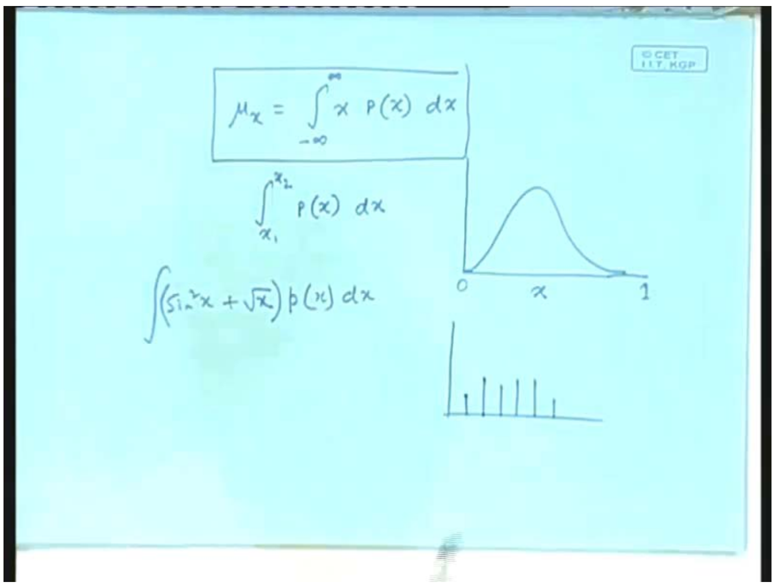
Chaos Fractals and Dynamical Systems
Prof. S. Banerjee
Department of Electrical Engineering
Indian Institute of Technology, Kharagpur
Lecture No # 29
Analysis of Chaotic Time Series (Continued)

In the last class two questions remained. One of you asked the question that when I was talking about the probability density function and said that you can obtain the average, you can obtain the mean square from the probability density function by the following expression. That is mu_x is equal to integral of the whole range so x px dx integrate over the whole range minus infinity to infinity.

(Refer Slide Time: 00:01:12 min)



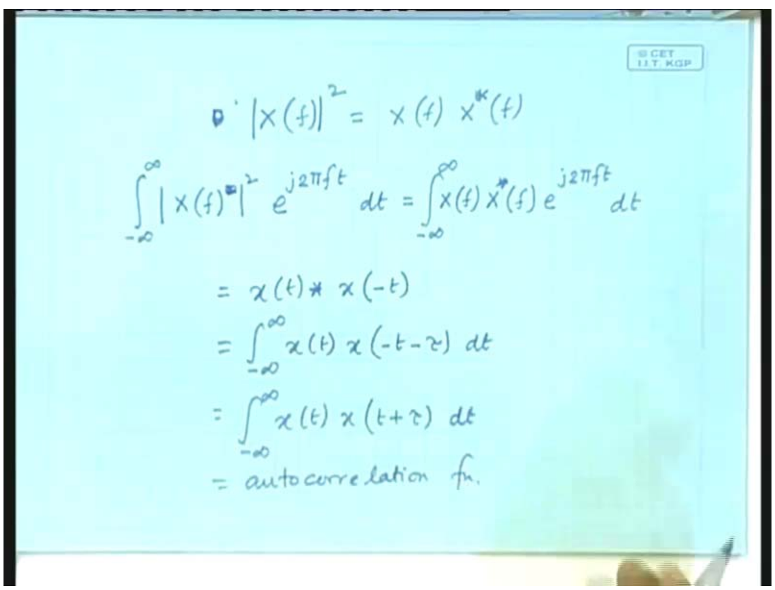
This gives you the mean value. One of you asked that this will have the dimension of a square. This will actually not because the way the p (x) has been defined, the probability of finding the state x within a certain range is given as integral of x_1 to x_2 which is the range p (x) which means that over the whole possible range this quantity will yield one. Because what is a probability of the state lying between minus infinity to plus infinity? It is one, it has to be inside that range. This integral yields one over the whole range. Over take over a smaller range it yields a number that is smaller than one.

Now in general the p (x) would be a continuous function for example you might say the p(x) is a function something like this which means say this is 0 and this is say some number say 1. For different values of x, it takes different probabilities and the area under this curve is 1. Now when we do this what exactly do we do? There is another way of looking at it. Suppose I want to find the average of some function like say sin square x plus something.

