**Path Integral Methods in Physics & Finance**

**Prof. J. P. Singh**
**Department of Management Studies**
**Indian Institute of Technology, Roorkee**

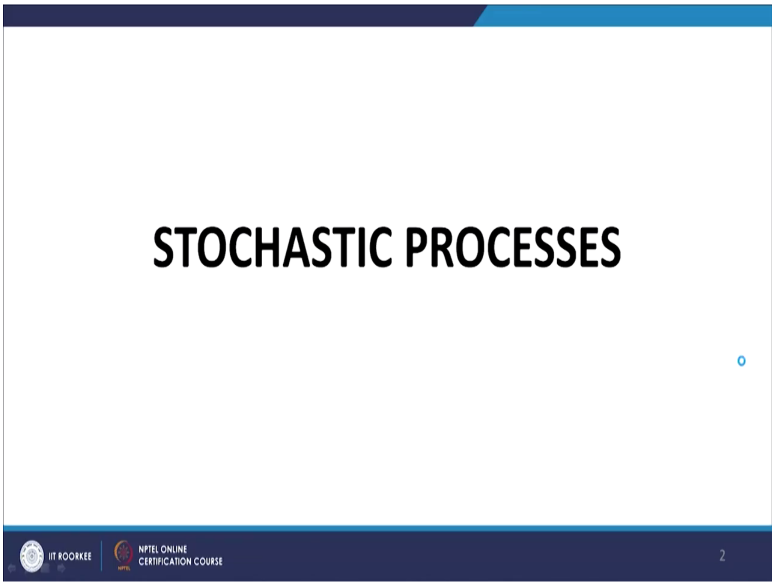
**Lecture - 07**
**Elementary Theory of Stochastic Processes**

Welcome back, in the last lecture I concluded the discussion with the central limit theorem, which is the cornerstone of lot of statistical applications. I wanted to take an example on the central limit theorem, however due to time constraints I will not start with that; maybe if towards the end of today's lecture if I get some spare time, I will take up that example.

It illustrates beautifully the real nuances behind the central limit theorem; in any case I will definitely keep it in the presentations and of course viewers may ask questions on it in the discussion forum, in case I am, I do not get the time to discuss it in detail anyway.

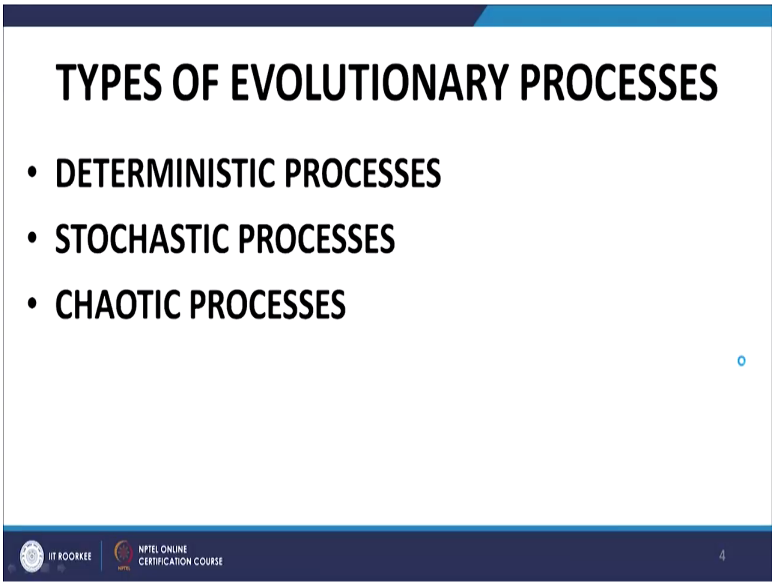
So, let us start with the framework in relation to stochastic processes.

(Refer Slide Time: 01:16)



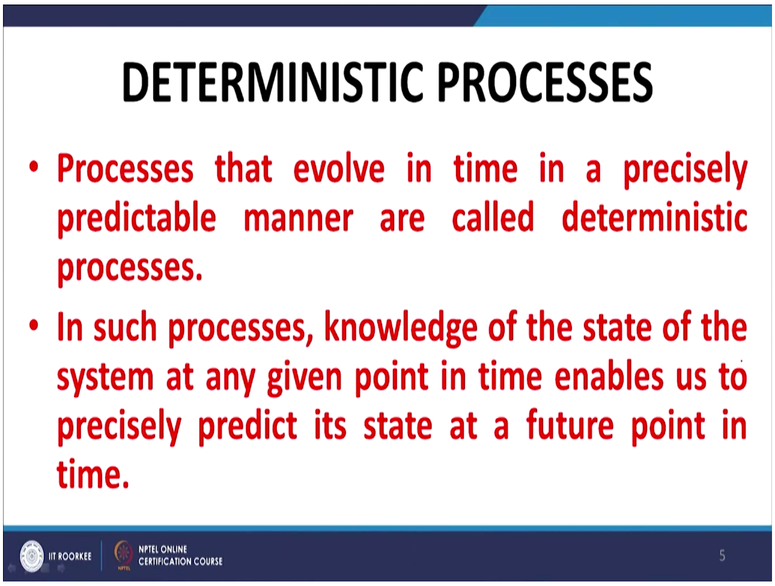

Before we, before I talk about stochastic processes; it is necessary and useful to understand the difference between the concept of stochastic process and the concept of a deterministic process. Physical processes are generally classified into two types; deterministic processes and stochastic processes.

(Refer Slide Time: 01:36)




Like a new version of relatively recent origin is the concept of chaotic processes, which essentially is a variant of deterministic processes. Now a system is said to evolve deterministically; if knowing the state of the system at a particular point in time, we are able to precisely predict the state of the system at any future point in time. So, if you know the state of the system at t equal to 0, then you should be able to completely determine the evolutionary dynamics of that particular system.

(Refer Slide Time: 02:26)




Clearly in the case of deterministic evolution it is necessary that, we should be able to identify a complete set of variables that provide a complete description of the system. And then secondly, we should be able to determine or we should have the laws, the physical laws under which those variables that we have determined for identifying completely the physical system evolves with respect to time.

Normally what happens is, these physical laws manifest themselves in the form of differential equations; there has, therefore a solution to this differential equations with the given initial conditions provides us with the necessary state of the system at a future point in time.

This is how classical mechanics normally works. The physical laws may manifest themselves in the Lagrangian formalism or in the Hamiltonian formalism; but at the end of the day they give us the differential equations which govern the evolution of the system with respect to time, and given the initial conditions we are able to precisely determine the state of the system at a future point in time that is how deterministic systems operate.

So, the important thing, the fundamental point here is that, a system can only be classified as deterministic provided, we have a complete knowledge of the variables that go into the description of the state of the system. Otherwise if there are some ignorance, some lack of knowledge about the variables that go into the description of the system; then obviously the variables that would actually impact the system dynamic, but which we are not able to identify, which we are not able to model will have their say and as a result of which our predictions may not turn out to be precise, that leads us to the concept of stochastic dynamics.

