Digital Voice and Picture Communication

Prof. S. Sengupta

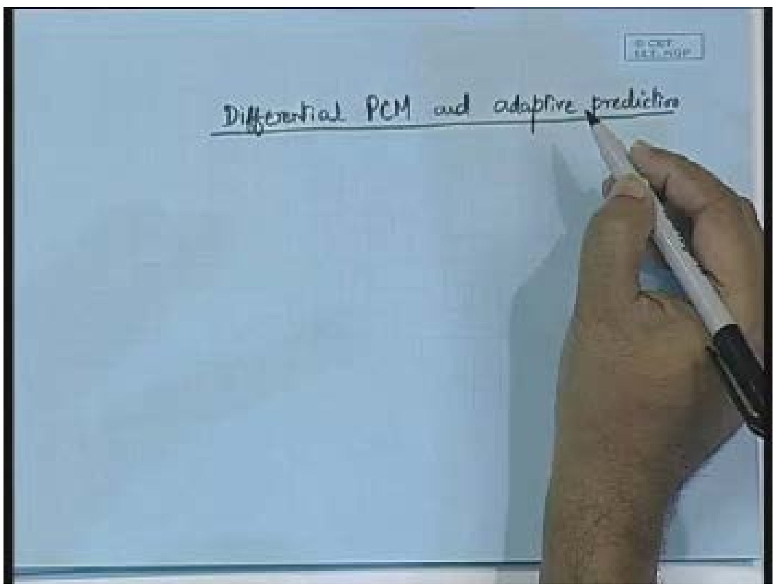
Department of Electronics and Communication Engineering

Indian Institute of Technology Kharagpur

Lecture - 09

Differential PCM and Adaptive Prediction

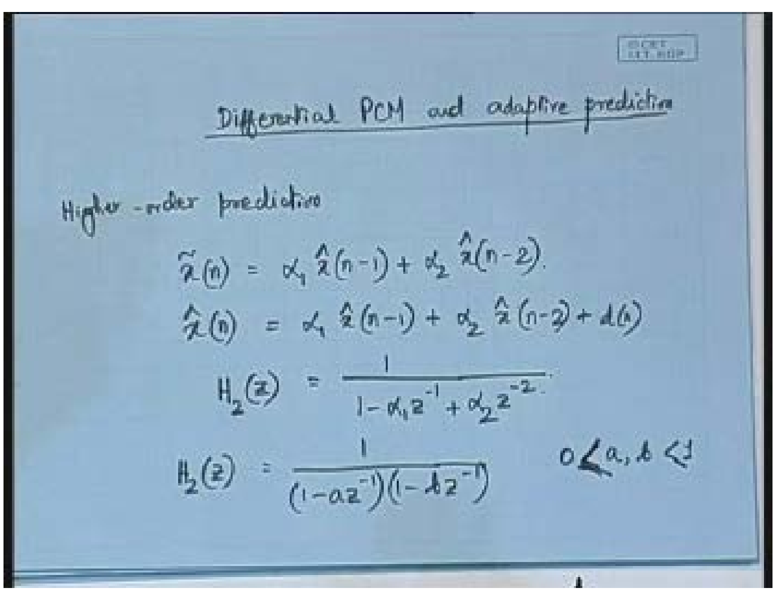
(Refer Slide Time: 1:13 min)



........................and also we are going to talk about adaptive prediction. Now you may be wondering that why i said adaptive prediction. Because we have been talking about adaptive aspects over the last two three lectures so what is new about adaptive prediction. Now mind you, what we have considered so far is adaptive quantization but there is also an aspect which is called as adaptive prediction whereby you can adjust the predictor coefficient. Because remember that we have got the alpha 1 and alpha 2 and all these things. I mean, if it is only a first-order prediction in that case only one alpha, but if it is a second-order in that case we are going to have alpha 1 alpha 2 so how to make that adaptive with respect to this; this also we need to study. This will be the focus of our present lecture. But before we move into this topic just one more small aspect of the adaptive delta modulator or also valid for the linear delta modulator that one should talk about is that using the higher order predictors in delta modulation. So before we go in to the topic let us discuss about the higher order prediction.

Instead of a first-order prediction if we go in for a second-order prediction, in that case we can write down the predicted signal x tilde n as alpha 1 times x cap n minus 1 plus alpha 2 x cap n minus 2, only two, this is a second-order prediction that is what we are going to have and in this case one can write down x cap of n as x cap of n as x tilde of n plus d of n the prediction error. So we can write down x cap of n as alpha 1 x cap n minus 1 plus alpha 2 x cap n minus 2 plus d of n and this is characterized by the corresponding transfer function corresponding Z transform function for the encoder is going to be H, let us call it as H suffix 2 because this is a second-order prediction. So, if we write H 1 as the first-order prediction we are going to write the second-order prediction as H 2; so H 2 as a function of Z can now be written as 1 minus alpha 1 into Z to the power minus 1 plus alpha 2 into Z to the power minus 2. And in fact it should be possible for us to write down this H 2 Z in this form; the second-order predictor can be expressed as 1 upon 1 minus az to the power minus 1 into 1 minus bz to the power minus 1 where this 0 where this a and they are varying between 0 and 1; so 0 less than a or b and this is less than 1.

(Refer Slide Time: 5:20)

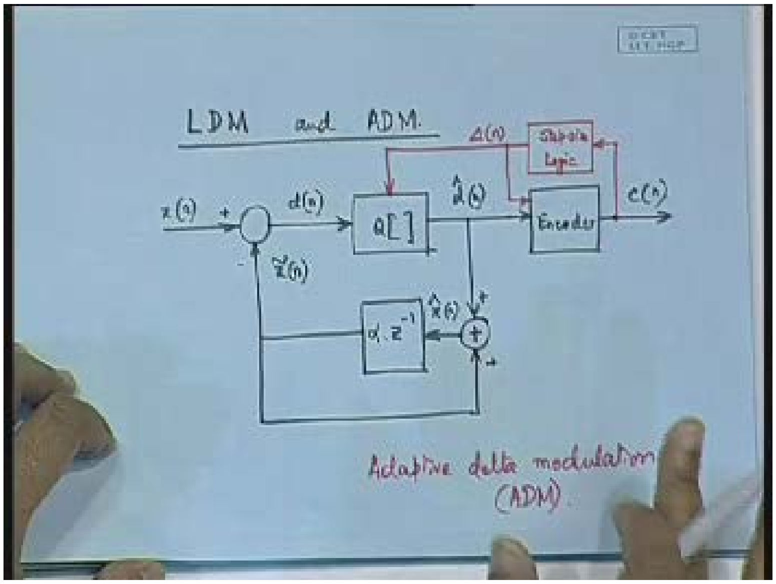


And in this case one can observe that we have got both we have got two poles in this H 2 Z function and if both the poles happen to be real in that case it gives an increased prediction gain. So instead of a first-order prediction, using a second-order prediction one can observe an increased gain definitely. But thing is that many a times in many of the systems you will be finding that despite the fact that second-order prediction will give you an increased gain, people prefer first-order prediction because first-order prediction uses a very simple circuit, only one unit delay is enough whereas in order to have a second-order prediction or a higher order prediction the circuit becomes more complex because one has to account for several past samples because in this case you can see that x cap of n is based on x cap of n minus 1 n minus 2 the higher order prediction would have meant that we would have got x cap of n minus 3 even.

