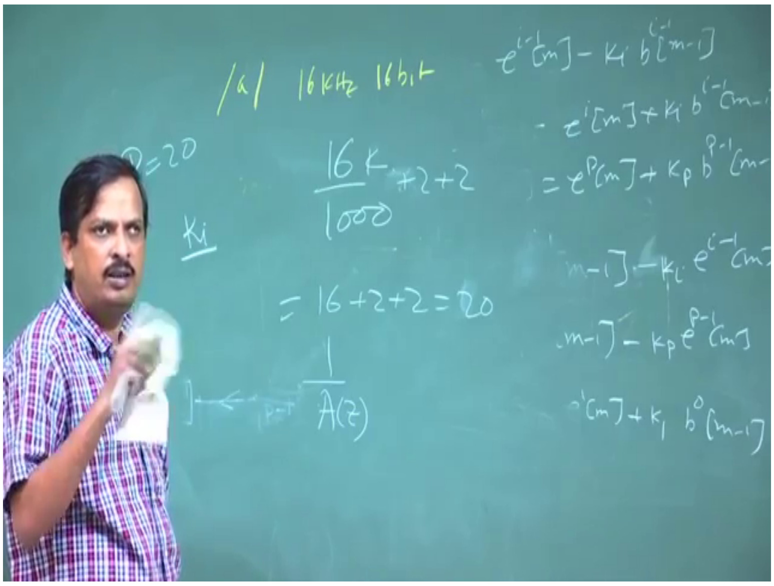
E 0 m is equal to b 0 m and if I want to calculate b 1 m; b 0 m, I require b 0 m minus 1 b 1 m minus 1; so, I can do that. So, this is the implementation of H Z; so, details you can see these slides and you draw this. Now, try to implement in a program also; I can implement this in a program also; so, if you see the algorithm; algorithm is given.

(Refer Slide Time: 21:23)



You can go through these algorithm and implement this in a program; so, I give you a task what is the task? Let us record in your own voice, record the vowel [FL] you record vowel [FL] how you record?

(Refer Slide Time: 21:40)

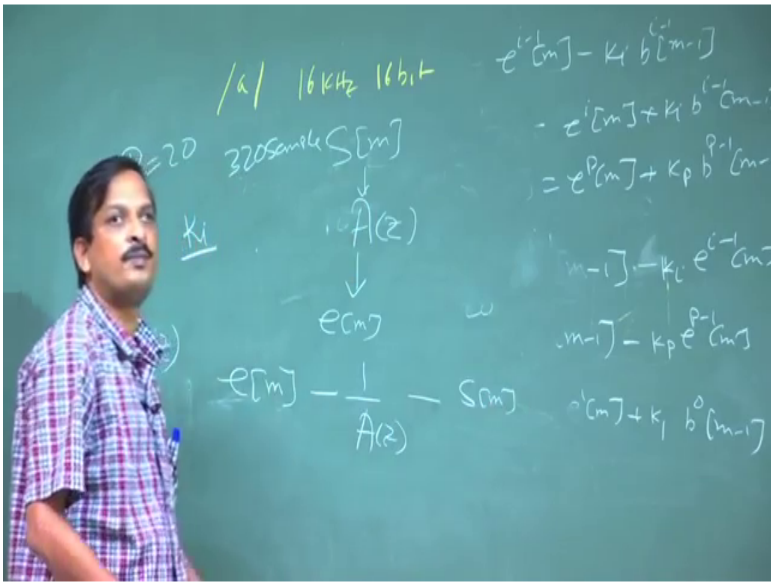


Let us for some time you say [FL] then you take the signal of 20 millisecond from the so, there is a [FL] signal is there. So, from the middle you cut 20 millisecond speech signal. So, record e [FL] vowel with 16 kilo hertz sampling frequency 16 bit encoding then cut 20 millisecond of speech signal. So, let us try for one frame only one frame; then what should the order of the LPC analysis 16 kilohertz.

So, 16 K divided by 1000 into plus 2 plus 2; 2 for glottal 2 for radiation. So, I can say 16 plus 2 plus 2 is equal to 20. So, I can get P is equal to 20 then implement A Z in a program and implement 1 by A Z using a program; in MATLAB or C; anywhere you can write the program implement A Z and 1 by A Z and calculate K i value.