Some function of x and say I am asking what will be the average value, there is a random process going on giving me the variable x. Now what will be the average value of this term? In the first go, it looks very formidable problem. But actually once you have the probability distribution function it is no longer a difficult problem because all you need to do is for every value of x, you multiply this by the probability. Integrate over the whole range and that's it. All you need to do is to integrate this times p (x) and then it will yield. That is one of the strengths of having the probability density function by… (Refer Slide Time: 00:04:13).

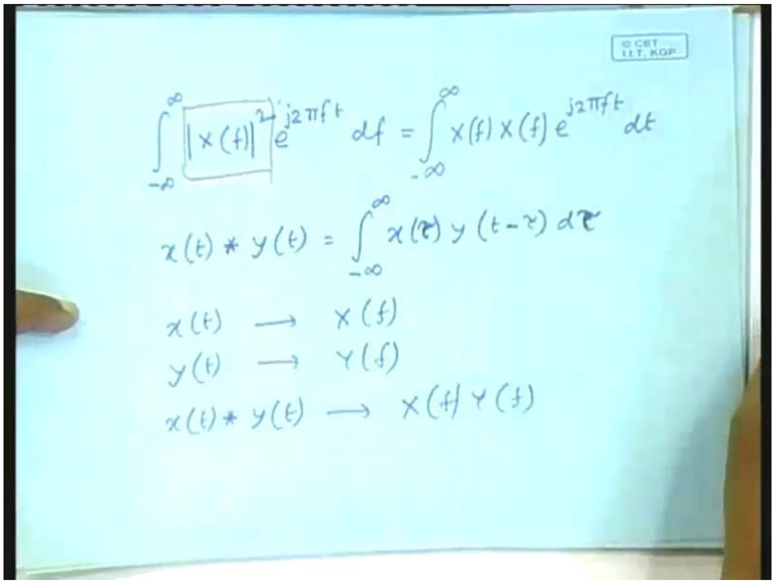
Now in this case what are we doing? Imagine for the sake of our understanding that this is not a continuous probability spectrum. It takes specific values. What is the probability of taking this value? This much (Refer Slide Time 04:32). What is the probability of taking this value? This much. What is a probability of taking this value? This much so on and so forth. Suppose it is given like this then what would this imply? It implies that take this value of x times this amount of probability. This probability is a number smaller than one. You get something times one, it's something times x, plus this value of x times this probability, plus this. You would notice that what it is yielding after all the summation that has a dimension of x. It actually yields and has the dimension of x_0 x square. The second question actually pertained to the way we were writing the power spectral density.

(Refer Slide Time: 00:05:33 min)



We said the power spectral density is X (f) mod square and then in the process of obtaining the theorem that we proved that day, what did we say? Do we have the pages there? Should be somewhere, yes it is here.

(Refer Slide Time: 00:06:08 min)



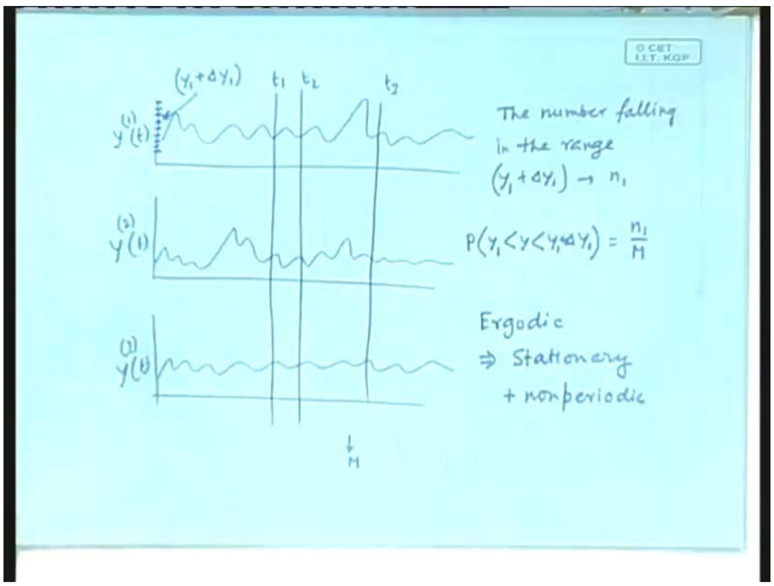
See we had written it as X (f) square, the power spectral density as X (f) times X (f). Now this is true only if X (f) is a real number, if it is complex then it would not be valid. Yesterday we did it. Basically under the special condition where the Fourier transform of X yields a real number but that would not be the general case of course. In a general case how would we go about it? We will say that this is equal to… (Refer Slide Time: 06:55) where this is the complex conjugate of this fellow. Then we have to go ahead with this formulation.

