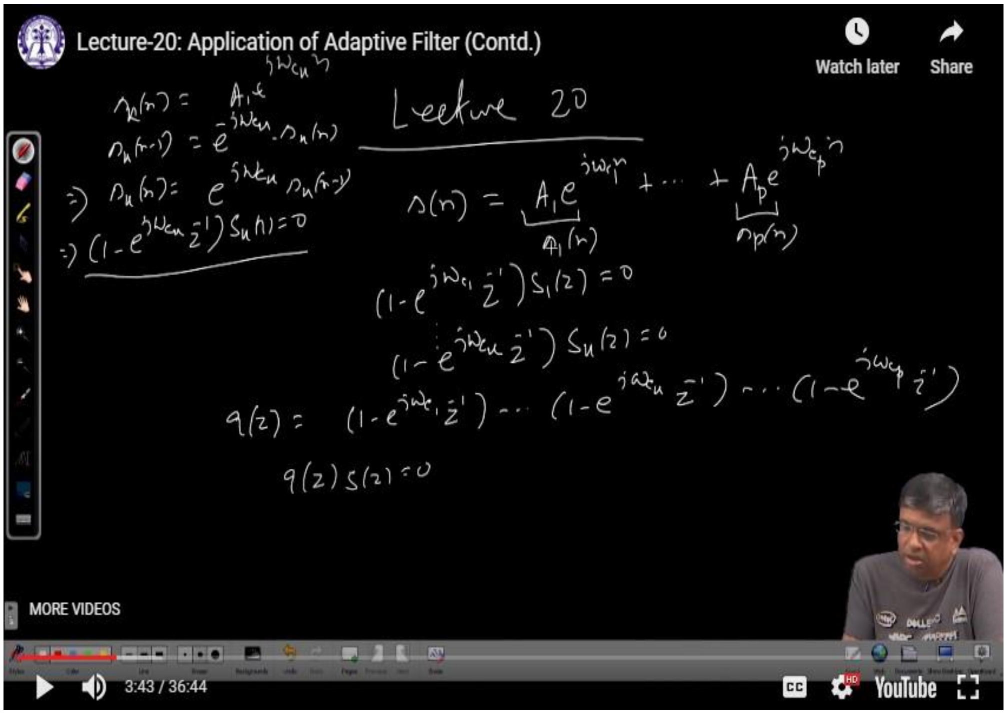
Introduction To Adaptive Signal Processing
Prof. Mrityunjoy Chakraborty
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture No # 20

Application of Adaptive Filter (Contd.)

So, in the last class we are considering this problem, we were given a single suppose Sn which is the summation of p number of complex sinusoids, this one is S1n and Spn all right. We have seen in the previous class that 1 minus e to the power j omega c1 z inverse this factor times S1z equal to 0 dot dot dot dot in general 1 minus e to the power j omega c may be k z inverse 1, Skz equal to 0 that we have seen I am not re deriving. And therefore, if you take the product of this you call it Qz, you call it Qz , the product dot dot dot 1 minus last one is. So, there are p such factors. So, if you take this Qz obviously Qz into Sz was 0, Qz into Sz was 0 because if you take this factor we have seen earlier or maybe just for real quick recalling you take this S1n was a1 say Skn is e to the power jckn.

