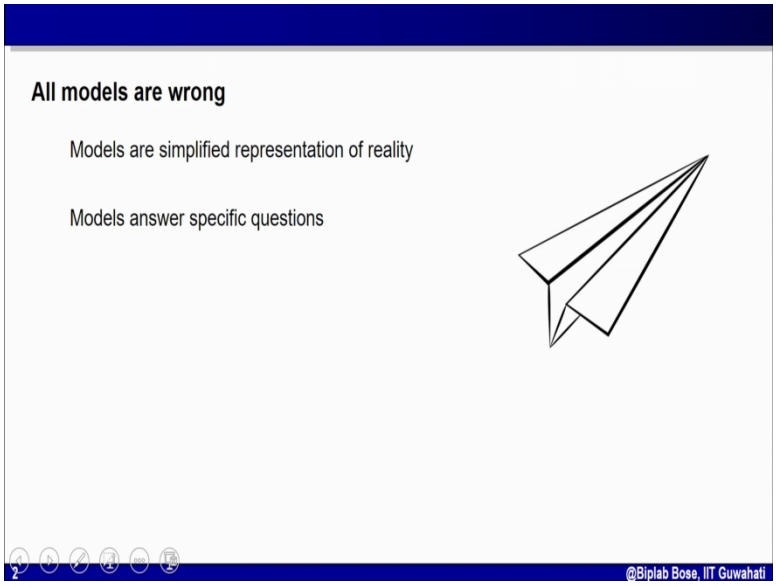
Introduction to Dynamical Models in Biology
Professor Biplab Bose
Department of Biosciences and Bioengineering
Indian Institute of Technology Guwahati
Module
Lecture No 2
How to start Modeling

Hello, welcome to module 2, week 1 of our course on Introduction to dynamical models in biology. In this module we will discuss about essential features or basic concepts that is required before you start making a model of a biological system. So let us start with some basic concept, while making a model we have to keep in mind that no model is actually exactly same as the reality. Every model you make is actually simplification of real phenomena, and the second key issue we have to keep in mind is that as every model is a simplified version of a real process or real phenomena, it cannot answer each and every question. So every model can only answer a specific set of questions that you ask to the model.

(Refer Slide Time: 1:45)

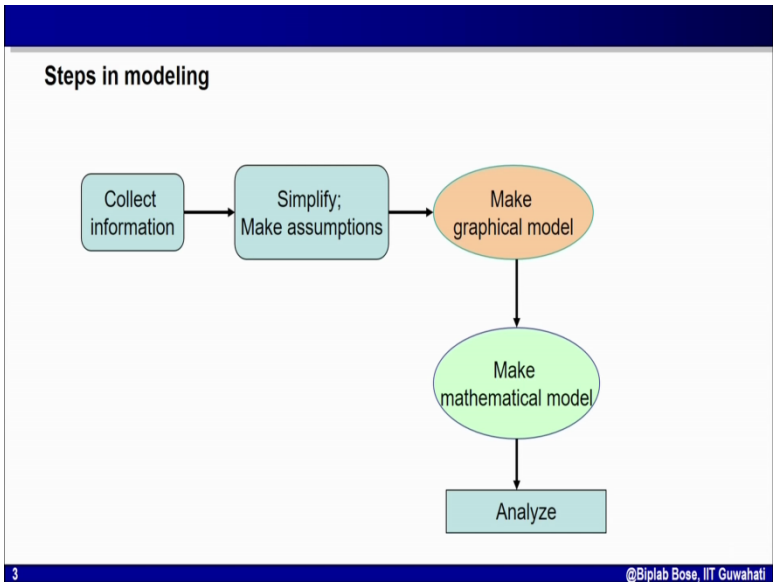


Let me explain this with a simple example. Suppose I want to give you the direction from airport to our institute and I am drawing map on a piece of paper. So obviously I will not draw each roads and by lanes of the city, I will not show each and every house, shops all around the road from airport to our institute, what I will do I will only draw the main street through which you should reach our place and also some important places which you can use as marker or flag. So the map that I am drawing on a piece of paper is actually a partial representation of reality and it is almost like a model of a city and the question that it can

answer is, it can only tell you how and through which route you should reach our institute from the airport. If you ask the map, how to go to the stadium from airport, obviously that map will not be able to answer that question.

