Modern Digital Communication Techniques

Prof. Suvra Sekhar Das
G. S. Sanyal School of Telecommunication
Indian Institute of Technology, Kharagpur

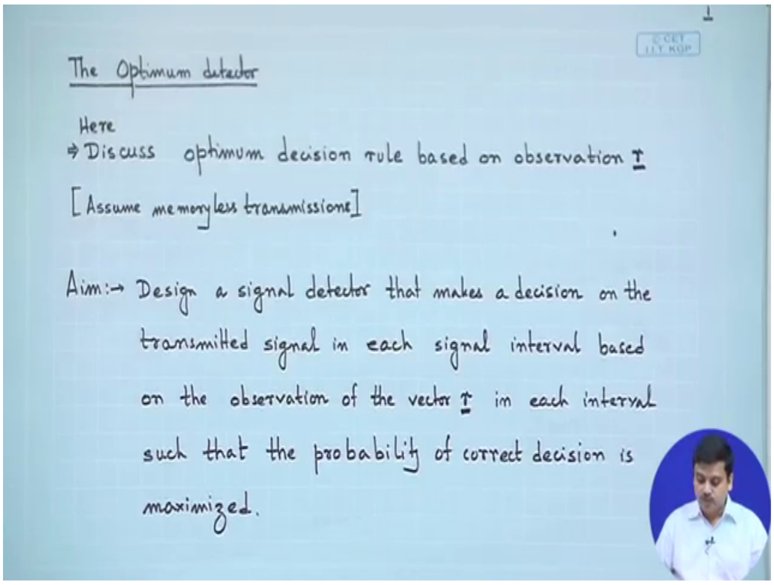
Lecture - 44
Optimum Receivers for AWGN (Contd.)

Welcome to the lectures on Modern Digital Communication Techniques. So, so far we have seen the transmission of signal passing through an AWGN channel as well as the receiver structures, where we have looked at the first part of the receiver; which consists of the demodulator. We are now going to look into the next part of the receiver which will complete AWGN receiver. So, the next part of the receiver will be the detector.

As discussed in the previous lectures that the de modulator could be realized in two possible ways: one is the correlator the other is the matched filter. In the correlator version we use the basis functions on n dimensions. That means the n different dimensions of the transmitted signals and correlates the incoming signals with each of the basis functions. Or in other words we project the signal on to the basis functions and we get the components of the signal on the dimensions.

In case of matched filter we say it is a filter realization and the impulse response is matched to the basis function; where if h t is the impulse response it should be equal to f k of capital T minus t small t- that is the flipped version of the signal. We also saw that matched filter with this kind of an impulse response which is matched to the signal produces maximum signal to noise ratio at the sampling interval of small t equals to capital T; where capital T is the symbol duration.

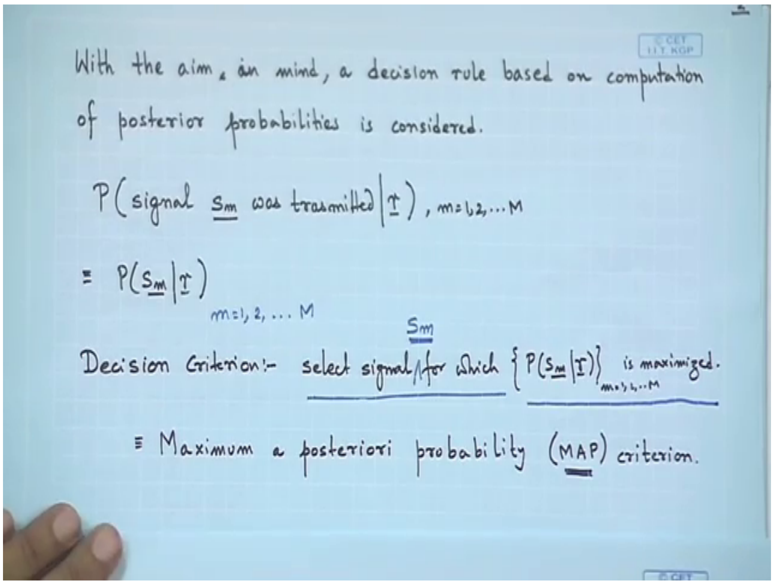
(Refer Slide Time: 02:58)



