Stochastic Structural Dynamics

Prof. Dr. C. S. Manohar

Department of Civil Engineering

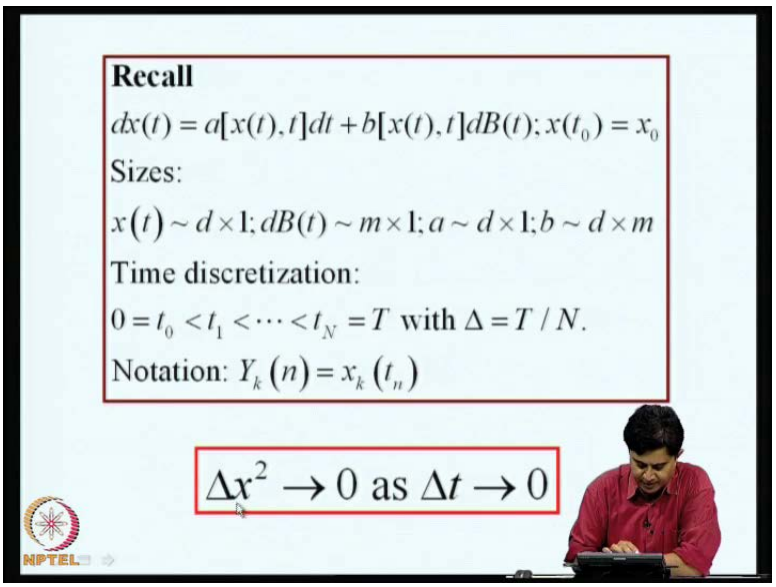
Indian Institute of Science, Bangalore

Module. No # 07

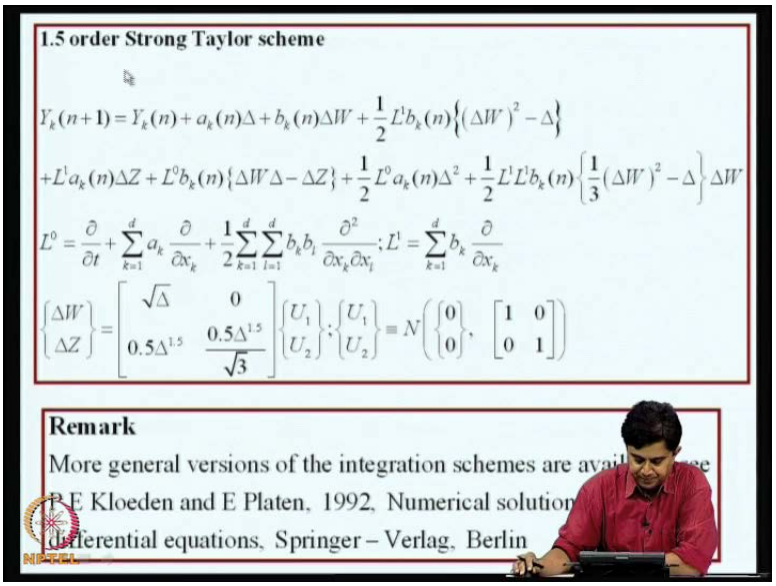
Lecture. No # 30

Monte Carlo Simulation Approach-6

(Refer Slide Time: 00:28)



(Refer Slide Time: 01:00)

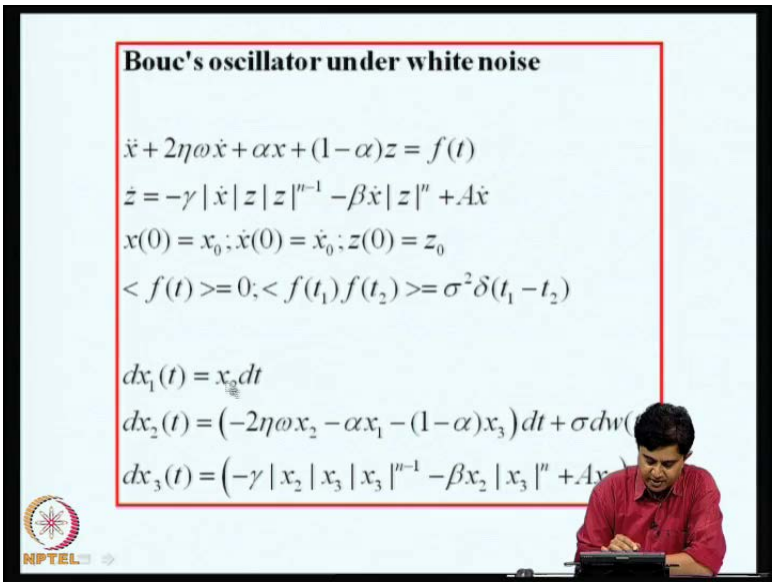


In this lecture, we will continue with our discussion on Monte Carlo simulation methods for analyzing dynamic response of randomly excited and randomly parametered systems. We will quickly recall what we discussed in the previous lecture. We considered systems governed by stochastic differential equations of this form, dx of t is a dt plus b dB t; b of t is the Brownian motion process. And we developed discretization schemes, which respect the peculiarities of sample process of Brownian motion process, where delta x square goes to 0 as delta t goes to 0.

And based on that discretization procedure, this was the one of the discretization procedure that we considered, we tackled a few problems, basically single and 2 degree freedom systems, linear and non-linear, and these integration schemes that I discussed in the previous lecture, is one among a family of such schemes, and more details can be found in the book by Kloeden and E Platen.

But for the purpose of this lecture, I will be restricting attention on this 1.5 order strong Taylor scheme.

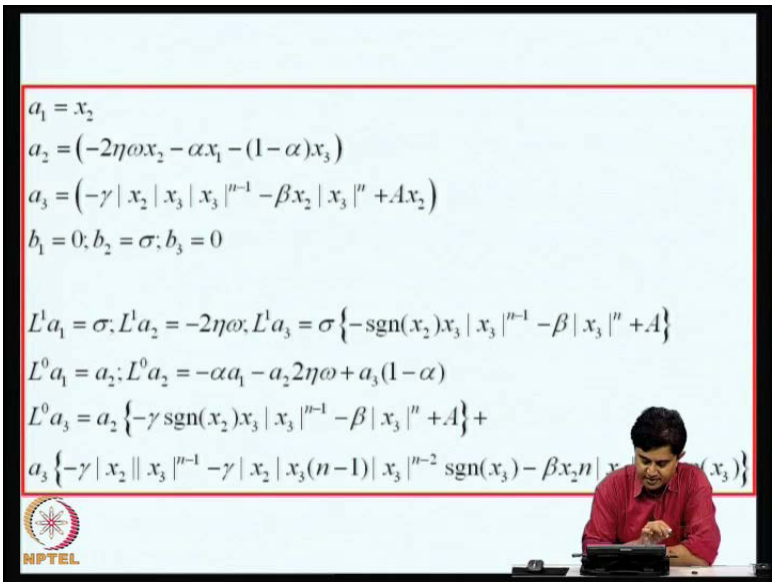
(Refer Slide Time: 01:37)



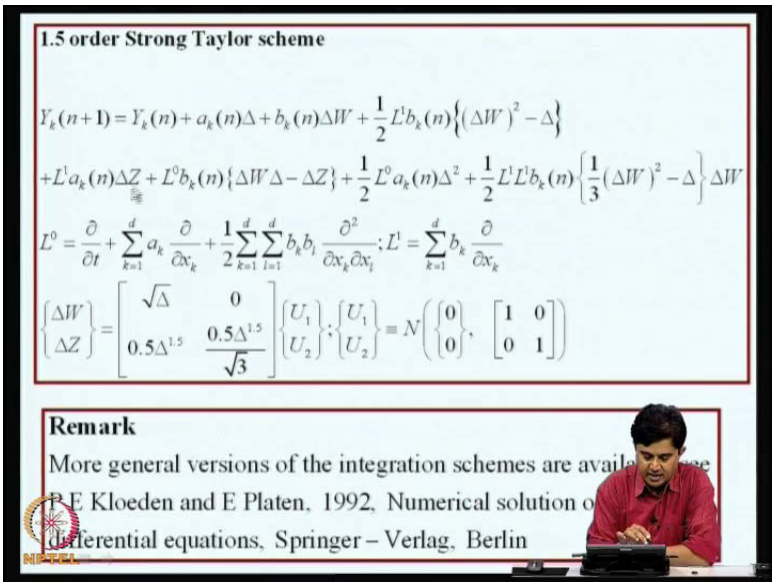
One more example to the set of examples that I considered, we will consider a system in which non-linearity is hereditary and one of the models for that is in terms of, what is known as, Bouc's oscillators. Then here, the non-linearity arises through an internal variable z, which is governed by an additional first order differential equation.

So, the resisting force at any time t depends on the response time history up to that time instant and these are known as elastic systems, typically associated with multi-level non-linearities. So, we will consider, now the problem of response of such an oscillator to a white noise excitation, f of t is 0 mean, stationary white noise excitation. To implement the solution scheme that we developed, we will first convert them in to set of first order differential equations. x 1, x 2, x 3 respectively, denote x, x dot and z and we can recast this governing equation in to this form.

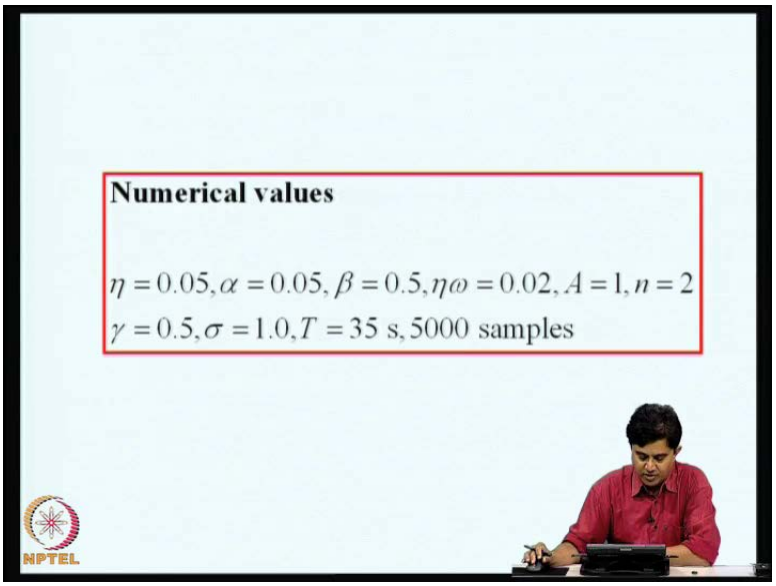
(Refer Slide Time: 02:50)



(Refer Slide Time: 03:06)

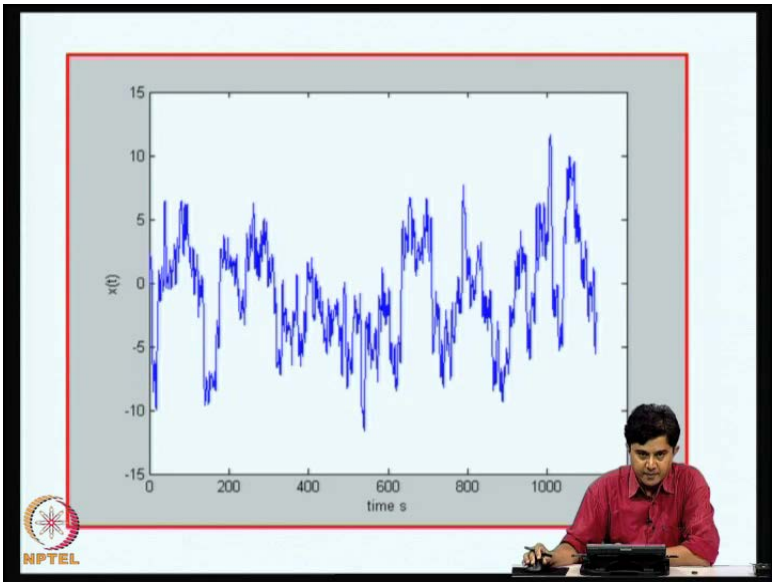


(Refer Slide Time: 03:15)

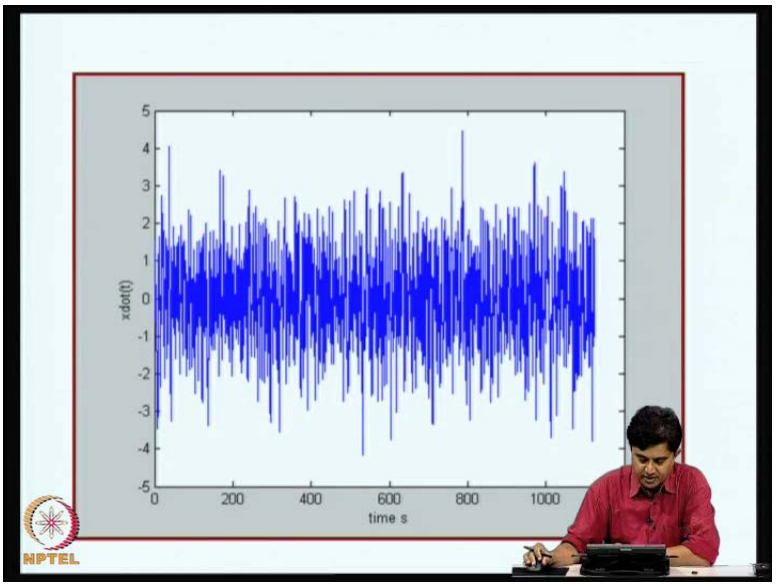


And identify the functions a 1, a 2, a 3, b 1, b 2, b 3, etcetera, in our standard form of stochastic differential equations. And we can compute the terms in the discretization scheme, that is, this operators L, L 1 acting on a, b etcetera, and that leads to a discrete map, which can be used for simulation. And the numerical values that I have used for illustration are displayed in this view graph; 5000 samples have been used for simulation.

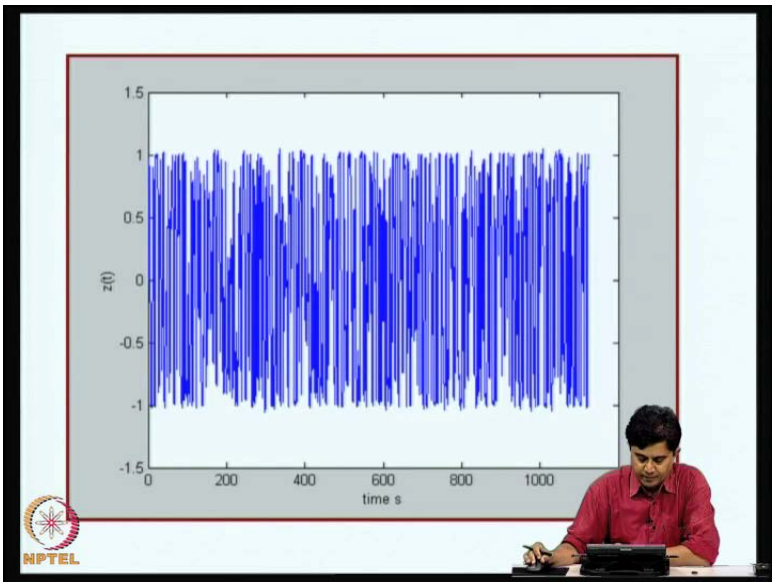
(Refer Slide Time: 03:28)



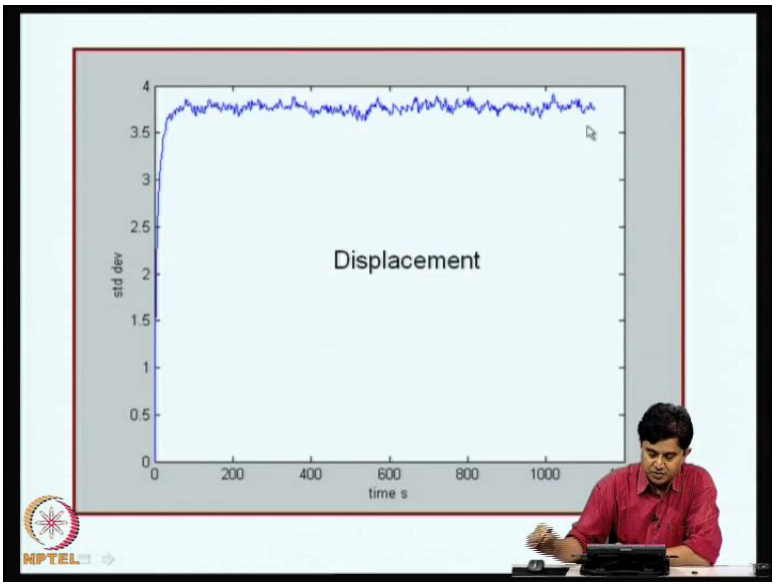
(Refer Slide Time: 03:35)