So while making a mathematical model of a biological phenomena you have to keep in mind that the model that we make will be a simplified version of the reality or the real phenomena or the real biological process that you are trying to model and you can only ask a specific set of questions to it, so that it will be able to answer. So let us look into the steps in mathematical modelling particularly dynamical modelling in biological system.

(Refer Slide Time: 3:11)



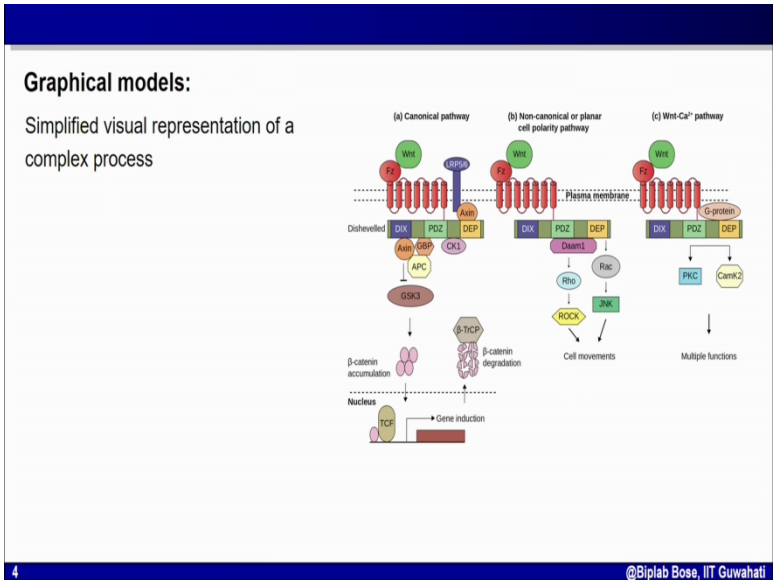
The first and foremost requirement is that you have to collect information, information about the process that you want to model. Suppose you want to model the dynamics of signalling by insulin in our cells, so you have to collect all the information required for modelling it. So you have to collect the information about the molecules involved in that signal transduction pathway, you have to know the processes involved so where do you get that information. You get that information from published literature, from other people's model, maybe from some databases or maybe you will be doing some experiments yourself to generate some data which can be incorporated into your mathematical model.

Once you have collected all this information, you try to make a some sort of simplifying assumptions. You make some assumptions so that your model becomes simplified, we will discuss about this simplifications and assumptions, when we will discuss about specific mathematical models in the subsequent modules. Once you have made those assumptions, the

next third steps comes, you make graphical models. This is where you make a visual representation of the problem that you want to sort out through mathematical modelling.

I will discuss about this graphical models in details in couple of slides. Once you have made the graphical model the fourth step is making the mathematical model that you essentially want to create. These mathematical models will be based on the information that you have collected at the beginning, and then based on the assumptions you have made and obviously the graphical model will help you to create those mathematical models. Remember, mathematical models are nothing but a set of equations that you write and we will discuss in details how to write down those equations. Last part of mathematical modelling particularly dynamical modelling in biology would be analysing that model.

(Refer Slide Time: 5:35)