A system is said to evolve stochastically or in a random manner when the given a state of a system, a future state of a system; future state of the system cannot be precisely predicted, it has a component of randomness embedded in it and as a result of which the evolution of the system at a future point in time is not precisely predictable.

There are a lot of views about how randomness manifest itself, indeed randomness itself can be a of two variants; one is classical randomness which I have just alluded to now, and the second is the concept of quantum randomness, which is more in the nature of indeterminacy. It is very interesting here to point out that, the Schrodinger equation in itself is a deterministic equation, and the solution that it gives us for the wave function is a deterministic solution; it is the interpretation of the wave function which is probabilistic.

So, here in the quantum world, in the context of quantum mechanics, we have a slightly different type of randomness; in the context of classical randomness, randomness is more in the nature of ignorance of the complete description of the system. For example, when you toss a coin, we are not able to precisely model each and every factor that goes into the evolution or the trajectory of the coin as it moves through the air and falls with its head up or tail up.

A consequences is that, if we were to be able to model, if we were to be able to model the coin tossing with all the variables, all the degrees of freedom duly identified which give which impact the trajectory of the coin. For example, the angle of flip, the viscosity of air, the temperature gradient and so many other things which would contribute to the to the trajectory of the coin.

And if you were able to identify all these factors and to model their interplay by some physical laws; then we would be able to predict the outcome of the coin toss experiment. This was indeed the view of Albert Einstein; he believed that the universe at the fundamental level was deterministic, and randomness as such was only manifestation of the lack of complete knowledge of a particular physical system.

However with the advent of quantum mechanics, randomness took a more important place in the physics literature, in the evolution of systems; and of course quantum randomness is as I mentioned earlier is different compared to classical randomness. Nevertheless the randomness, the concept of randomness got upshot by the discovery of quantum mechanics and its future development, we will talk about it later.

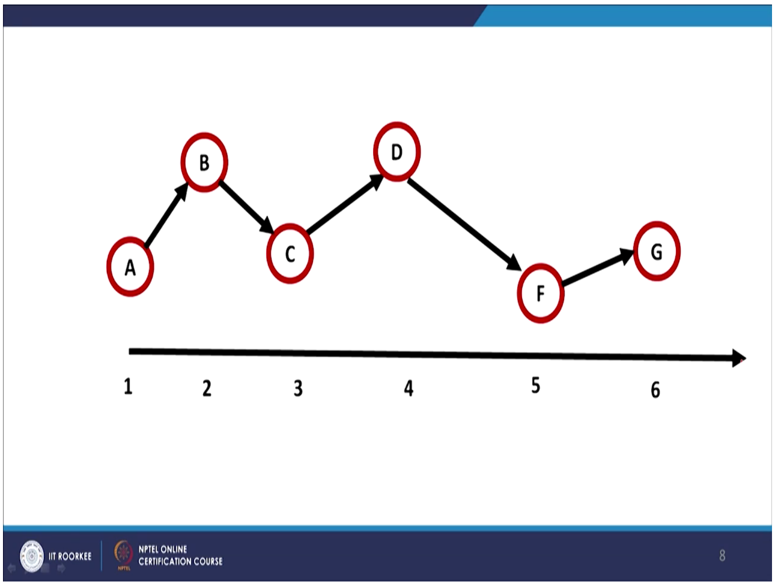
But that is the fundamental thing; deterministic systems are those systems whose future evolution is perfectly predictable; stochastic systems are those systems whose future evolution is not perfectly predictable. An element of randomness creeps into the future evolution, as a result of which we do not completely determine the state of the system at a future date.

Chaotic dynamics is a variant of deterministic dynamics, where the system evolves in such a way or the evolution map or the evolution laws are such that they are hyper sensitive, extremely sensitive to the initial conditions. Therefore, even a slight difference in the initial conditions, manifests itself as an exponential divergence of the future evolution of the system as time progresses. It is said in the context of chaotic systems that a flap of butterfly's wings in Bermuda, who could result in a hurricane occurring in New York.

So, that is what chaotic systems means; even the lightest difference in initial conditions, even a measurement error for that matter would manifest itself exponentially as time progresses. The nature of the map is such, the nature of the physical laws are such that this particular phenomenon occurs, right.

For our purpose we are more concerned with stochastic dynamics; stochastic dynamics is also relevant, because financial assets and instruments or derivatives of financial assets which are functions of the fundamental financial assets also are usually modeled in terms of stochastic processes. So, we need to discuss, we need to understand more about stochastic processes.

(Refer Slide Time: 10:12)



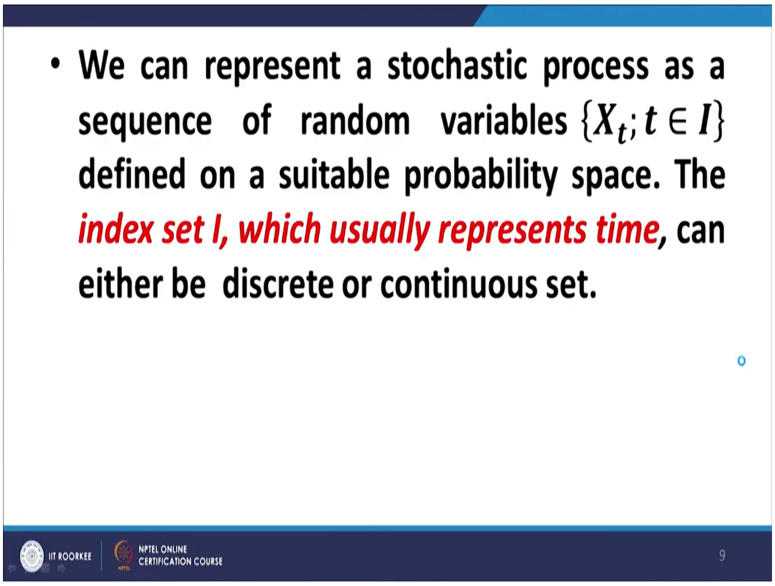

Now, let us take this example, we have got here on the slide. We have got process which is exhibiting a particular path, which is executing a particular path. This path is determine you see you can say it is a coin toss experiment or it is the price of a particular financial asset or whatever; at t equal to 1 the price is given by the point A, at t equal to 2 the point is given by the point B, at t equal to 3 the point, the price is given by C.

So, that these are random or they have a certain element of randomness in them in the revolution and as a result of which we get this fluctuating path. Now as you can see, in fact as you can see in this slide, it is possible and it is what is the actually done; this is a stochastic process, it is evolving randomly. So, let us look at this particular process. Let us say at t equal to 1 the system is in state A, at t equal to 2 the system moves to the state B, and at t equal to 3 the system moves to the state C and so on.

Now, as I mentioned the this evolution is random; therefore what we can do is, we can model this entire process, we can model this entire process as a sequence of random variables. At t equal to 1 we can use a random variable to identify the state at t equal to A; at t equal to 2 we can have a random variable to identify a state at a t equal to 2 and so on. In other words, the entire sequence or the entire process can be viewed, the sequence of random variables duly indexed by the time; time t equal to 1, t equal to 2 and in this forms the index set.