So, I for implementation I require K i value; so K i value can be calculated either using autocorrelation function from there you can calculate K i value or directly you can calculate K i value I will show you the slide how to directly K i can be calculated. So, if I know calculate the K i value then implement A Z and 1 by A Z using a MATLAB program try it. If you have any confusion write my in email or you can forum, you write it that I cannot implement it; then implement A Z; once you implement A Z.

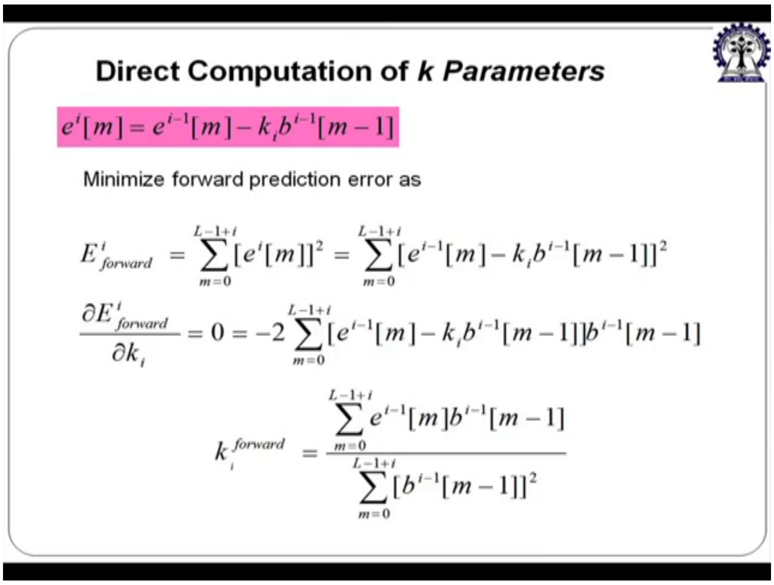
(Refer Slide Time: 23:40)



So, I implement A Z in a program; what is the input? Input is S n; which is 20 millisecond signal of A. So, if it is 16 kilohertz; so, there will be 320 sample value; S m is nothing, but a 320 sample value through A Z and find out the error signal. Now apply this error signal and implement 1 by A Z; same K value; K value will be same both function and see whether you get back the S m or not.

See whether you get back exact em, you apply input to the 1 by A Z and see that whether you get the drawback S m or not. For particular one frame; now if I want to synthesize whole [FL] then the frame will be repeated. So, that implement and told me in the forum report me in the forum implement it.

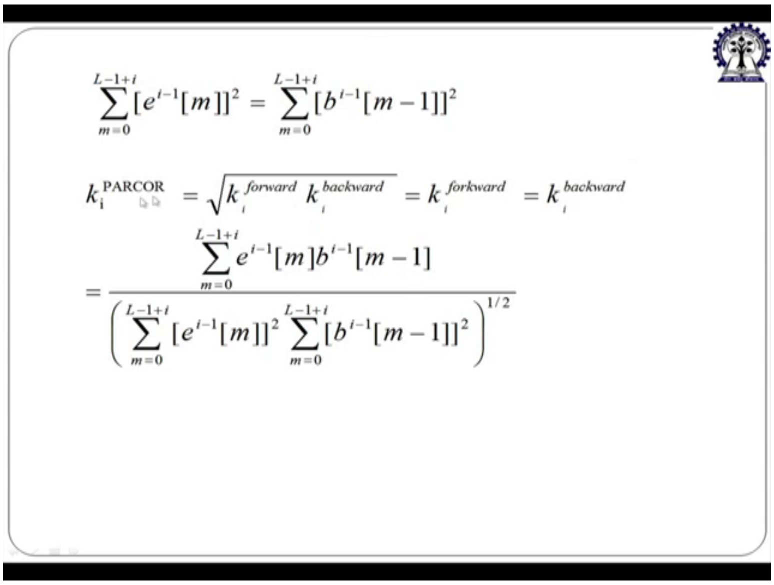
(Refer Slide Time: 24:51)



So, this is the direct computation of K parameter; what is the K parameter? How do you calculate the K parameters? So, K value actually minimize the error. So, if I say I am not writing in a board; if you see in the slides K forward. So, e forward prediction error min square forward prediction error this one. So, you know that eim can be written as ei minus 1 m minus K i into bi minus 1 m minus 1 that whole square. Now, how do you minimize? I differentiate with respect to K i and equal to 0.