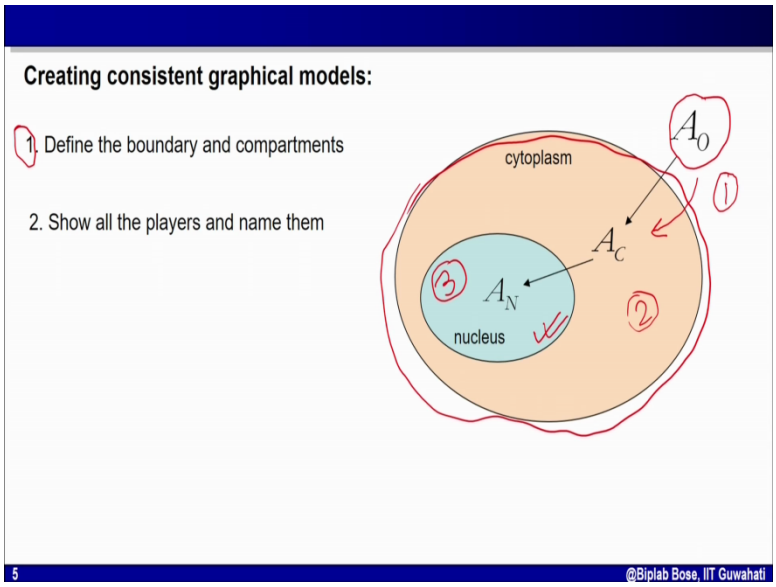


So let us start with graphical model, it is not essential for every modelling you create a graphical model, but usually I advice that you try to create a graphical model first so that you will not miss any particular property, any particular process or any particular molecules or any elements involved in the process while making the mathematical model. Essentially, graphical model will be a map based on which you will write down the equations for your mathematical models and graphical models are not something new.

As a biology student, we all are very accustomed to graphical model, most of your text books starting from bio-chemistry text books to molecular-biology text book; you will find this type of graphical model. For example I have shown one model here, a signal transaction path-way model; this type of model of graphical representation is very common in biology. We have to

remember most of the biological processes are very complex; explaining that by words to someone, by text to someone sometimes is very difficult. It is always easy if we make a plot or a diagram or a network map something graphical, which communicate the information properly to the other person. The same thing is here, you make a simplified visual representation of a complex biological process that will help you to create the mathematical model.

(Refer Slide Time: 7:26)



So when you are creating graphical models, we should try to follow some consistent rules for that, otherwise we may miss some critical issues in the model. First thing, the first critical point is that you define the boundary and compartment of the system that you are dealing with. For example, this diagram I have drawn here, suppose I want to simulate a model and simulate a process, which involves a particular process in a cell. So a cell can be represented by an external boundary line like this as I have drawn by the bigger circle, inside that I have shown the nucleus. What I am trying to model here? I want to model a process by which a molecule A gets transported from outside into the cytoplasm and then from cytoplasm it enters into the nucleus. So in this case I have to show three different compartments, the first compartment is outside the cell, second compartment is my cytoplasm and the third compartment is the nucleus.

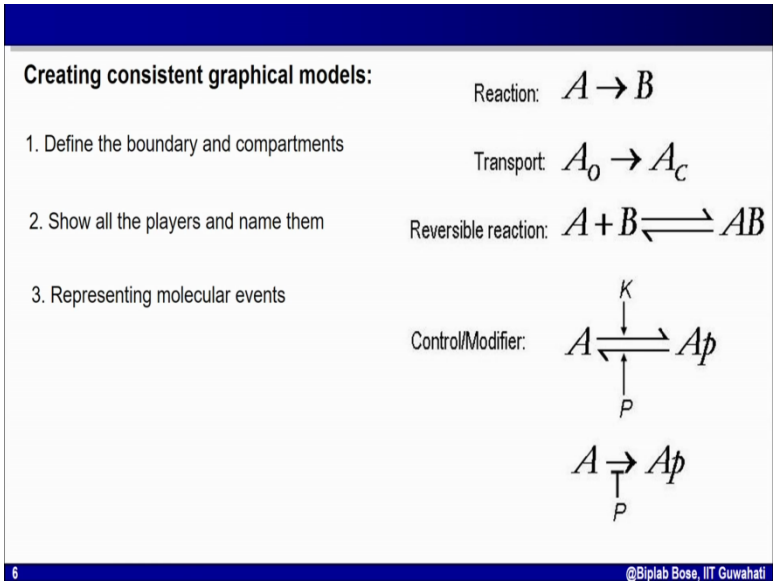
In this particular case I have to show the nucleus separately because the molecule is entering into the nucleus from cytoplasm. But in some other cases you may not have to show nucleus as a separate compartment because there is no particular process going on in the nucleus which you want to model. So in that case you have not to show the nucleus as a separate