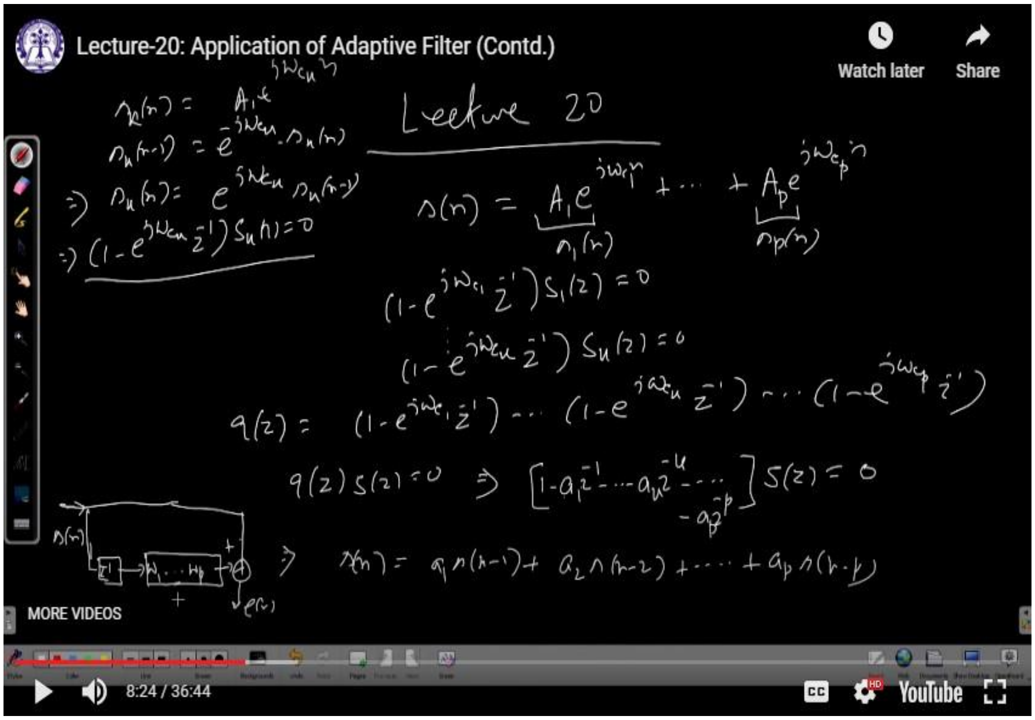
So, Skn minus 1 if you put n minus 1 here it becomes e to the power minus j omega ck n minus 1. So, one factor comes out rest is as it is. So Skn which means, Skn is e to the power j omega ck Skn minus 1 if you take z transform on both sides here it is z inverse time Skz it is Skz and then take all the terms to one side, it will be into Skz equal to 0. This for any Skz and now I have taken the product of all of them and this overall polynomial into Sz is 0 because suppose I am considering this then this is a product of the factors. So, I can push take the this factor out and put it in the front.



So, then this times S1n I mean S1z, so this first factor which is this into S1z that is 0 we have seen k equal to 1 so 1 minus e to the power j omega c1 z inverse S1z. So, that will be 0 if it is the kth sinusoid here then Qz from Qz I will take out the corresponding factor put it in the front. So, again Qz times that Skz here will be 0 and so on and so forth and that is why Qz Sz is 0, but Qz is a product of these factors. So, if you carry out the product it will be a pth order polynomial because p such first order factors multiplied and leading term is 1 1 1 1 1 so 1. So, it will be like Qz will be like 1 then some coefficient maybe a1 instead of plus a1 I am just putting a minus a1.

If you carry out the product it will be some coefficient times z inverse minus dot dot dot minus Ak z inverse k minus dot dot dot minus Ap z inverse p because pth order polynomial so high I mean highest degree of z inverse is p z inverse into z inverse into z inverse p times minus I just deliberately put minus a1 is the coefficient, minus Ak is the coefficient like this. So, this is my Qz Qz into Sz equal to 0. So, Sz into 1 that you keep on one side and remaining terms you take to right hand side then if you take the inverse z transform Sz into 1 will give you Sn and then a1 z inverse Sz which means in time domain Sn minus 1, then a2 z inverse 2 Sz so a2 Sn minus 2 and dot dot dot lastly Ap Sz Ap z inverse p Sz z inverse p Sz means is in time domain Sn minus p so Ap Sn minus p. It means if it is a summation of p sinusoids complex sinusoids these are called not cos and sin but complex sinusoids if it is a sum if the single is a sum of p sinusoids. Then any current sample will be exactly expressible you can write it exactly as a linear combination of past p samples 1, 2 up to p if it is p number of sinusoids you can express it as a linear combination of past p samples. That means in this case if you have Sn as a desired response in an optimal filter and you give a delay response sn, so it is Sn minus 1 then if you have a filter say w1 dot dot dot wp and this Sn you give as your desired response then the optimal filter will be like this is your desired response Sn and here you have w1 Sn minus 1 w2 Sn minus 2 because current input is Sn delayed so Sn minus 1 w1 then w2 Sn minus 2 w3 Sn minus 3 like w1 Sn minus 1 w2 Sn minus 2 up to wp Sn minus p and we find out these values so that the error variance is minimum then obviously the optimal weights will be w1 equal to a1 w2 equal to a2 wp equal to ap and so these output will be exactly equal to Sn and so filter error output will be 0.