So, what we have shown is that if you would set the impulse response which is matched to the signal you are going to get the maximum signal to noise ratio in both the cases the received signal is decomposed into its components and we have the vector r as components with r 1 r 2 up to r n. And we explain briefly in the previous lecture that these components consists of the component of the signal as well as component of the noise and we said that since the signal can be decomposed into the directions of the basis functions. So, once we have the components of the single we can reconstruct the signal or we can identify which is the signal.

So, now at the receiver we have the signal components which are corrupted by noise and the job of the detector would be to identify which of the possible m signals do we does this particular vector represent in the best possible way. So, we move on to our discussion on the optimum detector and we said that the optimum decision rule is to be based on the observation of the vector r and for this case we assume memory less transmissions. That means, we do not assume with memory transmissions and our aim is to design a signal detector which makes a decision on the transmitted signal in each signal interval based on the observation of the vector r.

(Refer Slide Time: 03:43)



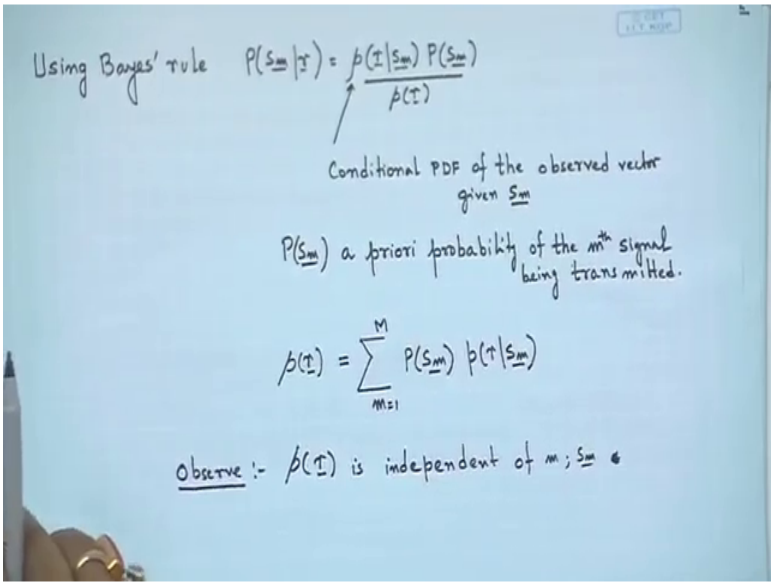
Such that the probability of correct decision is maximized and we explain qualitatively; why do we call it as the decision process, because you are choosing one of the possible outcomes one of the possible waveforms. So, with this we move on and we have the aim as stated that we have to create a decision rule and this is based on the computation of the posterior probabilities now why it is.

So, because we already have observed a particular signal based on the observation we want to know what has been transmitted. So, what was there at the source based on what has happened after passing through the channel that is why we call it the posterior probabilities? So, we basically have to calculate probability of signal S m was transmitted based on r where m could be one 2 up to capital M so; that means, you have to calculate this probabilities for each of the possible signals based on the observation r and. So, we indicate these particular probabilities as P S n given r.

So, this is the notation that we use and the decision criteria is select the signal select the signal S m. So, you have to select the signal S m for which this probability is maximized. So, what we are saying is you have observed r you have to calculate the probability of any of each of these signals given r. So, that is why we have this posterior probabilities what you could have done what you generally have is given a particular signal what is the probability of r. So, we are going in the reverse direction that is why you call posterior probabilities.