compartment, so you have to judiciously identify which are the compartments involved in your process and what is the boundary of it. The second point is that you have to identify all the players involved in the process that you want to model. For example in this case, this example that I have drawn here as a graphical model, the molecule is present initially outside as I have marked it as A_O, means molecule A present outside the cell, it enters I have marked the same molecule when it is in cytoplasm as A_C to remind you that this is the A molecule but it is present in cytoplasm.

The same molecule when enters into the nucleus, I have marked I am marking that as A_N representing that it is A molecule in nucleus. This type of naming the same molecule by different name when they are in different compartment has a particular advantage because now I can count and keep a track how many A are present outside the cell because I will record the number or the values of that as the value for A_O. I can also keep a record of how many A_C are there that means how many A molecules are in cytoplasm and I can separately record how many A_N are there inside the nucleus that means how many A are present inside the nucleus.

So note the crucial issue here, all the A present outside the cell, inside the cytoplasm and in the nucleus are same molecules, we have named them separately because they are present in 3 different compartment. Now suppose A_N comes back to cytoplasm, these molecules that comes up to back to cytoplasm will not be marked as new molecule but rather it will be same as the A_C that we have already drawn here. So in a sense what I want to communicate is that, once you have decided the boundary and the compartment then you start showing all the players involved in them and mark each of the molecules or elements or component present in each of these compartments uniquely so that you can keep a track of them.

(Refer Slide Time: 11:43)



Then comes the third point, till now we have represented compartments, we have represented molecules or elements, a player in the process, now we want to represent molecular events or processes. In most of the biological problems that we deal with, we will have certain common type of processes involved and I have only shown some of them in the slide that does not mean it is an exhausted list, but I want to give you an example how we consistantly draw certain molecular processes. The first one is simple reaction; A is becoming B so you represent that by A to B by a arrow with a arrow head in the product.

Suppose just like the previous example, A is present outside so I have marked as A_O that is going inside the cytoplasm so it is A_C which is just transport process. I can represent the transport process also by an arrow just shown here, so A_O is getting transported to A_C in the cytoplasm. Many chemical reactions in biological system will be reversible reactions the way I have shown here; A + B making a complex AB and it is reversible that means A + B will form AB, also AB will break down into A + B. As you know that many reactions in biology are enzyme catalysed. For example see the example given here; A is becoming phosphorylated by a Kinase, so this is my forward reaction. A is becoming A_P phosphorylated, who is controlling that? K, the Kinase is controlling that.

Notice, how I have written K, I have not written A + K, I am not writing A + K because in this process K is enzyme and by this process of phosphorylation of A, K enzyme does not get used up. Remember, your basic bio-chemistry knowledge that a enzyme or a catalyst does not get used up in an enzymatic reaction, it only controls or modifies the rate of the process. So I

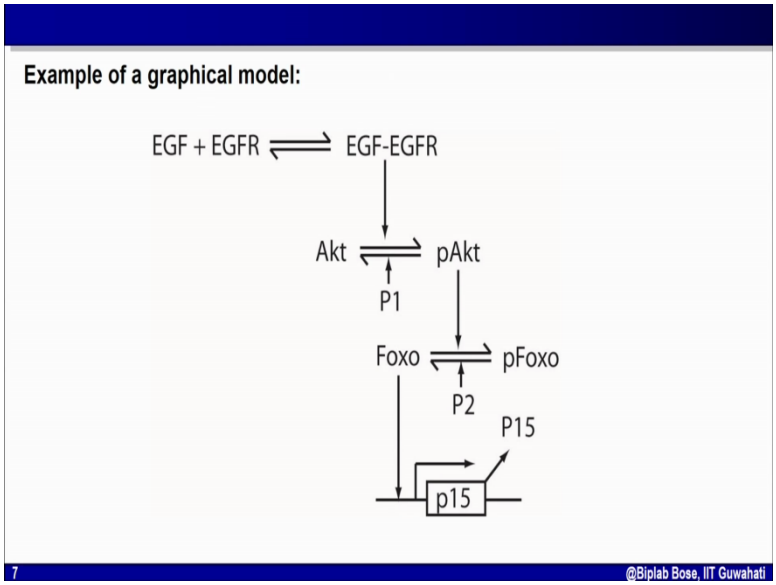
am not showing here it as A + K, rather I am showing that A is becoming A_P and that is controlled by K as represented here by an arrow.

