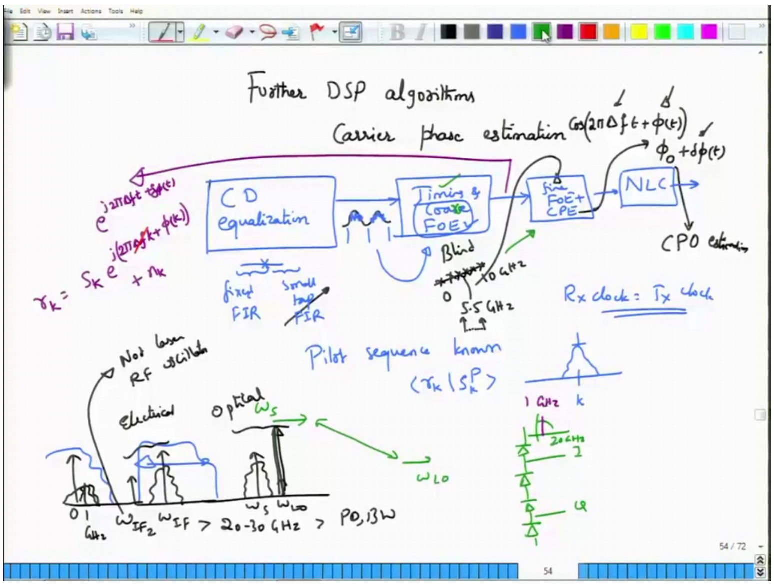
Fiber - Optic Communication Systems and Techniques
Prof. Pradeep Kumar K
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

Lecture - 57
DSP algorithms for Carrier phase estimation –I

Hello and welcome to NPTEL MOOC on Fiber Optic Communication Systems and Techniques. In this module, we are going to further look at DSP algorithms and this time we are going to concentrate on a specific problem called as carrier phase estimation.

(Refer Slide Time: 00:36)



As we have already motivated the D S P chain, you begin the first block by receiving the sample factor coherent detection. The samples that you receive will be de skewed aligned to make the in phase and quadrature component with aligned with respect to each other and then you perform C D equalization. The C D equalization usually is performed either in the time domain or in the frequency domain and it is usually also adapted in the sense that you have you know the total number of taps that you are going to use which your performing whether it is a sequential, the time domain does not matter.

We actually split the number of taps into two part one large taps will give a fixed F I R filter, ok. So, this should be a fixed F I R filter. In that case of a frequency domain essentially analogous to that one and then a small tap filter, this is also an F I R filter but this filter will be adaptive, ok. So, using adaptive technique actually allows us to converge this algorithms and actually to track the changes that are actually going to happening the channel which results in C D being or the C D effects being changing with respect to time not very fast as in wireless communication, but it does change with respect to time and you can use this adaptive filter to actually compensate for the complete chromatic dispersion. Again, please note that we are assuming load systems in which there is no in line dispersion compensation.

So, these are discussion un manage links as we would have as we have already discuss. So, after we have performed the C D equalization, the symbols now return back to their original slots of course, there will be lot of noise here that anyway if the contribution there are come from various stages, plus there are noises because of the laser having non zero line width which results in the phase errors and so on, but the important is that the symbols return to their original slot meaning that, you can now extract clock from these received symbol.

Usually, what you do is you actually send out a pilot sequence and detecting the pilot sequence or a frame header if you can think of, this will be important because once you detect this one wherever this pilot sequence you actually what you do is take this received sample and then correlate the received sample if the pilot sequence that is usually known to both transmitter and receiver.

So, as you take the received samples and then correlate with them; correlation is essentially taking the inner product. So, you take this correlation with the pilot sequence that I have already known. So, when you do this correlation and then look at the correlation between them, we will see that at some particular value of k, we will see that this would be actually maximum and then that could be the appropriate time for you to decode the or rather appropriate time to sample the signal. So, this will be used to obtain the timing or the sampling times and once you done that one, then your timing information is all right.

So, the receiver clock and then the synchronize to the transmitter clock ok. So, this is very important step and you have to do that step right after the C D equalization because C D equalization allows you to put the symbols back into their. Reasonably, now if there is some residual discussion, then this step of receiver clock and transmitter clock being a synchronize become to in complicated, but more or less if C D is completely equalized in the symbols are back into their time slots and now with the help of some pilot or training sequences, you can correlate the sequence it what we have received to actually begin the slots themselves.