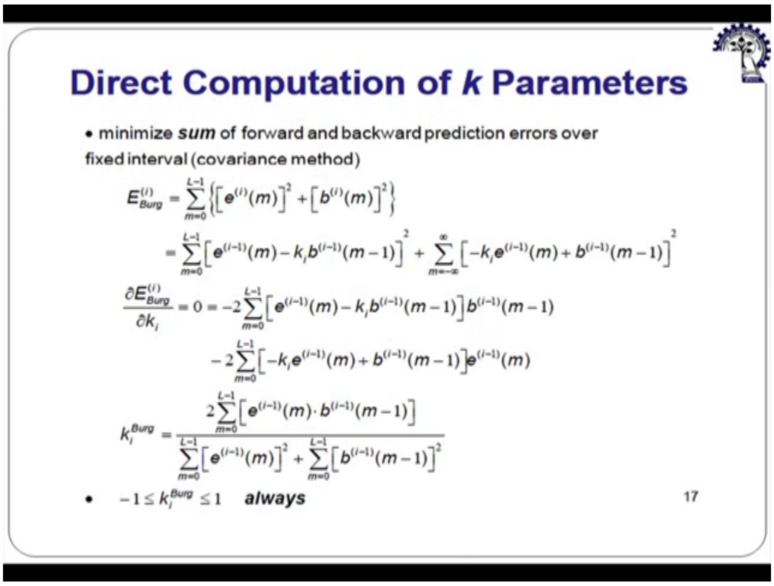
So, check the differentiation and find out the K i forward value. I can find out the K i forward value; once I get the K i forward value. Similarly, I can calculate the K i value for backward error then I can say the k is equal parcor; partial reflection coefficient or partial correlation coefficient.

(Refer Slide Time: 25:46)



So, if I say that partial correlation coefficient in k i; then it is nothing, but root over of K i k i forward equal to K i backward; then either K i forward or K i backward you can take.

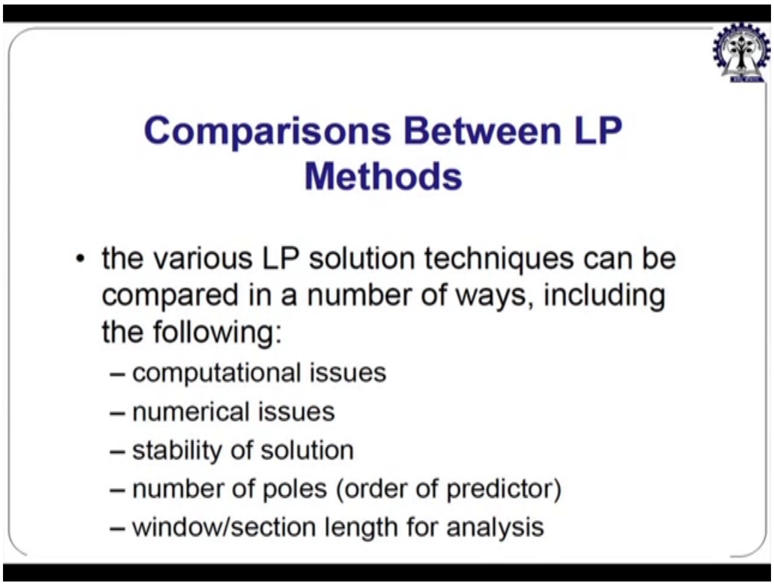
(Refer Slide Time: 26:12)



Or you can use covariance methods which is equivalent to call; I can say barge method. So, total error is forward error plus backward error; so, min square error forward error square plus backward error square and then take the differentiation with respect to K i and make it 0 and find out the value of K i burg. So, if you know K i you can implement A Z or if you know ei one m and ei. So, based on the previous error; I can find out present ki; so, using this equation you can implement A Z or 1 by A Z and record a vowel [FL] take a 20 millisecond window; pass through A Z find out em then pass the em with 1 by A Z implementation and get back that S m again.