What actually did we do, our line of argument? Look at the line of argument that we used in the last class. We said that if you have two time functions x (t) and y (t), if you take their convolution then this is defined as this and there is a property that the convolution of two time variables is equivalent to the product of their Fourier transforms. We had used this property but now the difference will be here. I have to look for it but we will not look for it. We will substitute this. We had started from the equation minus infinity to infinity. It started from here. Now we will have to write it at X (f) and X star f. We have to write it like this. Then this term, see we are again using this property. The product of their Fourier transforms, if you transform that you will get the convolution. We will have to do exactly the same way but then we have to consider the situation, the inverse Fourier transform X (f) is X (t).

What is the inverse Fourier transform of X star f? X minus t that is one of the theorems that you may have learnt while learning about the Fourier transform. Essentially what you get out of it is x of t and this is a convolution x minus t. Now this convolution if you write as the formal of our convolution, you will get minus infinity to infinity. This is x (t) times x minus t minus tau. At this stage, let me only comment that under certain conditions this is same as… That means x of minus times something is same as the x of plus the same time that will be through under certain conditions, I will come to that. That means under the condition in a shift of time, the character of the time series does not change in a statistical sense. No, actually it will not be. The issue is that different books give different definitions of this function and in some cases they take the t as the variable, in some cases they take the tau. You have to interchange these two. I mean you might as well interchange these two so that to bring this form. Essentially it is the same thing. I mean it will not be a big different if you interchange t and tau. This is same as the auto correlation.

Now this obviously brings the question under what condition this can be written. If it cannot be written, we have some difficult in writing like this. We will come to that question now. The point is so far we have been considering only one time series. A time series that is coming from some kind of a process going on. Now in practical situations there could be more than one time series available at the same time. That means that would be called on ensemble of time series. One time series is coming like this say and other coming and third one like this and so on and so forth.

(Refer Slide Time: 00:12:57 min)



Now what do these ensemble mean? You might imagine these as generated from the same process or in other words, if you have in your mind something like the Lorenz equation which you have come across, the same Lorenz equation but starting from different initial conditions. They will lead to different wave forms and in that sense they would be different time series and as a result you have an ensemble of time series. Now there are various situations where it is convenient for us to have only one time series to be able to talk confidently about the character of whole ensemble.

In some situations you have a finite length of an ensemble of time series but you won't to be able to talk about a large length of a single time series. What are the situations? One situation you have already come across. For example while talking about the Frobenius Perron operator, what was our essential logic? Our essential logic is that if we have ensemble of series, their property under certain condition would be the same as the property of a single time series even we cannot wait for an infinite amount of time to obtain the time series. We can still talk about the whole time series with confidence but that happens under certain conditions. In order to understand those conditions let us look at it this way. Suppose this is y, I say one of t. This is y_2 of t, this is y_3 of t and so on and so forth. I am putting it in bracket, otherwise it might look as y square and y cube so that should not be confused.

You have an ensemble of time series and we are trying to understand the property of this ensemble. Now we might say that I identify a particular time chunk and we might say that now we have not a chunk but just a t_1 say this is the time. We might divide the whole range on the y that means y is contained within specific range. You might divide into small boxes, the way we have already done. Say this particular box is y_1 plus delta y_1. Here is y_1 plus delta y_1, delta y_1 is the length of the box. Then at time t_1 we can count the number of the ensemble that fall within this box. Notice here we are counting across this, not counting across that. We are counting across the ensemble that means we have identified the time and saying that out of the whole number, how many are there. Suppose m number is there. Total m number of ensembles are there. So out of the total m say a small number n_1 was within this range y_1 plus delta y_1. So… (Refer Slide Time: 17:30) is n_1. Then we might say the probability of finding a member of the ensemble in the range y_1 to y_1 plus delta y_1 would be n_1 by m, at time t_1. We will say probability of y_1, y, y_1 plus delta y_1. So y to fall within the range of y_1 to delta y_1 probability of that would be n_1 by M.

So far simple stuff but now if n is large that means you have a large ensemble available and we take very small delta y then obviously we can calculate that for the whole range. What does it yield? It yields a probability density function at time t_1. If we take large value of m that means large number of ensembles and if we take small delta y which means it becomes smoothened into a continuous function. And what does continuous function then imply? It implies that at t_1 what are the relative probabilities of finding a member of the ensemble at different values of y but at t_1, remember that. Likewise we can calculate that at time t_2, another times t_3 and so on and so forth.