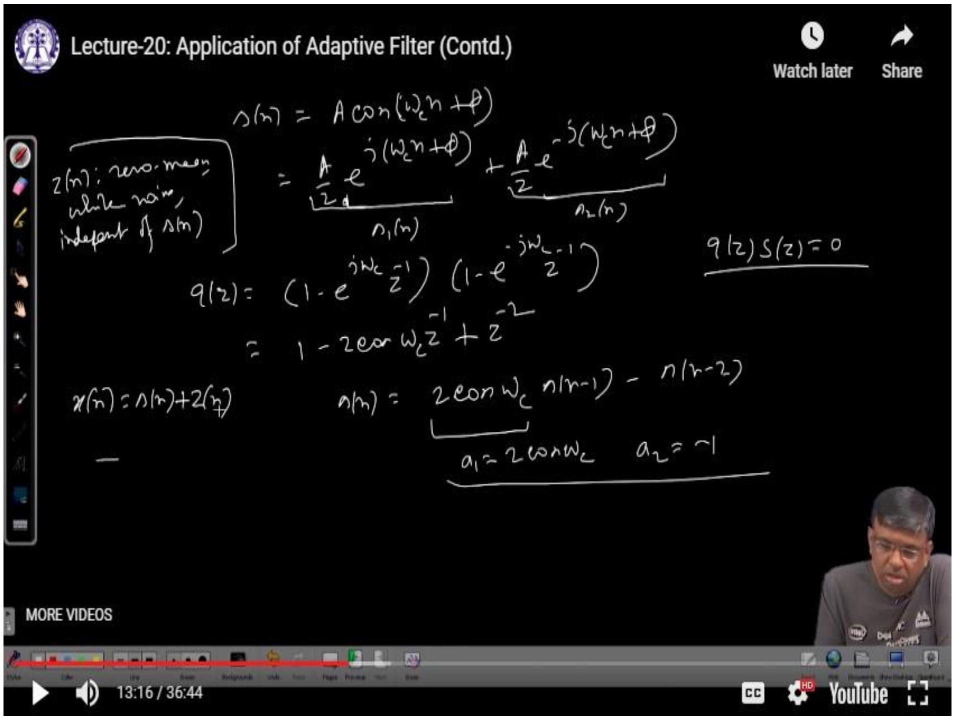
So that means current input is exactly predictable from the past p inputs this is the thing. So I am doing optimal filtering in a linear prediction point of view where current sample is predicted from as a sum of past p samples Sn minus 1, Sn minus 2, Sn minus 1, Sn minus 2, Sn minus p then the corresponding optimal filter coefficients will turn out to be a1 a2 ap because then we see from this expression error can be brought down to 0 therefore your variance can be brought down to 0 which is the minimum if are possible so optimal filter will be this all right this is a beautiful property.



As an example, we can consider we can consider suppose Sn as just some a times cosine some omega Cn plus phi okay. Now this is we can write as cos is we can write cos theta as 1 by 2 e to the power j theta plus e to the power minus j theta, so it is a by 2 e to the power j omega Cn plus phi plus a by 2 e to the power minus j omega Cn plus phi. So this is your S1n, this is your S2n all right this whole thing this whole thing is your S2 all right. One frequency omega C another frequency omega C minus omega C. So here the annihilating poly dual will be e to the power minus e to the power j omega z inverse and here frequency is omega C here minus omega C so here you get 1 minus e to the power minus j omega C because frequency is minus omega C here okay.