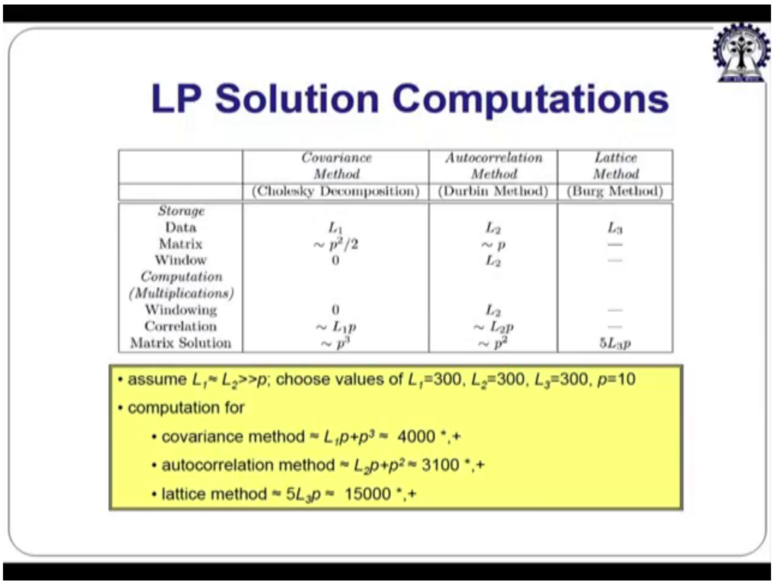
So error signal consist of excitation and gain; if it is noise signal if it is picketing sound then it is nothing, but a gain and noise.

(Refer Slide Time: 27:37)



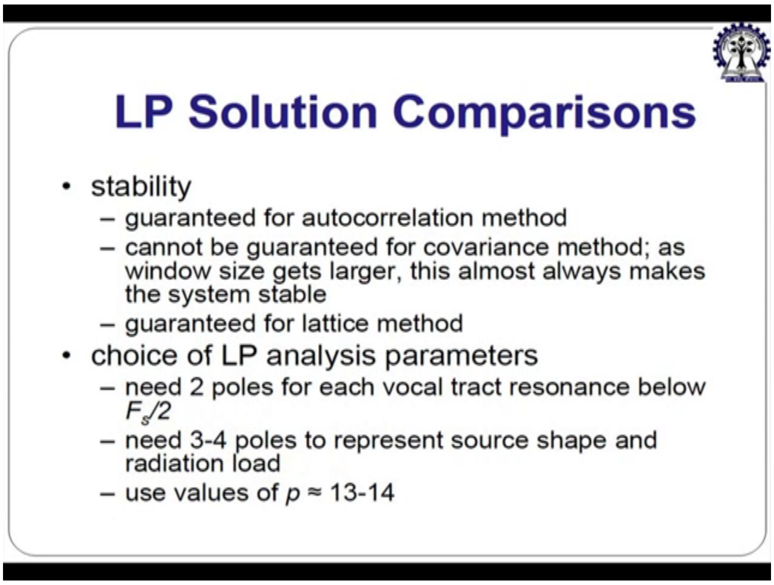
So, this is the summary now comparison; you may compare the different methods; what have the different methods for a LP extraction? If you see that lattice method only use, combine the 2 stage covariance that correlation or autocorrelation or correlation matrix and linear solution; they are combined both together and use single instance.

(Refer Slide Time: 28:04)



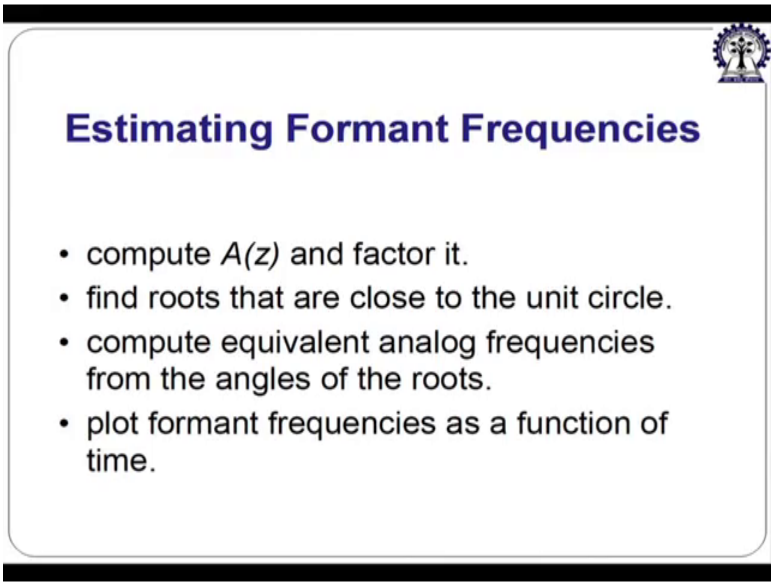
So, those are the computational issue is there, you can read it from that slides.