We can calculate a probability distribution function over the ensemble as specific times. Now if this probability distribution function over the ensemble does not change in time then the time series or the ensemble is called a stationary ensemble. This is the definition of something stationery. If it is changes with time then it would be non-stationery. The easy way of visualizing when a time series will be stationary and when a time series will be non-stationary is suppose you start from initial condition and attractor is here. After some initial oscillation or initial transient it goes into the attractor. If you take the time series before it reaches the attractor, it will be non-stationary it will change with time. But after at this though attractor it will be stationary but also the attractor could be a point attractor, a periodic attractor, a chaotic attractor, quasi periodic attractor all these possibilities are there. A stationary time series could be of all the different types.

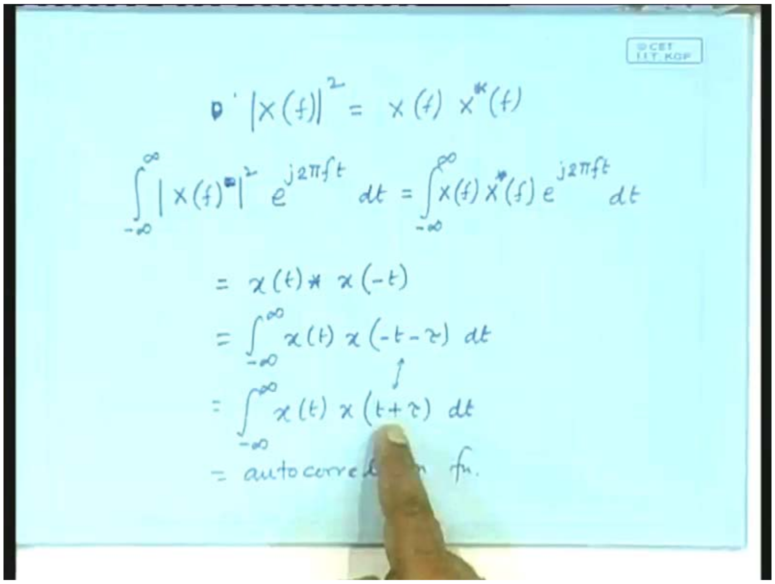
If a time series is stationary but its elements are not periodic then it is called ergodic. Ergodic means stationary. Why is this definition important? Because when we say that in some situations I may have a single time series over a long time but I want something of interest about the whole ensemble. In some situations I have a large number of an ensemble available but I want to infer something of interest about a long time series. These two different concepts will be equal, if a system is ergodic. That is why the importance of ergodicity comes. Practical importance of ergodicity that in some cases I want to infer something of interest about the average behavior of a system. For example whichever initial condition I may start from. The average current flowing through an ammeter will be this much. These are very practical considerations.

I want to know how much will be the deflection of the ammeter pointer. The actual current flowing through the ammeter coil would be a chaotic sink but ultimately you will see something. How much is that? In order to calculate those, you need the concept of ergodicity. If a system is ergodic that means it is stationary and non-periodic, obviously that will include within its fold quasi periodic orbits because quasi periodic orbit is stationary. It has already reached the torus and it is moving inside the torus. Obviously it is stationary and it's also non-periodic and that is why a quasi-periodic orbit is also ergodic. Not only the chaotic orbit but also quasi periodic orbit is ergodic.

There is an important thing associated with the concept of ergodicity. It's called mixing. Suppose there is a state space, there is a chunk of the state space inside the attractor and the attractor is ergodic. The moment I prove that it is ergodic, it immediately implies that starting from any initial condition it will visit neighborhoods of every other points in state space within that range. That means if I start from any point, I can target another point and I will reach there with confidence I can say that I will reach there some time or the other. In one of the future classes I will show you how to take advantage of this concept in practical conditions, practical problems. Ergodicity is this that there is mixing. The concept of ergodicity is used in say industrial processes for example where you want mixing. say a chemical reaction is going on, you want proper fast chemical reaction that means you want proper mixing of this state's and that can be ensured if you infuse ergodicity in the system. Essentially it is done by infusing chaosing system.