If I take the product of the two this will be my Qz down and this Qz Sz is 0 from my previous discussion all right. But here if I know product if you take the multiplication if you carry out the product 1 into 1, 1, then you take z inverse common e to the power minus j omega C and e to the power minus j omega C when added that will give you twice cosine omega C z inverse and then this will cancel e to the power j omega C e to the power minus omega C cancel z inverse 2 okay. So, this is what which means if you now multiply this by Sz Sz into 1 Sn that remains on one side other terms you take to the right-hand side twice minus because plus twice cosine omega C z inverse Sz that is Sn minus 1 and the minus z inverse 2 Sz that is Sn minus 2. So A1 is twice cosine omega C A2 is minus 1 you can verify yourself you really carry out this sum if you now replace Sn minus 1 by A cosine omega C N minus 1 plus phi here and again Sn minus 2 equals to the same thing with N by replace by N minus 2 okay. Carry out the summation by using trigonometry we will get Sn equal to this you can verify.

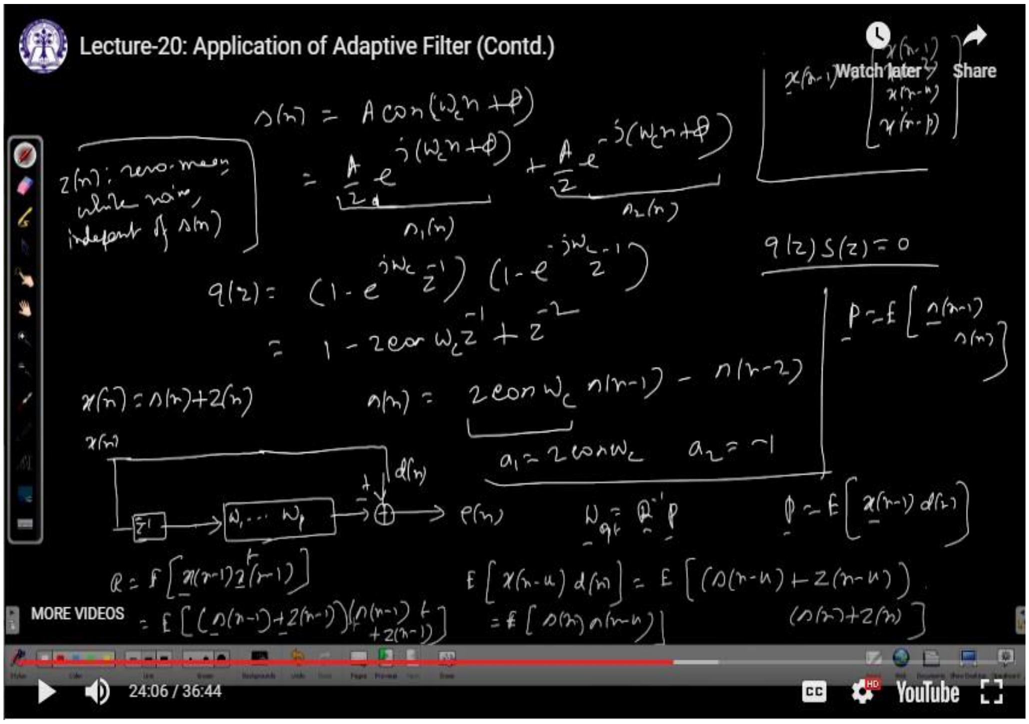
We are using DSP Z transform and all that but using pure trigonometry if you replace Sn minus 1 by A cosine omega C within bracket N minus 1 plus phi and Sn minus 2 by A cosine omega C within bracket N minus 2 plus phi then do some algebra of trigonometry you will get back summation equal to Sn which is this you can verify all right. So, in this case Sn can be predicted exactly if it is a pure sinusoid, it can be predicted exactly from its past two samples Sn minus 1 Sn minus 2. Ideally one optimal filter coefficient will be 2 cos omega C another minus 1 okay these are optimal filter or equivalently optimum predictor coefficients prediction coefficients all right. Now we have Sn we have the single Xn which has this Sn but not only Sn some white noise. It is 0 mean white noise independent of that is statistically independent of Sn. So, this is given to me and the noise power is reasonably high so it is not that noise is very small.