So, you can learn where the slots are supposed to be and then you can derive the sampling clock from them or the extract the clock from them and that is the major task of any receiver. So, you have the same sampling instants as at the transmitter. Of course, you can allow for a constant delay, but this particular process of using a pilot sequence and correlating actually allows you to know how much delay over all that you have to take into account for.

So, once you have taken that delay into account, in the transmitter clock will be running at the same rate as the receiver clock and there will be in phase with respect to each other. So, clock transmitter clock raises, the receiver clock also raises along with that except to some very very small differences which may not affect the system as much.

So, the first step would be to actually complete this timing offset. Now, one or timing offset of timing recovery after you have done this timing recovery. Then you can go and look for this coarse frequency of that estimation. Remember, you had your signal laser and then you have at the receiver a local oscillator and these two signals are usually of different frequencies, but if they are too wide, then the photodetectors that you are going to use, remember you would have used four photodetectors to extract the in phase and quadrature components correct; these photodetectors have a finite bandwidth.

So, let us say the band width are about 20 gigahertz. These are already very high band width there I am considering, but what it means here for our frequency offset is that the offset between the signal frequency or the transmitted optical signal carrier and the local oscillator should be well within this 20 gigahertz so that the recovered spectrum actually falls in this passband of the band width that you have for the photodetectors.

Please note that this is not really the restriction of the second stage. I mean stages after the photodetector, but it is actually the restriction of the photodetector itself. So, this photodetectors bandwidth itself is a limiting factor that tells you how much should be the maximum allowable offset for the system of course, 20 gigahertz is too high I mean may not going to make these a frequency offset the as high as 20 gigahertz. You will definitely want to reduce it as much as possible.

But if you are not able to reduce it, then let us say we try to reduce it to about 1 gigahertz of something by appropriate training sequences and adjusting them, but if this is an acceptable limit or whether you can allow for 1.1 gigahertz, 1.5 gigahertz depends on the application, but absolutely you cannot allow it to be more than 20 gigahertz because then the signals will fall outside the photo detector plus receiver filter band width and you cannot even detect anything right. That is not completely true because you may have omega S here let us say omega L O is larger than omega S.

And then, the different signal is what you are looking at and the different signal will actually fall at some intermediate frequency. So, because omega L O is not exactly equal to omega S actually we should put it in this way because omega L O is usually are stronger frequency and then you have the bandwidth of the data centered around omega S.

So, if omega L O minus omega s happens to be omega I F and this omega I F happens to be greater than say 20 or 30 gigahertz which definitely is greater than the photodetector bandwidth, then what do we do? So, you have your data recovered at the intermediate frequency, which is the difference between the local oscillator and the transmitted laser frequency omega S, but then this difference frequency or the centre frequency where this difference is happening which is the intermediate frequency is much larger than the photodetector bandwidth.

Then what do? I do well what you have to do then is to every put one more intermediate frequency that is called this as omega I F 2 and now you keep the difference between these two to be about the gigahertz ok. So, let us say this is your actual 0 frequency and then let us say this is your 1 gigahertz frequency, and then you centre your data around at this particular point or other because the data would be in this manner, right. So, you can use more than one stages and if you look at what is happening these two are essentially beating in the optical domain because you have a laser frequency omega S and then you have a laser oscillator frequency omega L O.

And therefore, this beating between the two signals is happening in the optical domain then converted to the electrical domain of course, but then this beating is happening purely in the electrical domain. This omega I F 2 is not a laser, but it is actually an R F oscillator ok. So, you can actually have an R F oscillator which would be say about 18 gigahertz or so while omega I F is about 30 gigahertz let us say so that the 20 gigahertz let us say, then the different frequency will be about 2 gigahertz which will be much more within the bandwidth of the photo detector for other which is much.

Now, bandwidth of the next step let us say, not exactly the photodetector, but the next step. Otherwise, if you do not do this one then your bandwidth would be center at 20 gigahertz and or 20, 30 gigahertz and then you are actually looking at this range of frequencies to actually work the next stages with. So, you have to sample the very high rate, you have to do lot of other things which is much more complicated if the I F is not located very close to 0 and any deviation from I F from 0 is what we call as frequency drift.

So, in addition to offset which is kind of a static phenomena, so you have an l o and then you have an omega f and the difference let us say whatever means that we have tried is about 1 or 2 gigahertz and then this 1 or 2 gigahertz will change slowly because omega L O changes slowly and omega F changes slowly. So, there is a drift between the two laser which then translate into a frequency drift itself.