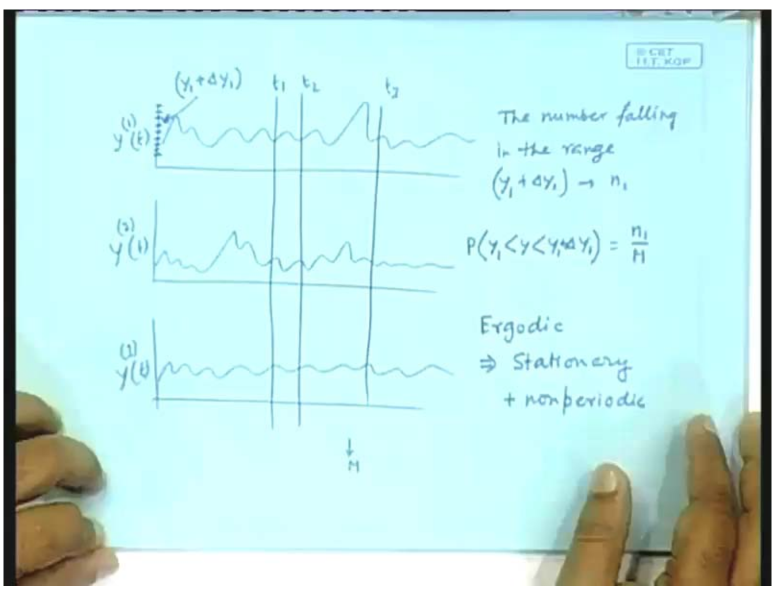
I told you about the motion on the torus, it could be a periodic orbit, it could be a quasi-periodic orbit and you have understood that the quasi periodic orbit is an ergodic but the periodic orbit is not. That is why in some literature when you want to go back to books and want to study, you will find that in some cases the periodic orbits occurring on a torus is called a resonance torus. That means two frequencies are in resonance. That torus itself is called a resonance torus and the situation where the two frequencies are incommensurate, as a result the orbit is quasi periodic it is ergodic. So that torus is called an ergodic torus. Remember these words so that if you encounter these words in any literature, don't be confused. An ergodic torus means quasi periodicity, resonance torus means not just any periodic orbit but periodic orbit happening on the surface of a torus, mod locked behavior.

(Refer Slide Time: 00:26:29 min)



Now refer back to this problem. You would immediately realize that these two are statistically true in stationary processes, ergodic processes.

(Refer Slide Time: 00:26:47 min)



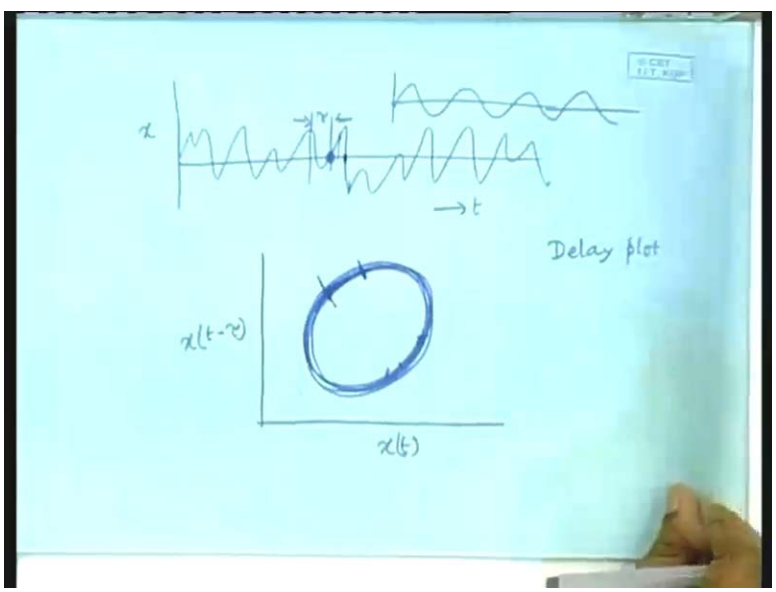
This particular theorem is particularly applicable to ergodic processes, chaotic processes, quasi periodic processes where you want to use this in order to either obtain the auto correlation function or you want to obtain the prospected density.

Let us come to another issue now again related to time series. There are a huge number of situations, practical experimental situations where it is only possible to place one probe in the system. Imagine that a chemical reaction is going on. That means there is a huge amount of things going on in different places and there are adequately large number of state variables in the system and if there is a mixer in mixing the thing, there will be a motor. A motor means inductances, a motor means windings, motor means current there are also state variables. There are huge number of state values in systems, hopelessly large to observe all of them. Similarly there are often situation where it is hopelessly inadequate to observe all the state variables. For examples say weather, how many state variables are there? Infinite or where you can at least say that it is a finite dimensional system but there are too many state variables to observe. Say the power plant, a huge power plant that means there will be something like a thousand state variables but you do observe.