We can minimize as I told you the very beginning of this course, I will give an example that ideally if you want to eliminate this noise and this is just a cosine sequence so it is in the I know discrete type Fourier transform we will have two impulses one at small omega C another minus omega C. So ideally you should design a band pass filter with very narrow pass band okay one at omega C another minus omega C and pass this sequence Xn through that so this signal Sn will come through the pass band it will go through as it is but almost all of the noise most of the noise will be filtered out by the narrow band filter because it is a white noise so it is present in all frequency ranges thereby the noise will be eliminated when it will large extent to a huge extent actually almost all noise will go except whatever little we have in the pass band or the filter band pass filter. But the problem is if the frequency changes omega C after some time, then the band pass filter will be rejecting this Sn also, because new omega C will be outside the pass band of the band pass filter you designed which has you know pass band around the old omega C. So, it will not work so we have to make it adaptive okay. So, what we do here we give this Xn and we take its delayed version and we pass it through a filter ideally this should be two coefficients but we make it larger does not matter I mean it will be W1 dot dot dot some Wp. This is my desired response Dn so Dn is Sn okay so here Dn if you are considering an optimal filter what will be Wopt here Wopt will be what I mean Dn Wopt is R inverse p, R is the auto correlation matrix of the input here okay and p is the cross-correlation vector between the input vector here and Dn. So, p first starts with p, p is expected value of Xn minus 1 this vector here means current input is Xn minus 1 right, so data vector at n minus 1. So, it will be Xn minus 1 vector times Dn now Xn minus 1 vector is current sample current bits at n minus 1, then n minus 2 dot dot dot n minus p, p terms because p coefficients this is Xn minus 1 and Xn minus 1 what is Xn? Sn plus Zn, so Xn minus 1 sample is Sn minus 1 plus Zn minus 1 Xn minus 2 is Sn minus 2 plus Zn minus 2 like that. So, it will be X that is S actually sorry Sn minus 1 plus Zn minus 1 let me instead of doing it elaborately let me do one thing let me take a general term here general term means a general term of this vector into Dn say you say Xn minus k I have got some of Xn minus k general term so expected value of a particular term Xn minus k into Dn what happens there? So, E Xn minus k and Dn okay what happens to that. Now I know it will be Sn minus k plus Zn minus k this Xn minus k and Dn what is Dn? Dn is Xn itself which is Sn plus Zn. Now this will have one term Xn minus k into Sn Sn minus k into Sn that is fine the other cross term is Zn minus k Sn the correlation between Z the noise term okay and single term with a gap of k okay. It is N it is N minus k but the two terms they are statistically independent and therefore uncorrelated and therefore expected value of the product is product of the expected value of this into this and it is 0 mean so this will go all right because Z and S they are statistically independent in fact that is a strong condition you can even assume they are uncorrelated because statistically independent means uncorrelated. You can even just take it to be uncorrelated only but even in that case the correlation at a gap of k will be you take a product Zn minus k into Sn n minus k here n here so gap is k now multiply an expected value there is a correlation. If uncorrelated it will be expected value of Zn minus k into expected value of Sn which is 0 because each has 0 mean so therefore this cross term goes by the same logic. Expected value of Sn minus k into Zn okay this also correlation at a gap minus k this goes and then Zn with Zn minus k now Zn is Zn is what Zn is a white noise white means the sequence where every sample is uncorrelated with every other sample from nearest neighbor to farthest neighbor okay, that is white we have seen earlier.

So Zn and n minus k and k can be 1 2 so whether it is n minus 1 n minus 2 n minus 3 whatever that with Zn multiplied expected value it will be a correlation autocorrelation of Zn with itself at a gap of k but Zn is a white sequence so it is autocorrelation at any k greater than 0 then k can be 1 2 3 anything is 0, because in a white sequence every sample say Zn will have 0 correlation with its neighbors neighbor can be n minus 1 so immediate neighbor or n minus 2 n minus 3 like that okay so these also goes. So essentially it will be E of n into Sn minus k okay. So, but as though as though there is no Zn you have only Sn so in that case Sn and Sn minus 1 there is no Zn okay. If there is no Zn if you give Sn here, we have seen already if you give Sn here and delay here and Sn minus 1 Sn minus 2 then you get the optimal filters such that the error is finally 0 when there is no Zn. So in this case even when Zn is present we find that the cross correlation vector p is same as what we would have got if there is no Zn present because it is just E of Sn into Sn minus k okay this is one term so this was just one term now you can carry out with n minus 1 n minus 2 n minus k minus p. So that means cross correlation term p is E of just Sn minus 1 vector into Sn itself okay because n minus 1, n minus 2, n minus 3 and all that Sn fixed okay. I took n minus k Xn minus k arbitrarily could have been n minus 1 could have been n minus 2 so general term okay dn here and now there is no Z present so as though when I compute p for this setup, I get the same p when there is no noise present so just Sn and Sn minus 1 here and when there is no noise present this predictor coefficients give me an output which is exactly Sn and error is 0 okay. So, for that optimal filter so I mean should have a p vector which is this because that time Sn is the desired response and Sn minus 1 Sn minus 2 they are the input terms okay. So, Zn has no effect on changing the p, p vector cross correlation vector when I have no noise case to a noisy case because there is no Zn present but changes come here in the autocorrelation matrix.