(Refer Slide Time: 03:37)

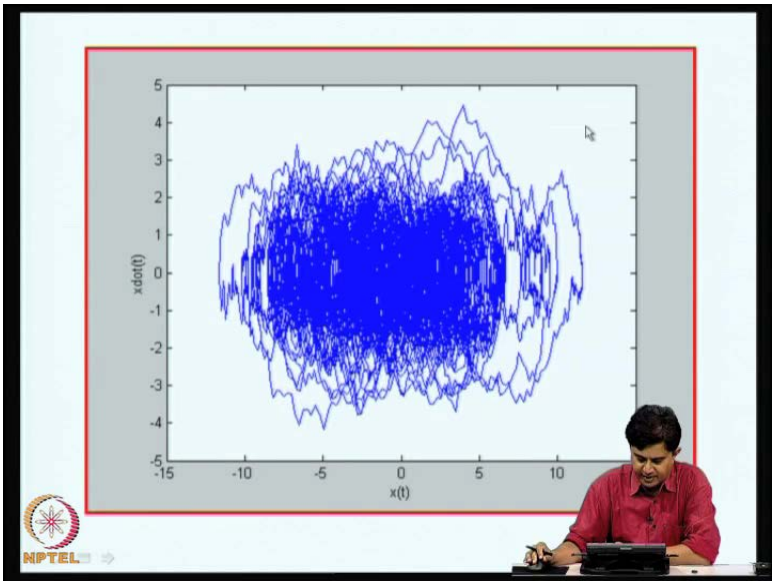


(Refer Slide Time: 03:41)

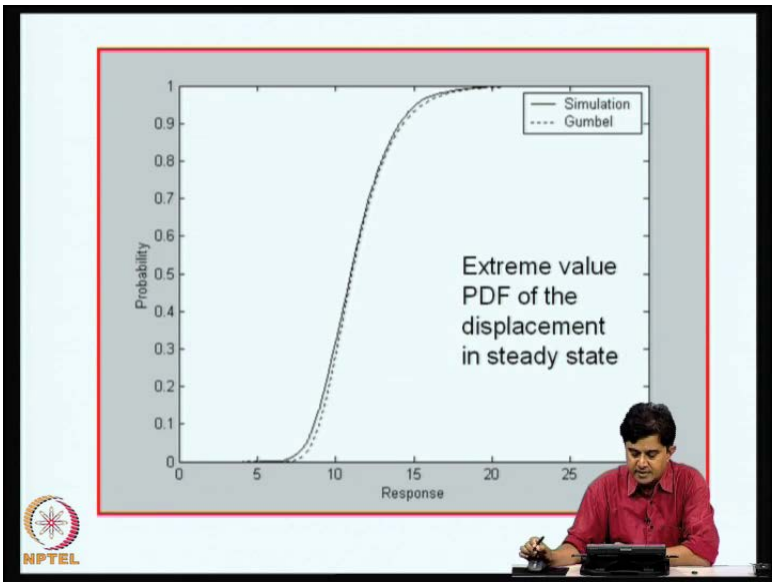


So, this is one of the time histories of response or response displacement; this is velocity; and this is the internal variable z of t; this is the time evolution of standard deviation of the displacement process. In the earlier examples, that I considered, some of the examples were such, that one could determine the steady state response exactly using Fokker Planck equation approach. But the Bouc's oscillator model does not belong to the class of problems, which are amenable for exact solution through Fokker Planck equation approach.

(Refer Slide Time: 04:17)

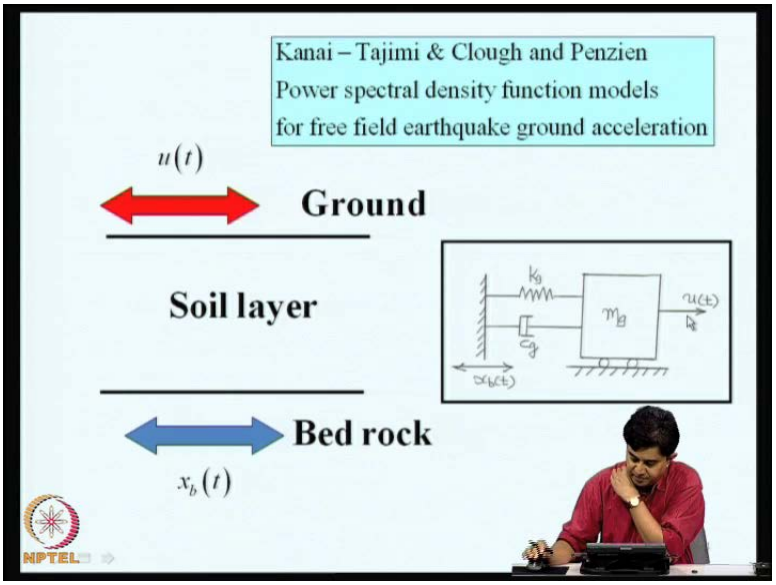


(Refer Slide Time: 04:21)



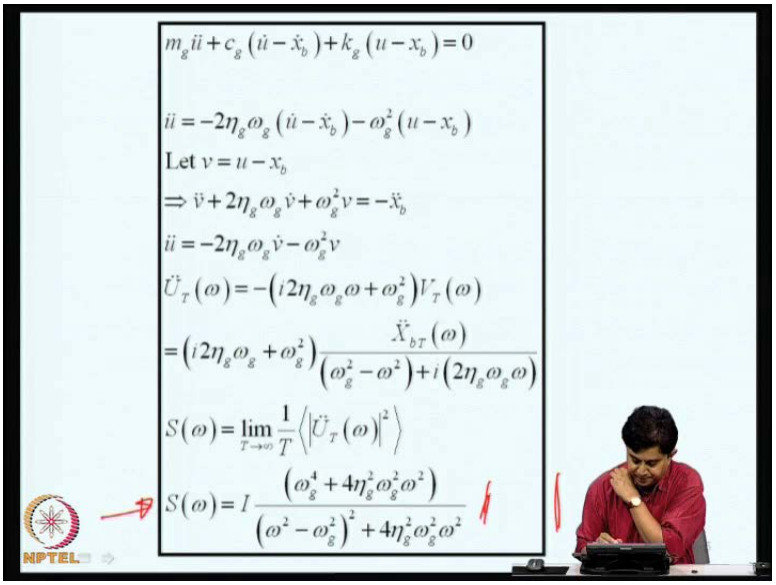
So, therefore, there is no result on analytical solution here. This is the phase plane plot, x versus x dot, and this is the simulation on extreme value distribution with 5000 response, response displacement in steady state, where a period of 35 seconds is simulated using 5000 samples. And the full line is the simulation results and for sake of comparison, a Gumbel model is fitted to these extreme responses, and that is also displayed through a dotted line. There is no reason why they should match, but this is only for purpose of comparison, that the results are shown on the same graph here.

(Refer Slide Time: 05:04)



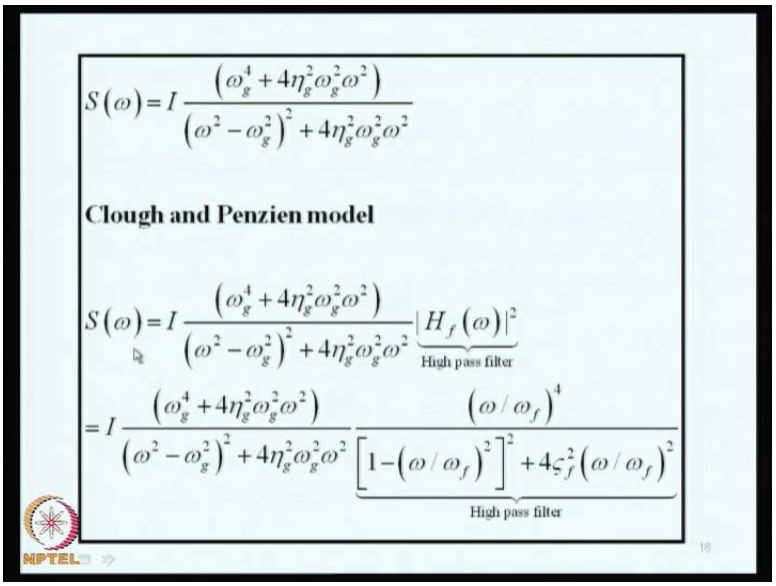
Now, I will consider another example. Here, the simulation, the motivation for considering this example is to demonstrate, how simulation methods can be used to generate ensemble of load time histories? So, for the purpose of illustration, we consider the simulation of earthquake ground motion and one of the simple parametric model, that we discussed earlier, is to consider the ground acceleration at the bed rock level to be a Gaussian white noise. And this excitation at the bedrock level gets filtered through the soil layer and it manifests as ground motion, and our interest is basically to model the ground displacement velocity and acceleration.

(Refer Slide Time: 06:10)



A simple model for that would be to model the soil layer as a single degree (( )) system. As shown here, m g, k g, c g are the ground mass stiffness and damping parameter, this is the base motion and u of t is the absolute response. So, if x b of t is modeled as the random process, we could develop the power spectral density function for the response and in the particular case, when the ground acceleration, the bed rock acceleration modeled as white noise, the absolute acceleration is modeled, the power spectral density of absolute acceleration can be shown to be given by this. This is something that we have already discussed, I am quickly recalling, therefore, it is not necessary for us to get in to all the details at this stage. So, the issue is that the power spectral density function of the ground acceleration is given by this expression.

(Refer Slide Time: 06:53)

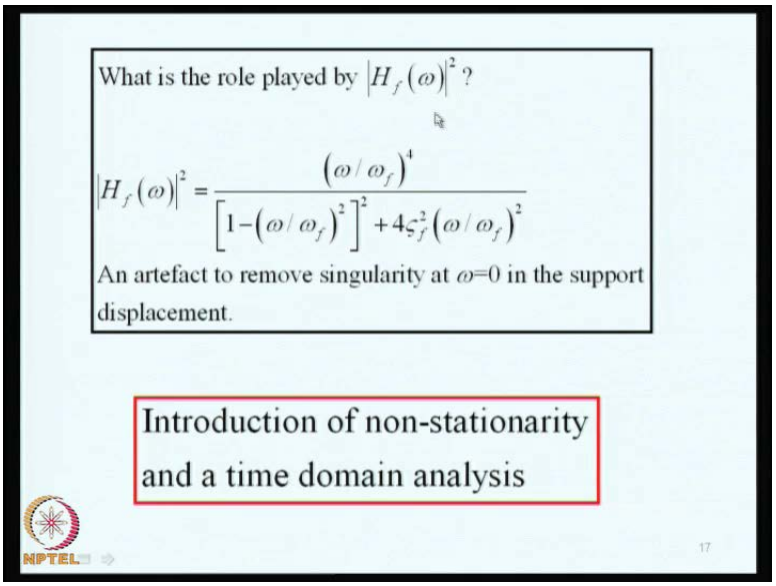


Now, there is a modification, this is the celebrated Kanai-Tajimi power spectral density model and a modification to this was suggested by Clough and Penzien, in which certain problems associated with the Kanai-Tajimi power spectral density function, when one considers ground displacement and velocity, as omega goes to 0, one gets a singularity and the displacement and velocity are not well behaved in a Kanai-Tajimi power spectral density function model.

So, to circumvent this difficulty, Clough and Penzien propose, that this ground acceleration can be passed through a high pass filter, where low frequency components are suppressed and high pass filter is suggested, a high pass filter is, as shown here. This is a model for again, when ground acceleration is a stationary random process.

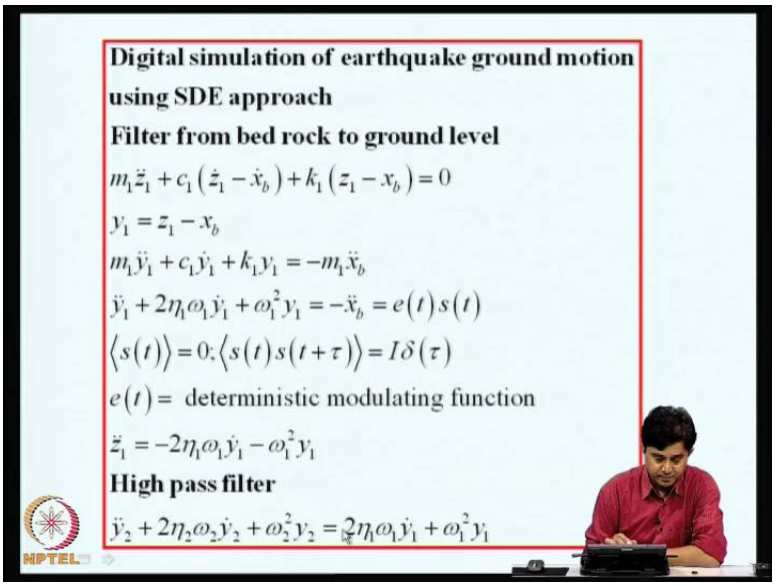
Now, a given power spectral density, we have already discussed how to simulate samples of Gaussian random process, one could utilize the Fourier series approach and obtain samples of the ground acceleration, whose psd - power spectral density - function will be compatible with this target prospect density function models.

(Refer Slide Time: 08:07)



This is the clarification on the high pass filter, this, as, as already mentioned, this is an artefact to remove singularity at omega equal to 0 in the support displacement. Now, what I would like to do is, consider this problem of modeling ground acceleration as basically, as a problem of seismic propagation through ground layer using single degree freedom models, but let us consider the problem of introducing the non-stationarity in the ground acceleration. We know, that earthquake ground accelerations are non-stationary because the earthquake is a transient phenomena, so it is important to see, how we can introduce the non-stationarity into the model and how we could perform a time domain analysis?

(Refer Slide Time: 08:59)

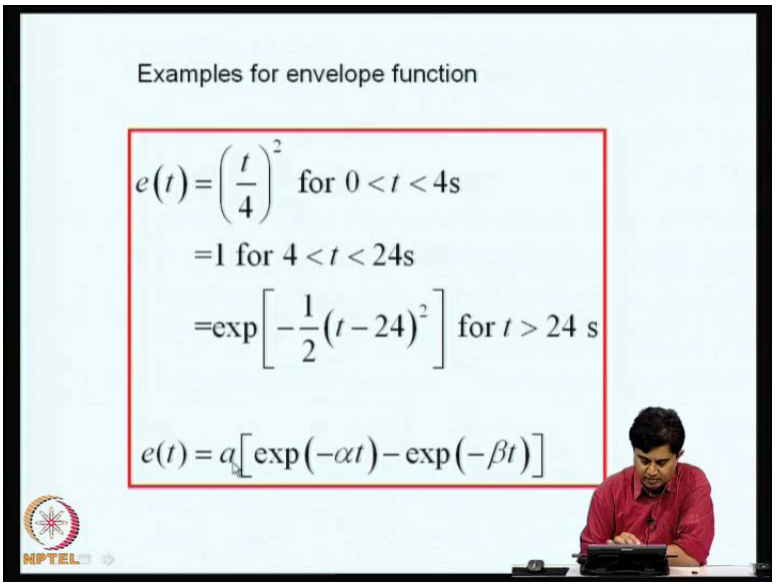


So, this is where we can consider the application of the stochastic differential equation approach to simulate samples of ground accelerations.

So, here, we have used slightly a different notation - x b of t is ground displacement, bed rock level displacement; x b dot is the velocity at the bed rock and z 1 is the absolute ground acceleration, prior to the application of high pass filter. So, we introduce the notation y 1 as z 1 minus x b, where y 1 is the relative displacement of the ground with respect to the bed rock and we can get this equation, where on the right hand side now, we get ground bed rock acceleration.