I am now setting up a situation where there is some kind of experiment but logically speaking there would be a very large number of state variables and if you want to model that say a convection process, say the process where slowly turbulence develops in a fluid. You need to really measure the state at every places, the motion, the pressure and stuff like that at every point and ultimately then only you are able to write down some equations. In such situations there are some state variables that you can possibly observe but that is finite in number and less than the actual number of state variables that you can logically identify.

One way to represent such systems would be to start from the basic first principles then try to write down the equations and try to develop the model but in many cases that is a hopeless pursuit because the system is so very complicated. Now much of science has developed on the idea that even if it cannot write down the exact model, we can still mathematically represent this in terms of relatively simpler models. What often people do is suppose you have got just one state variable coming and you have the time variable with x. Now from it how much can we infer about the character of the system?

(Refer Slide Time: 00:30:56 min)



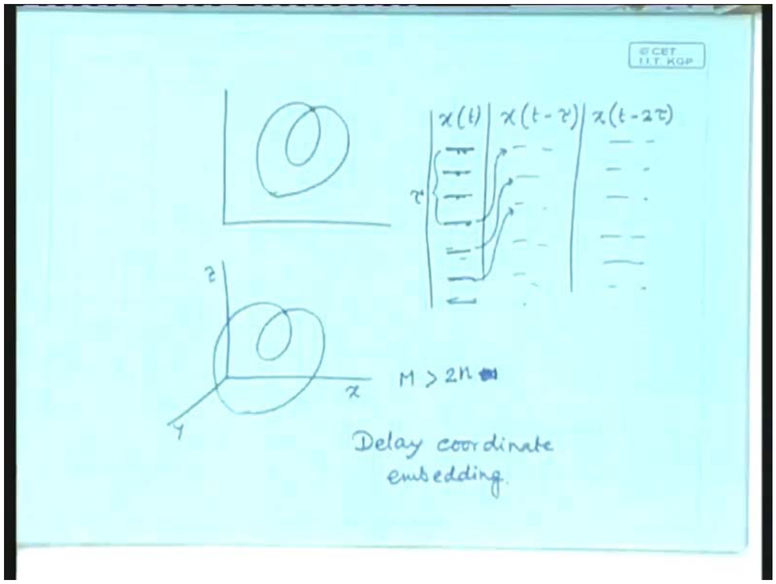
The moment you say that what do you mean by the character of a system? What we really want to do by writing a model? We want to predict if the system is doing such a thing now, what will it do say 10 seconds later? That is what we want to do. From this kind of a waveform, can you predict that. Say I am here, I observe the variable value and say right now it is 0. Then can I predict what will happen say once seconds later? Obviously we cannot, because here also it is zero and here also it is 0. If this much is one second later, after that it went up. If this much is one second later after that it went down and therefore simply saying that the state value is 0, we cannot infer anything of interest. We cannot really say what it's going to do in the future. Scientist, especially the nonlinear dynamics people has formulated an extremely useful method to overcome this problem. What they do is, it is not difficult to see that the difference between this point and that point. What is difference between this point and that point.

If we plot in a two D state space, they will be seen as different because you have drawn it in one D, we are not being able to distinguish them. So you need it somehow draw a two D state space. There is still something but the problem is that you have actually accessed only one variable, not two. How to construct the other variable? The nice way is what is known as the delay coordinate, delay plot. Delay plot is where, if define a given amount of delay say from here to here tau and plot x (t) minus tau in the y axis. Now this delay could be advanced or backward both, the difference between the two is what really matters which is tau.

If this is a periodic waveform say it had been suppose like this. What will be the result of the delayed thing? It will actually be like this which means that this is generated from a two dimensional system with a closed loop. That is clearly seen, the moment you make a delay plot.
You might ask how much would I take tau. It doesn't really matter mathematically because whatever the tau, the topological character of the orbit that you get on this two D plane will be the same. Topological character of the orbit will remain the same. Only if tau is too small, it will be almost like a flat thing and the more you increase tau it will become more and more roundish.
It becomes easily seeable but mathematically it is still a closed loop, whatever the value of tau. That is why, often we said the value of tau simply by visual impression.