This particular time line forms the index set and the values that the random variables take will determine the path that the particular system evolves in the realization of the stochastic process.

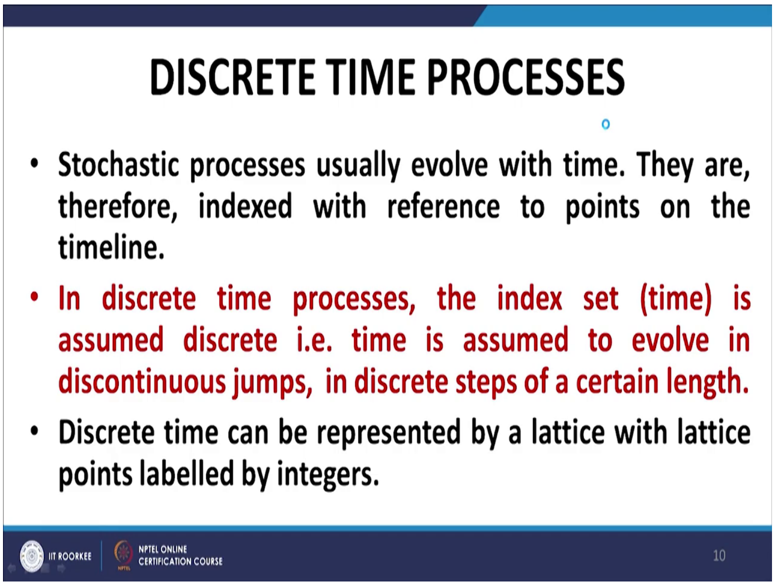
(Refer Slide Time: 12:34)




And therefore, we can represent the stochastic process as a sequence of random variable X t; X 1, X 2, X 3 all these are random variables remember, but they are indexed by the time t at which they become relevant.

So, they are different random variables, it is not they are not the same random variable, they may have the same distribution, they may be a in independent identically distributed random variables that is fine, but they are different random variable; X 1 is different from X 2, X 1 models the system at t equal to 1, X 2 models the system at t equal to 2, X 3 models the system t equal to 3 and so on.

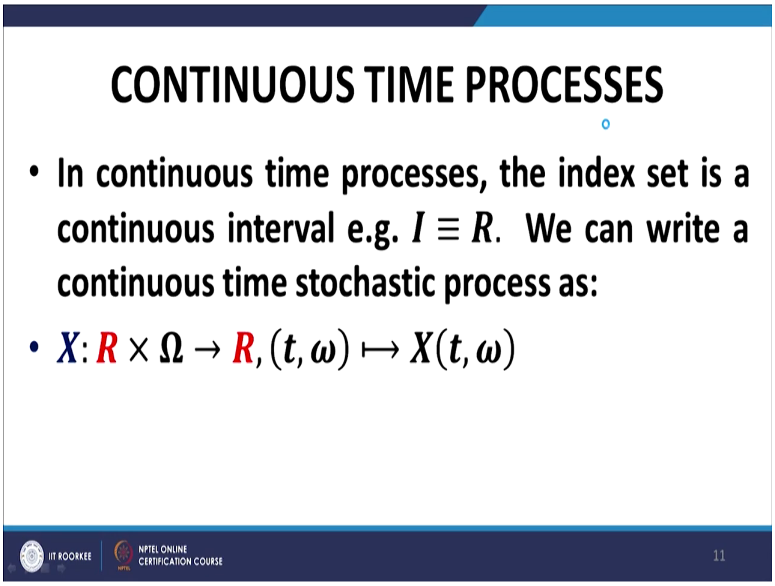
(Refer Slide Time: 13:13)




Now, the index set that is usually taken as the time, not necessarily so; but usually the index set is time, because we are talking about evolution with respect to time in general. And so, the index set is taken as time, you can model the time by discrete points; that is we can modulate as points shown a lattice, and in that case time takes discrete integral values t equal to 1, 2, 3 so on. And the index set therefore is a discrete set, in such situations we call them discrete time processes.

So, time, in that case time jumps discontinuously and usually it is of constant length and usually these, usually these points are taken to be those points at which the system makes a transition.

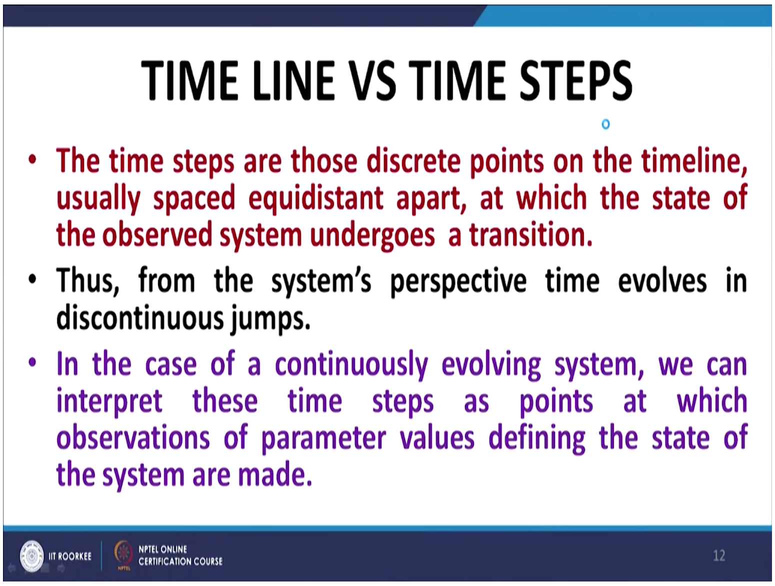
(Refer Slide Time: 14:08)




Then we have continuous time processes which is more general, where time is assumed to flow uniformly and continuously. And in other words the index set that is representing time as I mentioned; index set is a continuous set is an interval on the real line or a collection of intervals on the real line.

So, in that case the stochastic process can be viewed as in mapping R cross omega; R is the index set which is now real interval, cross omega, omega remember is the sample space of the random variable that collection of which represents the, represents the stochastic process. And this goes to R which is the, which is the what you call the output, which is the output of the random variable the real number that the random variable associates with every element of the sample space.