So, decision criteria is select the signal S m for which these probabilities are maximized so; that means, we have to calculate these probabilities for m equals to one 2 upto capital M. So, to calculate all these probabilities; so, for that value of m for which this probability is maximum we will be choosing that value of m as the solution. So, this kind of criteria is hence known as the maximum a posteriori probability criteria and this kind of receiver is denoted by the M A P rule maximum a posterior probability maximization criteria. So, this is the rule that is what we are going to follow. So, to proceed with this we have to start looking at what does this particular expression tell us?

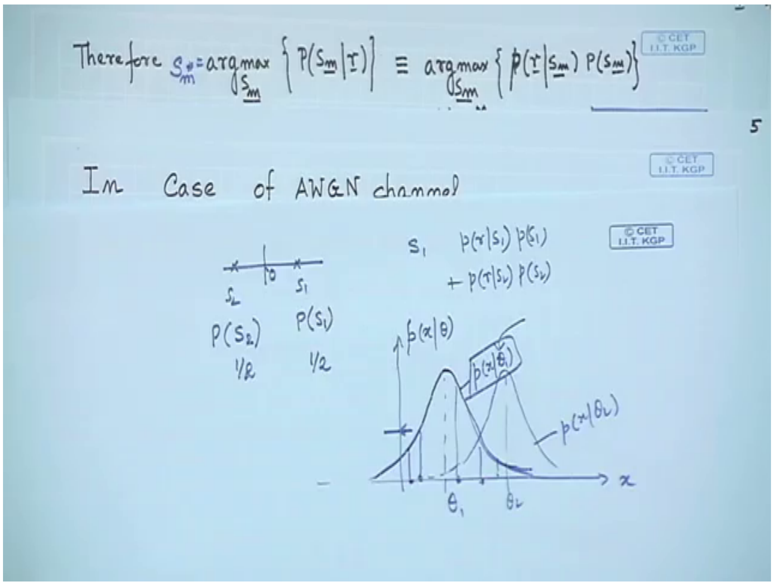
(Refer Slide Time: 06:15)



So, we use the Bayes rule. So, we have on the left hand side P S m given r because r is observed. So, you have to find all possible probe all possible S m find the probability for all possible S m and then choose the best of it. So, to evaluate this we expand this expression using Bayes rule where it P probability of S m given r can be expanded as probability of r given S m multiplied by probability of S n. So, this gives the joint probability of r n S m upon probability of r. So, in another way you could also see P of S m given r times P r is the joint probability of S m and r and this is also joint probability of S m and r. So, that is a very short revision.

So, anyway this is by the Bayes rule we have this expression and where P of r given S m is the conditional P d f of the observed vector given S m and we have calculated this given the expression when we computed the output at the at the output of the matched filter because r is the vector at the output of the matched filter and P S m is the prior probabilities of the m-th signal being transmitted.

(Refer Slide Time: 07:30)



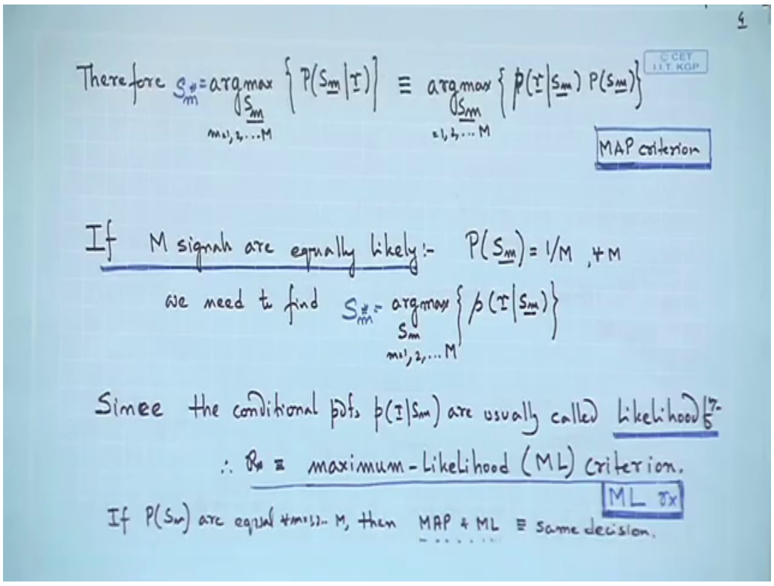
So, if we have a constellation let us say with 2 symbols S 1 and S two. So, all we are saying is that probability of S 1 is P S 1 and you have P S 2 as probability of S 2 now generally as we said earlier if you want the mean to be 0 new to be 0 in one of the earlier discussions we said the constellations have to be symmetrically placed and they have to be chosen with equal probability you can make 0 with some other probabilities also in that case the probabilities would be having non equal values.