Similar thing for a phosphatase, once your A has become phosphorylated and becomes A_P it can become de-phosphorylated by a phosphatase enzyme, again P is representing here phosphatase and it is controlling the reverse reaction. So remember what I am doing here, I am representing the control or the modifier or the controller of the reaction by these two enzymes K and P, and they are drawn as vertical arrows on the reaction that is controlled by these individual molecules.

Now here in this example A_P is getting de-phosphorylated to A, so that means P, phosphatase which is removing this Phosphate group from A is actually working opposite to the Kinase reaction phosphorylation. So sometimes some people may want to represent it as if P is a molecule or enzyme which is inhibiting phosphorylation of A. So that can be represented here as I have shown A to A_P that is a normal process by which A gets phosphorylated to A_P and P is inhibiting it. Notice that we do not have an arrow head here, we have a hammer head, usually these types of hammer head represent inhibition.

So what I want to represent here, I want to represent that P is inhibiting the phosphorylation of A to A_P. So in brief you use an arrow to represent an activation or reaction or transport whereas, you use a hammer head to represent inhibition. We are drawing these unique processes, we will club these unique processes like reaction, reversible reaction, inhibition control by enzymes to create a larger map of particular processes, let us see that example.

(Refer Slide Time: 16:56)



Suppose we want to model how each EGF controls cell cycle. We know by cell cycle, cell initially remains in G1, and then enter in S phase then into G2 phase and M phase where cell divides. This cell cycle which controls division of cells is activated by molecule signalling molecule like EGF- epidermal growth factor and we want to model that. And from literature we know EGF eventually inhibits and activate multiple processes and here in this particular case we want to model how EGF controls a particular inhibitor P15. P15 is an inhibitor of cell cycle, so if you have to promote cell cycle, you have to inhibit that inhibiter and EGF- epidermal growth factor just does that.

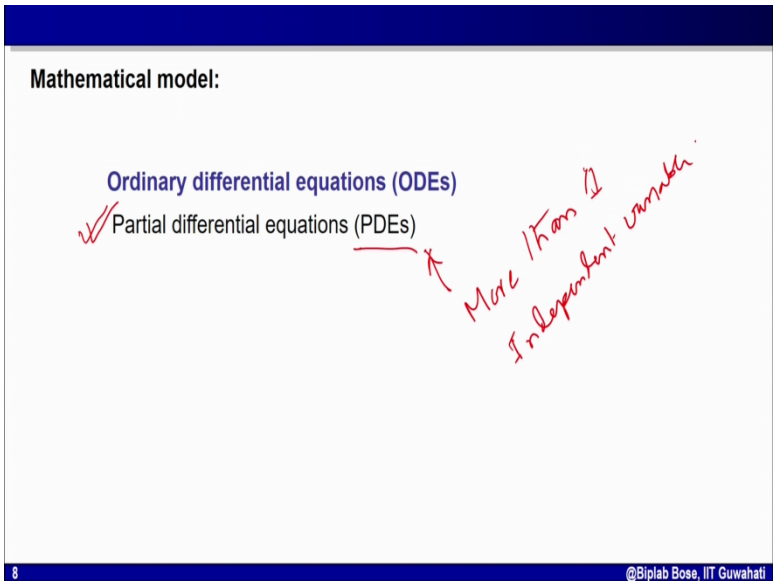
So here I have made a graphical model of that process. What is happening? The first step is EGF is binding to EGFR the receptor and they make a complex called EGF-EGFR complex and obviously it is a binding process so it is a reversible one. Once this complex have been formed these complex trigger multiple processes, we have not shown all these processes. What happens after the activation of the multiple processes is that AKT which is a Kinase gets phosphorylated, so rather than showing all the processes involved in this I have just shown a arrow and here AKT is a substrate which becomes phosphorylated to pAKT, so this is the process by which AKT is getting phosphorylated to pAKT, whereas this complex EGF-EGFR is actually controlling this process. Once pAKT is formed, pAKT is the active form of AKT and is the Kinase. It can phosphorylate Foxo to pFoxo, pFoxo is nothing but phosphorylated form of Foxo and pFoxo is inactive.