In fact the improvement by using the second-order prediction could sometimes be of the order of even 4 dB; it can vary from speaker to speaker and can vary from one speech content to the other but 4 dB improvement is sometimes quite typical for the case of second-order predictions. It is only beyond some........ for P is equal to 4 or so beyond that point we do not have much of an improvement in the performance but for second-order 4 dB improvement is enough so this is what we should consider for the higher order prediction and higher order prediction can be used in the linear delta modulator, it can be used in the adaptive delta modulator or it can be used even for the differential pulse code modulation which we are going to talk of now.

For the differential pulse code modulation the basic block diagram that we are going to have should be on a very similar line as that of the delta modulator. In fact, remember that this is the block diagram that we considered for the delta modulator.

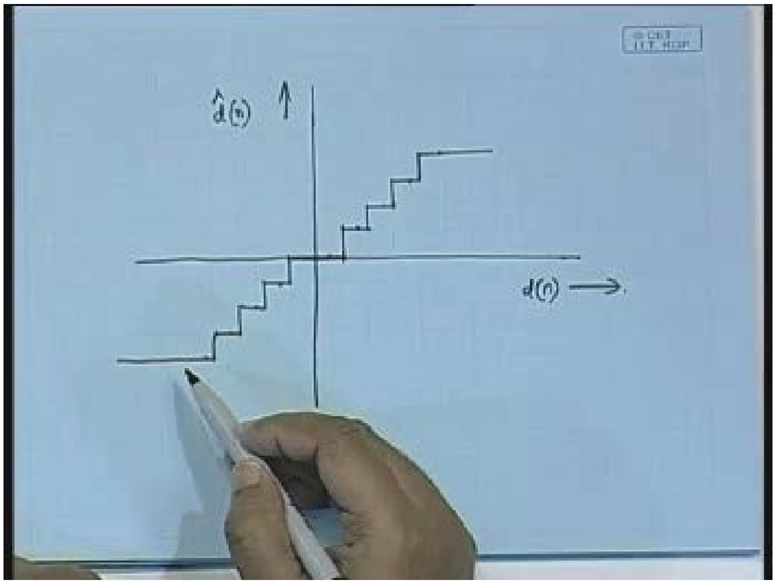
(Refer Slide Time: 8:11)



Later on we had added this step size adaption component also (Refer Slide Time: 8:06). But if you leave aside this step size adaption component, this basic block diagram remains valid even for the differential pulse code modulation also. There also if you follow a first-order prediction then x cap of n is going to predict x tilde of n d of n and then d of n will be quantized to d cap of n. It is only in this quantization where the DPCM is going to differ from the delta modulator because delta modulator is only a 2 level quantizer whereas the differential pulse code modulator will be a multi-level quantizer.

In the case of DPCM our d cap n verses d(n) characteristic would be different. On the y axis if we show d cap of n and on the x axis if we show d of n; now let us say that if we are having a mid-rate quantizer then one can have that the quantizer performance could be like this:

(Refer Slide Time: 10:33)

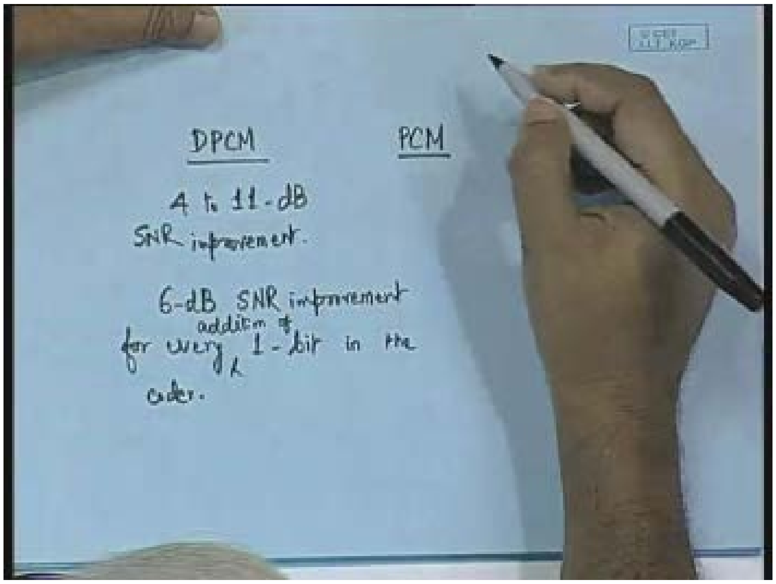


Now you see that there are how many levels 1 2 3 4 5 6 7 8 9 nine different levels we have got unlike only 2 levels which we had used in the case of the delta modulator. And if it is a if it is instead of a mid-rate we use a mid-riser quantization then the number of step sizes would have been an even number; this you already know. So in the case of differential pulse code modulation we have several sub-steps so now we are in a position to finally quantize this d of n, we were very coarsely quantizing it. So delta modulator is in fact a special case of the differential pulse code modulation where the number of levels is equal to 2. But in general in a DPCM the number of levels is larger as compared to 2.

But if we compare a differential pulse code modulator system with a simple pulse code modulator system............ so if we compare DPCM against a PCM; PCM is one which directly quantizes the signal, no differential mode, it just takes the signal and directly quantizes it to one of the permissible levels. Now, if you use the same numbers of levels that is to say that let us consider that it is a 3-bit PCM means 8 level PCM and 8 level DPCM if we are comparing then we will be observing that DPCM has got something like 4 to 11 dB of improvement over that of PCM vis-à-vis direct quantization DPCM gives you 4 to 11 dB of advantage with same number of bits mind you and then if you are allocating more number of bits to the DPCM then with every bit increase in the DPCM that will follow the same rule what we had followed for the case of PCM.

In PCM also we had seen that for uniform quantization, for every bit allocation the signal to noise ratio increases by 6 dB and even here also for a uniform quantization DPCM we are going to have 6 dB SNR improvement so this is 4 to 11 dB of SNR improvement over PCM and 6 dB SNR improvement for every 1 bit so every addition of 1 bit I should say so for every addition of 1 bit in the coder meaning that we improve we can improve it further then instead of the uniform quantization if we use a mu law quantizer then for mu law quantizer also it will result in a 6 dB improvement over a uniform quantizer 6 dB improvement over the uniform quantizer would be obtained.