Autocorrelation matrix is now is of this so E of R is now E of Xn minus 1 vector Xn minus 1 right transpose. So, it will have terms like S plus Zn minus 1 into X transpose this is what we have now again these been vectors okay

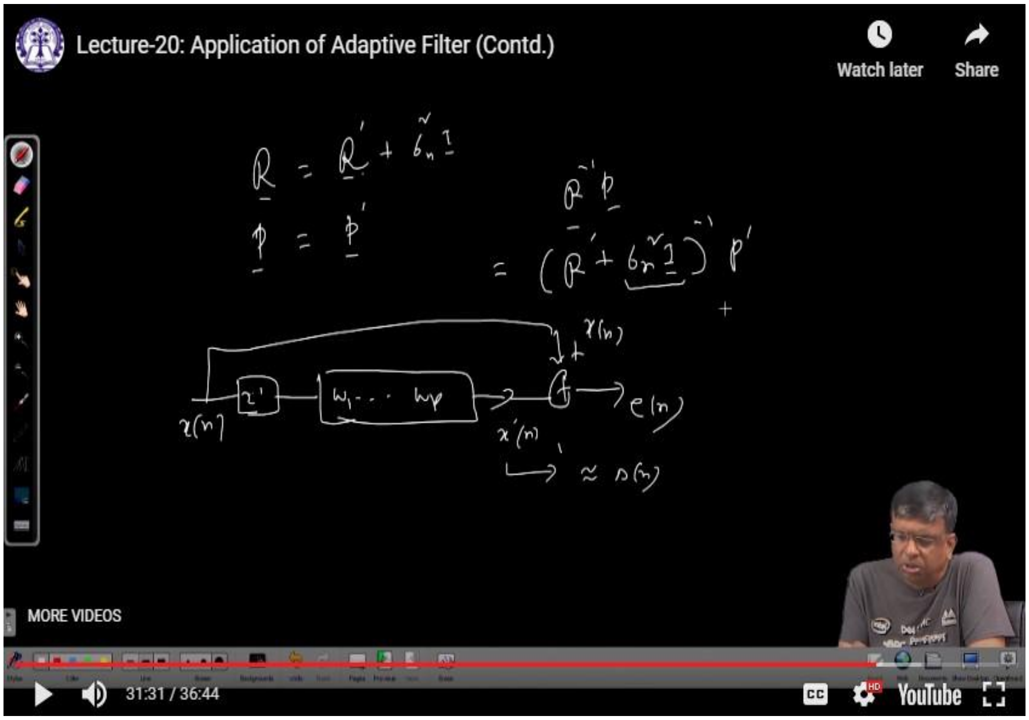


Oh, it is gone so let me go to the next page yeah okay. so, Zn is the noise vector like Xn minus 1 it has two components Sn minus 1 vector Zn minus 1 vector that is very simple Sn minus 1 vector means small s n minus 1, small s n minus 2, small s n minus 3 dot dot dot up to Sn minus p Zn also Zn vector means Zn minus 1 vector means Zn minus 1, Zn minus 2 dot dot dot Zn minus p a vector right this is just that. Now Sn minus 1 Sn minus 1 transpose fine no problem it is the correlation of the signal part okay and again because of whites and stationarity n will go it will just be autocorrelation of the input Sn all right. cross terms Sn minus 1 vector and the column vector and another term here is Zn minus 1 transpose is a row vector. So, every term of some element of S and some element of Z. They are uncorrelated expected value of that will be 0 because they are uncorrelated because they are statistically independent so expected value of some term of this vector is n minus 1 which will be some Sn minus k and some term of this row vector Zn minus 1 transpose okay that will be this cross term will be 0 because one term consisting of S from here just Sn minus k another may be Zn minus L okay from here their product and expected value which is a correlation between the two terms which is 0. Again, similarly Zn minus 1 vector from there if I pick up any elements Zn minus k because Zn minus 1 is a column vector and then Sn minus 1 transpose which is a row vector you take any element say maybe Sn minus L.