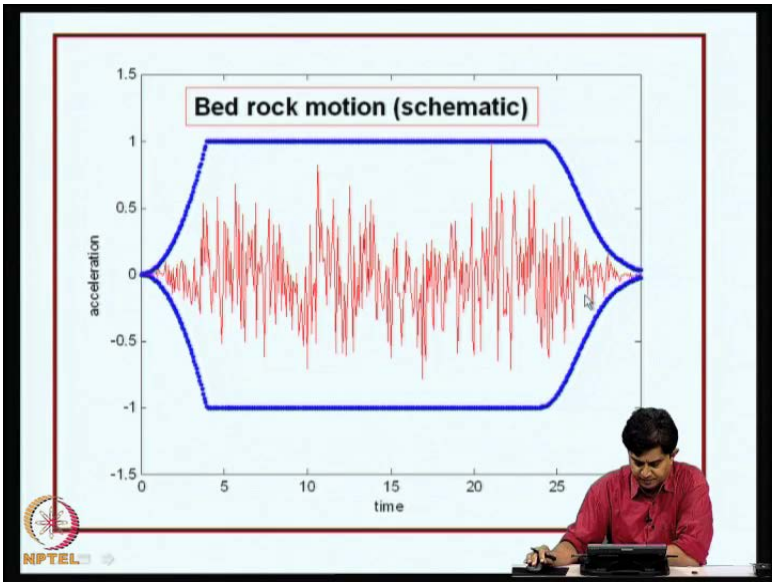
So, one, the upper one dividing by m 1, we get the equation in the standard form as shown here and we could model x b double dot of t as e of t into s of t, where s of t is taken to be a 0 mean stationary white noise random process, Gaussian and e of t is a deterministic modulating function. So, the absolute ground acceleration here can be given by minus 2 eta 1 omega 1 y 1 dot minus omega 1 square y 1; because if you look at this equation, you take the terms associated with damping and stiffness to the right hand side and divided by m 1, we get essentially, this equation.

(Refer Slide Time: 10:38)



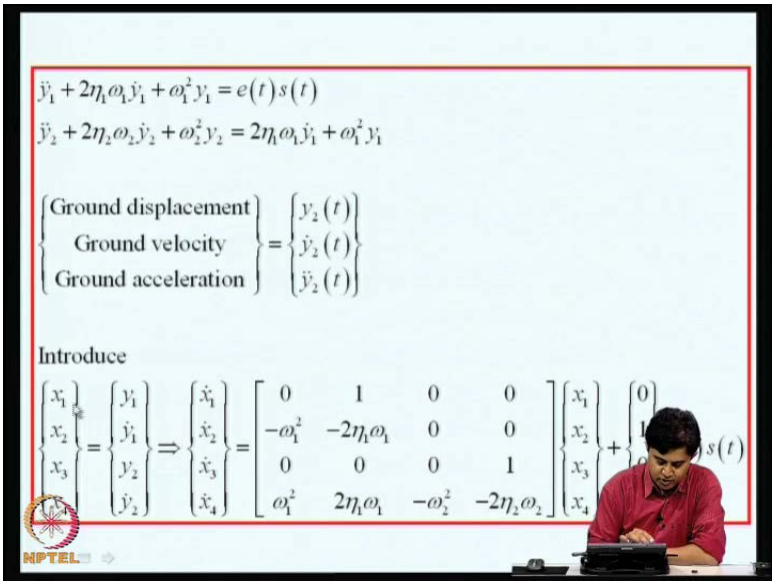
Now, we will pass this acceleration through the high pass filter to remove singularity at the 0th frequency and this is the filter equation for that in the time domain. Now, the envelope that I was mentioning e of t, there are various models proposed, one of the models is as shown here. There is a rise time upto about 4 seconds and the envelope rises, as rises quadratically, and it remains constant for about 20 second, which is a strong motion phase and then there is an exponential decay.

(Refer Slide Time: 11:10)



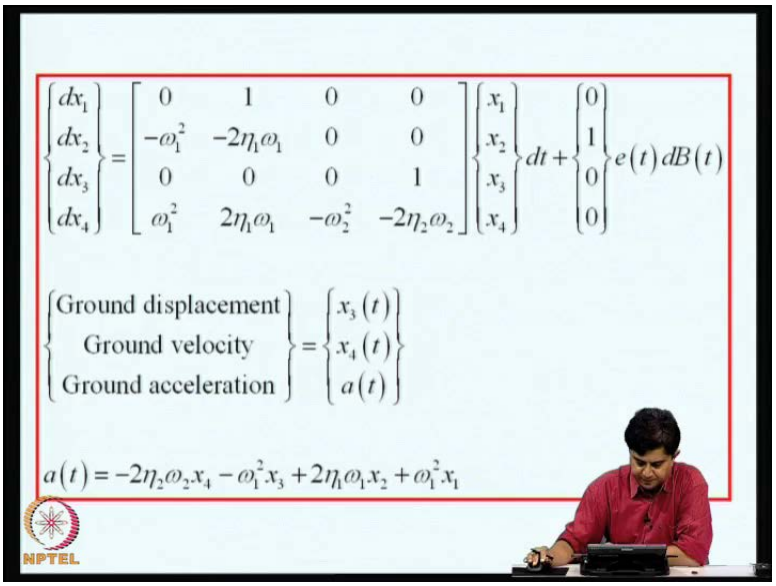
Another model is a difference of 2 exponential functions, which can also be used as a model for e of t. So, schematically, for example if you adopt the first model, the blue line here is envelope, this plus e of t and this is minus e of t and this is the bed rock level ground acceleration, which is now modulated by this envelope. This is, please note that this is schematic, because at the bed rock level we are modeling the ground acceleration is a white noise, and white noise does not truly exist. So, it is not strictly correct to show a sample of a white noise, which is continuous in time because such samples do not exist; this is the schematic representation.

(Refer Slide Time: 11:50)



Now, what we will do is we will consider these 2 differential equations and recast them into the form of a stochastic differential equation. So, the ground displacement, ground velocity and ground acceleration after filtering, that is, after it passes through a high pass filter is given by y 2 y 2 dot y 2 double dot. So, our objective is to simulate samples of ground displacement, velocity and acceleration when the ground bed rock acceleration is a modulated white noise process. So, the envelope, the non-stationarity is introduced through the envelope e of t.

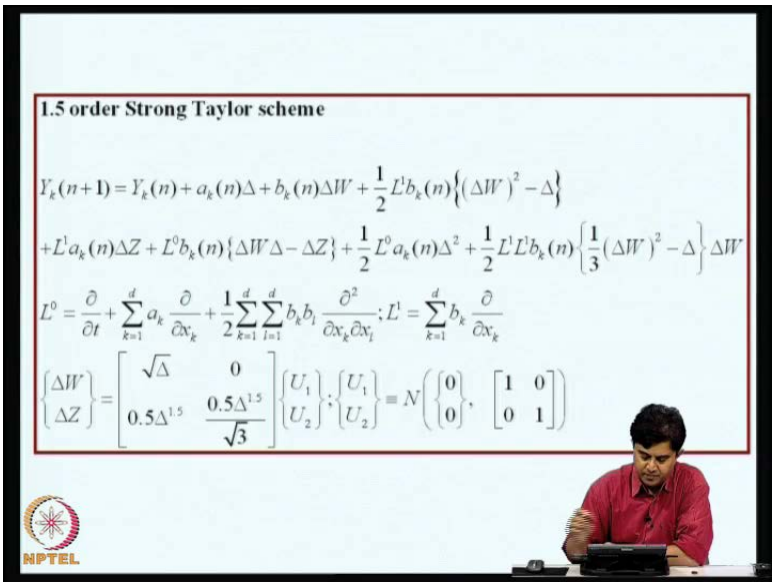
(Refer Slide Time: 12:33)



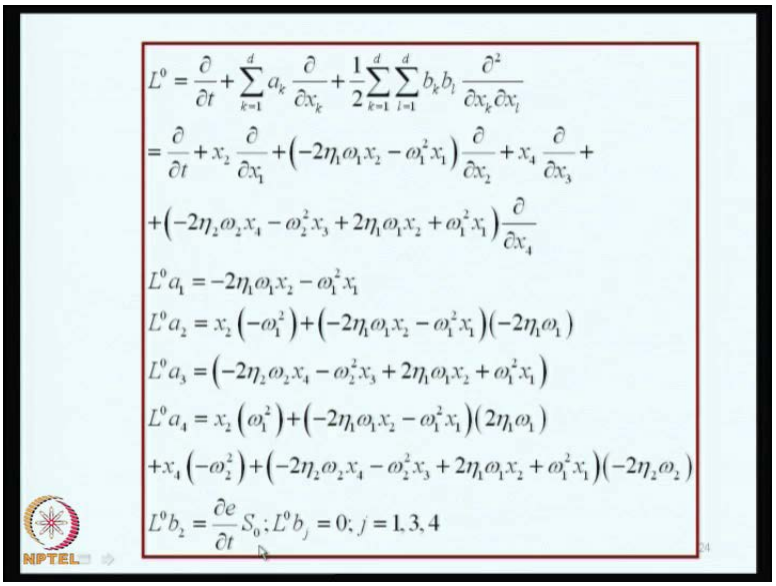
So, by using a vector x 1, x 2, x 3, x 4 as y 1, y 1 dot, y 2, y 2 dot, we can get the equation in the stretch phase form and this can be further simplified, further recast into the SDE - stochastic differential equation – form, in terms of increments of Brownian motion process. There is only 1 white noise driving component here, therefore the b vector is simply 0 1 0 0 and this, a, is the matrix, this matrix into x 1, x 2, x 3, x 4.

So, the, here the stochastic differential equation is linear and we are interested in ground displacement, which is x 3; ground velocity, which is x 4; and ground acceleration a of t, which is given in terms of x 1 x 2 x 3 x 4, through this relation. It should be noted, that we can actually perform a random vibration analysis of this system and get exact solutions for moments of x 1 x 2 x 3 x 4 say, variation of covariance etcetera, but the objective of the present discussion is, how to simulate samples of ground displacement, velocity and acceleration? Because these simulated samples can further be used for analyzing non-linear dynamical systems, and so on and so forth.

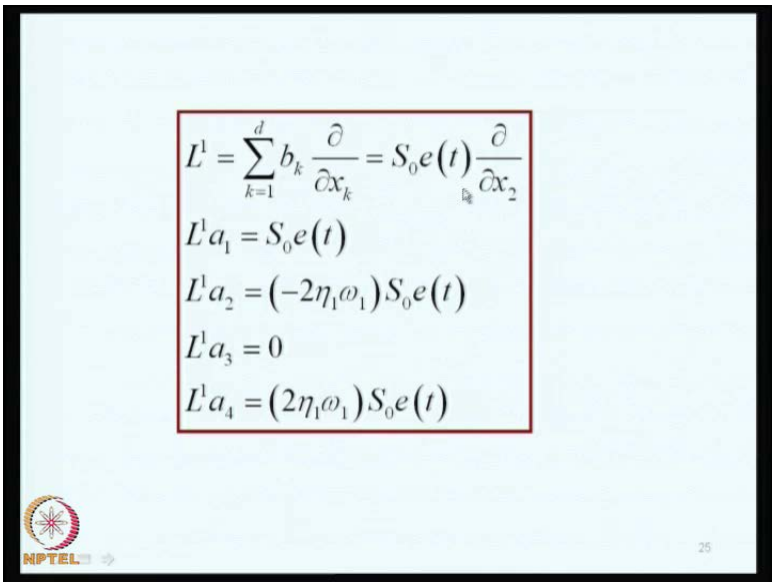
(Refer Slide Time: 13:52)



(Refer Slide Time: 14:03)

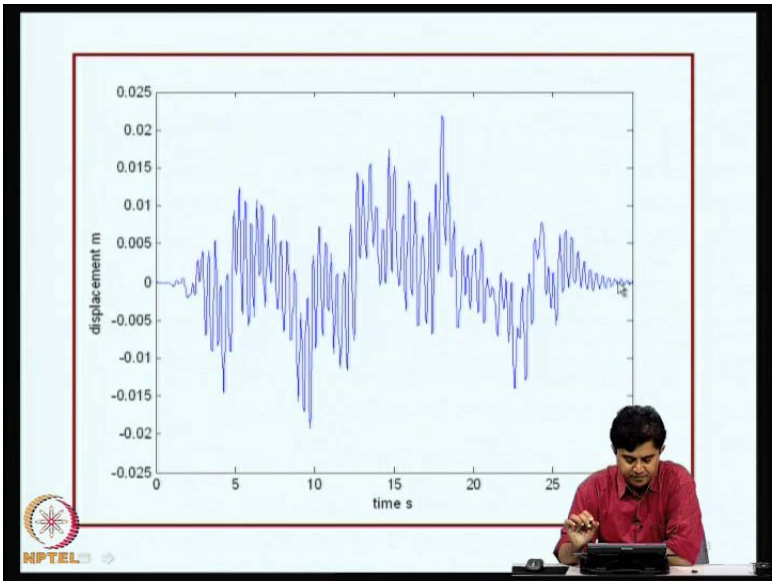


(Refer Slide Time: 14:22)

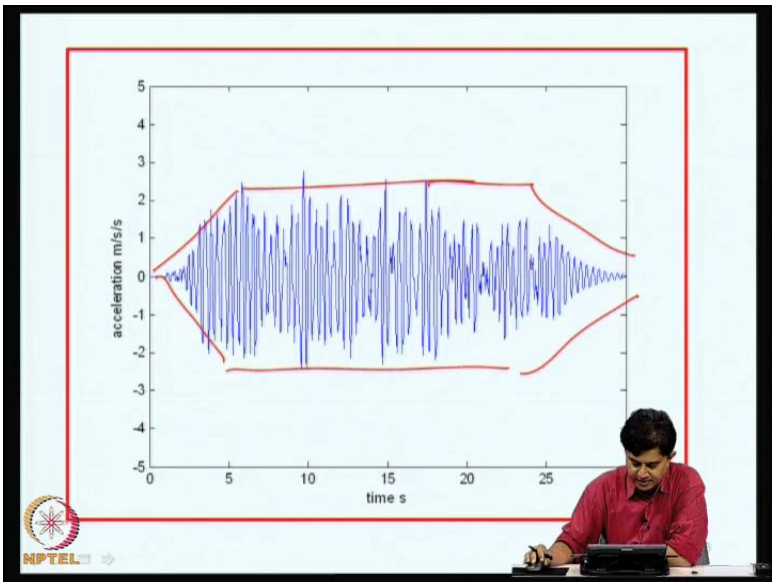


So, that is the objective here. So, we used the, the 1.5 order strong Taylor scheme. I am just flashing this to show the various operators - this is L naught, which is dou by dou t x 2, etcetera, as shown here and if L naught acts on a 1, a 2, a 3, a 4 for this problem, I have shown here the details of the various terms and similarly, L naught b 2 and L naught b j etcetera, and the operator L 1 is given by this and it acts on a 1, a 2, a 3, a 4 to produce these terms. And based on this, we can now go back to this map and write a simple program to implement the recursive relation.

(Refer Slide Time: 14:42)



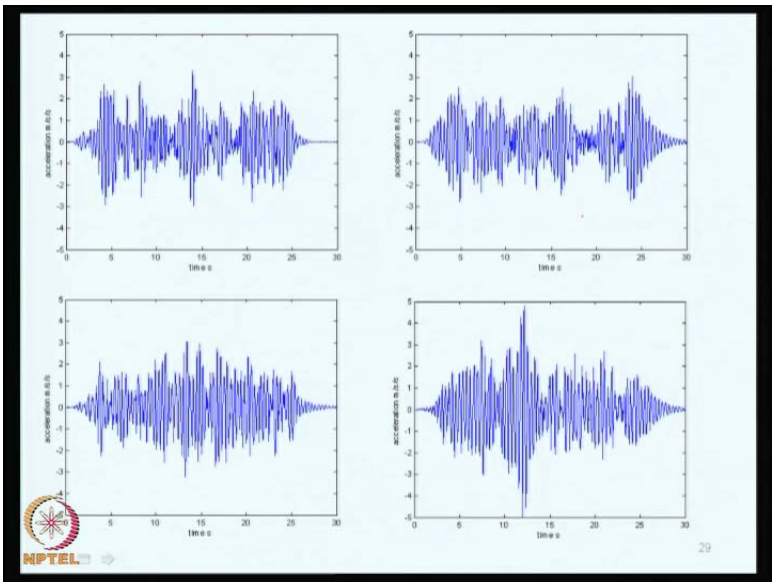
(Refer Slide Time: 14:49)



And I show here a few samples. This is the sample of ground displacement; this is a, this, a sample of ground velocity; and this is a sample of ground acceleration. So, you can see here, the non-stationarity trend of the one that we were trying to implement is captured here and you can also see that this is predominantly, a narrow band process, and that is to be expected because a white noise is being passed through a single (( )) system. And this is acceleration, this is velocity and this is displacement.