For different values of tau, we generate a 2 D time series and plot it and see whether it looks very nice. What exactly do you mean by very nice? Here there is a point and here there is a point, how far can we distinguish these two? The more it is roundish, the more easily it will be possible for us to distinguish. Then you see from here up to a point it goes here, from here up to a point it comes here. We can distinguish. Then in the absence of the exact equation, how do we really do the prediction? This is often done by simply analogy means if the state were here, there is a large time series available so essentially you have something like this. If you go on plotting, it will something like this. If the state is here, from one of the points you know that if the state were here where it went one second later because that is already there in in the time series. Then by analogy you can say that if the state now reaches here, once second later it will go there and for that it is necessary to distinguish between all points in the orbit. In order to say from analogy, it is necessary to ensure that all points in the orbit are unique.

(Refer Slide Time: 00:37:31 min)



Suppose after having done this, you end up in an orbit like this. What would you infer? The inference is obviously that here we have a problem because we said that in order to predict by analogy, we will need to have each point unique. This point is not unique (Refer Slide Time: 37:59). This is a cross over so we need to do something in order to remove this crossing. Logically it cannot have this crossing. We had generated the time series by saying, there one is x of t. This is one set of data that is what we had. We had started with then we said that x (t) minus tau is another set of data, generating this. The moment you have intersection, you know there you need to increase the number of coordinates that you generate. You do it as another set of data which means what are you doing? Suppose you have got a data's data file in which x (t) are plotted, x (t) are given. So there will be one line here, another here, another here, another here so on and so forth. This is x at time 0, the time at next observation instant here.

For example if this is generated by some kind of a data acquisition card then this time difference relates to the speed of the data acquisition card. Suppose we have identified tau as four lines of this. This is our tau suppose then what we will do? We will put this point here, the next point there, the next point there and so forth. Thereby generate a time series and plot.

Similarly here, the moment we do that this no longer remains a cross over point or in other words even after doing so if it is a cross over, by slight tiny perturbation you can eliminate the crossing. Essentially this becomes two distinctly different points and therefore by using the same analogy technique, you can from the time series itself obtain sufficient information to be able to predict, what is going to happen. Now there is an important message here. The important message is that the system itself may be eight dimensional system. You are representing it in 3 D. What is the guarantee that it is a useful, meaningful representation? That is a serious theoretical question. Is it a meaningful representation? It would be a meaningful representation if somehow we could prove that for each point in the original eight dimensional orbit, there would be unique point in this orbit. Then only it would be mathematically meaningful. Listen let me repeat again. I cannot really draw eight dimensional thing.

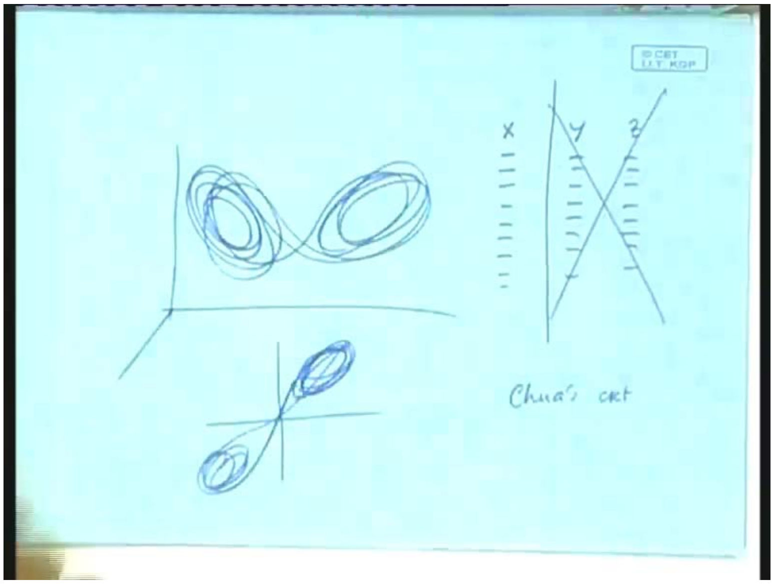
So try to understand. The actual system is eight dimensional say that means at every instant eight state variables are being generated and you might imagine some kind of an orbit going on in the eight dimensional state space which is undrawable. But now we have drawn something by accessing only one state variable out of that. Out of all the eight we have just taken one data and we have generated a state space something like a fictitious state space by this delay coordinate method. Is it useful? There exists a theorem due to… which says that yes, in this case of one dimensional object, so long as the object that you are trying to represent is say n dimensional. If you are drawing just a curve, it is a one dimensional object but you might also like to draw a two dimensional manifold or a whole attractor. Whole attractor means what is the dimensional of that object?