So, that is also something that you need to worry about, but drift being a time varying phenomena, the traditional way of you know solving the problem of drift or estimating the drift and compensating for the drift is to use an adaptive technique. So, this adaptive filter that you are going to place will first estimate the offset frequency and that is done by a static filter also and after that, you then track the drift overtime by letting the adaptive filter coefficients vary with respect to times. So, this tracking is an absolutely important phenomena in almost all of the receiver D S Ps that you are going to use for high speed coherent optical system.

Anyway, I will not cover much of frequency offset technique because most of this techniques are quite similar to the next technique that we are going to talk about which is carrier phase estimation or the algorithms we are going to discuss for carrier phase estimation with some adjustment can be use for frequency offset estimation as well. Usually, the coarse frequency offset is done by the blind methods or basically to stepping the ranges.

So, you have difference of say 0 gigahertz and up to let us say 10 gigahertz, you select or you select about 10 or 20 values in between and then look at each value. So, most of the time your offset will be around some let us say for your 5.5 gigahertz. So, let us say just for this purpose, let us say this is about 5.5 gigahertz.

So, as a step through different frequencies you start looking at the matrix the received matrix, I would not tell you what those matrix are because that would require as go to a separate coarse itself, but there are certain performance measures that you are going to monitor as the frequency of the local oscillator is kept along let us say with resolution of 1 gigahertz.

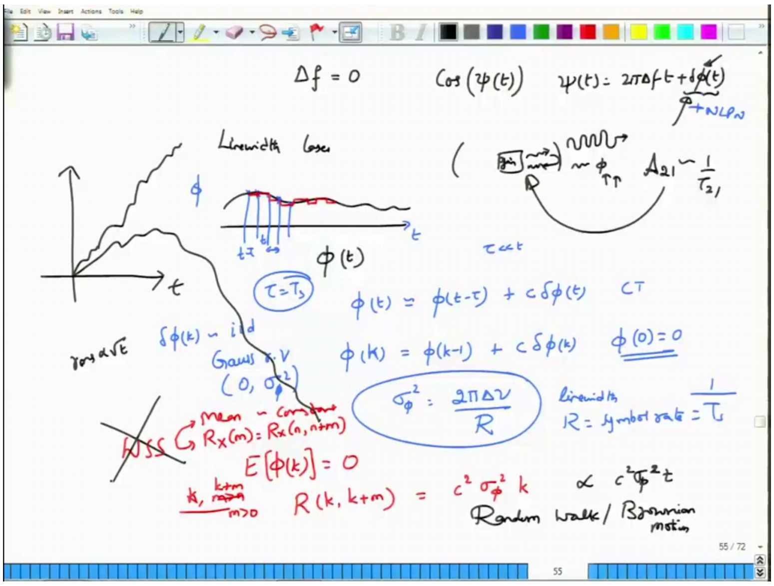
At 5 gigahertz and 6 gigahertz, you get a maxima right because this is or the maximum will get 5.5, but then by looking at the metric you can then decide that the frequency offset lies somewhere in between. And in between this is about 0.5 gigahertz and that will go into a next estimator called as fine frequency of that estimator. So, you can actually combined the fine frequency offset which we normally denote with delta F and then with the unknown carrier phase which we would have denoted already by phi. So, the received signal that you are looking at is something like this.

So, this is what you actually are looking at the carrier. So, the carrier is offset in terms of delta F and please remember this delta F is not very large offset because that large offset is already been taken care of or estimated. So, whatever that small offset that you see which is the fine frequency as we would say that is what goes into the carrier. So, the two lasers ah; the transmit laser and receive laser after accounting for the coarse frequency offset, you are now separated by a fine frequency of delta F.

And then, there is of course, the phase which is time varying again the space could be in offset phase meaning that there could be us kind of an average phase phi naught plus there could be a variation after phase around that phi naught. Again, you can actually perform an offset estimation. So, that is called as carrier phase offset estimation and once you have done that phase offset estimation, then you can go and then look at that time varying small variations in the phase delta phi of T or the fine phase as you could call them rights.

So, your essentially looking at 2 phi delta F where delta F is fine plus delta phi which is also fine quantity. How fine and coarse, these are depends on the application and the algorithms that you have used ok. What are we now going to look at in terms of carrier phase estimation? Let us first understand a little bit more about the statistics of this carrier phase estimation.