So, in either case P of S m is the prior probabilities; that means, in general cases generally this probabilities are equal, but they are could be situations where these are unequal and if you look at P of r P of r could be expanded as P of r given S m multiplied by probability of S m. That means, we are taking suppose you choose S 1 and then you get P of r given S 1 right multiplied by P of S 1 plus P of r given S 2 times P of S 2 so; that means, you have averaged it over all possible transmitted signals. So, that is P of r. So, you could also see that P of r is independent of S m because you are taking over all possible S m you have already averaged over all possible S m right.

So, P of r is not dependent on S m it is just the probability of observing a particular a particular output so; that means, if we have to maximize this probability we can see that the denominator term is almost e relevant for us all we are left with is the numerator term a similar situation we had used when we discussed the properties of matched filter. And there also we had found that if we could hold the denominator constant we could maximize the signal to noise ratio at the output by maximizing the numerator.

So, although these are 2 different context, but we have a similar a situation of operation over here. So, in this expression we do not need to consider the denominator part because it will be common for all S m, so, will concentrate on the numerator part only.

(Refer Slide Time: 09:59)



So, going ahead so, therefore, what we have with us is maximize. So, what we want what we have represented over here is Argmax; Argmax means the argument which maximizes this particular expression. So, the argument which maximizes is we are taking S m as the argument. So, our outcome will be S m star you could say S m star you could also say m star that is the index right. So, we want to choose this. Now, instead of maximizing the whole expansion of this we just concentrate on the numerator part what we discussed over here. So, we are now concentrating only on the numerator part and going by the M A P criteria. So, you have to maximize P of r given S m which is relatively easier to do and multiplied by P of S m right because you have considered the prior probabilities of S m as we just said that it is highly possible that P of S m are equiprobable with this particular example we said that probability of sending S 1 is half probability of sending S 2 is half.

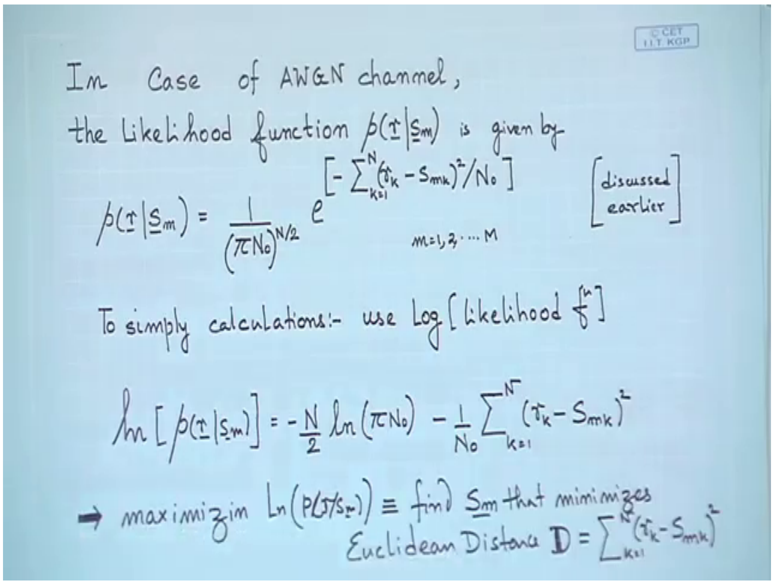
So; that means, they are equiprobable. So, P of S m should be equal to 1 upon 2 if there are m such possible waveforms if there are m such possible signals and if there equal equiprobable then P of S m is 1 upon m and then again you take a look at this particular expression. So, what you will find is that this portion of the expansion is equal to 1 upon m which is not dependent on small m. That means it is not dependent on S m. So, you have to maximize this choose the argument which maximizes this; this part is no longer dependent on S m if the signals are equiprobable. That means, they are equally likely, so, in such a situations if you proceed with the case that if they are equiprobable we would simply put P of S m as 1 over m and we are left with this term.