(Refer Slide Time: 15:20)

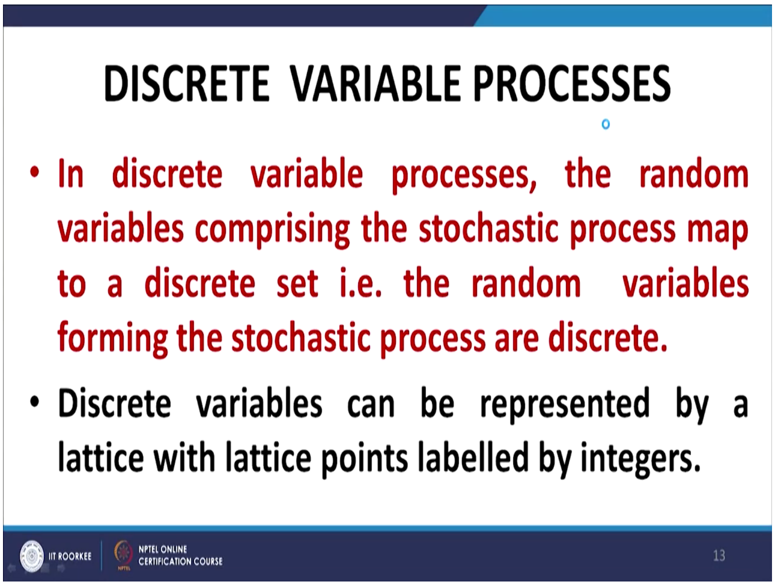



Now, it is important to note the difference between time line and time step; timeline is the uniform continuous line that time evolves, in time steps are usually when we have discrete numbering of time steps are those points in time at which the process makes a transition t equal to 1, 2, 3, 4, these are the points at which the process makes a transition.

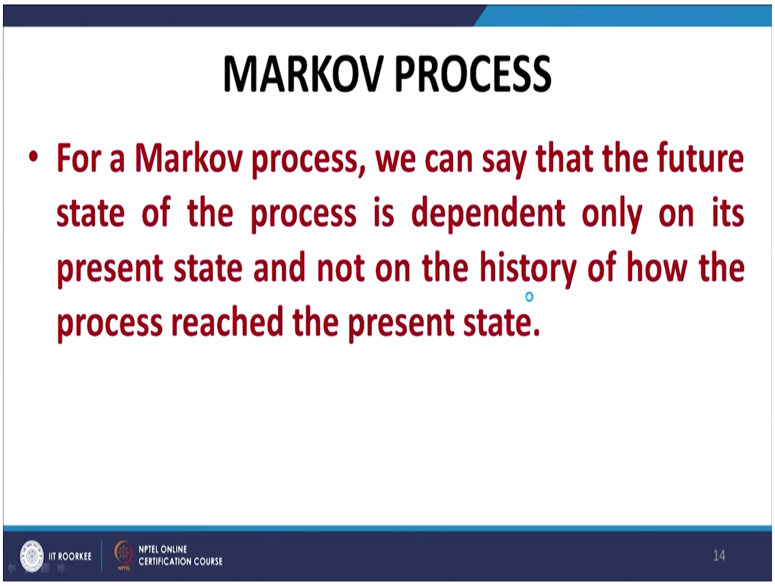
Or there can be points at which we are making an observation on the process in, even if the process is evolving continuously it is the; these are the points at which observations on the salient parameters, on the parameters being tracked are being made, these are points constitute the time steps. Time line is the uniform underlying line that is continuously flowing.

Similarly, we can have discrete variable processes, where the random variable, the co domain of the random variable is discrete; and if the co domain of the random variable is continuous, then we have a continuous variable process.

(Refer Slide Time: 16:13)

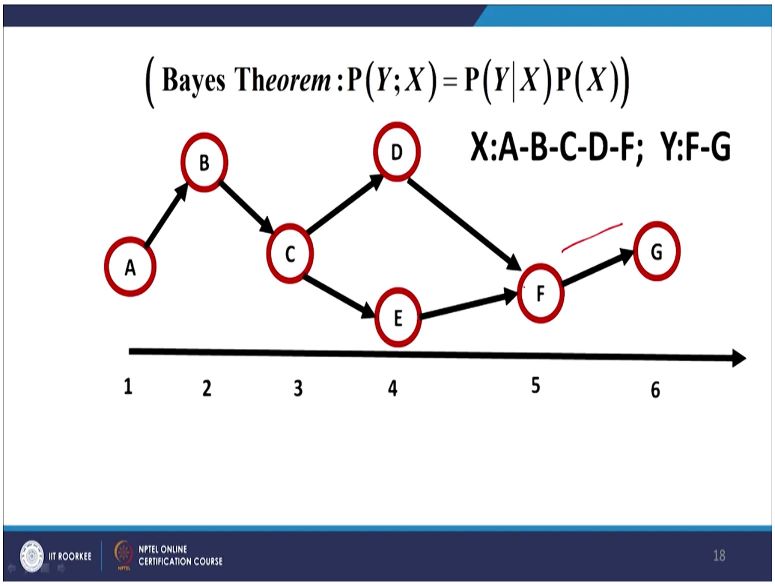


(Refer Slide Time: 16:17)




Now we come to a very important sub category or a class of stochastic processes which are called Markov processes. In the case of a Markov process what happens?

(Refer Slide Time: 16:36)




Let us look at this diagram in this diagram if you look at it is carefully there are two paths, two realizations; A, B, C, D, E, F and G that is path number one and the second path is A, C, E, F and G.

Now, the important point is, if in a particular stochastic process; the evolution from F to G is independent of how the process reached the point F, then it is said to be a Markov process. In other words it has no memory; it has, it does not know which path it followed in reaching the point F.

From F to G it is the evolution of the process from the point F to the point G is completely determined by the state of the process at the point F, it is not at all related to or influenced by how the process or what path the for a process followed in reaching the point F. For example, there are two paths by which this process could have reached the point F, but that does not matter in the end; how it evolves from F it is dependent on the state F, what the value that the process takes at t equal to 5 which is F. So, that is called a Markov process.

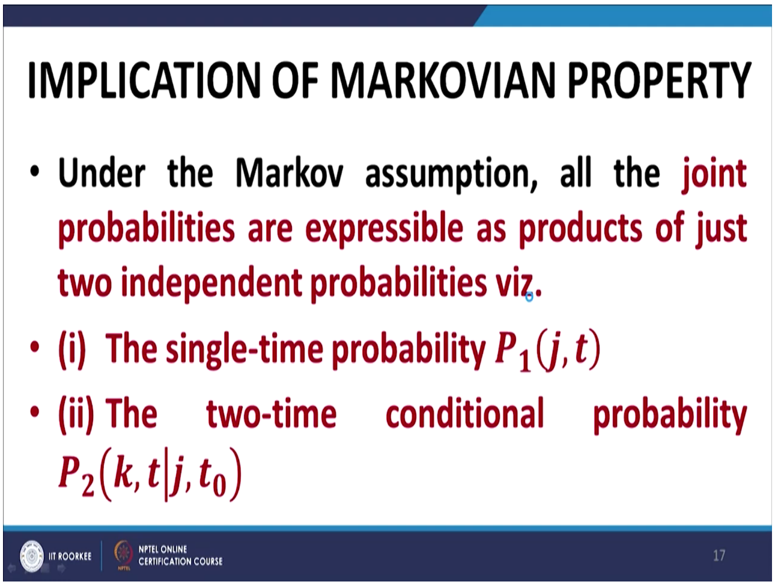
So, let us go back to the definition; for a Markov process we can say that, the future state of the process that is from F to G is dependent only on the present state; that is the state F and not on the history of the process, how or what path or it actually transgressed in reaching the point F or reaching the present state. And therefore its memory is restricted at any instant to the immediately preceding time argument and not beyond that.