(Refer Slide Time: 13:37)

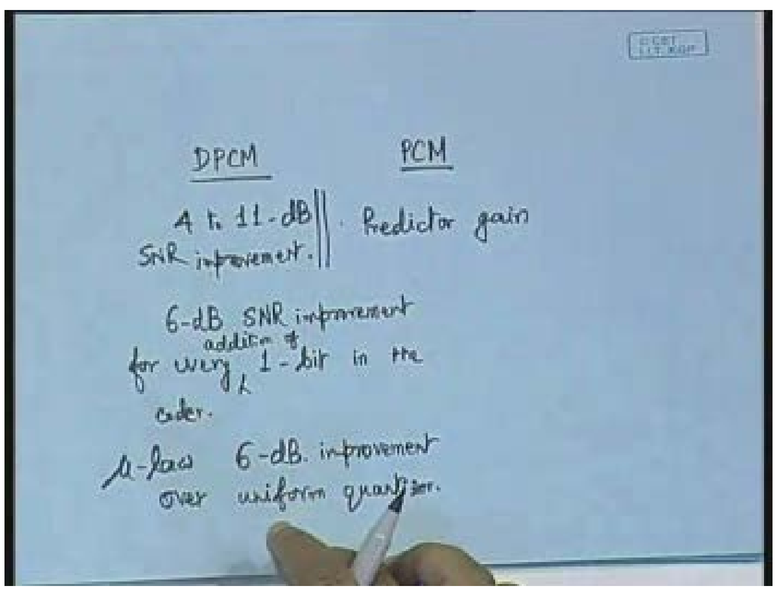


Therefore, what you can try to do is that you can make use of a differential pulse code modulation, that itself gives you 4 to 11 dB of SNR improvement and if you are and if instead of a uniform quantizer you use a mu law quantizer that gives you further 6 dB improvement and now if you allocate more number of bits you can increase the signal to noise ratio further. So definitely DPCM will result in a better performance. If you compare the differential pulse code modulator system against any standard PCM system, you can obtain some remarkable improvement in the performance at the identical bit rate.

Any questions?

[Conversation between Student and Professor – Not audible ((00:15:10 min))] 4 to 4 to 11 dB of SNR improvement is simply because of the predictor gain. So this is arising from the predictor gain. [Conversation between Student and Professor – Not audible ((00:15:25 min))] No, no......let me let me let me clarify this point. See, predictor gain is a different philosophy. Predictor gain we are getting because of the intersample redundancy. Predictor gain is nothing but the G p what we were talking of in the last class. The G p improvement is because of the correlation between the adjacent samples. So that is the phenomenon that results in the predictor gain.

(Refer Slide Time: 16:14)



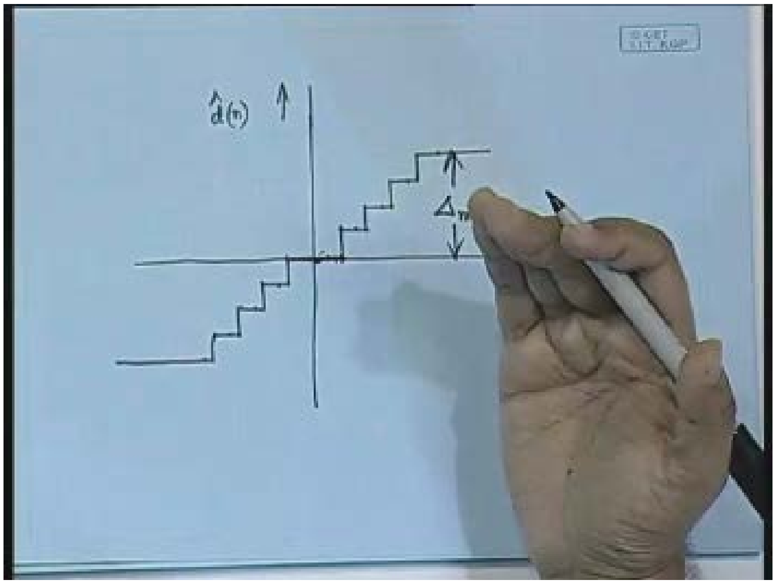
Now this 6 dB improvement with every 1 bit in the coder means that now you increase the quantization level. Supposing you have 8 level quantizer; instead of 8 level quantizer now you decide to have a 16 level quantizer, so earlier it was 3 bit and now you are having 4 bit. So with the addition of extra 1 bit into your encoder you are getting a 6 dB improvement as compared to the 8 level coder. So these are the independent factors, these are not the dependent factors.

Predictor gain is because of the signal correlation between adjacent samples whereas this improvement is because of increasing the number of quantization levels and mu law gives you another 6 dB improvement especially whenever you are having............. I mean; mu law, as you know, that because of its very basic characteristics we are allocating we are having a finer quantization for the more probable portion of the signal.

For more probable values of the signal we are having finer quantization and for less probable values we are having coarser quantization, so that gives us a 6 dB improvement which we already talked of. Now again the DPCM, if I ask that can DPCM suffer from the disadvantages that we mentioned about the delta modulator that is to say the slope overload and the granular noise, yes or no?
[Conversation between Student and Professor – Not audible ((00:18:13 min))] See, I can understand the psychology of the class. That means to say that you are now not very much sure that whether it should or should not but you are at the same time not able to emphatically say that no, there will not be any slope overload or you are not emphatically saying that there will not be any granular noise which means to say that in your mind you are convinced that yes, DPCM is definitely going to be better as compared to the delta modulator but it may not be able to reduce the......... or it may not be able to eliminate the slope overload or granular noise fully it may be still there but to a lesser extent.

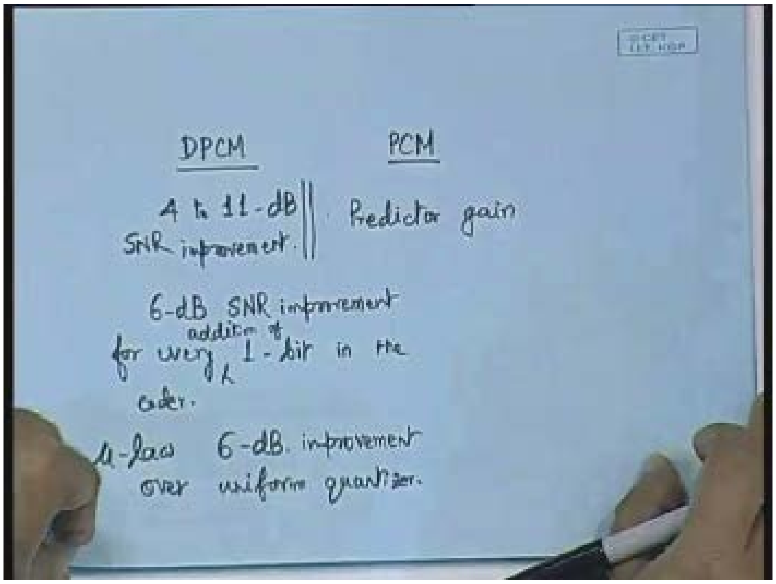
Let us see that if we have a d cap n verses d(n) characteristic somewhat like this in that case if the signal varies, if d(n) varies from this level to this level (Refer Slide Time: 19:30) sudden change in the signal or sharp change which means to say that which will result from a sharp change in the value of the analog waveform itself in that case because of that sharp change we are going to alter the quantization level from here to here and if the signal change is sharp, I mean if I say that this is nothing but a value equal to delta max and if the signal change if delta max upon d that happens to be lower as compared to the rate at which the signal is changing, the slope overload can still happen. But again granular noise in this case we can expect a little lower granular noise because at least this step size is not delta max but this step size is lower than the much lower than the delta max. But even to some small extent it should be there.