So, Zn minus k Sn minus L if you take their product again, they are uncorrelated and therefore 0. So, only other term will be Zn minus 1 and Zn minus 1 transpose that is the autocorrelation matrix of the noise and noise is white, so again it will be a diagonal matrix because their cross-correlation terms of the autocorrelation matrix of Zn minus 1 will be 0. cross terms will be 0 because it is a white sequence only diagonal terms will give the variance okay. So, and because of weights and stationarity of the process all the variances are same they are all dependent on the location okay. So, it will be a constant times identity matrix okay constant times constant that constant is the noise variance for every time index N because it is WSS so that does not change with time index. So, it will be altogether R will be having R let me call it R this thing this is R without noise okay.

So, maybe call it R prime plus some constant sigma N, N for noise sigma N square times identity matrix this is my R and P we have seen it is same as new P same as original okay when there is no noise. So, now, new R inverse P we will have P prime if sigma N square I is very small or it is not present then R inverse P will be same as R prime inverse P prime that is if there is no noise then R inverse P or if the noise power is less very less R inverse will be very close to R prime inverse P, P prime that is which is the optimal filter or the no noise case and that will give me the correct SN and the filter output so that error is 0. But because of the presence of this term there will be you know some deviation from optimality as a result if you have if you have XN here and then you are passing it through the filter this will be and this is your linear response XN EN they will no longer be the original optimal coefficients when there is no noise and therefore, there are no coefficients given by R prime inverse P prime and if you put the coefficients these output is same as XN so that error is 0. Because if there is no noise, I have got a pure sinusoid which is predictable exactly from its past P samples you have seen. But we do not have just R prime alone we have this extra factor and that is why these filters will no longer be same as that original optimal filter so output here will not be exactly XN but still close to that X prime N.

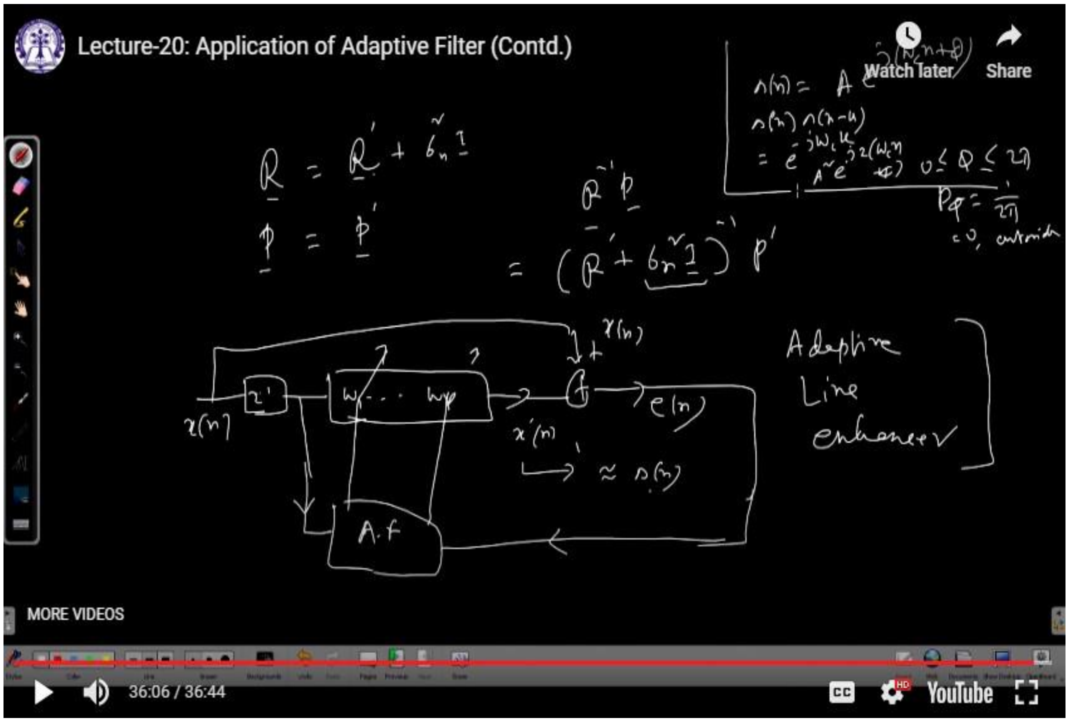
So, and these elements the diagonal matrix it only attacks or changes the diagonal elements of R prime which are the variances of the noise input and usually input variance is large. So, only they get perturbed or somewhat modified because of this addition by sigma N square factor other terms are not other terms of R prime they do not get changed. So, what all the change in its effect is not much and therefore, X prime N is close to still XN and error is not ideally 0 now it is not 0 ideally should have been 0 but it is no longer 0 but still close to 0 and in that case, this will be close to XN and XN is the noise free noise XN. Alright and what is X prime N? X prime N is almost same as I mean the output which you would have got if there is no noise that is X prime N will be close to S N, why S N because if this is not present if the sigma N square is not present obviously the output would have been same as S N is not it XN is S N and this is S N minus 1, S N minus 2, S N minus p there you combine them by the optimal coefficient. So, that S N is given exactly by the filter output it is exactly equal to S N at that time your optimal filter coefficient should have been R prime inverse P prime.