At any point, we only are concerned with or the process is concerned with immediately preceding time point not beyond that; beyond that means earlier than that, earlier than that whatever is happened has no influence on the evolution of the future evolution of the system. The system has no history, it has no; it has history, but it has no memory, it about what actually happened. So, that is a very important class of process which is called Markov process.

In the case of Markov processes, you see now let us understand; stochastic processes can be you see. If a process Q is going to evolve in a few in the future in a random manner; indeed it is a collection of random variables, the random variables can take variable various values. So, the examples of a the variable X 1 at t equal to 1 can take a set of values in the sample set; the variable X 2 at t equal to 2 can take as a value from the sample set and so on.

So, there are multiple paths which can be followed by the stochastic process. The probability of a particular path is the joint probability of the random variables taking the respective values that are and compares to that are contained in the particular path.

(Refer Slide Time: 19:46)



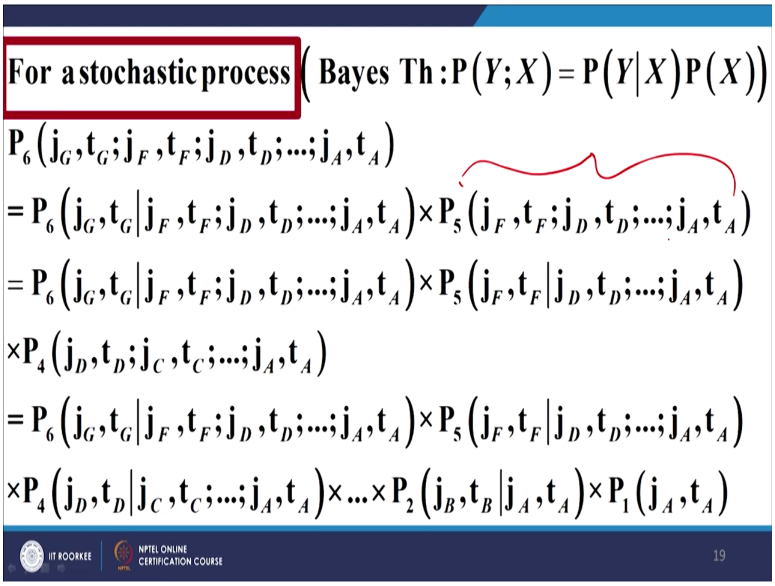

So, in other words the probability of a particular path being followed by the stochastic process is the joint probability of various random variables taking various values that constitute the stochastic process. Now in the case of a Markov process we can simplify that and the entire path probabilities can be expressed in terms of the one-time probability, P 1 j, t; and the two-time conditional probability P 2 k, t is subject to j at t 0. Remember P 2 k, t subject to j t 0 means; that the this is the probability of the path or of the system being in a state k at time t provided that or subject to the condition that it is in state j at time t 0.

Now, let us look at this, as I mentioned; suppose I want to determine the probability, I am sitting at t equal to 0 and I want to determine the probability of this particular path being followed by the system, by this stochastic system. Let us say the stochastic, let us say I want to determine the probability of path A, B, C, D, F and G. Now by Bayes theorem, let us say I split this into two parts; I write it as this thing A to F as an event X, and F to G as the event Y.

A to F I write it this, this path A, B, C, D, F, I write as event X; and F to G I write as event Y. Then by using Bayes theorem, what do I get? The probability of the entire path from A to G; A, B, C, D, F, G the probability of this entire path which is the joint probability, which is the joint probability, which is the joint probability of X and Y is given by the probability of Y subject to X. The probability of this, this movement subject to the system being at the point F, and the product of this with the probability of this particular path A, B, C, D, F.

So, in other words let me repeat, the path probability of the path A, B, C, D, F, G is equal to the path probability A, B, C, D, F multiplied by the probability of the system moving from F to G, subject to the condition that the system is at the point F. So, now, this can be iterated backwards and in other words the joint probability, the total probability of the path; joint probability of this entire path can be expressed in terms of the one point probability and the two point probabilities as you can see here.

(Refer Slide Time: 23:24)




Now this is this analysis is true for any stochastic process. We have got six points here. So, P 6 j G, t G; j and j F these are joint probabilities please note this; j G, t G; j F and t F and j D, t D up to j A, t A is given by using this Bayes theorem which I explained just now. We have the probability of the last step that is F to G and in into the probability, into the probability that the system is in state F. So, this is the probability of, this is the probability of the system reaching the point F; and this is the probability that being at the point F, it now moves to the point G.

This is the conditional probability; the first term is the conditional probability and the second term is the joint probability of one point earlier in the path, in the complete path. Now you use this recursively and you can break down the second term here, this term into again another conditional probability and a smaller joint probabilities. For example, we started with a six point joint probability, we now have a five point joint probability; we converted a four point joint probability and we end up with a one point probability and a set of conditional probabilities.

(Refer Slide Time: 25:07)

**For a *Markov* process**

$$P_6\left(j_G,t_G;j_F,t_F;j_D,t_D;\ldots;j_A,t_A\right)$$
$$=P_6\left(j_G,t_G \middle| j_F,t_F;j_D,t_D;\ldots;j_A,t_A\right)\times$$
$$P_5\left(j_F,t_F \middle| j_D,t_D;\ldots;j_A,t_A\right)\times P_4\left(j_D,t_D \middle| j_C,t_C;\ldots;j_A,t_A\right)$$
$$\times\ldots\times P_2\left(j_B,t_B \middle| j_A,t_A\right)\times P_1\left(j_A,t_A\right)$$
$$=P_2\left(j_G,t_G \middle| j_F,t_F\right)\times P_2\left(j_F,t_F \middle| j_D,t_D\right)\times P_2\left(j_D,t_D \middle| j_C,t_C\right)$$
$$\times\ldots\times P_2\left(j_B,t_B \middle| j_A,t_A\right)\times P_1\left(j_A,t_A\right)$$

Complete set of conditional probabilities multiplied by a one point probability or one time probability. So, for a Markov process, it can be simplified further; for a Markov process what happens is, the probability that the system is at state, the probability that the system is at state G subject to the system being a state F depends only on the, does not depend on the earlier path. So, it does not depend on this particular stuff; this stuff does not becomes irrelevant redundant when we talk about a Markov process.

So, this all the stuff goes out and what we are left with is and what we are left with is j F j G, t G subject to j F, t F. And you can look at it here j G, t G subject to j F, t F; and similarly the other terms can be simplified for example, j F, t F subject to j D, t D and these earlier terms will become redundant, right. So, we have j F, t F subject to j D, t D; the only the preceding argument is relevant, other arguments become redundant, because it is a Markov process.