(Refer Slide Time: 20:14)



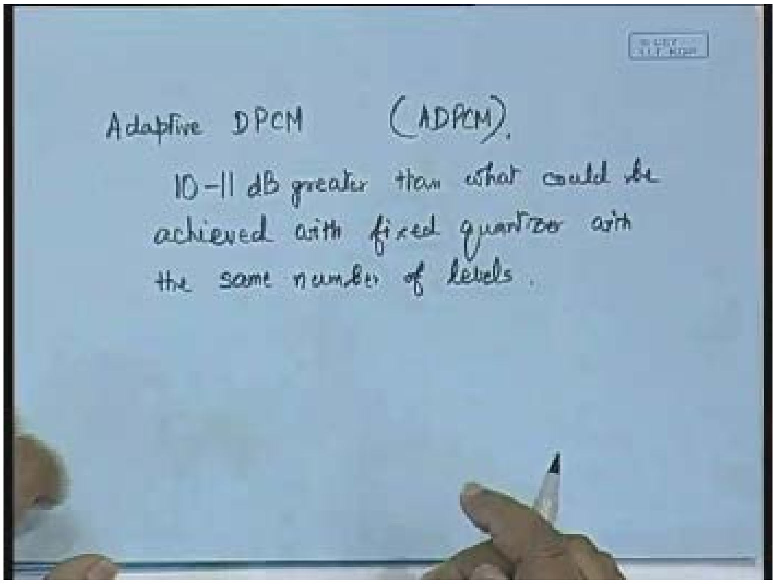
So, in order to have in order to extract the best performance out of the DPCM, one should think of making the DPCM also as adaptive which means to say that whenever you have the slope overload just give it a boost, just increase this delta max do not alter the number of levels and all these things, just increase the delta max by some factor so that your step sizes are larger and whenever you are dealing with small signals, you are dealing with unvoiced portions of the speech or you are having some duration of silence in the speech waveform there you reduce the value of the delta max so that your step size also accordingly reduces and you can reduce the granular noise. That is why instead of preferring the normal DPCM; one goes in for the adaptive DPCM.

(Refer Slide Time: 00:21:57 min)



Hence, we have adaptive differential pulse code modulation or we call it as adaptive DPCM which in short form is referred to as ADPCM and it is seen that the adaptive differential pulse code modulator, this results in a gain; the adaptive DPCM's gain happens to be 10 to 11 dB greater than what could be achieved with fixed quantizer, fixed quantizer with the same number of levels. So, 10 to 10 to 11 dB is the performance improvement between adaptive DPCM and DPCM.

(Refer Slide Time: 23:21)

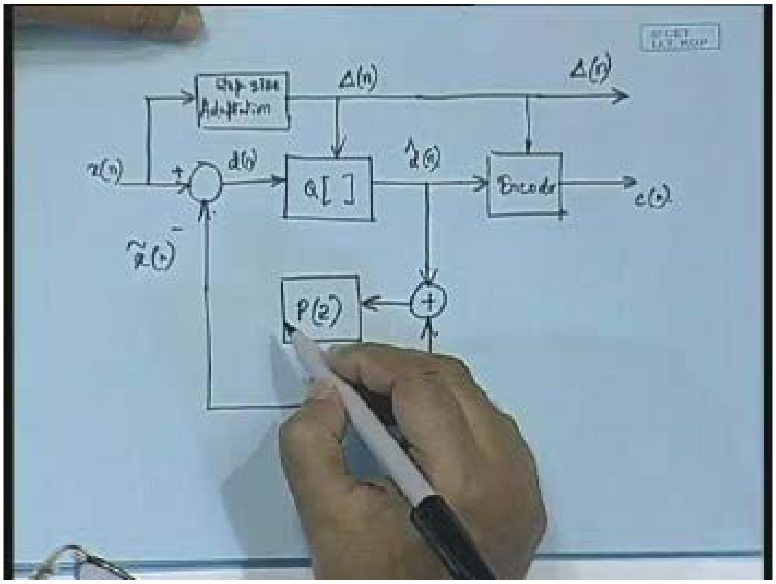


Now you see that.......... if you now compare the adaptive DPCM performance with that of a PCM the improvement is many fold and for speech waveform coding, adaptive DPCM is going to be a very attractive type of a waveform coding approach.

How do we make that adaptive?

We can........ Again likewise we can adapt it using the feedforward or feedback mechanism. So we can have a block diagram like this and for the feedforward as you rightly understand there is going to be.................. so we have to transmit the delta n also, not only c(n) but we have to transmit delta n so it should be a block diagram like this that here x of n should be the input and here we are going to have d of n that is to say the differential signal and then the quantizer Q and here we get d cap of n the encoder and the output here is c of n (Refer Slide Time: 25:02) whereas for the step size adaptation what we should do is that using this x(n) itself we are going to have the step size adaptation control and this will generate the delta n so that using the delta n we can control the quantizer and the encoder and then we have to give this delta n to the channel so delta n and cn both goes into the channel and then this d cap of n, this we have to add up with the predicted value of the signal which is x tilde of n.

(Refer Slide Time: 26:02)

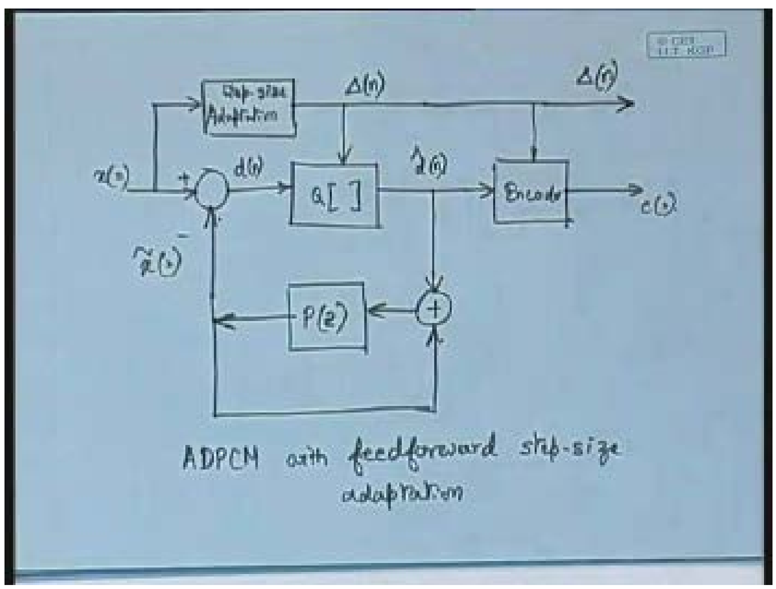


This is x tilde of n, so x tilde of n if we add to d cap of n this gives us x cap of n and then we will be having a predictor block. So let us write it as P of z because we can implement this P of z either as a first-order predictor or we can use it as a higher order predictor or as I told in the title of this lecture that we could also have it as the adaptive predictor. But anyway...... So this going to be the block diagram, so this is ADPCM with feedforward feedforward step size adaptation.