(Refer Slide Time: 13:58)



And for that reason, I am going to say delta f equal to 0 if delta f is not equal to 0, then you can write your carrier as cos psi of t and psi of t can include only that part which is phi right. So, it could include only these parts which is 2 phi delta f t plus delta phi of t; of course, if you do not estimate delta f what will happen maybe you would have transmitted these constellation point, but then because delta f times t and time is increasing linearly each of these point start to drift away in terms of this space.

So, you will eventually receive a constellation it is completely closed assuming that no other noise is there in the constellation is completely closed. In addition to it, there is a small delta phi of t change as well. Now clearly a close constellation is absolutely useless to you. So, you will have to remove this 2 phi delta f t term. When you remove the 2 phi delta f t term, what you will find he is that there is some amount of phase noise in each of those.

Now, as long as the phase noise is small then you are all right, but if the phase noise results or becomes larger, then you will it will again close the constellation and then give you lot of errors which are not at all desirable ok. So, you have this phase offset and frequency offset both rotating the constellation and that is what we need to avoid anyway. So, we are good assume delta f equal to 0. So, which means that we have not really looking at the frequency offset at this moment, but then we have to look at this delta phi of t.

Now, delta phi of t arises because of the line width of the laser. Why this occurs is because of the spontaneous emission and if you go back to our studies in terms of the laser, we said that there is a cavity then there is and gain medium. In traditional laser where there is a cavity and again, there is always spontaneous emission and this spontaneous emission could be reflected back into the cavity or will be filled inside the cavity itself most of the time and that will add on to the stimulated emitted photon.

So, photons that are emitted by stimulated emission are nicely coherent in their own way, they have you know no almost no fluctuations in terms of the average photon number, but this little spontaneous emission will have their own phase and they will also have they will contribute to the photon number fluctuation because in a given time, you do not know how many such photons are occurring or arising because of the spontaneous emission. Remember their actually denoted for their decided by the spontaneous emission parameter or the Einstein coefficient A 2 1 and the rate was usually given by 1 by A 2 1, right.

So, or rather rate was given as 1 by A 2, I mean A 2 1 and then the time was given by 1 by A 2 1. If you go back and look at this from this was actually, 1 by A 2 1 if you remember. So, these are the main results are reasons why you have line width of a laser being non zero and what is that mean for us? What it means is that, the phase of the carrier emitted from the laser will be random ok, but I am not looking at the continuous version of the phase change right, so continuously yes phase of the laser is changing as phi of t, but if I am now going to look at just two small instance of this phase.

So, this is with respect to time this is phi, if I am looking at two small instants which is a t minus tau and t where tau is very very small compared to t. So, your actually looking at two closely a related point, then it can be shown that phi of t can be approximately written as phi of t minus tau plus a constant c times delta phi of t. Please note that this delta phi of t is slightly different from the delta phi of t because this delta phi of t will also include the non-linear phase noise which we have not decide I mean discussed yet. We will going to discuss that were later on, but then you have the c times delta phi of t. If you are to fix these time instant differences right, then what you can actually do is you can convert this continuous time equation into a discrete time equation by writing this as phi of k equals phi of k minus 1 plus some constant c delta phi of k with an initial condition that phi of 0 b equal to 0 because there is nothing but has happened yet. So, as soon as a laser start, that the initial phase we taken as the reference and that phase will be taken to be 0 right.

What about this delta phi of k? Delta phi of k is considered to be independent and identically distributed Gaussian random variable. So, this a Gaussian random variable with 0 mean and a variance of sigma phi square ok. So, this variance is actually dependent on the time difference in the sampling that we do for the actual laser continuous time laser phase and it is given by 2 pi delta nu divided by R, where delta nu is the line width of the laser.

So, you can clearly see that poor lasers with large line width you going to have more uncertainty in them because sigma pi square will be larger in that case. R is the symbol rate because we are taking one sample every symbol time right. So, this is a symbol rate it is also equal to 1 by T s where T s is the symbol time.

So, if you set tau equals T s in this one and then start sampling the carrier phase of course, you do not really have access to this phase this is kind of a model that we are actually building up. So, if you imagine that the phase has been sampled every T set T s seconds then the variance of that random variable of the phase jumps that you are going to get will have a variance of sigma phi square. So, the correct way of thinking the phase noise would be to have the jumps here because that is how we can actually obtain the samples right. So, you are actually approximating a continuous time function with discrete functions but these functions are not deterministic function, they are random function, ok.