(Refer Slide Time: 28:08)

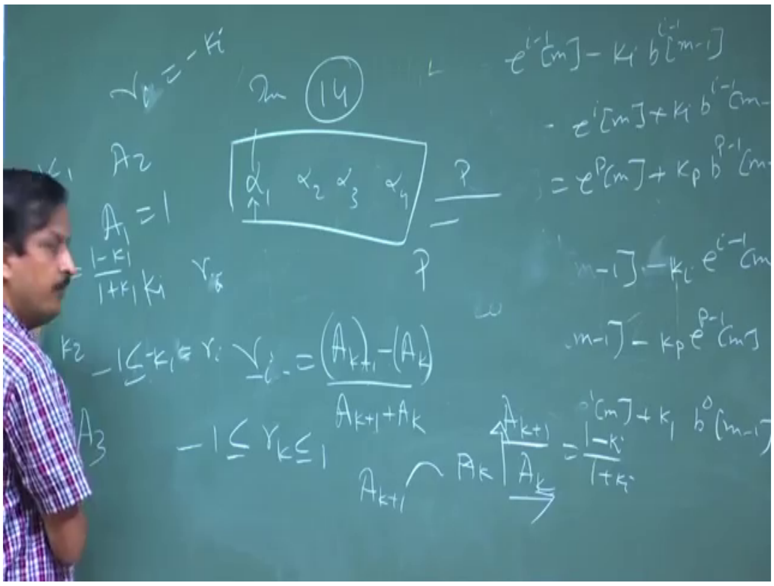


Stability issue is there; generate for autocorrelation guaranteed for autocorrelation guaranteed for lattice method. But covariance method it is not guaranted because if it is a large matrix then you know the upper triangular matrix, lower triangular matrix. So, they are have some problem; so, you can say the covariance method is not guaranteed, but autocorrelation method and lattice method stability is guaranteed. Choice of LP analysis already I have discussed F S by 1000 or F S by F S by 2, F S by 1000 plus 2 for glottal excitation and 2 for the radiation.

(Refer Slide Time: 28:52)



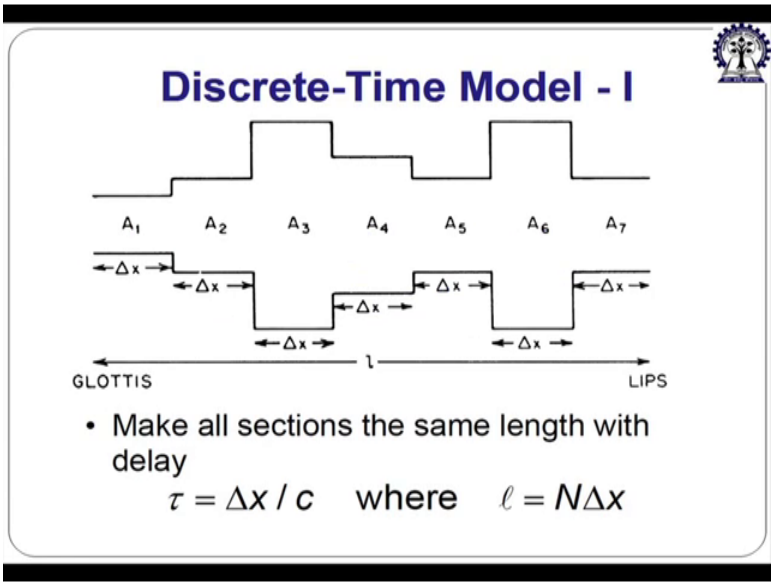
(Refer Slide Time: 28:59)



Estimation of for so, if I know A z; I have already discussed that if I know alpha 1, Alpha 2, alpha 3 and all LPC coefficient alpha 4; those represent the transfer function at that position and each pole; if it is those are implemented the pole position equivalent to the pole. So, each pole represent the formant frequency actually alpha 1, alpha 2, alpha 3. should represent the formant frequency is the frequency plane.