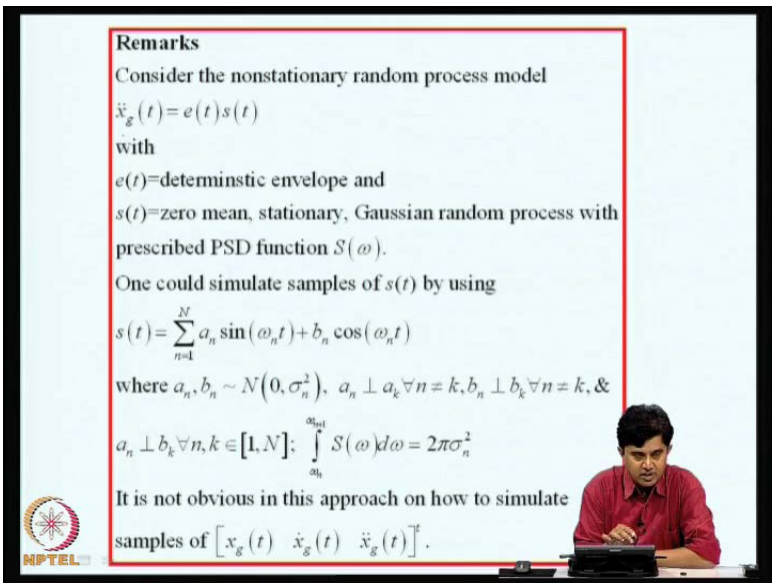
So, this algorithm is quite helpful in simulating simultaneously, the 3 required components.

(Refer Slide Time: 15:30)



So, this is a few more samples of ground acceleration. Of course, one can do an analysis of this ensemble and determine the properties of moments of this ensemble, but that is not the objective of the present study as I mentioned, because we are interested in using these sample ground acceleration in further analysis, that is the objective of this illustration.

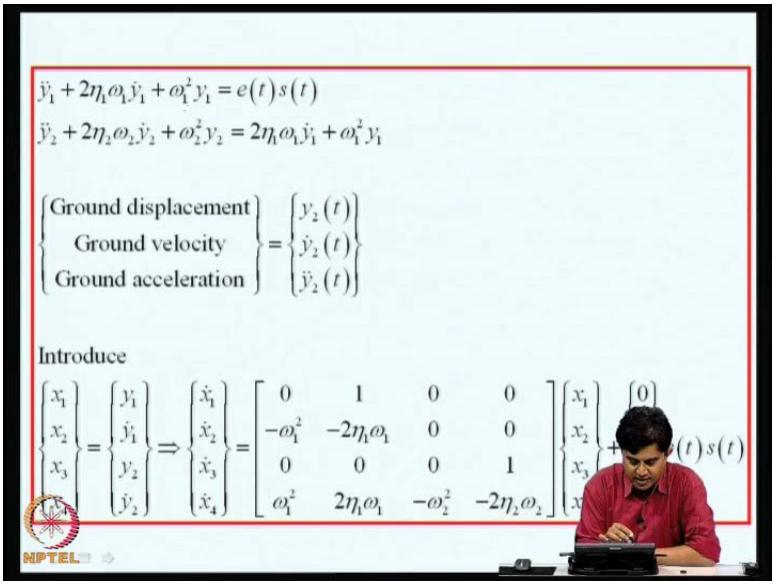
(Refer Slide Time: 16:00)



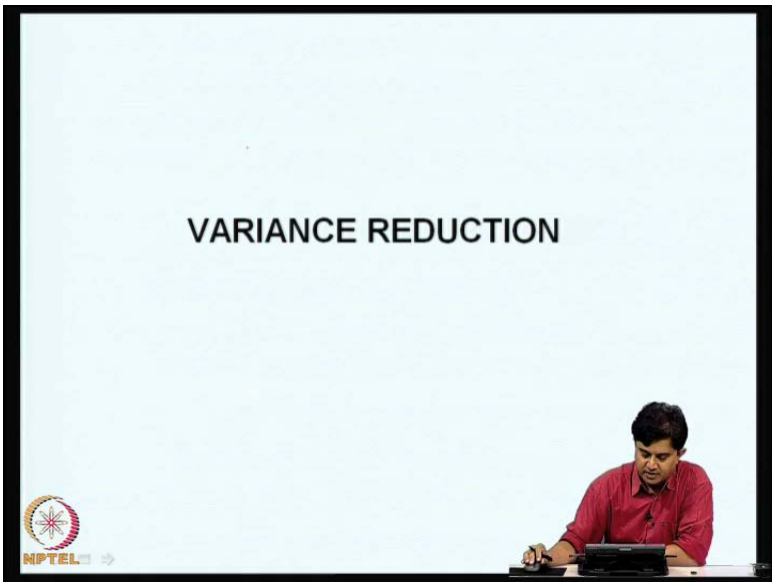
Now, just I would like to emphasize 1 important factor here, suppose we consider x g double dot of t to be e of t into s of t, this is one of the proposed model for introducing

non-stationarity, where s of t is a stationary random process with a given power spectral density function. So, this is an alternate mode, alternative model, small difference from one that I just now described. This is what is widely used in literature, but there is a problem with this in the sense, if you are interested in, only in ground acceleration, this is alright, but if you would like to get ground displacement and velocity also, which is needed, for example in analysis of multi-supports, multiple supported structure, as we saw in one, some of the earlier lectures, there is no obvious way that we can evaluate the velocity and displacement compatible with this non-stationarity in the ground acceleration.

(Refer Slide Time: 16:52)

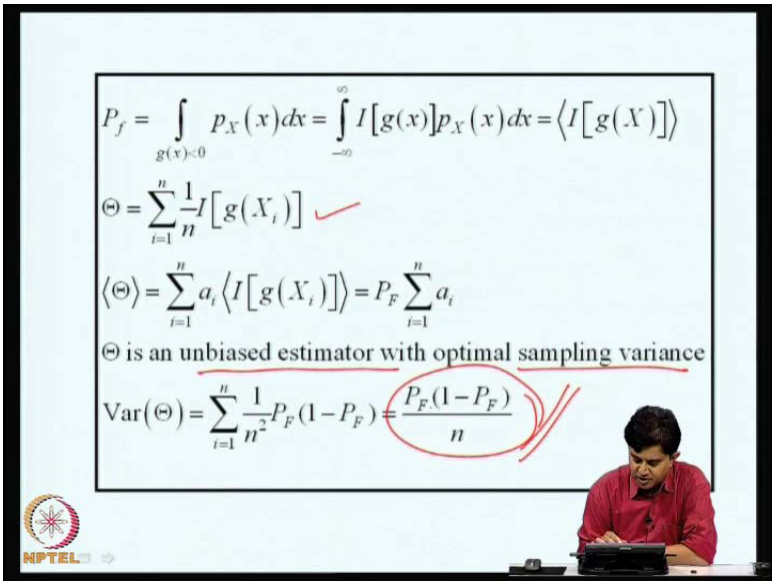


(Refer Slide Time: 17:16)



So, this difficulty is overcome in this approach, where we are simulating simultaneously, the displacement, velocity and acceleration in the time domain. Now, I move on to an important aspect of simulation procedures namely, the problem of variance reduction. In one of the earlier lectures, I have briefly touched upon this; now, let us see in slightly greater detail, what are the, what is the meaning of this problem?

(Refer Slide Time: 17:37)

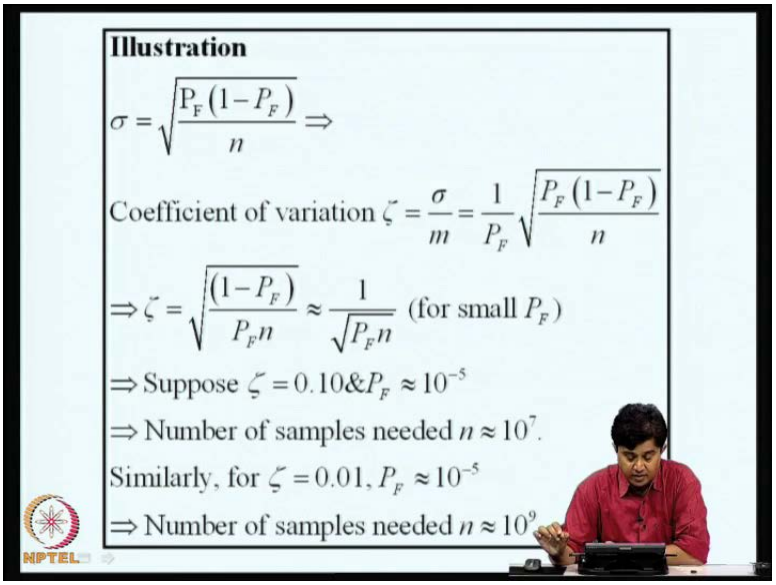


Let us consider the problem of evaluating the probability of failure with respect to performance function and this is written as, an integral minus infinity to infinity I of g of

X P x of x dx, where I g of X is an indicator function, which is equal to 1 if g of X is less than 0, or it is 0 otherwise.

So, the probability of failure therefore, can be thought of as an expected value of the indicator of g of X. Consequently, we can get an estimator of the formula I equal to 1 to n 1 by n I g of X i, where X i are simulated samples from P x of x. Now, this estimator, we have already discussed the properties of this, we have shown, that this is an unbiased estimator with minimum sampling variance and the optimal sampling variance is given by this quantity. This we have already seen in earlier lectures.

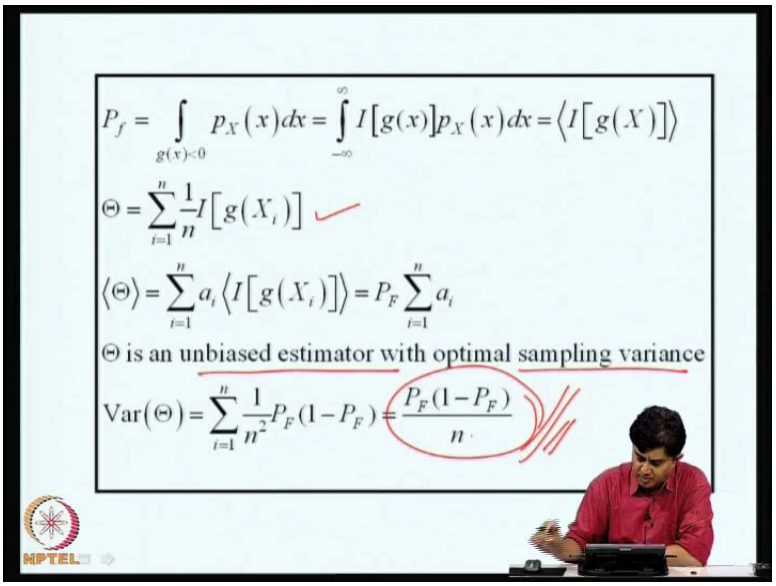
(Refer Slide Time: 18:49)



Now, our attention is to be now focused on variance of the estimator, which is P F into (1 minus P F) by n. Now, let us look at the implications: the standard deviation, the associated with the sampling variance, the square root of the sampling variance, which is given by this and if we define now, a coefficient of variation, where P F is a mean, we get this as approximately as, 1 by square root of P F of n, especially when P F is small.

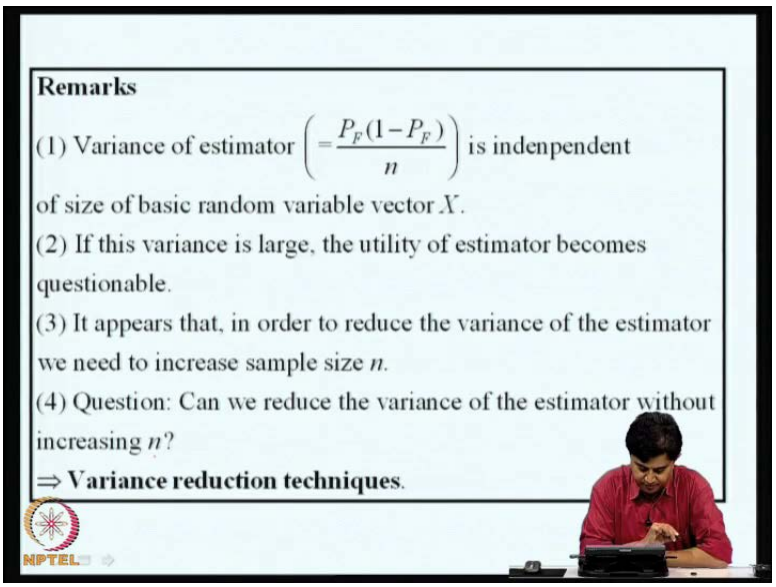
So, if coefficient of variation is 0.1 and P F is of the order of the 10 to the power of minus 5, we can compute from this the number of samples needed for simulation and this transfer to be about 10 to the power of 7. Similarly, if you look at coefficient of variation of point naught 1, that same level of probability of failure, the samples now become, number of samples needed to achieve this coefficient of variation is 10 to the power of 9.

(Refer Slide Time: 19:48)



Now, in a good simulation work, we do not want the sampling variance to be very high. So, our objective is to use sample size sufficiently large, so that this variance is small enough. If variance is large, the estimator will not be utility of the estimator, will be compromise. Now, if you examine this expression, it is immediately evident, that to reduce the variance we are required to, reduce the, increase the sample size.

(Refer Slide Time: 20:17)

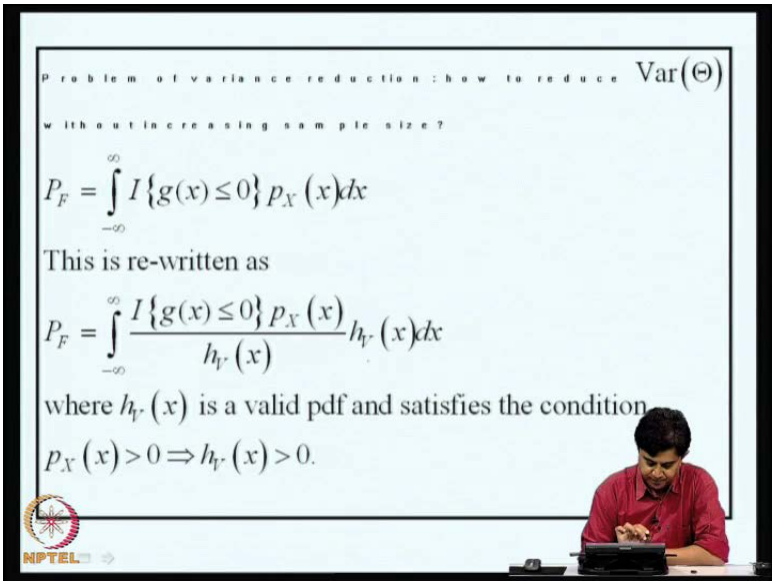


But the question that we consider now is, can we reduce the variance of this estimator without increasing the sample size? Before we get into that, we can observe certain

features. This variance, that is, P F into (1 minus P F) by n is independent of the size of basic random variable vector x. This n is a sample size, but x is, can be any dimensional, I mean, it can be large dimensional.

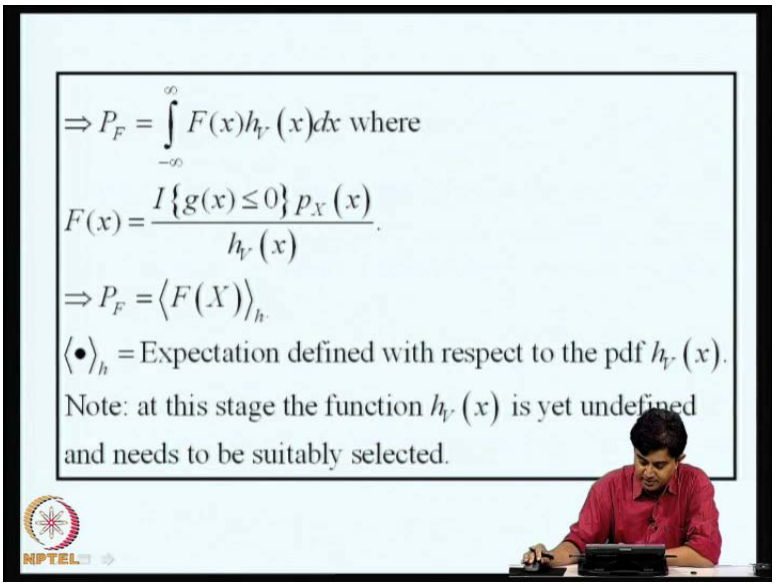
Now, if this variance is large, the utility of the estimator becomes questionable. Therefore, it appears that in order to reduce the variance of the estimator, we need to increase the sample size n. The question now is therefore, can we reduce the variance of the estimator without increasing n? So, the techniques, that, which answer this, this question are known as variance reduction techniques.