Again since we are talking about maximization over the argument S m and this is not a function of S m we need not considered this term we can drop this particular term. So, what we have is P of r given S m. So, all we need to find is P of r given S m over all possible S m right. So, whichever maximizes this, whichever is a maximizes this would be our choice of our answer or our solution. So, since the conditional P d f P of r given S m are usually called the likelihood function I can briefly tell why do we call it the likelihood function.

Therefore, the receiver is known as the maximum likelihood criteria receiver. So, also you can write it as the m l receiver right and just at this point I would find it pertinent to remind you that there is a M A P criteria which is distinct will of course, come across this again the difference between the M A P criteria and the m l is clearly visible over here in the m l the maxim likelihood criteria we have all signals with equal probability; and therefore, P of S m is no longer used as you can clearly see in this expression, where is an M A P criteria P of S m is used.

So, if the signals are equal equally likely equiprobable in that case these 2 expressions will be the same would give the same output the expression will not be the same, but they would give the same output so; that means, the M A P and m l would result in the same solution. So, if we know that signals are equiprobable we do not need to use this part we only use this part of the expression. And therefore, we have the m l receiver. So, with this we proceed with our discussion about the m l receiver. So, we have been discussing the AWGN channel.

(Refer Slide Time: 14:26)



So, let us continue with this in case of AWGN channel I opened this part because we have P of r given S m the likelihood function P of r given S m is given by this expression which you have seen before. So, the difference or probably you slight change that you might see is we are using P of vector r given vector S m. Now this consists of P of r k given S m k that is what we had seen before and we had seen we had calculated that P of r k is independent of P of r m; that means, for 2 different components we found that they are independent.

So, the joint distribution we found it as the product of the marginal conditional distributions and therefore, its a product of those distributions if it is the product of n marginal conditional distributions then we have in the denominator pi n naught to the power of n by 2 where for each case it was just a square root and for each case P p of r k S m k we had e to the power of minus r k minus S m k squared by n naught. So, now, since you have multiplied all this r k because of joint distribution in the numerator we have got the summation, because e to the power of minus alpha times e to the power of minus beta times e to the power of minus gamma comes out be e to the power of minus alpha minus gamma minus beta and so on and so forth. So, this is the expression of the P of r vector given S m vector remember we had used underline to denote the vector notation where S m is equal to S m 1 up to S m n.

Similar for r k this is equal to r 1 r 2 up to r n right that is what we had used. So, now, if you look at these expressions this expression is looks a bit cumbersome. So, to reduce the computation what we can do is we can use the log of this function. Now briefly we said that we can discuss the likelihood. So, suppose this is x and we have some P d f. So, this is P of x condition on let us say some parameter theta. So, this is the P d f and let theta with the mean of the P d f suppose.

So, there we have the mean of the P d f. So, if I changed theta to some other value. So, this is P of x condition on theta one to some other value I can get the P d f which goes like this as theta 2 and this can be P of x conditional theta two. So, now, what we have is in this particular thing x is already observed if you look at P of r given S m r is the received vector. So, we already have some of the values. So, we already have some of the values right.

So, when I put it back into this P d f P d f is no longer a probability is it is it can be now read as a function because probability would mean that what is the probability; that means, we are reading on the x on the y axis, but now we say that we have already observed these values that a whole set of values and I want to treat this as a function. So, when I feedback all these values into the function it is now a parameter of theta. So, I can I can have this function of theta of all this average values and my objective would be to find the theta which is the best fit for all these set of values of x that I have got given this kind of a function. So, generally these kind of expressions are known as the likelihood function because they are constructed out of the P d f.