It will be a fractal dimension but suppose that whatever it is, that dimension is n of the object. If you choose a number of time series like this of dimension 2 n +1 then the theorem guarantees that each point maps into a unique point. Which means that if you are just trying to represent an orbit, orbit is a one dimensional thing then you need just three dimensions. If you plot it in three dimension that means you generate three time series, it is sufficient. If you want to generate a unique surface, surface means 2 D thing then you need five dimensions at least. If there is a chaotic attractor which means the attractor has a dimension.

Suppose you have measured the dimension and you have found that to be 2.8. Then what is the minimum dimension you need in order to represent that in this delay coordinate method. 2.8 means 2 n, let me put it this way, plus one is minimum, just to be greater than 2 n. 2.8 which means it's twice is 5.6. The whole number greater than that is 6. 6 is sufficient which means this technique by which you generate pseudo time series which actually does represent the real thing in the sense that if the real thing becomes a periodic orbit, you also get a periodic orbit here. If the real orbit is period two, you also get a period two orbit here. If the orbit is period three, you also get a period three orbit here. If the orbit is quasi periodic, you also get a quasi-periodic orbit here, chaotic also you get a chaotic. Means the exact property of the orbit is uniquely represented by means of this. This is called delay coordinate embedding.

The dimension that you need in the minimum to represent that particular dynamical system is called the embedding dimension. It may not be, it is not the same as the actual dimension of the system. This means the minimum number of dimensions that you need in order to properly represent the dynamics of the system. It will be nice exercise for those of you who have obtained the Lorenz systems. You have the equations and you also obtain this orbit.

(Refer Slide Time: 00:46:22 min)



This kind of orbit you have obtained. Now the way to taste what I have just said is this orbit if you have generated in mat lab and plot it, there was a stage in which you have generated the time series and then given the plot statement. After you generate the time series simply eliminate the vector two and vector three. Retain only the first vector. What you have obtained? This is the three D plot, so you have x y and z, x had this numbers coming, y had these numbers coming and z had. That is what you have plotted here.

Suppose you eliminate this and keep only this. Choose a delay, how much should the delay be? This is not difficult to see that should be of the same order of magnitude as the period of rotation of these. That you can easily identify and accordingly choose an amount of delay. If it is too short in comparison to the period of rotation, it will be almost like a very thin thing along the 45 degree line. No in the Lorenz model you had… no, Lorenz model did not relate to voltages. Lorenz model was atmospheric convection model. You are talking about probably the chua's circuit. Chua's circuit, fine no problem. You had the differential equations available to you, solve it and you can always plot the behavior as something like this.

In case of chua's circuit it is a double scroll like this, it will come like that fine no problem. This is actually a three dimensional system, there are two capacitor, capacitor of voltages, one inductor, inductor current. These are x y and z so x variable vc_1, y variable vc_2, z variable it is only one l. Generate that and eliminate that, keep only this. Now give a suitably chosen tau, the delay length. Suitably chosen means you might play with it. No problem, I take 10 steps as my tau let's see what happens. Take 50 steps let's see what happens. Take 100 steps let's see what happens.

Essentially that will be related to, you have to take a tau that is of the same order of magnitude as the period of rotation around this lobes or here. Then if you do this that means you regenerate y and z by means of the delay coordinate and then plot it, then you will be able to convince yourself that it really generate the same orbit. Not exactly the same orbit but topologically the same orbit. If your problem is to identify when a bifurcation occurs, when a period two to period four change over happens. Obviously the surfaces. Because a periodicity is retained, just from one having obtained the continuous time state place you might place a Poincare section. You might obtain the discrete type description no problem. Yes you can do that.

In fact a large number of experiments have been done in which the bifurcation diagram has been obtained by this. Simply take one state variable, obtain the continuous time dynamics by delay coordinate plot, place a Poincare section, obtain the discrete time description and then on that basis plot the bifurcation diagram. That is how most of the experiment work in the field of nonlinear dynamic system because in most cases unlike the cases in electronic circuits where you have two variables, three variables, easily indefinable variables. In most practical system it is not and that the delay coordinates plotting is absolutely a must. Let’s stop here.