So, this you have to keep in mind if you probably little further into the space noise to see whether the space noise you know equation that we have written phi of k, is it constituting a discrete time stationary random process or a weak sense stationary process at least, then you can show that the mean of the sequence phi of k will be equal to 0 because a process is said to be W S S or weakly stationary process or stationary in the weak sense, then you have to satisfy two conditions; one that it is mean be actually equal to constant because this is 0.

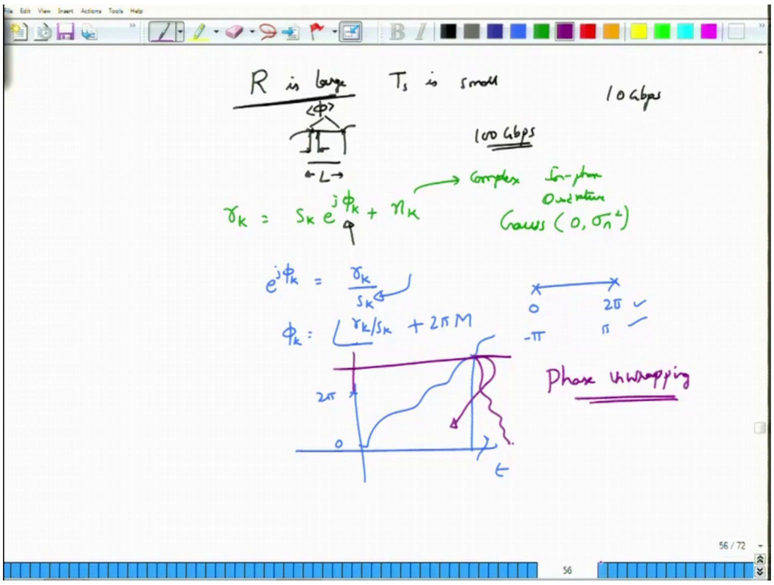
So, this first condition is satisfied and second the autocorrelation of any random sequence x right be a function only of the difference between the two time instant. So, let us say the time instant are n 1 n plus m, then the autocorrelation function between the 2 must be a function only of the time difference for the index difference and not exactly depend on both n and m, ok.

So, this is a second condition autocorrelation of the space sequence, you can show that for n and m greater than n or rather n greater than this one, then the autocorrelation function or rather k right. So, we are using k and if you assume that k plus m, m greater than 0 such that k plus n is actually greater than k, then the autocorrelation function r of k k plus m can be shown to be c square sigma phi square because that is the variance of this one times k itself ok. So, please note that if you were to write this one in the continuous time version, then this would be something like c square sigma phi square times t.

So, that phase noise sequence is actually proportional to square root of t, the variance is proportional to t and the R M S value is proportional to square root of t. This is also called as random walk or Brownian motion which was again analyzed first by Einstein who derived all these statistical idea.

So, what you can see is that, if you were to actually somehow access the phase, then the phase would start at 0 and then it is variants keeps on increasing. It may not only increasing this way it may do it something like this, but overall it will the R M S value will be proportional to square root of t for as t grows larger and larger as k increases and increases, then this variance also increases. So, clearly this is not a process which is W S S but luckily for us, we do not need to worry about these changes especially when R is large.

(Refer Slide Time: 22:49)



When R is large, T s is small which means the samples that are being taken are much more closer to each other and there is sufficient correlation between those samples when you take them to be very close with respect to each other.

So, even over several symbols let us say some L symbols long, the phase can be approximated to be a constant or you can actually find out the average phase here and then work with this average phase because the deviation from the average will be quite small for large symbol rate.

So, if you are looking at 100 Gbps when things are going to be much better than things are going to be a 10 Gbps because 10 Gbps smaller, sampling time t s is larger. So, this in fact, gives your problem what is the problem that we are addressing? Well if you go back to this receiver that we have at right at this point ok, right at the input of the fine frequency and carrier phase offset we have said that there is a phase of e power j 2 pi delta f t plus delta phi of t. This is my continuous domain what is the discrete domain this would be e power j 2 pi delta f k plus delta phi of k or rather phi of k because that is what we have used.

So, this would be plus phi of k but at every k th sample, we actually have also transmitted a signal S k and because there is also noise in the system where it has come from the amplifiers and other thing, this will be the actual received sample at the k th instant. And since we have assumed delta f to be 0, then we do not need to worry about it and then go ahead and then write down that the received sample at that k th time instant is given by whatever that you have transmitted S k times e power j phi k where phi k is the phase sample phi of k that I have used. I am just using instead of using a bracket; I am just using this as a subscript so, but please understand this one plus n k.