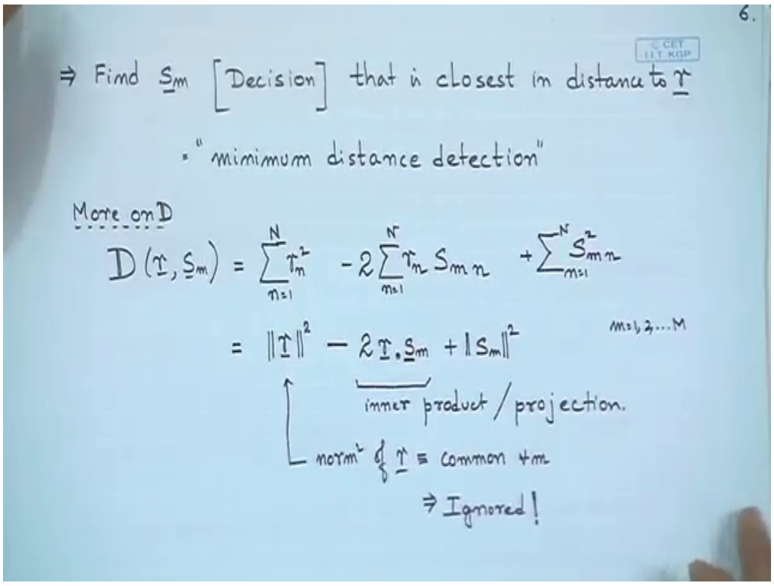
So, because given the function we know that the curve should look like this and therefore, we would like to find the value of theta for a given set of x values which is closest matching to that function. So, that is why it is called the likelihood function and when we take the log of it log of this likelihood function this called the log likelihood function. So, instead of maximizing the likelihood function you can take the log of it because log is a monotonic function. So, maximizing this would give the argument of it that maximizes this would be the same over here. So, taking the log of it natural logarithm you going to get minus n by 2 that is over here in the denominator l n by n naught which is straight forward and since this is e n we are taking a natural logarithm this would come straight down. So, you have expression which is in the exponent of E.

So, now we said we want to maximize this. So, if we have to maximize this; that means, started with maximization of this therefore, maximization of the log likelihood function and now let us look at the log likelihood function in the log likelihood function we again see that this particular part is not influencing our result this particular part is constant for all S m whereas, this part has S m.

So, if we want to maximize this since there is a negative sign and we take this particular part separately it is as good as minimization of this section this part of the expression; that means, if I could minimize this part again since n naught is constant not dependent on m. So, I could take minimization of this particular part right. That means, we could write maximization of the log likelihood function is same as find the vector S m that minimizes this particular expression that we have.

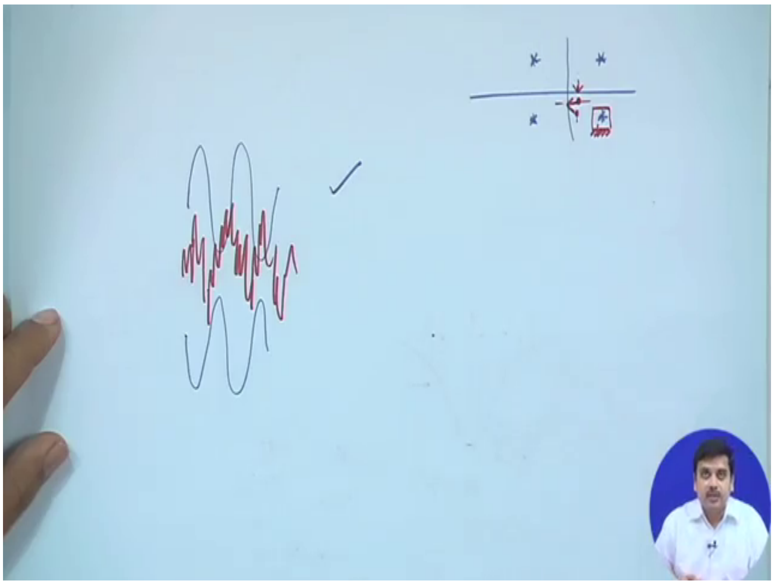
Now, if you take a careful look at what this 2 expression means it is component twice difference squared and some over all components. So, clearly it is the distance of the vector r from the vector S m. So, if we say I want to maximize this which is maximization of this which is produced by minimization of this term. That means minimization of this term is minimization of the distance of the vector r from the vector S m. So, what we are trying to say is that we want to find the vector S m which is closest to the vector r.