So, what is happening? If you look at this carefully, these are all two-time probabilities, two-time conditional probabilities and this is the one-time initial probability; that is sufficient in the case of a Markov process to describe the path probability of a particular path.

(Refer Slide Time: 26:41)

**For any stochastic process $(GENERAL)$**

**$\left(\text{Using Bayes Theorem } P(B;A) = P(B|A)P(A)\right)$**

$$P_n(j_n,t_n;j_{n-1},t_{n-1};\dots;j_1,t_1) = P_n(j_n,t_n|j_{n-1},t_{n-1};\dots;j_1,t_1)\times$$
$$P_{n-1}(j_{n-1},t_{n-1};j_{n-2},t_{n-2};\dots;j_1,t_1)$$
$$= P_n(j_n,t_n|j_{n-1},t_{n-1};\dots;j_1,t_1)\times P_{n-1}(j_{n-1},t_{n-1}|j_{n-2},t_{n-2};\dots;j_1,t_1)$$
$$\times P_{n-2}(j_{n-2},t_{n-2};j_{n-3},t_{n-3};\dots;j_1,t_1)$$
$$= P_n(j_n,t_n|j_{n-1},t_{n-1};\dots;j_1,t_1)\times P_{n-1}(j_{n-1},t_{n-1}|j_{n-2},t_{n-2};\dots;j_1,t_1)$$
$$\times\dots\times P_2(j_2,t_2|j_1,t_1)\times P_1(j_1,t_1)$$

(Refer Slide Time: 26:43)

**If the stochastic process is Markov**

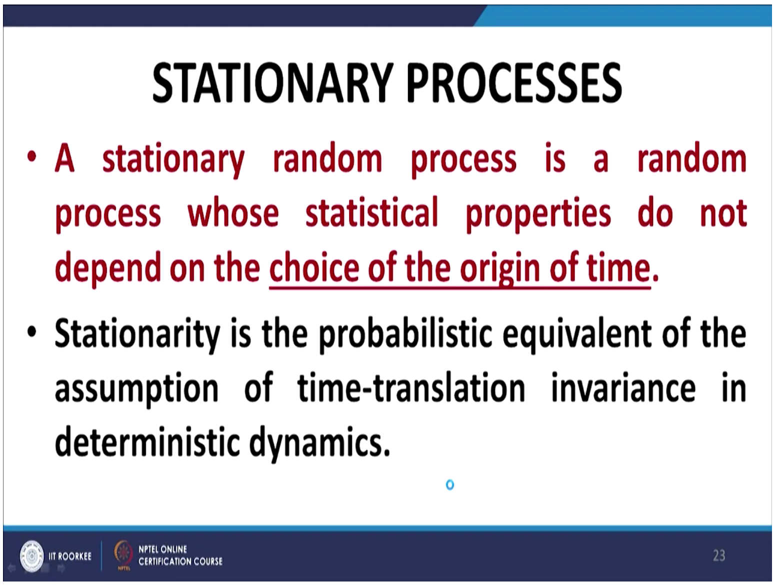
$$P_n\left(j_n, t_n \middle| j_{n-1}, t_{n-1}; \dots; j_1, t_1\right) = P_2\left(j_n, t_n \middle| j_{n-1}, t_{n-1}\right) \; \forall n \geq 2$$

**Therefore, for a Markov process**

$$P_n\left(j_n, t_n; j_{n-1}, t_{n-1}; \dots; j_1, t_1\right) = P_2\left(j_n, t_n \middle| j_{n-1}, t_{n-1}\right) \times$$

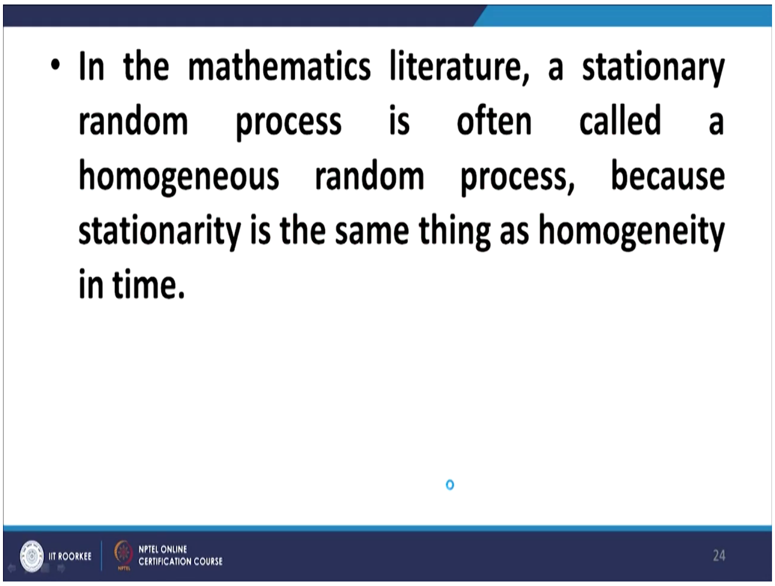
$$P_2\left(j_{n-1}, t_{n-1} \middle| j_{n-2}, t_{n-2}\right) \times \dots \times P_2\left(j_2, t_2 \middle| j_1, t_1\right) \times P_1\left(j_1, t_1\right)$$

(Refer Slide Time: 26:44)



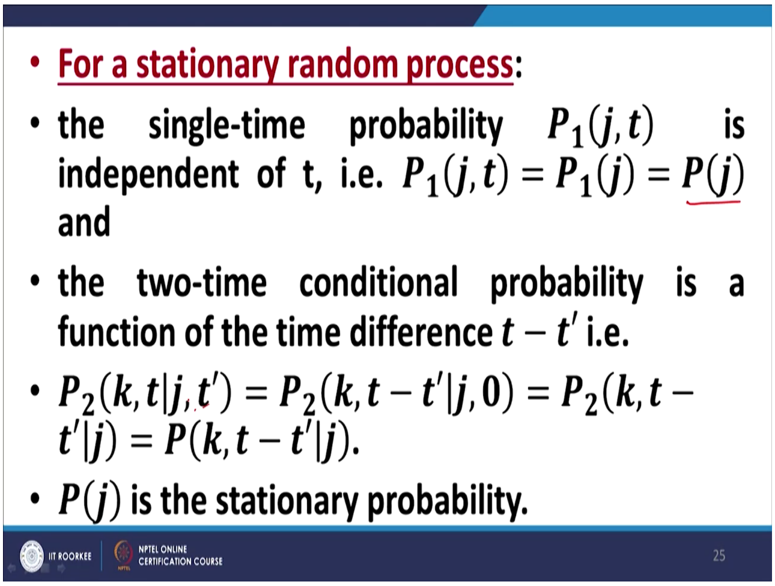

So, this is a generalization of what I have explained just now. Now we come to a concept of stationary processes. A stationary random process is a random process whose statistical properties do not depend on the origin of the point of time. What is the point taken as the origin of time, does not make any difference to this statistical property, properties of the system.