(Refer Slide Time: 21:09)



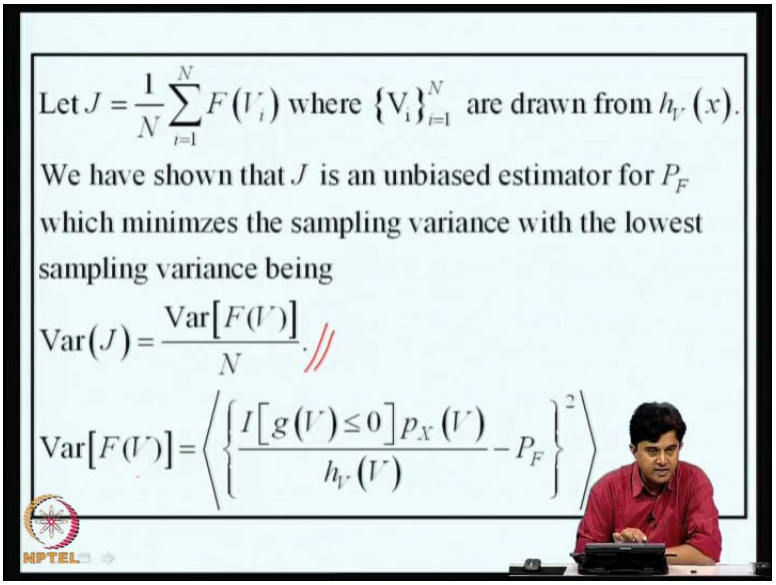
Now, we can consider this problem of variance reduction by again considering this integral P F minus infinity to infinity I of g x P x of x dx. What we will do is, now we will multiply and divide by h V of x, where h V of x is a valid probability density function. It satisfies the requirement, that P x of x is greater than 0 h V of x, is also greater than 0.

(Refer Slide Time: 21:38)



Now, if we now look at probability of failure, we can rewrite the integral integrant in the slightly different form. I will now call this indicator g of x into P x of x by h V of x is F of x and I will interpret P F as an expected value of F of x with respect to the pdf h. Therefore, now I can estimate P F by drawing samples from h. This h is known as the important sampling density function, we will come to that shortly.

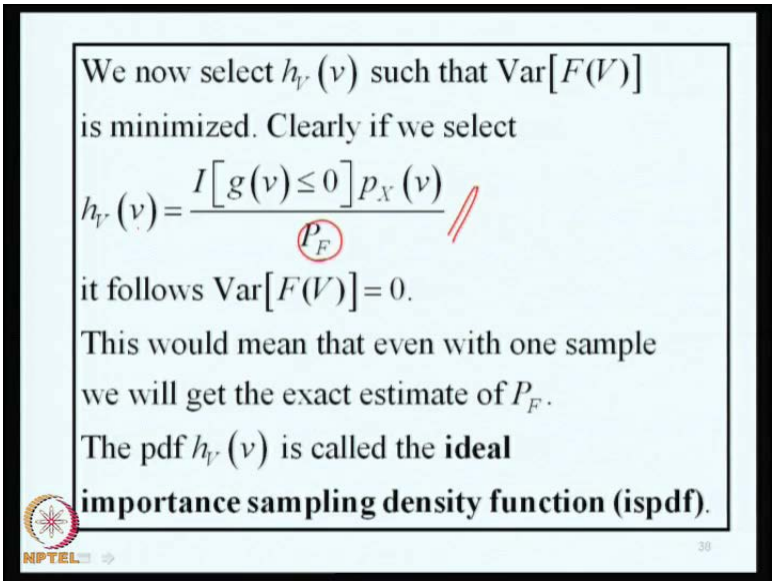
(Refer Slide Time: 22:15)



So, if we now take J to be 1 by N i equal to 1 to N F of V I, where V i is drawn from h v of x. Again, J will be an unbalanced estimator, which minimizes the sampling variance with the lowest sampling variance given by this.

Now, what is the variance of F of V? It will now involve the function h V of x, which we have not yet chosen. So, now, we can select this h V of x, so that this variance is minimized.

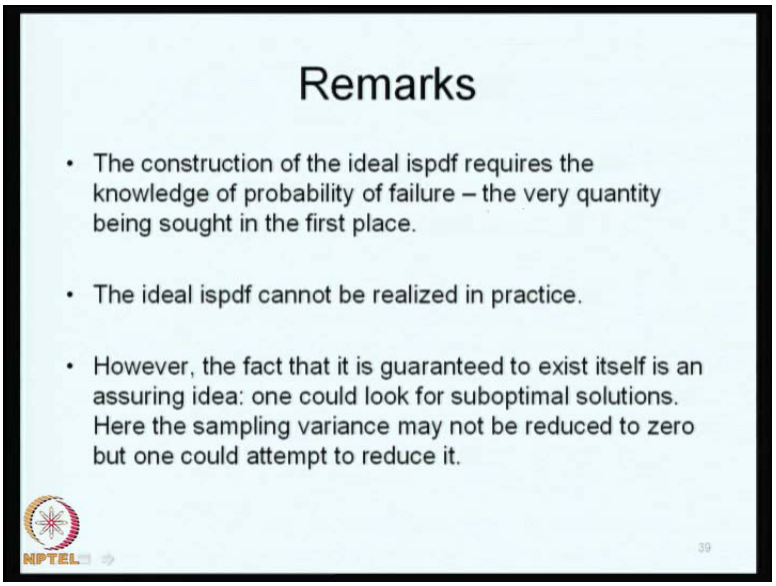
(Refer Slide Time: 22:48)



Indeed, an exact solution to this problem is possible. We can verify that if h V of v is taken to be given by this function, there, in terms of that variance, it becomes 0. This would mean that even with 1 sample, we can get the exact estimate of P F. This looks amazing, but the mystery will be resolved if you notice, that the definition of h V of x involves P F, which is the very quantity that we are trying to find out.

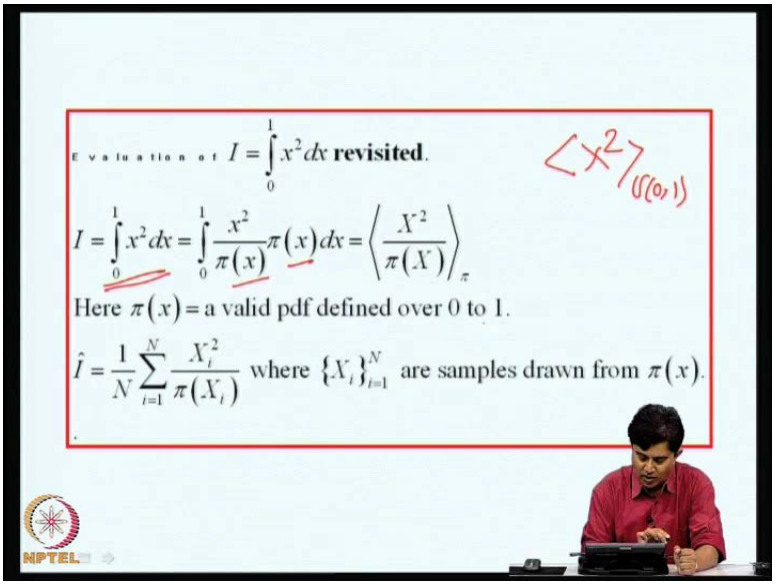
So, in that sense this is an ideal importance sampling density function. If you know P F, there is no reason why we should do simulations at all. But we will see, what is implication?

(Refer Slide Time: 23:32)



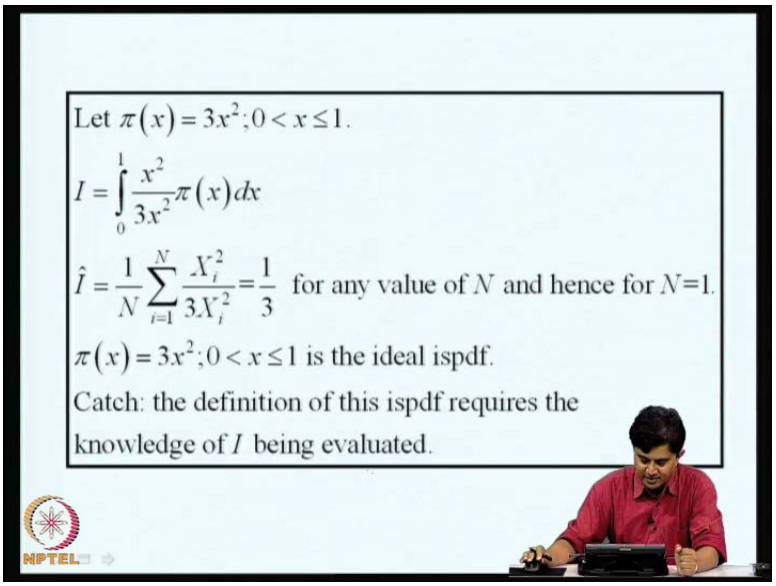
We can make some remarks now. The construction of the ideal ispdf requires the knowledge of probability of failure - the very quantity being sought in the first place. Therefore, the ideal ispdf cannot be realized in practice. If you know P F there is no reason why you will do Monte Carlo simulation, but the significance of these results lies in the fact, that it is guaranteed, the important sampling density function is guaranteed to exist. You know, this fact in itself is an assuring idea; one could look for suboptimal solutions, which may not be ideal, but they can still reduce the variance.

(Refer Slide Time: 24:23)



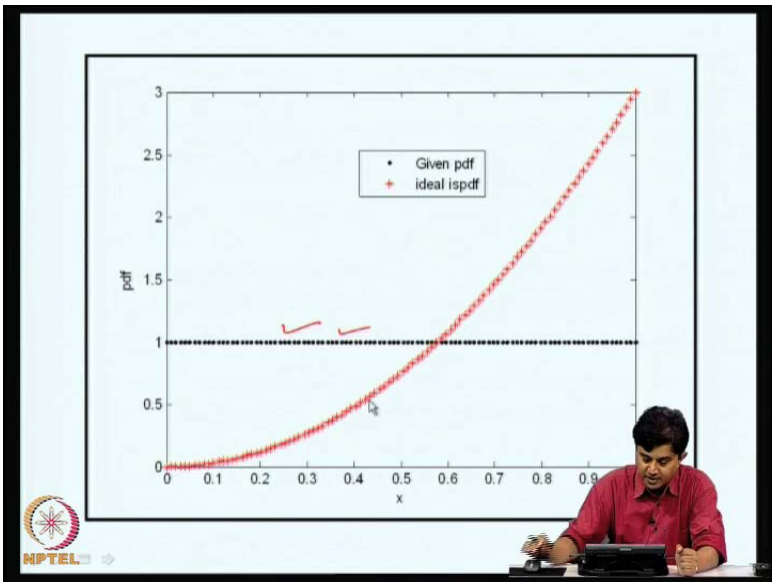
So, in fact, this is the one of the themes in Monte Carlo simulation approaches for reliability analysis, that we will consider some of it shortly. We can reconsider this problem; revisit this problem of evaluating this integral 0 to 1 x square dx. This we have seen earlier shortly, the answer is 1 by 3, as you can quickly verify. Now, what I will do is one approach, that we discussed is to interpret this integral as expected value of x square with respect to random variable, which is uniformly distributed in 0 to 1, that is one approach. But what we will do now is we will multiply and divide by pi of x, where pi of x is an important sampling density function and then interpret I as the expected value of x square divided by pi of x, where we are drawing samples from pi. The expectation of, operation, operator is with respect to pi of x and not this uniformly distributed random variable.

(Refer Slide Time: 25:20)



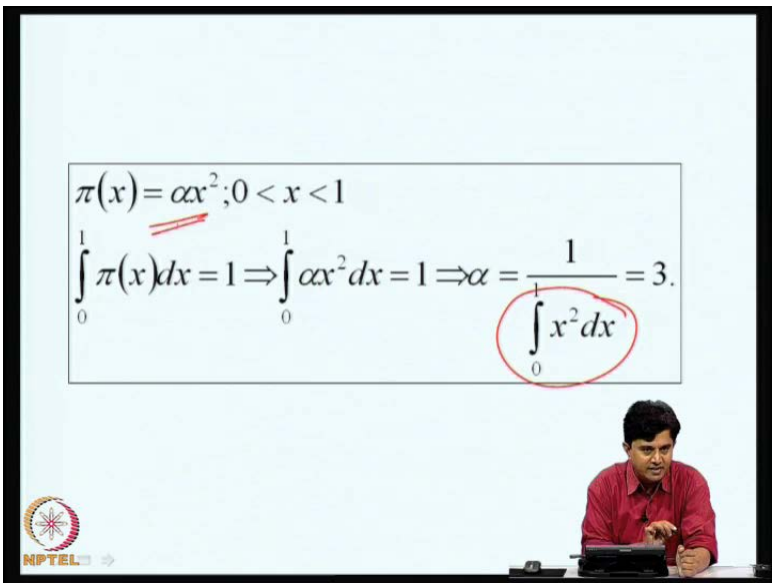
Here, pi of x can be any valid probability density function over 0 to 1. Now, if we select pi of x to be 3 x squares, where pi of x is, x takes value from 0 to 1. Now, you substitute into this, I will get I tilda as 1 by 3 for any value of N and hence, for N equal to 1 also. That would mean, with one sample I am getting the exact solution, provided you draw the sample from pi of x.

(Refer Slide Time: 26:06)



So, what is the catch here? I seem to be getting exact solution with only one sample, but the catch here is that the definition of this important sample density function requires the knowledge of I being evaluated, that is, this is my uniformly distributed probability density function, instead of sampling from this we are advocating, that we should sample from this red line, which is the ideal important sampling density function.

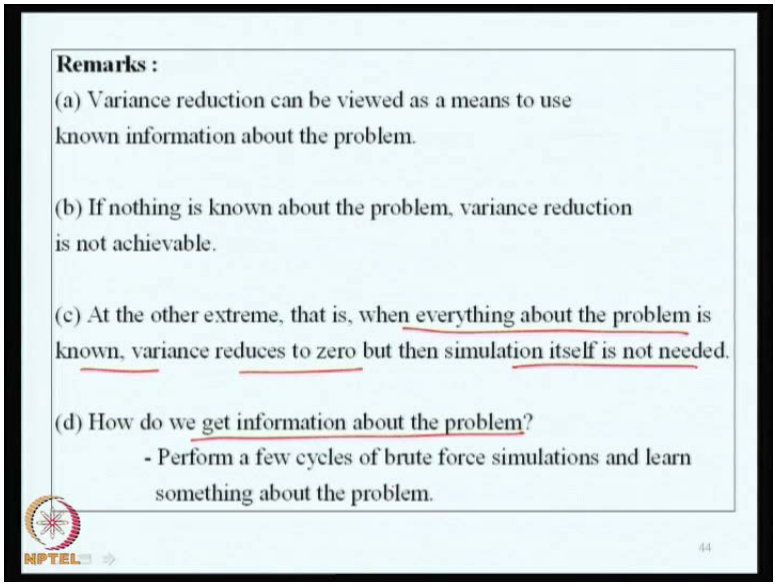
(Refer Slide Time: 26:23)



But how do you get it in the first place? If you analyze that problem, if you take pi of x to be alpha x square and for condition of normalization requires, that area under this

integral should be equal to 1, and if you impose that alpha will be 3, but to find that that it is indeed 3, we need to evaluate this integral, which is the very integral that we are trying to evaluate in the first place. That would mean, you cannot really construct the ideal important sampling density function without knowing the answer to the question that you are posing at the outside.

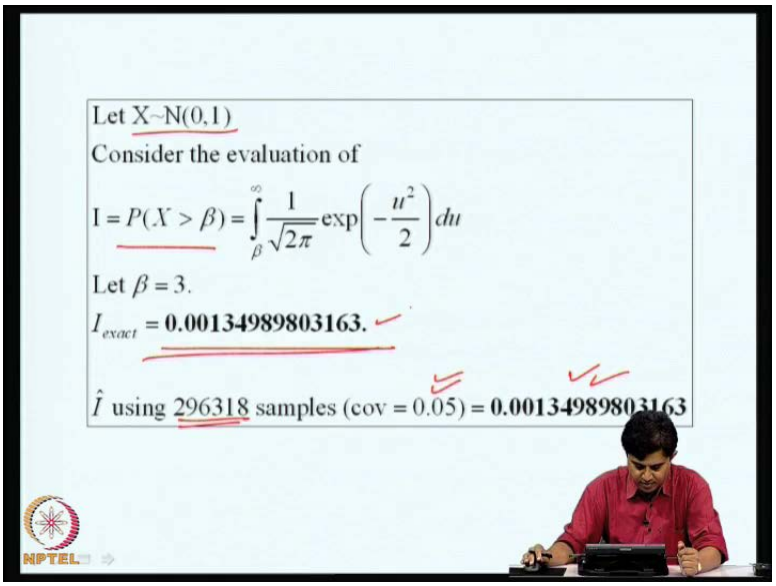
(Refer Slide Time: 27:00)



So, if you remarks, variance reduction can be viewed as a means to use known information about the problem, this is how we should interpret. An ideal solution is not possible, but if you know some of the aspect of the solution, how can we reduce the variance, this seems to be the question to ask. If you do not know anything about the problem, variance reduction is not achieved. At the other extreme, that is, when everything about the problem is known, variance reduces to 0, but then the simulation itself is not needed.