(Refer Slide Time: 21:45)



So, that is why we call this particular expression the Euclidean distance and our algorithm or our result is to find S m that is closest in distance to r. And this is could also be claimed as the minimum distance detection accordingly, so you still in the detection problem.

(Refer Slide Time: 22:04)

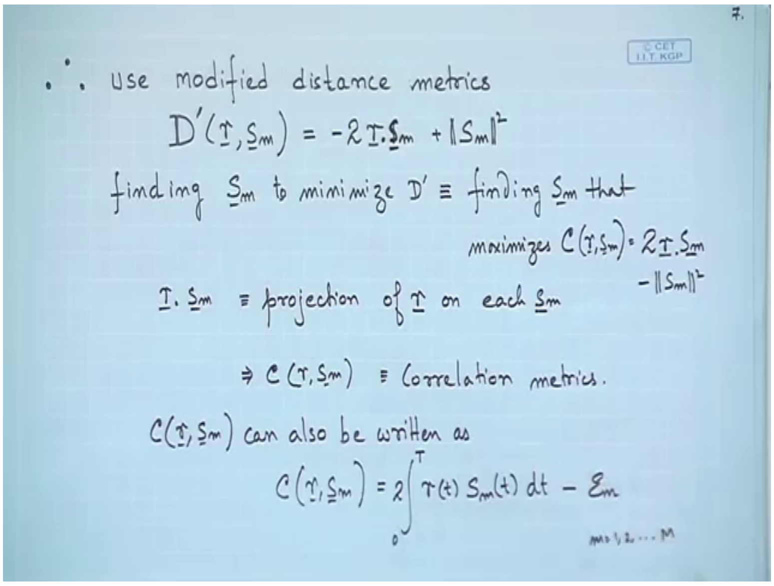


So, we have let us say 4 constellation points and whatever we have received in the vector form is suppose here right and it has produced a component there and it has produced another component which is there if you take this.

So, now you have to find that which of these points are closest to this clearly in this particular case we find that this constellation is closest to this and hence we would choose this as our possible transmitted signal. So, that is what it means right. So, more further moving ahead what we could also see is that this d that we have discussed here that we have discussed here d expression we expand this d expression that is a distance and it turns out to be the length of the vector minus the inner product you could see this as the inner product plus the length of the vector of S m. So, this is written in the vector notation length squared and what you see is that this is common to call S m this is not dependent on S m.

In the same manner we may ignore this while computing the distance component and what is left with us is the projection of r on S m. So, your projecting r on S m and there is the energy of the signal right. So, what we have is projection and energy of the signal which is relevant for us.

(Refer Slide Time: 24:10)