So you see that being feedforward we are also transmitting this delta n (Refer Slide Time: 27:10) and why we decide to have this as feedforward; one can have it either as a feedforward or as a feedback. but feedforward we know the advantages that there it is less susceptible to the error and this delta n; so in order to make it less susceptible to the error, whenever you are transmitting this delta n into the channel then this delta n also should be error free which means to say, to the corresponding decoder if we draw the decoder then at the decoder input we call that as the delta prime n and this as c prime n.

Now if there is a there is an error in c prime n in one of the samples it really does not create much of a problem. May be again if it is a mistake in the least significant bit then that may not affect the performance very significantly. But delta n is very crucial because in case if there is any mistake in delta n, in case there is any error in delta n and delta prime n is significantly different from that of delta n in that case the step size would change drastically; the step size would make an error and as a result of that the signal reconstruction itself goes wrong to a very significant extent. That is why, when this delta n is transmitted through the channel, we should transmit this delta n with a better error protection as compared to c of n because delta n happens to be a more sensitive information as compared to c of n and in that case even in presence of high error rates, one can have good performance if one just encapsulates this delta n with better error protection mechanism.

(Refer Slide Time: 29:26)



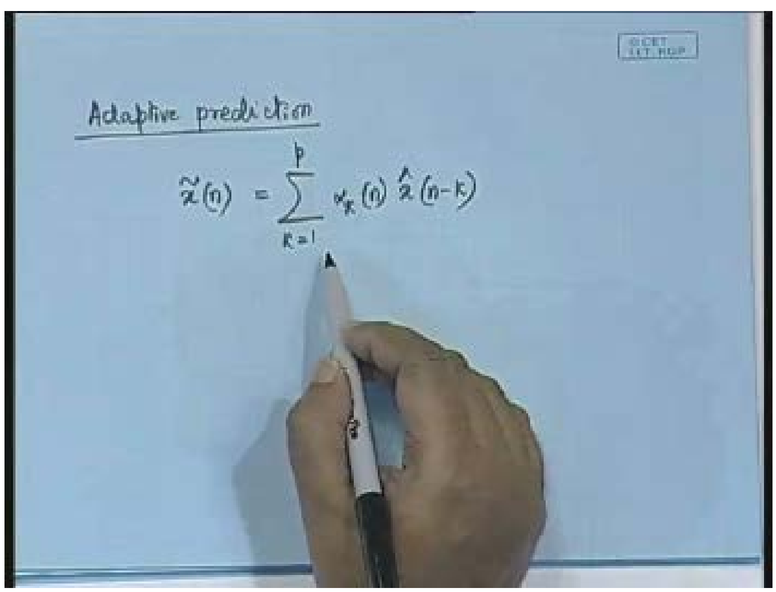
So definitely the adaptive quantization is helping to a very significant extent as compared to the regular differential pulse code modulator. So, first improvement is because of the differential coding which is because of our G p sample to sample correlation and the second advantage is by trying to make the step sizes adaptive, we can really track the analog waveform to a much better extent whether it is varying fast or whether it is changing very slow in the case of the unvoiced part of the speech. So both voiced as well as unvoiced speech requirements are catered for very well with the ADPCM system and that is why in the waveform coding philosophy this ADPCM is very often a preferred solution for the speech encoding.

Now we should talk somewhat more about this P of z. P of z, all that we have talked of is that, in the simplest kind of situation P of z could be a first-order predictor but one can make it even a higher order predictor also by having second-order or third-order and all such kinds of things. But the ideal situation would be to have this P of z to be adaptive. If we can have the predictor coefficients, that can change dynamically in accordance with the signal; that should be the ideal kind of a solution.

Now how we can formulate that problem?

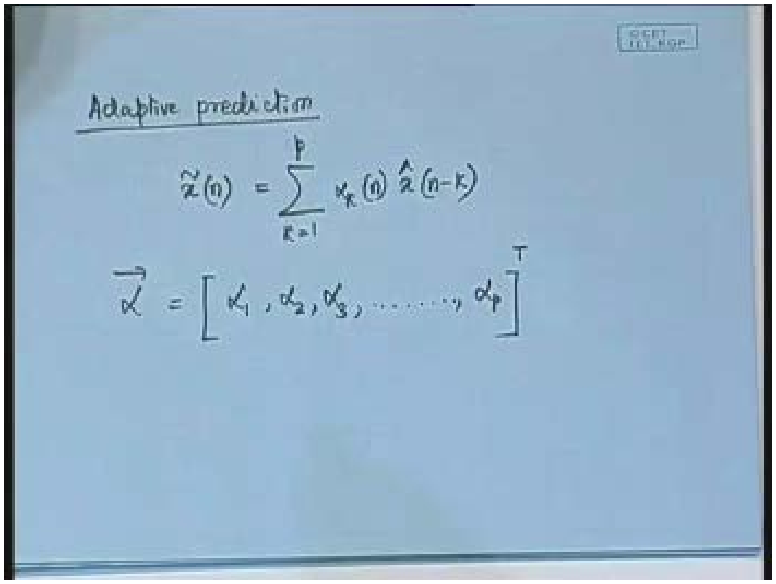
For the case of the adaptive prediction; for adaptive prediction one can write down a general expression for the predicted waveform that is x tilde of n, this could be written as a summation series K is equal to 1 to P alpha K of n into x cap n minus K. This is a general expression which means to say that, as per this expression we are having a P order predictor. And mind you, you also observe that in this case we are not keeping fixed alpha Ks but we are keeping alpha K as a function of n which means to say that we should prepared to vary this alpha K from sample to sample so that is what is making the prediction as adaptive. Adaptive prediction means that adaptively controlling the prediction parameter alpha K or the prediction coefficient alpha K.

(Refer Slide Time: 32:54)



In fact what we are going to have is that these predictors will be alpha 1 for the first-order; and for the sample which is for the two sample the old thing the coefficient will be alpha 2; for three sample old we have alpha 3 etc so alpha 1 alpha 2 alpha 3 up to alpha P, P is the coefficient which is to be used for the sample which is P sample old and this whole thing we are calling as the alpha vector. so alpha vector is something which we can write down as alpha 1 whose elements are alpha 1 to alpha P and we should normally write it down in the form of a column vector so that is why if we write it as row let us put that as the transpose of this so this becomes the alpha vector which is a vector of all the predictor coefficients.