So, the question therefore we should ask is, how do we get the information about the problem on hand? The answers to this lies in performing a few cycles of brute force simulation and learn something about the problem from the first few simulations that we perform. So, as a simulation proceeds, we learn something about the problem by performing the few, first few simulations and based on that knowledge we can alter the way we find the, perform the simulations subsequently because we have already learned something about the system.

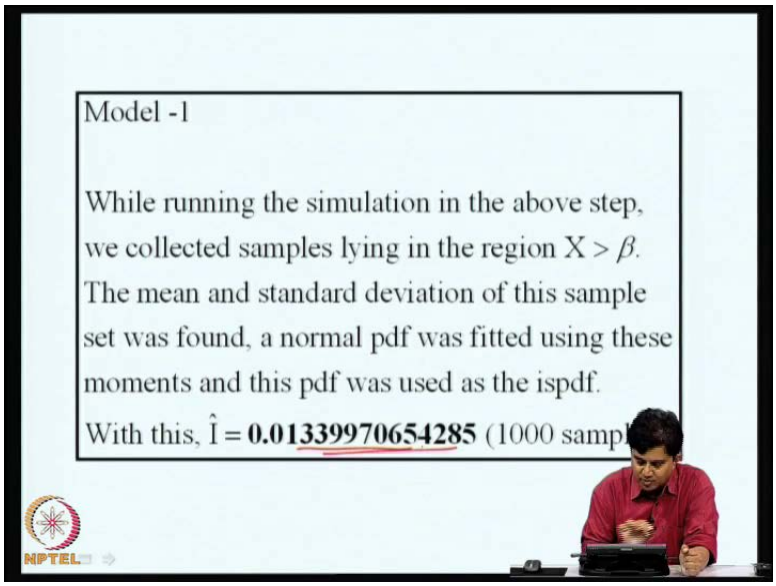
(Refer Slide Time: 28:13)



So, this is the essential idea of problems of variance reduction and we will come to some details shortly. Before that we will consider another simple example, suppose X is a normally distributed random variable with 0 mean and unit standard deviation. We are interested in evaluating the probability P of X greater than beta, where beta is 3, that is given by 3 to infinity 1 by square root of 2 pi (( )) as to minus u square by 2 d u. This is an exactly, we can evaluate this integral exactly and this is the answer that we already know.

Now, if I now use a brute force Monte Carlo by selecting samples size, so that coefficient of variation is point naught 5, I need to use these many samples, you can verify that and I can get this answer. So, this seems to match reasonably well with this because this coefficient of variation is point naught 5, even if you were to repeat these calculations, they will be reasonably close to this number.

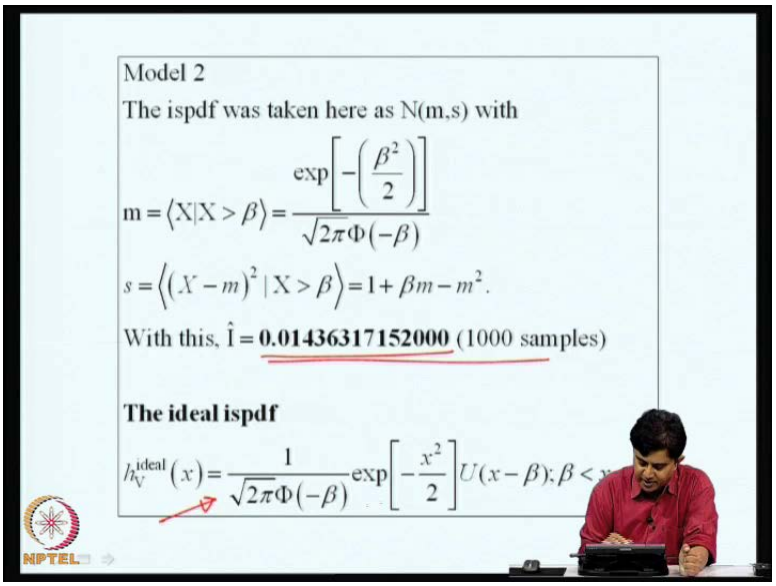
(Refer Slide Time: 29:22)



Now, I am not willing to spare so many samples in my work. So, what I will do? We will consider different models. Suppose, while doing this simulation, this, 2009, 29, 296000 samples, we can collect few samples, which are lying in the region x greater than beta and using only those few samples, we can fit a normal probability distribution function for those samples and that normal probability density function can now be used as the important sampling density function.

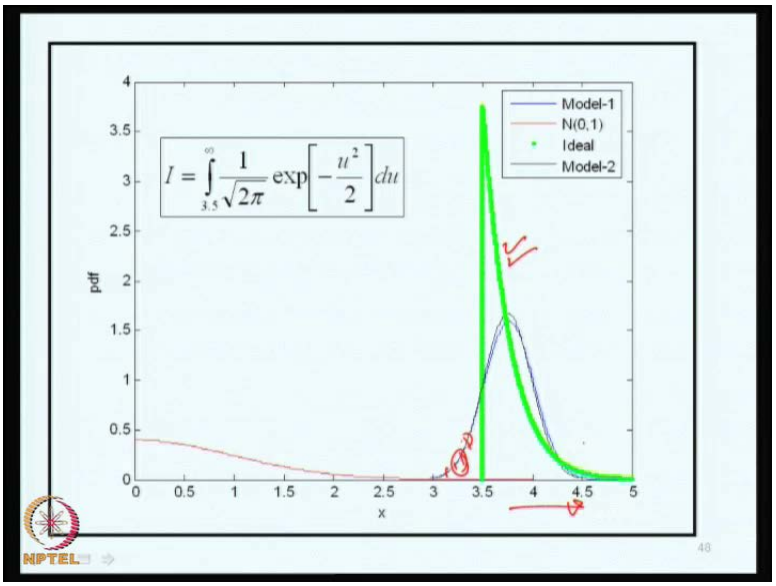
So, what we have done? We have, in brute force simulation, we have been able to simulate few samples in the failure region, which is of interest to us and therefore, you have learnt something about the failure region that I am now using in fitting a probability density function; normal probability density function is, I mean, somewhat adhoc, but that is what I choose. I am not interested in ideal solutions, some optimal solution, so if I use this with thousand samples, I get this number as my estimate for i.

(Refer Slide Time: 30:23)



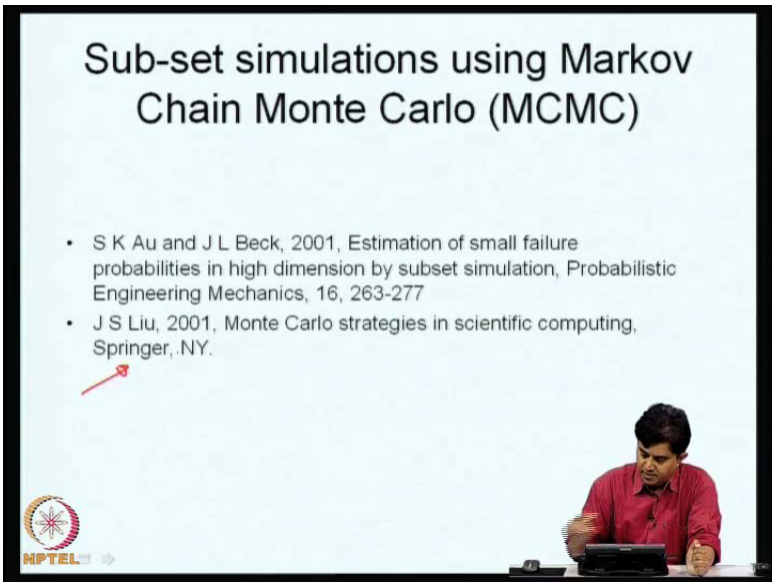
We can do something different; what we could do is, we could take the important sampling density function to be a normal random variable with mean m and standard deviation s, where mean m is taken as the conditional expectation of X condition on the fact, X is greater than beta. And similarly, variance can be taken as conditional variance, where X is greater than beta and we get this answer using this as the important sampling density function. Of course, we will know the exact, the ideal important sampling density function because we know the exact solution and in this case, this, this we can show that, I leave this as an exercise, you can show that this is the ideal important sampling density function.

(Refer Slide Time: 31:07)



So, let us see how they look like? So, this red line is the P x of x, that is, this integrant and we are interested in area under this curve from 3.5 onwards, what is the area under that curve. This green line that you are seeing is the ideal important sampling density function because it is so happen, that for this problem, we know the exact solution, we can construct this. The blue, these 2 lines, red and, black and blue are the sub optimal solutions. So, when I draw samples from these probability density functions, I will be drawing samples from this region, which is not in the failure region. Some of the samples should be not put to proper use, but it does not matter, it is the using this important sampling density function, we are able to get some reasonable estimate for the integral.

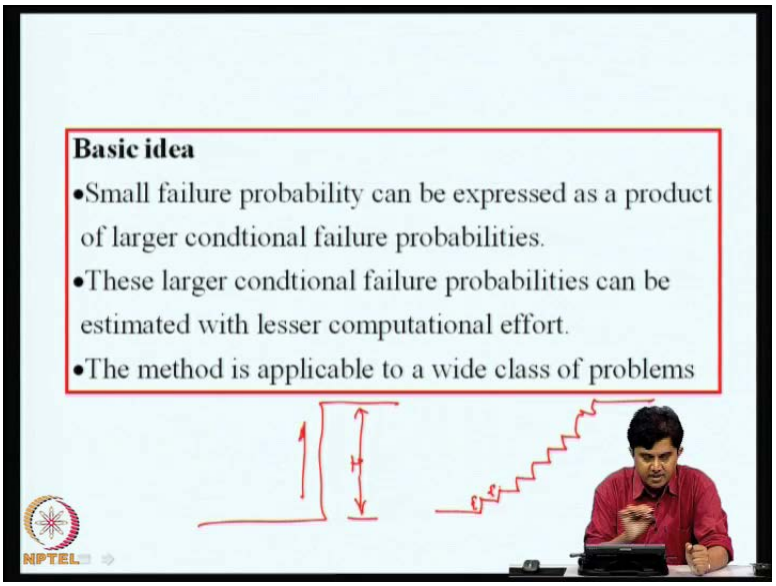
(Refer Slide Time: 32:18)



So, the idea is, we learn something about the failure region by doing few simulations and use that knowledge in constructing the important sampling density function. Now, there are various techniques in existing literature on using this variance reduction technique, where we perform few simulations and learn about the failure region and go along evaluating the required probability of failure or certain expected values with the aim to reduce the variants.

So, one of the techniques that I will briefly discuss is what is known as subset simulations using Markov chain Monte Carlo. This is the relatively a reason, technique, developed somewhere 2001 and there is a paper by Au and Beck, which gives the details of this subset simulations and the details on Monte Carlo, Markov Chain Monte Carlo can be found in the book by Liu.

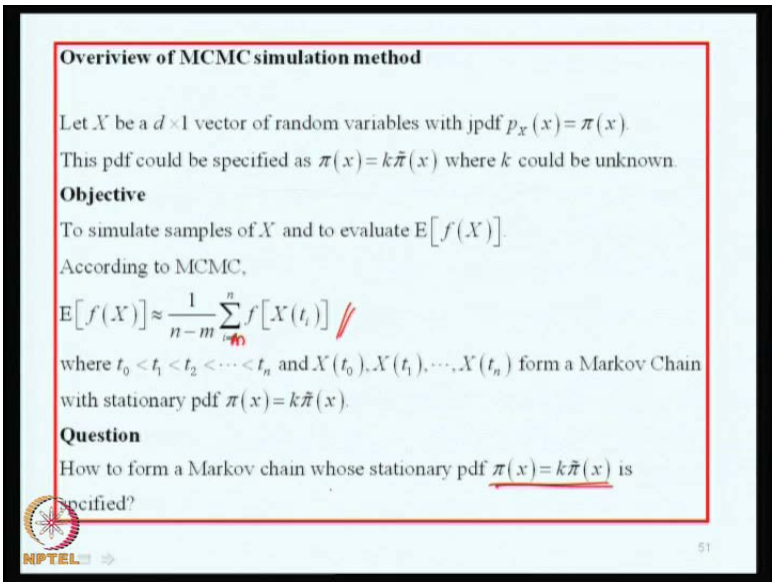
(Refer Slide Time: 33:22)



So, what I will do is I will begin by explaining the, out, basic idea here. The basic idea here is that we are interested in evaluating a small probability failure with, call it as failure probability, and the idea here is, this small failure probability can be expressed as the product of larger conditional failure probabilities. These larger conditional failure probabilities can be estimated with lesser computational effort.

So, the subset simulation method employs these basic ideas and this method is fairly generally applicable to wide class of problems. So, to schematically explain the approach, suppose we need to climb a step of this height and if you want to do it in one go, you should learn how to climb this? This is quite a difficult task; what subset simulation does is we construct the problem where we go in steps and reach the height.

(Refer Slide Time: 34:35)



So, at every stage we take small steps. So, this is what essentially the subset simulation does, so I will clarify this as we go along. Before that, we will quickly review what are this Markov Chain Monte Carlo simulation methods. To illustrate that, let us consider X to be a d cross on vector of random variables with a joint probability density function P x of x. So, let us call it as pi of x, this probability density function could be specified as k into pi tilda of x, where k is the normalization cost and which could be unknown.

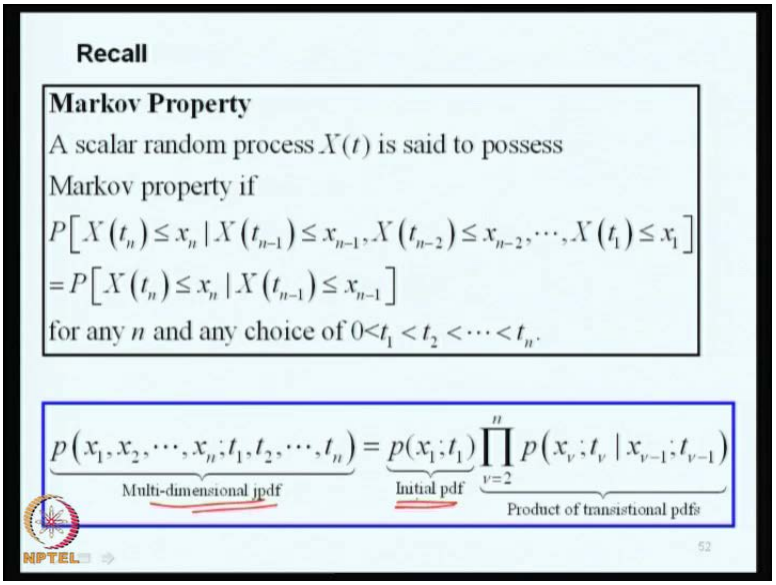
So, the MCMC simulation method essentially enables us to simulate samples of X, probability density function which can be, which may be known only to the extent of an arbitrary that there may be arbitrary multiplying constant, which is unknown. So, the objective is to simulate samples of X and to evaluate, say for example, expected value of some f of X.

So, according to this MCMC procedure, this expected value of f of X is obtained as an average over n minus m samples, where this is m X of t i is a Markov chain with a steady state distribution given by pi of x. That is the basic idea, that means, we consider a time sequence t naught, t 1, t 2, t n and form the random variables X of t, t naught X of t 1, etcetera x of t n, which is a Markov chain with its stationary probability density function pi of x

So, if we can construct this Markov chain, then we can simulate samples of X following this rules of this Markov of chain and evaluate this expectation, so that is a basic

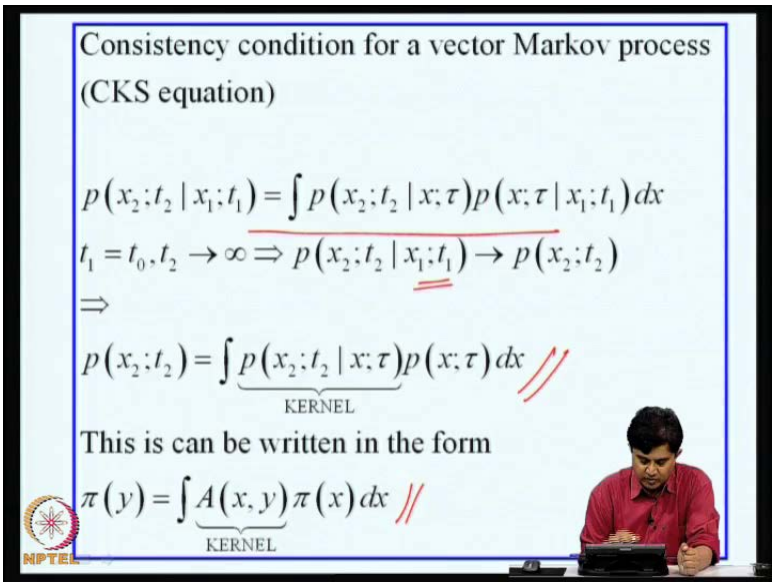
aspiration of this method. The essential question here is, how do we form a Markov chain whose stationary probability distribution density function is specified, which is that Markov chain; so, that is a problem.