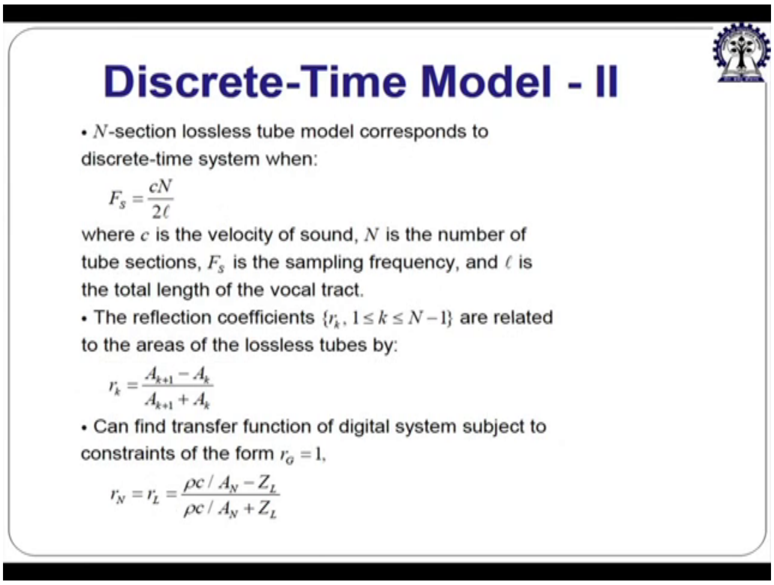
Now, if I take fourteenth order; then the alpha value which is not close to unit circle not give you the formant frequency. So, the alpha value which root value close to the unit circle; give you the formant value.

(Refer Slide Time: 29:47)



Then I have already said in my tube class this LP analysis is related to the loss less tube model. If you remember that this is the lossless tube model, if I consider all tube has same dimension; then I know L equal to N into letter X.

(Refer Slide Time: 29:58)



So, you know that rk; value of r K is nothing, but A K plus 1 minus reflection coefficient A K; plus 1 plus A k. So, if I consider rg is equal to 1; then rn equal to rn; rg is equal to 1; then rn is equal to rl and rn equal to rl; rl equal to row c pi a an divided by Z L. So, let us Z L is equal to 1 also 1; forget about that. So, this is the reflection coefficient.

Now, I give you the problem; the problem is here, can I estimate this area function from the speech signal? I want to estimate that suppose I produce [FL] I want to half the area function is changing during the production of [FL]. So, how this area function is changing; let us this tube is model. So, if it is a pth order LPC analysis. So, I can say the tube is divided into P number of section.

So, if you relate that K i value and r r K value has a relationship or ri value has a relationship. So, minus 1 minus K i is equal to ri; so, you know that the r K value is minus 1 r K plus 1. So, r k is nothing, but a minus K i or you can say ri; let us say this ri is equal to minus ki. Now, if I know r K value or ri value then I can express A K plus one in term of A K or I can say that A K plus 1; divided by A K is equal to 1 minus, it will become 1 minus K i divided by 1 plus K i if I put the K i value.

So, if I know the ri value; if I know the K i value I know ri value, once I know ri value; I can find out the A K plus 1 divided by A K. So, if I know the previous tube cross sectional area; then I can find out the; what should the cross sectional area of the next tube. So, if I say beginning tube let us A 1 is equal to 1; then I can estimate A 2; if I know A 1; K 1.

So, if I say A 1 is equal to 1; then I can find out A 2; if I know K 1; 1 minus K 1 divided by 1 plus K 1; then if I know K 2 and A 2; I can estimate A 3. So, that way I can say if I know first consequence relation 1; then I can say how the cross sectional area is different cross sectional area is created during the production of that vowel; from where I estimated the K i value.

So, for a given vowel I can draw this is given taken from the professor Rabiner books that if this is my vowel signal; those of the area function. So, area function I can draw if I know the speech signal. So, reversely if I know the area function I can generate the speech signal that is a tube synthesis. Now, if I know the speech I can find out the what kind of constriction; what kind of different cross sectional area is made during the production of that sound.

Thank you.