(Refer Slide Time: 27:19)




It is very similar to what you, what we have in the case of physical systems, where we have time translation invariance. And it is the statistical equivalent of that; but the implication is very interesting.

(Refer Slide Time: 27:22)

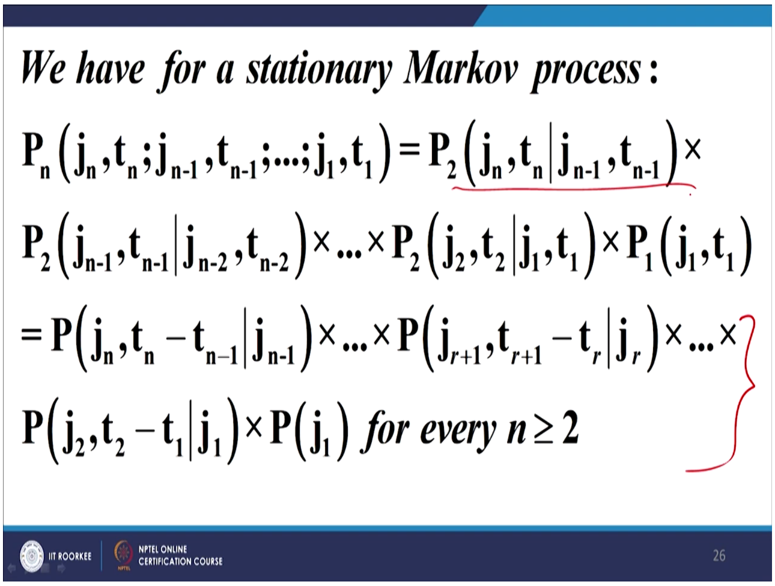



You know for a stationary random process what happens? P 1 j, t probability of state j at time t is if; we write it simply as P 1 j, because t becomes a redundant quantity, if it is the initial state, t would be 0. And so, we simply ignore it and we write it as P 1 j; even we ignore the one factor and as a matter of convention we write it as P j, P j means actually P 1 j, 0.

The two-time conditional probability now this is important; P 2 two point two-time conditional probability, k, t subject to j, t dash. You start the system in state j at time t dash, and you work out the probability of the system reaching the state k at time t. And that is given by P 2; because of time translation invariance, we can simply shift this t dash to, we can take the origin of time at t equal to, t dash equal to 0. This t dash we can shift as the origin of time scales and therefore, this t dash becomes 0; and when we make this shift of time, this t will become t minus t dash.

So, we have P 2 k, t minus t dash j, 0. So, that being the case, we write it as k, t minus t dash j 0 that is equal to P 2 k minus t dash j; because by convention that I mentioned earlier 0 is usually ignored and we simply write it as t minus t dash j, even this 2 is ignored and we simply write it as P k, t minus t dash j.

(Refer Slide Time: 29:13)

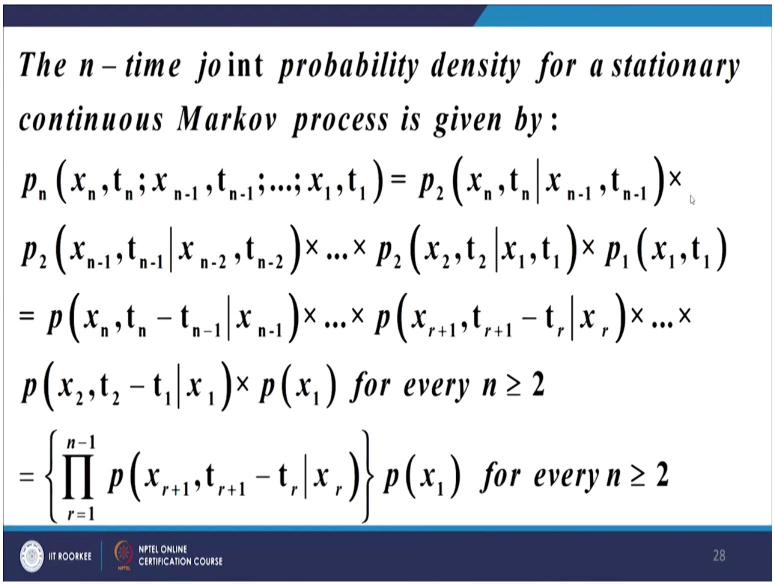



So, for a stationary Markov process there is a lot of simplification now we have; as I as in the earlier case, the Markov process can be expressed in terms of two point conditional probabilities and one point probability at the end at the initial probability.

Now, because of time translation invariance, because of time translation invariance; this whole thing gets simplified to this particular expression. I and just to illustrate one particular term by shifting the origin to t n minus 1, at the point t n minus 1 and the first term becomes j n, t n minus t n minus 1 subject to j n minus 1 0, which is simply written as j n minus 1.

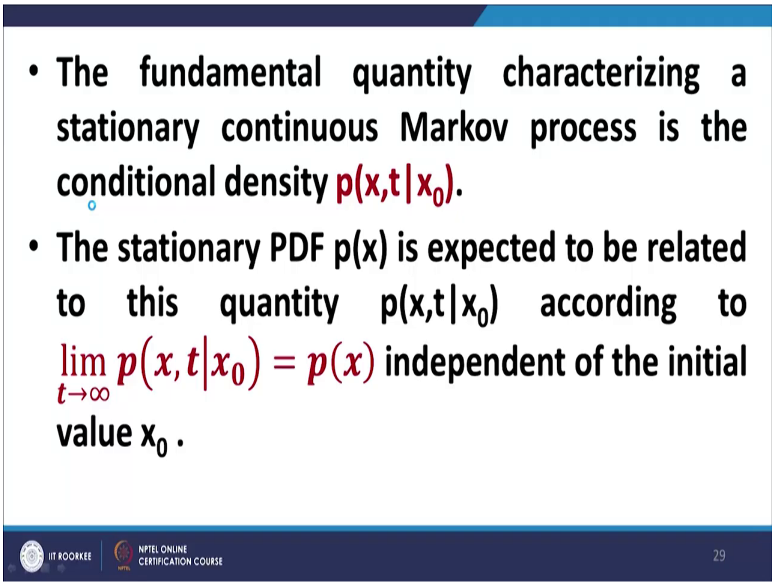
Similarly, the other terms are treated and we have this expression for the stationary Markov process path probabilities. Now in the case the important thing is, so far we have been talking about discrete processes; in the case of continuous processes, there is a slight variation, instead of having path probabilities, we work with probability densities. So, we make the substitution or we make the change; instead of using probabilities, we make use of probability densities, of course summations are replaced by integration, so that is another important part.