(Refer Slide Time: 36:47)



Now, we quickly recall some properties of Markov processes. We have seen already, that if a scalar random process X of t is said to possess a Markov property, if the probability of the event X of t n less than or equal to X n, conditioned on a sequence of events related to its pass namely X of t n minus 1 is less than or equal to x n minus 1, etcetera, etcetera. x of t 1 less than equal to x 1 depends, is equal to the conditional probability of X of t n less than equal to x n condition on only the most recent observation for any n and for any choice of t 1, t 2, t 3, t n, this is the definition of Markov property.

(Refer Slide Time: 37:41)

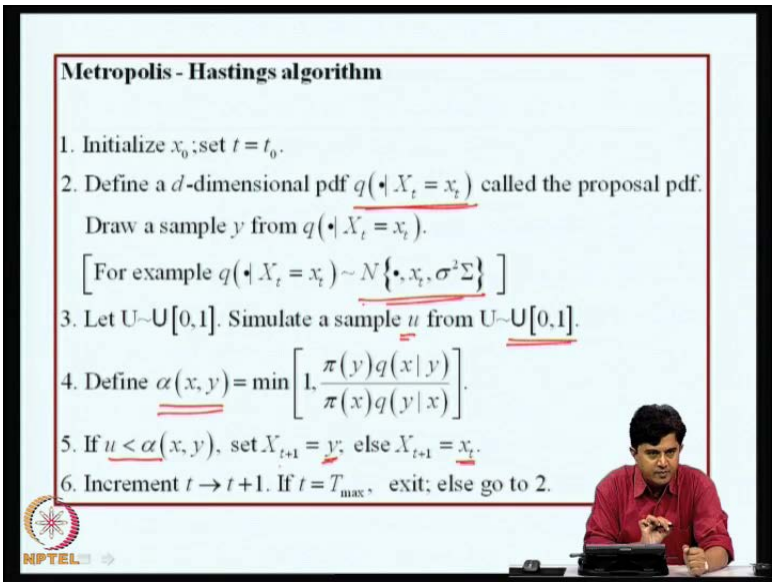


If your process is Markov, the nth order probability density function can be expressed in terms of products of transitional probability density functions and the initial probability density function, this we have seen. Also, what we have seen is the so called Chapman-Kolmogorov small choose k condition, which is a consistency condition for a process to be Markov, where this transitional probability density function has to satisfy this integral equation.

Now, let us consider these integral equations slightly carefully. There are 3 time instance - t 1, t 2 and tau and tau, and tau lies between the t 1 and t 2. If we now take t 1 to be t 0 and take t 2 to infinity, we can show that this transitional probability density function could become independent of the initial time t 1. If that happens, we say, that the Markov process has reached a steady state.

So, in the steady state, I can now omit the dependency on t 1 and write this integral equation in this form, where p of (x 2 : t 2) is p of x 2 condition on (x : tau) and p of (x : tau) dx. This can be written in a slightly different simplified form, which is the standard form pi of y is integral A x y pi of x dx and this A, which in this case is P of x 2 condition on x; here it is A (x , y), is known as a kernel. So, this we need to consider.

(Refer Slide Time: 39:09)



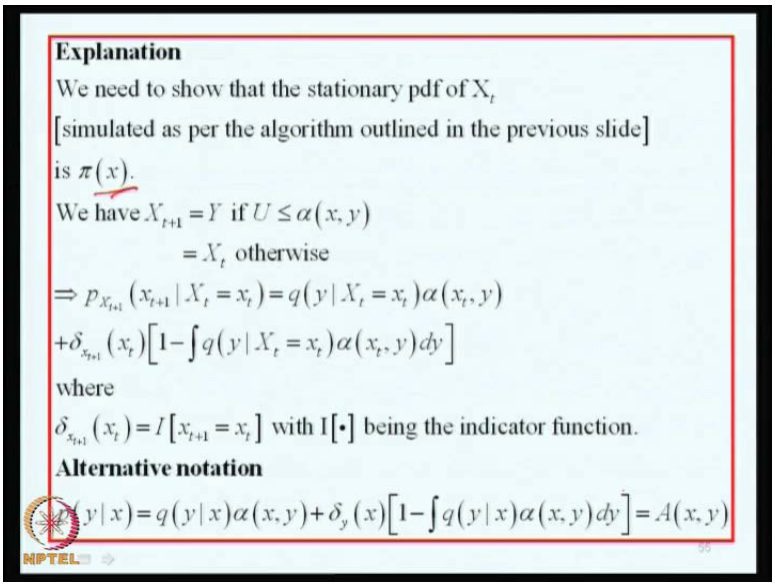
There is an algorithm by the name Metropolis-Hastings algorithm, which helps us to form the Markov chain, which has the target stationary probability distribution function. So, the algorithm itself is quiet simple; so what we do is, we initialize x naught and set t to t naught. Then we define a d-dimensional probability density function denoted by q and this is known as proposal probability density function, and we draw a sample from this proposal density function.

This proposal density function, for example, could be a multi-dimensional, normal probability density function with the mean x t and covariance sigma square into sigma. This, we will see, how to select this in due course. We simulate, we consider another random variable U, which is uniformly distributed between 0 and 1; we simulate a sample of U from this random variable.

Then we define alpha as minimum of (1, pi of y) into q x condition y divided by pi of x into q of y condition on x. Now, if U is less than equal to alpha, we move the chain to the value y, otherwise the chain stays at x t, where x t is the value to which we have set the initial, you know, x.

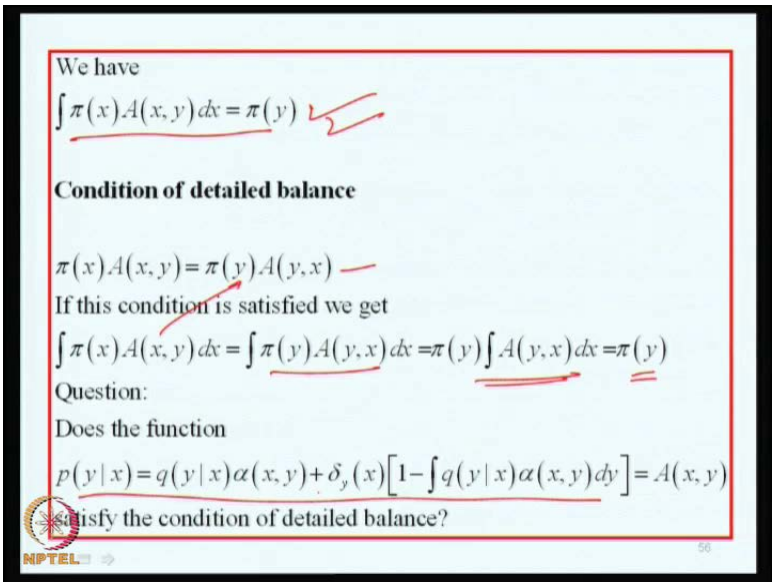
Now, we increment t; if t is equal to T max, we exit, otherwise we go to step 2. So, this is the very easily implementable algorithm, the only need, only thing that we need to satisfy is that we should be able to draw samples from this proposal density function. If that is possible, then this method can be implemented.

(Refer Slide Time: 41:05)



Now, we need to explain the meaning of this algorithm. So, what we need to show, that is, that the stationary pdf of X t, simulated as per the algorithm just outlined in the previous slide is pi of x; that is what we, need to, need to show. Now, we can construct the simple arguments, we have X t x of t plus 1 is y if U is less than or equal to alpha or X t otherwise. Therefore, I can construct the probability density function of X t plus 1 conditioned on X t equal to X t as q into alpha. Recall, that U is uniformly distributed between 0 and 1 plus an indicator function, I will come to that, into 1 minus this probability. There are only 2 states, this, this state has a probability of, probability of U less than or equal to alpha, which is nothing but alpha because U is uniformly distributed and 1 minus of that is this. So, this is, this is an indicator function, which is equal to 1 if X t plus 1 equal to x t, otherwise it is 0.

(Refer Slide Time: 42:28)

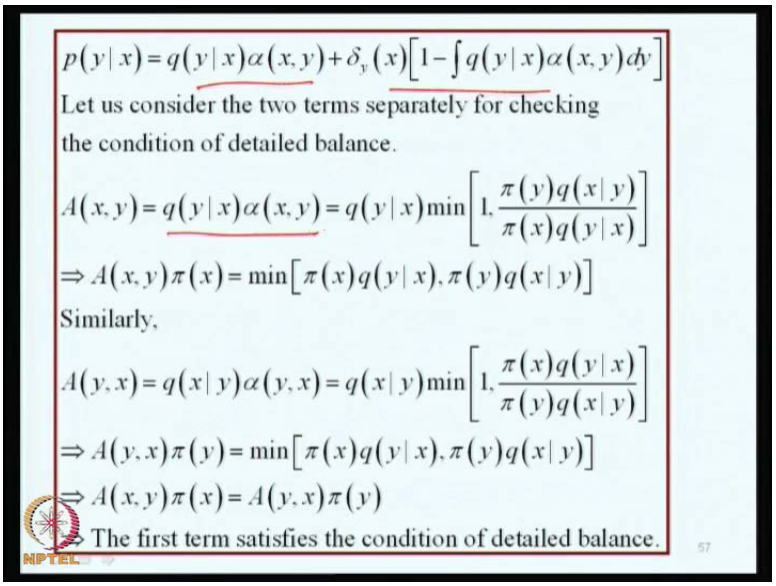


Now, in the alternative notation we can write this as p of y condition on x as q y condition x alpha plus delta y x plus 1 minus this and this is the kernel. Now, we have, if stationary condition is realized, this condition to be satisfied. Now, there is a condition known as condition of detailed balance, which helps us to verify if steady state is possible, that is, pi of x is A (x, y) is pi of y A (y, x). If this condition is satisfied by pi with respect to the kernel function A, then we say that there is state of detailed balance.

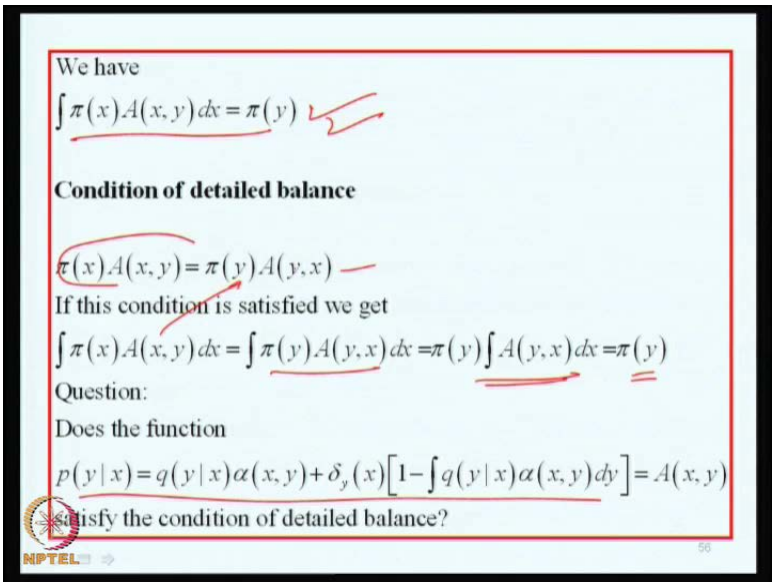
Let us assume that this condition is satisfied, then you can verify, that we can now use this and show that integral pi of x A (x, y) dx is nothing but, because this is equal to this, this is equal to this, I will now substitute this and integration, still with respect to x pi of y can be pulled out and I have integral A (y, x) dx and this is A is a valid density function. Therefore, area under that curve is 1; therefore, this is pi of y. So, if condition of detailed balance is satisfied, then in turn, we are satisfying this requirement.

Now, the, therefore the question that we should ask is, the kernel that we have right now is this, does this kernel satisfy the condition of detailed balance? If condition on detailed balance is satisfied, then we are proving that the stationary pdf is pi of x.

(Refer Slide Time: 44:06)

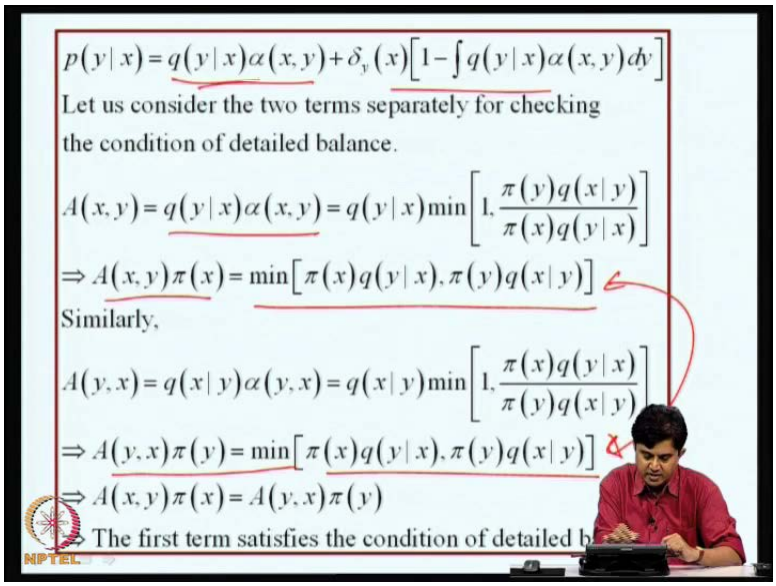


(Refer Slide Time: 44:20)



So, we can consider now the 2 terms, this term and this term separately. Suppose, if you start with the first term, q into (y, x) alpha (x, y), the condition for detailed balance is that this should, this should be equal to, this should be equal to this. So, let us see how, what happens for this specific example.

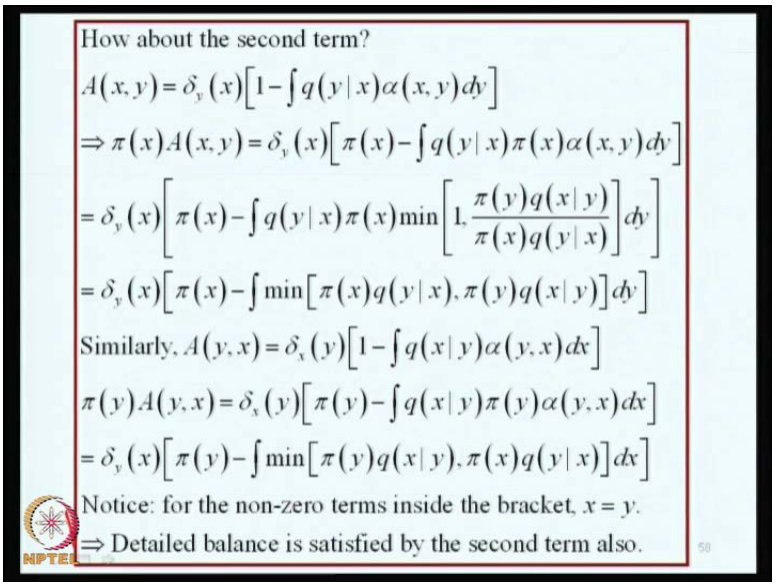
(Refer Slide Time: 44:23)



Now, I have now interchanged x and y, I am writing this, no, I am writing for alpha, alpha is minimum of this terms inside the bracket. Now, I can take this denominator outside and multiply on the left hand side; this q of (y, x) and this q of y of x cancel and pi of x comes to this side and we get A of (x, y) is pi of x, which is this.