(Refer Slide Time: 34:00)



Now how to obtain these predictor coefficients?

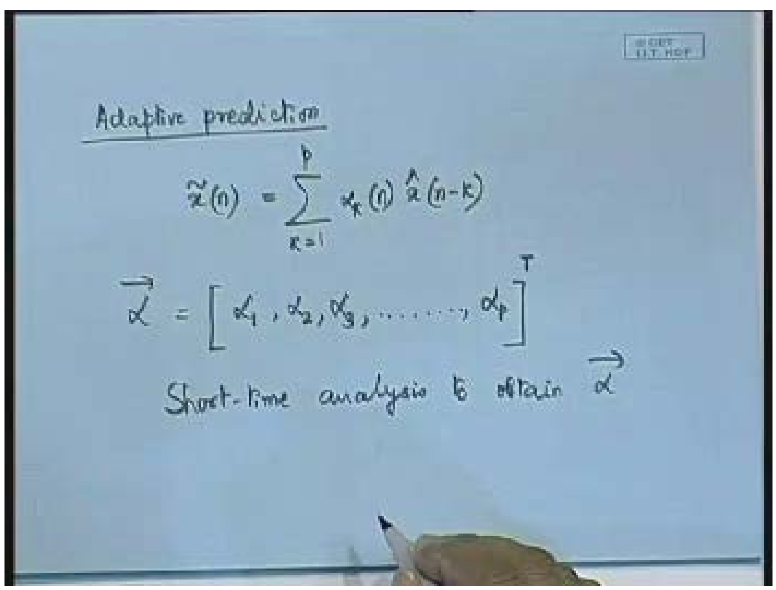
Definitely one has to optimize, one has to optimize the choice of the predictor coefficients and how one can optimize that choice is by simply reducing the mean square error. So you compute the mean square error and you differentiate with respect to all these individual predictor coefficients; you differentiate with respect to alpha 1, you differentiate with respect to alpha 2 and so on and by equating that to zero, equating the derivative to zero, one can obtain the solutions for the optimal predictor.

But mind you, whenever you are solving such equations and equating the derivative to zero, you are making one inherent assumption. And what is that inherent assumption?

That is to say that during your period of analysis your alphas are not going to change. That means to say that you are having definitely a short time analysis. So you are doing a short time analysis to obtain the elements of the alpha vector because during that short time you are assuming the properties of the speech waveform to remain constant.

See, because the speech waveform is a quasi-stationary process; it is not a stationary process. If it is a stationary process then we would not have done any short time analysis, the long time analysis could have been sufficient. but because we are having short time analysis because it is quasi-stationary that is why we have to do short time analysis that within that short time the property of the speech signal does not vary and only then we will be able to choose these coefficients alpha vector so that one can obtain the optimal values of this adaptive prediction.

(Refer Slide Time: 35:52)

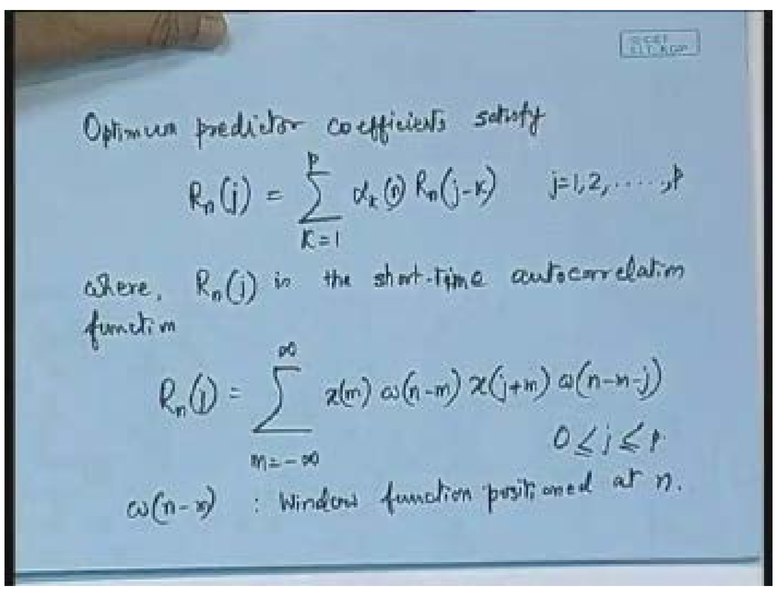


Now it can be shown mathematically that the optimum prediction coefficients that satisfies this kind of a relation; that the optimum prediction coefficients............... this can be written as R n of j the optimal predictor coefficients satisfy this equation and what is that; R n of j is equal to summation K is equal to 1 to P alpha K of n into R n (j minus k) for j is equal to 1, 2,................. up to p) and what is R n? R n is not the autocorrelation no........... it is autocorrelation definitely but short time autocorrelation; R n, this letter R we are specifically reserving for short time autocorrelation function where R n(j) is the short time autocorrelation function.

And how did we define this R n of j?

R n of j we defined like this: R n of j was m is equal to minus infinity to plus infinity x of m into w (n minus m) and I mind you, it is this w which makes it short time because w is going to be a finite window and this multiplied by x of j plus m which means to say that you are taking the signal with a lag of j and then multiply it by w n minus m minus j for zero less than or equal to j less than or equal to P. So this is the definition of the short time autocorrelation function (Refer Slide Time: 39:06) and w n minus m this is the window function, this is the window function which is positioned at n.

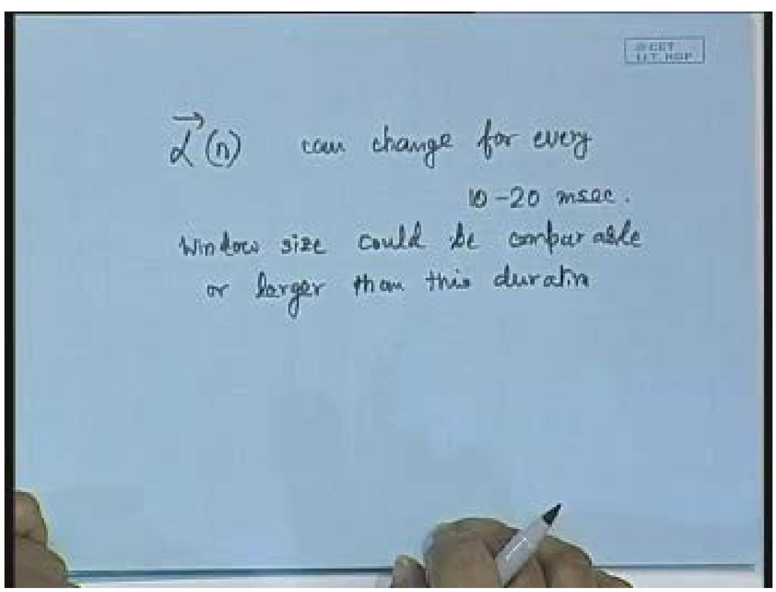
(Refer Slide Time: 39:24)



But since it is a short time autocorrelation function how frequently are we going to estimate these values of alpha? Are we going to estimate too frequently because short time analysis.............. so short so short time analysis means every short time analysis should lead to some values of alpha and then again in the next frame we are going to alter that? But normally what happens is that the speech parameters that do not change very frequently and because the speech parameter remains...................; rather, it changes but it changes quite slowly that is why even these estimates of this alpha vector elements what we are obtaining this also could change in a relatively slower way.