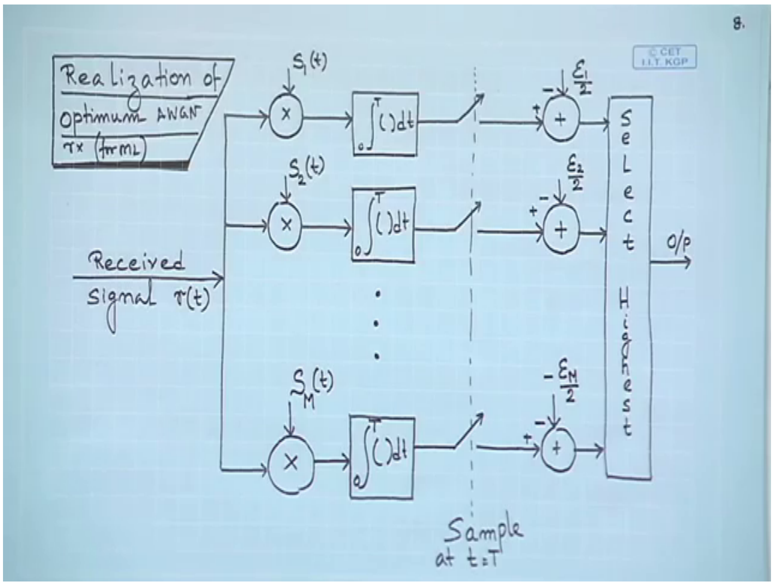


So, since we do not need this we modify the expression of d and presuppose mark it as d prime d prime of r and S m. So, which could be read as minus 2 r dot S m; that means, the inner product and non S m squared; that means, find we need to find S m that minimizes this metric. So, we have to find S m that minimizes d which is same as find the one which minimizes this expression right.

So, d minimization since is a positive term is the same as minimization of the rest of the term because this is not influencing. So, if we look at d prime we could also say that instead of minimizing this we could also write it in by changing of notation of sin we could say you could maximize 2 r dot S m minus non S m squared right because we have simply multiplied by a minus sign. So, instead of minimizing something it is a maximizing the negative of it. Now, if we interpret this, this projection of r on S m this project r on each S m, so, we have to try for different values of m and hence r dot S m can be also read as the correlation metric.

Correlation of r on S m we have to do over each S m. So, the correlation metric in the signal notation form you could write it as whatever we have c over here as 2 times integrate r 0 to capital T r t times S m t that is projection of r on this particular vector takeaway e m which is the energy of the signal right and this particular expression now you could visualize just inform of a block diagrammatic representation similar to the earlier structures that we had seen.

(Refer Slide Time: 26:15)



So, what we can have with us is the received signal will use the expression that is present with us and just for the sake of seeing it on the same screen we have the received signal coming in received signal being split and in the first part it is getting multiplied with S 1 that is what we have r with S 1 and then there is the integrator. So, there is the integrator and then you sample it at small t equals to capital T, because you have to read it up to capital T because this S m is valid for capital T.

So, remember we had always stated that our decision rule is for this interval and our signal S m is also valid for this interval and then I would take away the energy now since we have a 2 over here. So, in this particular realization we do not put a two. So, therefore, when we take away the energy we take away half the energy. So, you could have multiplied here 2 and you could have taken away the energy itself this would have resulted in the same expression. So, just try to follow how we have converted from vector notations and finally, reach the integral form and which can be realized in this way. So, again as you see it is for S m: That means, we have to do for all m and find that m which would maximize the c.

So, we could represent this diagrammatically in the form that the received signal gets correlated with all the different S m right the m possible signals you may note this may be a bit confusing with what we had studied before where we had taken projection of r on the different on the different dimensions of the signal here we are projecting on the different signals themselves directly and we have these components. So, you could do it in the vector form as here or you could do it I mean as here or you could do it in integral form when you read in the integral form you get it in this form and then what you have are the different outcomes of the correlation of the received signal with the different possible waveforms and then I would select that particular outcome which is the maximum. So, which is represent over here as the highest possible signal and that will be the decision output.

So, if this branch produces the maximum output my selection would be the signal S 2 was sent and as we would like to see in the constellation diagram if there is or a waveform diagram suppose I have waveform like this. And let us say I have another waveform like this and what I receive is a noisy version let us say I receive something like this. So, what it tells you is you have to correlate the signal with S 1 and with S 2 whichever correlation give the highest output you choose that particular signal as your desired signal.

So, that is how a detector would work. So, up to this point we have at least taken a complete view of a transmitter when it the signal goes out of the transmitter passes through a channel enters the receiver exits the matched filter of the correlator receiver. And then goes into the detector and you have seen the detailed architecture of how the correlator on matched filter works as well as a detector works.

Thank you.