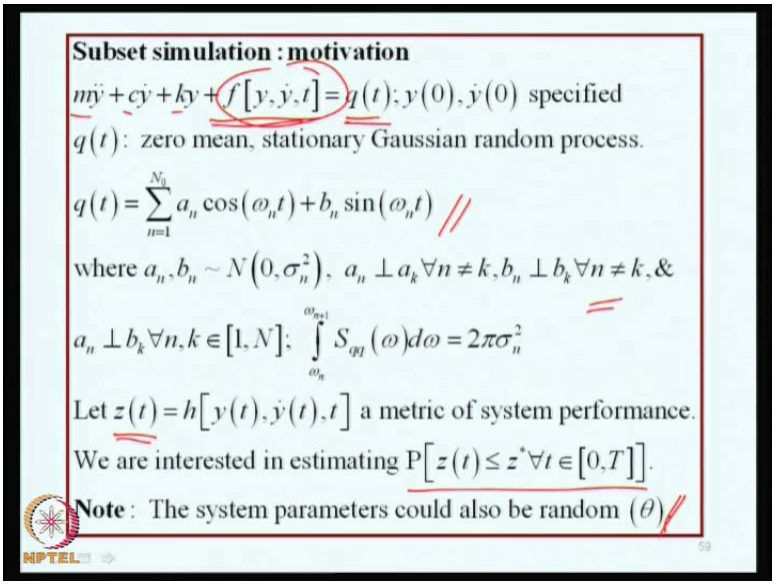
Similarly, if I now consider A (y, x), I can rewrite this and follow the same steps and I can get A of y of x into pi of y to be this. And if you compare these two, we can verify that the first term in this kernel satisfy the condition of detailed balance. The same logic can now be used for the second term.

(Refer Slide Time: 45:23)



So, again, I, I write this and substitute for alpha and take out these terms and I leave this as an exercise, you can verify, the proof is displayed here and you can show, that the detailed balance condition is satisfied by the second term also. Therefore, the Metropolis-Hastings algorithm is, will deliver what we want.

(Refer Slide Time: 45:53)



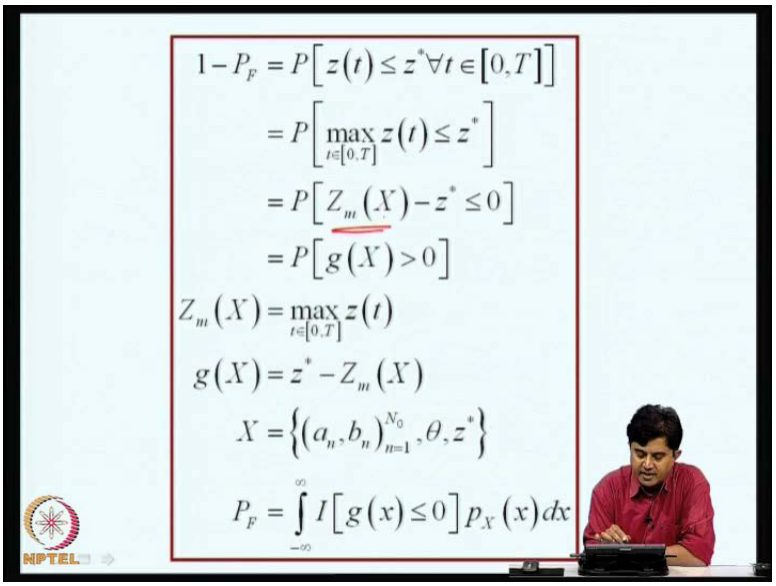
So, now, let us come to the problem of subset simulation. So, let us consider a dynamical system - my double dot plus cy dot plus ky, which some non-linearity and an excitation q of t where q of t is 0 mean, stationary Gaussian random process. It is not necessarily

white, it can be anything. So, for purpose of discussion, we can write this q of t in terms of Fourier series, where a n, b n are Gaussian random variable. This we have described in the earlier lectures and we are assuming that q of t is mean square bounded and therefore, as Fourier series representation of this kind is possible.

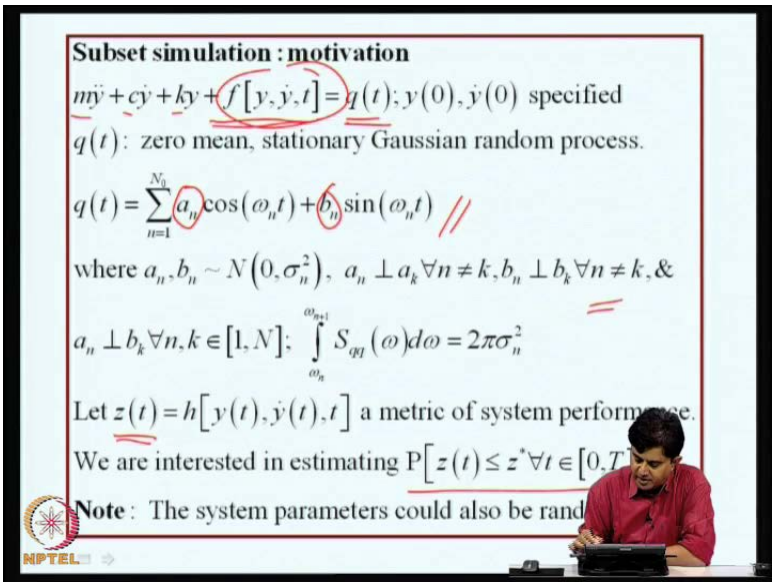
Now, let us consider z of t to be a metric of system performance which is a non-linear function of y and y dot. This could be for instance 1 mises stress or say, reaction transfer to a support or so on and so forth. Now, z of t is my metric of performance and what we are interested is in estimating the probability that z of t stays below z star. Assume that z star is a safe limiting value for all t during a time duration 0 to capital T.

Now, we can assume in this problem, that system parameter is like m c k or any other parameters that are buried in this non-linear terms, could non-linear term, could be random in nature and we will call that, all this setup random variables as theta.

(Refer Slide Time: 47:22)



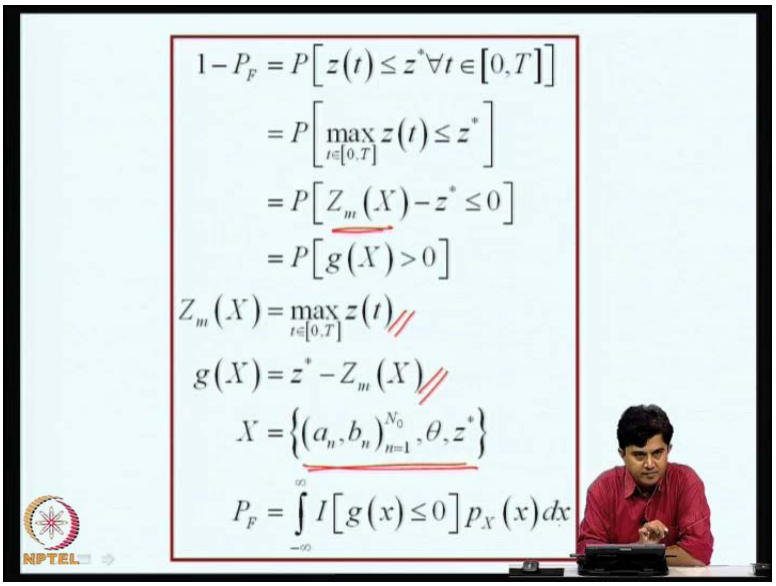
(Refer Slide Time: 48:01)



So, what we are interested is, evaluating probability of z of t stays less than or equal to z star for all time in 0 to T.

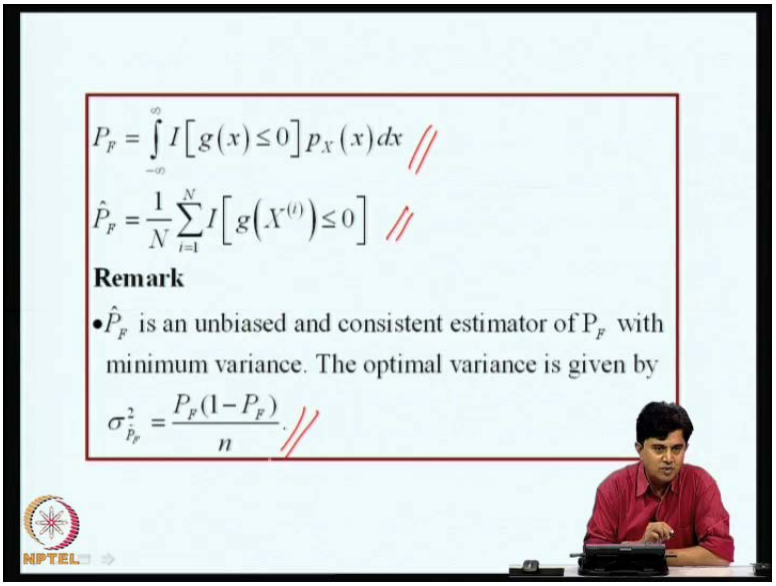
Now, this is known as a problem in time variant reliability analysis. This can be converted into a problem in time invariant reliability analysis by replacing this by the maximum of z of t over a duration. If z of t is a random process, maximum of a random process, where time duration is a random variable, so that random variable I write it as Z m and I, I assume that Z m is a function of x, where x is the set of all random variables present in this problem, that basically include by random variables a n b n and the random variable theta.

(Refer Slide Time: 48:11)



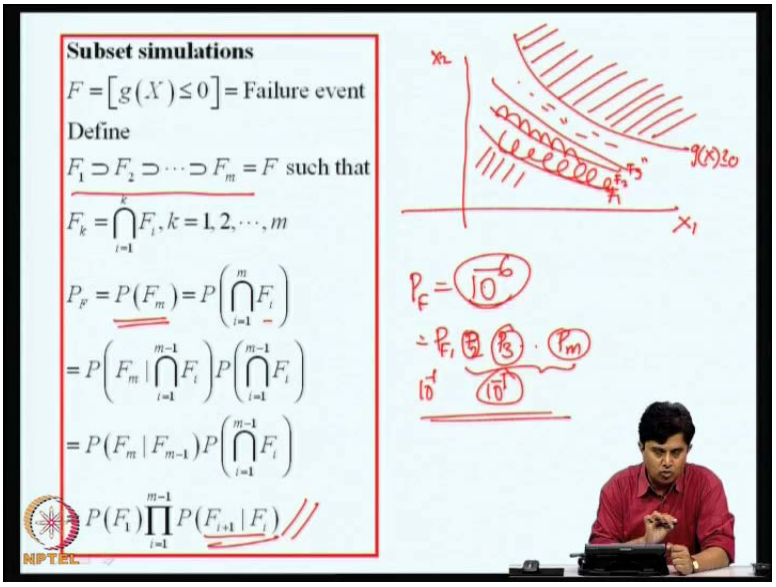
So, I will now define z star minus Z m of X as my performance function and we are interested in probability of g of X greater than 0. So, where Z m of X is maximum were the time 0 to t of z of t and g of X is this and this X is the collective, the collection of all the random variable a n, b n theta, even z star can also be a random variable. We can put that also here and we get the probability of failure as now the integral of minus infinity to plus infinity integrator of g of X less than or equal to 0 P x of x dx.

(Refer Slide Time: 48:56)



Now, I can estimate P F using this estimator, which is unbiased estimator with minimum variance, which we discussed shortly with the variance given by this. The subset simulation, now you would like to apply on this problem with an objective of initiating the simulations and learning something about the system behavior and based on what we learned, we would modify our simulations steps subsequently.

(Refer Slide Time: 49:29)



So, that is a basic spirit of variance reduction and let us see, how we proceed. So, what we do here is, we consider the event F g of X less than or equal to 0. So, imagine that this is x 1, this is x 2 and this is g of X less than or equal to 0.

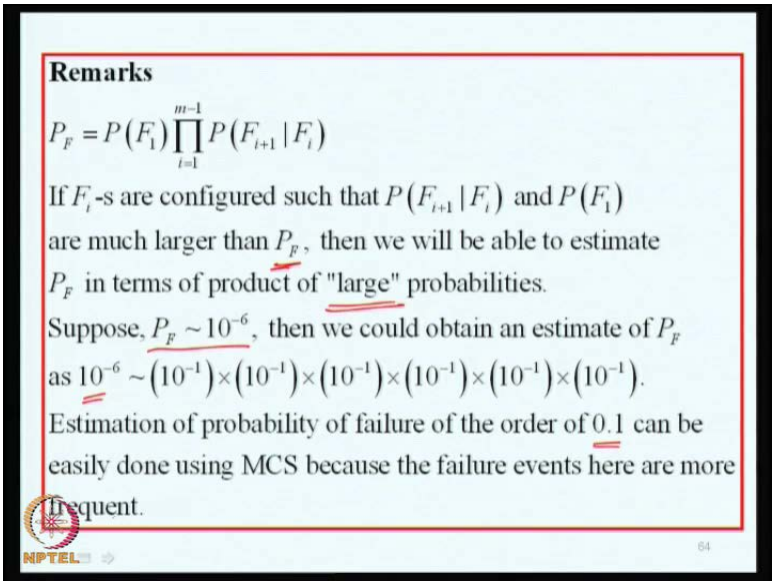
So, what we do here is we form a sequence of events F 1, F 2, F m, where F m is F, such that F k is intersection of i equal to 1 to F i. That means we form events like this, a sequence of events. So, this is F 1, this is F 1, this is F 2, F 3 and so on and so forth. So, F 1 is subset of F 2, F 2 is subset of F 3, F 3 subset of F m and so on and so forth.

Now, probability of failure is actually the area under the joint density function over this region; that can be quite small. So, for example, this P F can be as small as say, 10 to the power of minus 6. Now, what I do is, this is the probability that I am looking for, I will now write this as probability of intersection of the events F i. This itself, I will write it as F m conditioned on intersection of F i up to m minus 1 into probability of intersection of i equal to 1 to m minus 1 F i.

So, I can repeat this and finally write the P F as P F 1 into product of these conditional probabilities. That means, I will evaluate the probabilities, the first one, what I will do is this and the second one is this, third one is this and 4th one is this, so on and so forth. Then, the final probability, I will simply multiply all these probabilities. So, P F for example, can be obtained as P F 1 into P, some P 2 into P 3 into P m. So, this can be for instance, 10 to the power of minus 1 and all this could be of the order of 10 to the power of minus 1.

So, if I, if I device now a simulation procedure, such that at every stage I only evaluate one of these probabilities, I will be evaluating a small probability. In the end, I will multiply all these small probabilities to get larger probability, to get a smaller probability. So, estimation of a larger probability of the order of 10 to the power of minus 1 is not easier than estimating a probability of failure of the order of 10 to the power of minus 6 in one go.

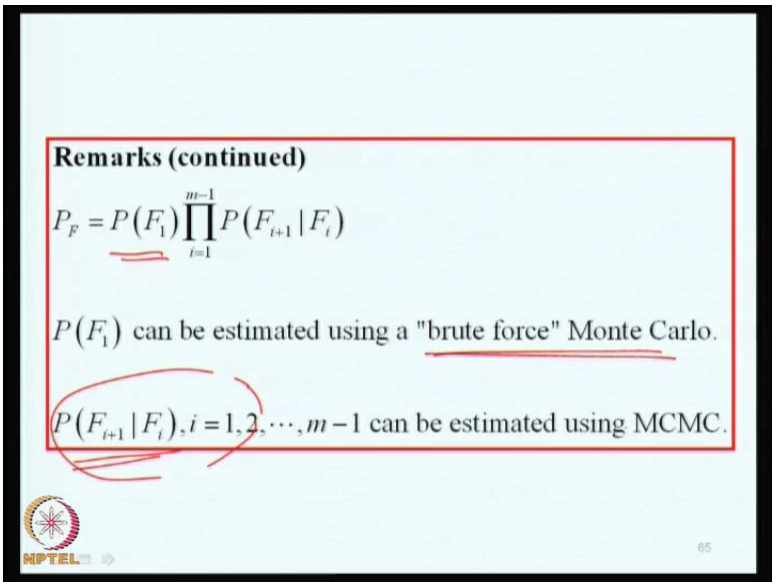
(Refer Slide Time: 52:16)



So, how do we do that? So, what we are doing is we are expressing P F as P F 1 into the product of P of F i plus 1 into condition on F i where I runs from 1 to m minus 1. So, if F i's are configured, such that P of F i plus 1 condition on F i and P of F i are much larger than P F, then we will be able to estimate P F in terms of product of large probabilities. Suppose, P F is 10 to the power of minus 6, then we could obtain an estimate of P F 10 to the power of minus 6 as product of 0.16 number of times.

So, estimation of probability of failure of the order of 0.1 can be easily done using Monte Carlo simulations because the failure events here are more frequent.

(Refer Slide Time: 53:08)



So, what we do is, we start with, we have to find P F of F 1 and we what we do is, we follow a brute force Monte Carlo simulation to find P F of F 1. Subsequently, these conditional probabilities, we evaluate using MCMC tools.

So, what I will do in the next class is to describe this step of how to find samples, I mean, how to find this probabilities using MCMC and illustrate this method using few simple examples. So, we will conclude the today's lecture at this stage.