(Refer Slide Time: 30:48)



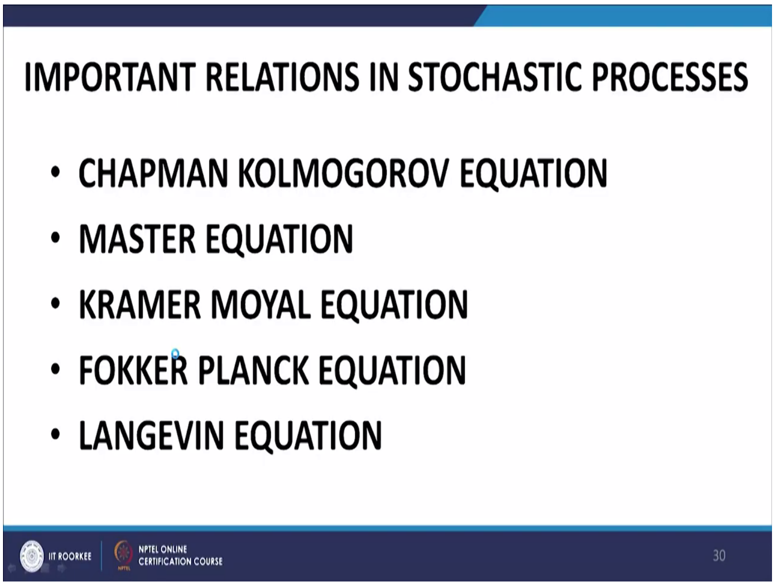

And therefore, the n time joint probability of a Markov, continuous Markov process, stationary Markov process, continuous process can be written in this form. And these are now, the small p's are now pdf's; they are not probabilities, they are probability density functions, the rest of it is more or less the same is. In the states are now written in terms of the small x is which are the values that the continuous random variable capital X can take and the rest is more or less parallel to what we discussed just now.

(Refer Slide Time: 31:21)



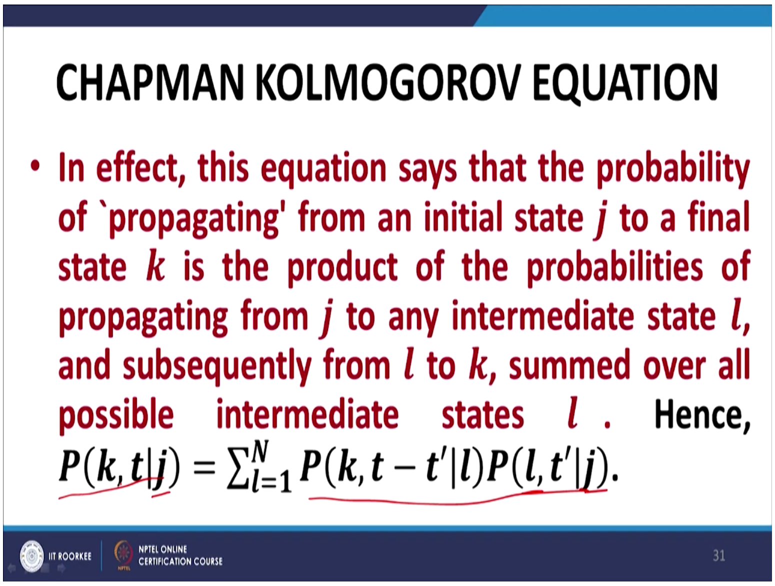

So, the fundamental quantity in this cases is the conditional probability p x, t subject to x 0; x 0 is the initial state that is the state at t equal to 0 and this is the two point conditional probability with p x, t subject to x 0. And in particular in most in many physical applications, we make the assumption that as t tends to infinity, as t tends to infinity, this particular conditional probability, the steady state is achieved. And a because of the steady state this, this conditional probability approaches p x which is independent of t; that means any extension of t further would not affect the probability once the system is settled down to a steady state.

(Refer Slide Time: 32:18)




Now, there are some very important results in stochastic processes; we got the Chapman Kolmogorov equation, the master equation, the Kramer Moyal expansion, Fokker Planck equation and Langevin equation. So, let us take them one by one, let us start with a Chapman Kolmogorov equation. This equation has a very important relationship with the path integral formalism; the solutions of this equations under certain conditions yield the path integral.

(Refer Slide Time: 32:45)



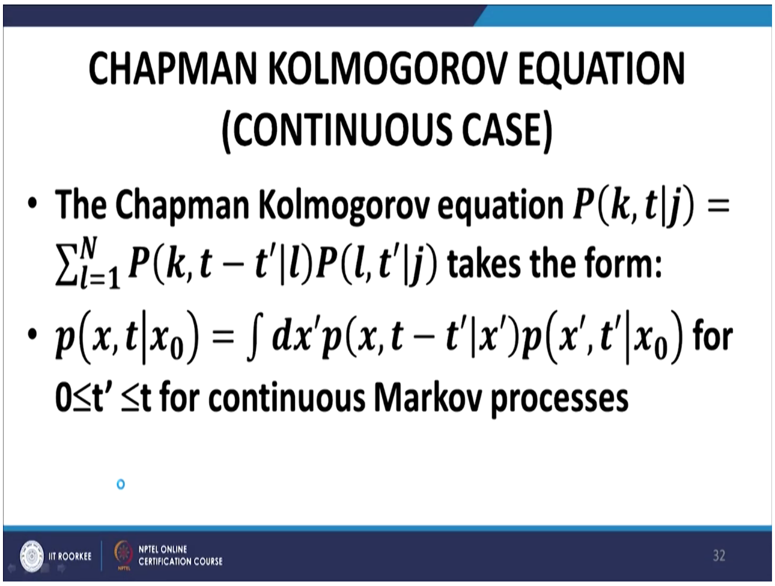

So, what is this equation let us look at this. This this is you, see you are given a certain initial state; let us say you are given a certain initial state j and k is the final state at the time t. There is an intermediate point in time, let us say t dash; j is the initial state, k is the final state at time t and there is an intermediate state at time t dash at any point time t dash.

Where the intermediate state let us say, let us call it the state l; then what the equation says, what this condition says is that, is similar absolutely parallel. If you recall the first lecture, this is something which is absolutely parallel when I introduced the concept of path integrals in the context of quantum mechanics and the Feynman two slit, a double slit argument following the Davisson Germer experiment.

Now what does it say, it says that, this probability P k, t is subject to j is equal to the product of the probabilities of the system moving from j to this state l, which is at time t dash; and then moving from there, moving from this state l to the state k in the remaining time t minus t dash. So, it starts with this state j, it reaches some intermediate state at the time t dash; and then from that state l, it moves to the state k in the remaining time t minus t dash. But, the this particular expression has to be summed over all the intermediate states, in other words l equal to 1 to N.

Whatever are the possible intermediate states l, whatever states l represents, whatever values l can take as the random variable all these have to be summed over, only then we can get this probability on the left hand side. So, this is absolutely similar to what we have for the path integral. The state, I repeat the state system starts from the state j and ends at the state k at time t; the probability of this happening is equal to the summation over all intermediate states of the system moving from the state j to an intermediate state and then jumping from that intermediate state to the final state k.

(Refer Slide Time: 35:53)



So, that is the Chapman Kolmogorov equation in the continuous case; it is pretty much the similar thing except that the probabilities are replaced by probability density functions and the summations are replaced by integrals. So, we will continue after the break.