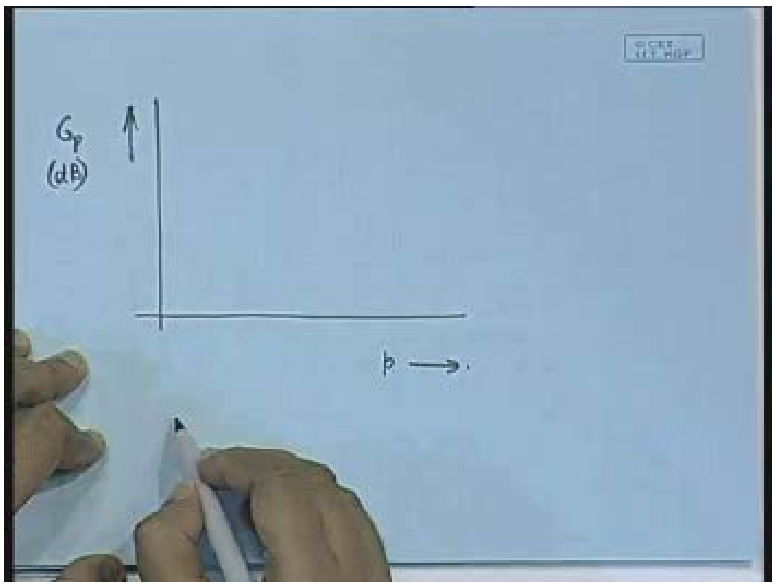
Therefore, it is observed that this alpha n, this alpha vector elements as a function of n can change for every 10 to 20 milliseconds which means to say that it is a quite slow change; every 10 to 20 milliseconds is definitely resulting in a slow change. And the window duration, it may be equal to the interval between these estimates. So one can obtain the window duration either greater than this alpha n's sampling. So alpha n's samplings sampling itself is going to be much slower and one can chose the window size for the autocorrelation computation to be either equal to this duration or maybe somewhat larger than this duration. Therefore, window size could be comparable or larger than this duration.

(Refer Slide Time: 42:22)



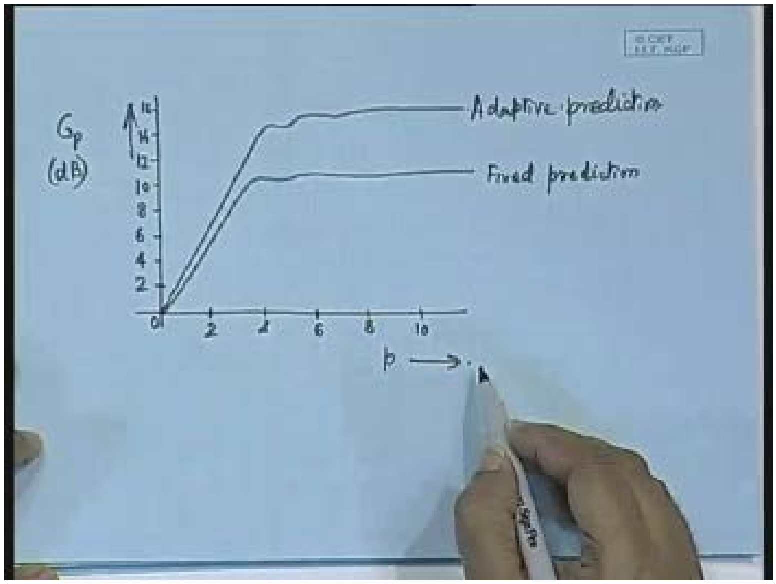
Hence, now we can show some performance graphs that, with the value of P how the adaptive predictor should improve; if at all to what extent does it improve the predictor gain. So we make a plot now as the predictor gain G p versus the number of coefficients and we will be plotting for two situations: one is using the fixed prediction where the alpha values are fixed once for all and the other is with the adaptive prediction what we just now described. And the curves are obtained somewhat like this that on this axis it is P that is to say the number of coefficients and on this axis we plot the predictor gain G p and this we are going to plot in dB.

(Refer Slide Time: 43:10)



So, number of coefficients one can have this as 0 so definitely 0 is a very hypothetical case but 2 4 6 8 10 like this so these are the number of coefficients and the predictor gain we plot this as 2 dB 4 dB 6 8 10 12 14 16 and so on. So one observes that for adaptive prediction it goes like this (Refer Slide Time: 44:00) and for fixed prediction it is slightly lower. So this may be for fixed prediction and this is the nature of graph one may obtain for adaptive prediction. So you see that there are two observations that we can make from this.

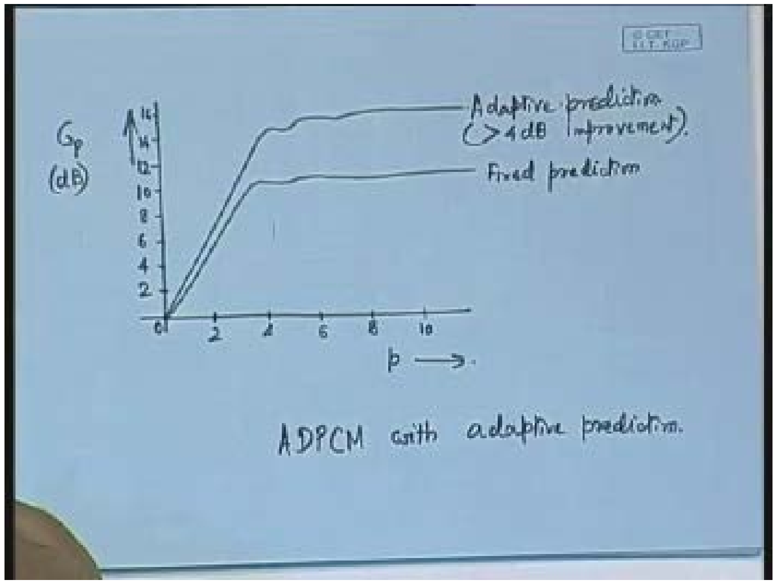
(Refer Slide Time: 44:48)



First is that, with the increase in the value of P, just to say the order of predictor, if we increase the order of predictor to 2 definitely there is some improvement in the gain, I mean in the differential in the gain G p. But you see, that beyond P is equal to 4there is hardly any improvement; may be 0.5 dB or less than 1 dB which will not be really a which will not make any perceptible difference. So there is not much of change in gain after P so that is why we did not have to complicate our circuitry by having very high order predictors.