This n k is complex meaning it has both in phase and quadrature component we are going to assume that this is also Gaussian random process or a Gaussian random sequence which has 0 mean and some variance which we will call as sigma n square ok.

Now, what is the objective here? The objective is to find phi k. How do I go about it? If there was no noise in the system, if n k was completely 0, then e power j phi k could be written as r k divided by s k correct. I could have done this one and then said phi k is the argument of r k divided by psi k and everything would have been fine as it is right.

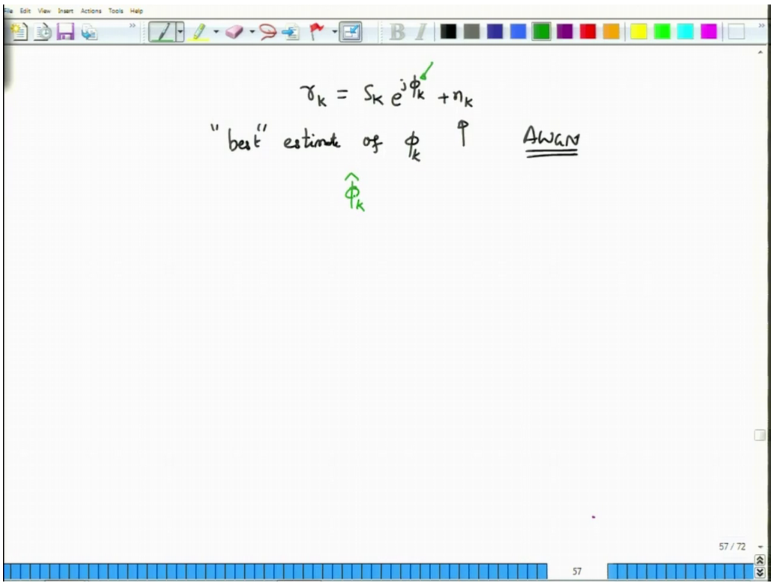
Unfortunately this is not correct. In the sense that I do not know what is s k and there is a chance that s k if I do not know what is s k, then and if s k is very small and r k is large in this argument may not give you the correct value within the appropriate bounds and therefore, that will be a problem in itself. Moreover, whatever the angle that you are going to find out will always be you know you can always add value of 2 pi to this one or in fact 2 pi m to this one because changing the argument by a factor of 2 pi m where m is an integer does not really change the equations.

So, you have to also take into account what is the fundamental range over which you want this argument to fall, and that fundamental range could be either the first cut 0 to 2 pi or minus pi to plus pi ok, both are widely used. So, you can use whichever that is you want, but then if you are actual phase you look at it the actual phase may actually exceed this right.

So, what so, if you actually plot this phase with respect to time the phase may actually exceed more than 2 pi. So, this is a 0 to 2 pi at this time the phase is actually greater than 2 pi. So, what you are actually doing here is to bring this back onto this range ok. So, we have to keep subtracting whenever you see that the phase actually exceeds this range, you have to subtract 2 pi and then pull it back into the fundamental range of 0 to 2 pi ok.

So, this problem is called as phase unwrapping problem and it is an extremely important problem not only in this communication system, but in many many other optical systems as well as in other systems as itself ok. So, all that happens because of the 2 pi ambiguity that any argument of a trigonometric function can with stand, so, you change the argument by 2 pi nothing happens to the function sin cos and tan they are all periodic with that particular value of 2 pi. So, that is the source of this phase unwrapping problem.

(Refer Slide Time: 27:38)



Anyway, just to because we are about to finish this module, we have just laid down what is the problem that we are discussing and we are only discuss the solutions of this problem in the next class, starting from looking at how to obtain the best estimate of the phase phi k and we need to define what is best, because best in one aspect may not be best in the other aspect and we will see that if there is noise it actually has to be the best estimate you have to change depending on the noise, but we are going to keep all are noises to be additive white Gaussian noise.

Therefore, we do not need to worry about what is the statistics of this noise anything more than requiring the mean of the noise to be known which anyway we have taken to be 0 and the variance which were going to assume to be constant and all these noise samples are assumed to be independent I know with respect to the time.

So, n k will be different from n j, they will not be correlated with respect to each other. So, remember from this equation the objective is to find this phi k and when you find that phi k you call that as the best estimate phi k hat and we are going to see two method of finding this or estimating this phi k in the next module.

Thank you very much.