R prime is the autocorrelation matrix of S the corresponding cross correlation vector is P prime. Now there is a change small change caused by the sigma N square I which is changes the diagonal elements of R prime. As a result, the optimal filter coefficients will no longer be same as original S N but will be close to S N because the effect of this may not be much. So, I can be happy with this component okay. I can be happy with this component which will be largely noise free close to original S N. So, I can be happy with this and now I can make it adaptive because if the frequency changes and all that you know it has to be tracked so I can make it adaptive. This is the filter input adaptive filter and then these coefficients maybe by elimination algorithm will be adjusted okay.

So, this is called adaptive line enhancer all right. Remember if the input frequency changes if the input frequency changes correlation also changes okay. As a result, R new R new autocorrelation matrix will be defined from old autocorrelation matrix. You can take one example suppose I give you one signal S N A e to the power j omega c N plus phi, this is random because suppose phi is random if the phase is random it varies from minus 0 to 2 pi with some probability okay. In that case the cross correlation suppose phi so p phi is given to me suppose I mean phi is from this actually you should know phi is in this range okay and it within that suppose it is given to me uniform as an example uniform so this is 1 by 2 pi.

So, 0 to 2 pi within that probability density is constant so 1 by 2 pi so 1 by 2 pi into 2 pi area is 1 that is what probability density should have if you integrate it should have 1 okay. So, 0 to 2 pi is where the phase is at that time probability density is constant for all phi is called uniform distribution and outside that 0 equal to 0 outside. If this is the case you can verify cross correlation S N S N minus k if you do then you do you get one term only what S N is this S N minus k. So, it is but what will come out is just a into this omega c N okay minus k and then S N minus k so basically one term will come out and you will have a square so one term will come out N minus k and otherwise you will have e to the power j omega c N plus phi again so it will be twice so s square all right. If you now take the expected value so multiply this by 1 by 2 pi integrate from 0 to pi, so this term will come out okay. This term will come out as it is because it does not have phi this term e to the power j 2 is omega c N plus phi this is phi. So, if you integrate after multiplying by 1 by 2 pi and integrate from 0 to 2 pi this integral will be 0. Anyway, I think I will consider this I will consider in the next class all right.



What I am trying to change this if the input frequency changes cross correlation and autocorrelation values will change so you need a new optical filter and then that is why this adaptive filter is necessary okay it will track and will move to or you know that adaptive filter you know will take you to the new optimal filter by learning from the input and that is what will happen that thereby it will track the input frequency okay. Let me erase this part if time permits next class, I will just show this but this you can try your hand on with this a very simple textbook exercise okay see you then thank you very much.