And another observation what we can make is that, as compared to the fixed prediction the adaptive prediction is giving you almost......... I mean, by looking at this graph, one can conclude that it gives almost 4 dB of improvement so around 4 dB improvement, greater than 4 dB improvement over that of the fixed prediction. We have already reached a good amount of sophistication. So what we should now attempt for a good quality speech encoding is that we should use an ADPCM; and not only ADPCM with a fixed prediction but if we use ADPCM with adaptive prediction then that is going to give us the best performance.

(Refer Slide Time: 46:41)



Now again this is one aspect which has been very well researched, so in the journals one can find a lot of good quality research papers pertaining to all these theories on the adaptive delta modulation and the adaptive ADPCM and these adaptive predictions and all these things. But there are always further and further improvements which the researchers try to do.

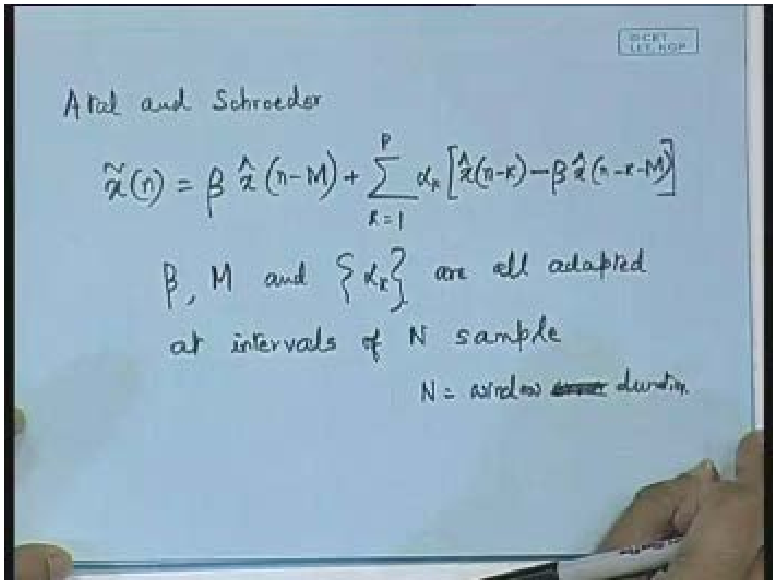
Now, in the case of the adaptive prediction one makes use of the short time autocorrelation. Now short time autocorrelation, this R of j, (Refer Slide Time: 47:39) what we have obtained is that we are always saying that this j should be much smaller because we find that beyond P is equal to 4 there is not significant improvement in this. But again you see that, if you look at the autocorrelation values, short time autocorrelation values, you will see that beyond j is equal to 3 or 4 there will be a sharp dip in the autocorrelation value but the autocorrelation value will be again attaining some peak at its periodicity interval.

Remember, speech signal is quasi-periodic. So during the quasi-periodic time, I mean, because of the quasi-periodicity at the value of that periodicity we are going to obtain another peak. So what we can think of is that, okay, it makes no sense in taking more than past four old samples but making use of the periodicity of the signal why cannot we use the sample which was one period old; because the sample which is one period old that can have a better prediction effect as compared to a sample which is five samples or old six samples old. Therefore, from this consideration what one should do is that, in the prediction mechanism one should not only invite the past three or four samples but also in addition include the periodicity.

Therefore, the predictor mechanism should be something like this: this was in fact proposed by Atal and Schroeder. So Atal and Schroeder's philosophy is that you can obtain the predicted value x tilde n as a coefficient beta times x cap of n minus capital M. This is this capital M is going to be what? Capital M is going to be the estimated period and beta is the coefficient corresponding to that. So capital so the x cap of n minus capital M is a fairly old sample; it is a sample which is capital M sample old and capital M is the periodicity so we are making use of the correlation that exists because of the periodicity and this plus (Refer Slide Time: 50:36) we can have the regular prediction term: K is equal to 1 to P alpha K and alpha K we normally have alpha K times x cap of n minus K but in this case we can make use of the periodicity property and say that it is x cap n minus K minus beta times x cap n minus K minus n. So what we are essentially doing is that we are taking the difference between this n minus K and n minus K minus M. So, that times alpha K is going to dictate that what is going to be the predicted signal x tilde of n and in this case one has to adapt many parameters.

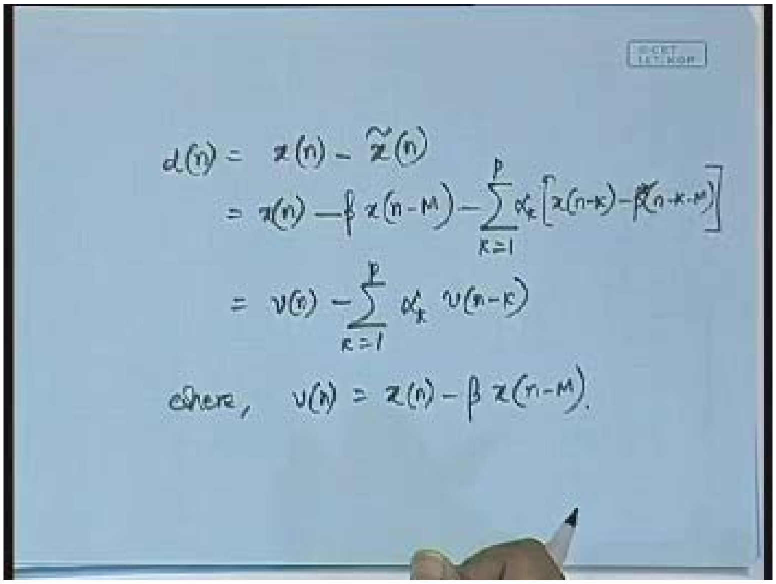
Therefore, the predictor parameters are beta, then capital M and the coefficients alpha K they are all adapted at intervals of all adapted at intervals of capital N number of samples where capital N is nothing but the window duration window duration.

(Refer Slide Time: 52:11)



Thus, now, neglecting the effect of quantization error, one can write down the prediction error as d of n; one can write down as: x of n minus x tilde of n which is equal to x of n minus, you simply substitute this expression (Refer Slide Time: 52:37) and you can write it down as: x(n) minus beta times x(n) minus capital M minus summation K is equal to 1 to P alpha K into x of n minus K minus beta x of n minus K minus M. I just happen to omit this x so please include beta times x of n minus K minus M. So this one can write down as: v(n) minus summation K is equal to 1 to P into alpha K times v(n) minus K where v(n) is nothing but x(n) minus beta times x(n) minus M. So this is by including the periodicity into our prediction process.

(Refer Slide Time: 53:43)



This was shown by these two persons Atal and Schroeder that this results in an improved predictor performance. All these developments have really led to high quality waveform encoding techniques for speech. From the next class we will be beginning the new chapter and that is on the LPC vocoders; so linear predictive coding vocoders that we will be taking up as the next topic, thank you.