Interestingly Foxo can control or activate expression of P15, so what is happening here? EGF is forming EGFR complex that is triggering phosphorylation of AKT to pAKT. Once pAKT is formed, pAKT is active form so that is a active enzyme, that phosphorylate Foxo to pFoxo, now pFoxo is not active, so transcription of P15 will decrease. So the whole process is represented by arrows and reversible arrows. Once I have this graphical model, I will try to create mathematical equation to represent each of these processes. Just remember one issue here, this P1 and this P2 are phosphatase, so P1 is controlling de-phosphorylation of pAKT to AKT, whereas P2 is another phosphatase and it is controlling de-phosphorylation of pFoxo to Foxo.

Once you have this graphical model we will move into making mathematical model. Remember, the thrust of this course is building dynamical models for biological processes or phenomena. By dynamical model means we want the model processes which are changing with time, which are time dynamics, so we will use primarily ordinary differential equation.

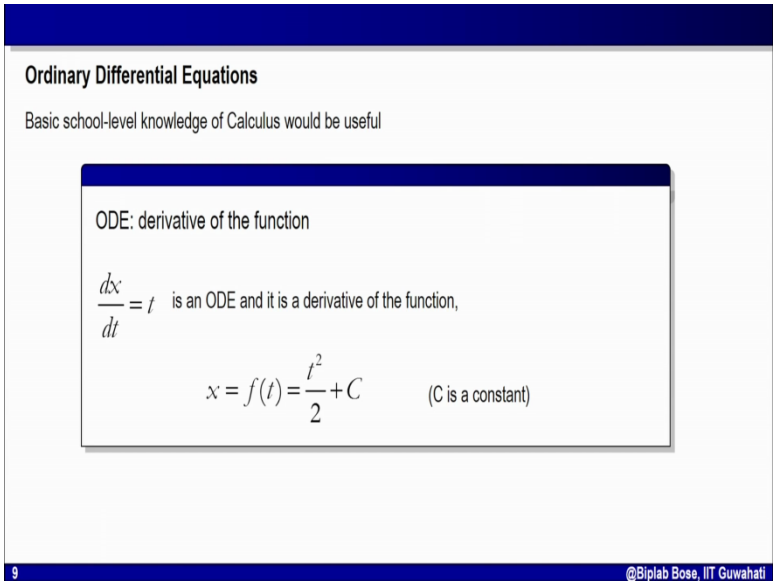
In ordinary differential equation, variables that we are measuring are changing with time. Sometime in some dynamical models we have to use partial differential equation. Partial differential equations are used when you have more than one independent variable. So you have to use partial differential equation when you have more than one independent variable.

(Refer Slide Time: 21:42)



For example suppose you want to simulate movement of protein on a cell surface, so with time the molecules are moving. So time is changing also the position of the molecules are changing. So if you have some reaction happening on cell surface involving those proteins, so with time concentration of this molecule will also change, at the same time the position of the molecules will also change. That means I have to track or keep record of those molecules with respect to time as well as space, so I cannot use ordinary differential equation in that case, I have to use partial differential equation. In this course we will focus primarily on ordinary differential equation based model where time will only be the independent variable. That means the processes are such we will keep track of changes with respect to time.

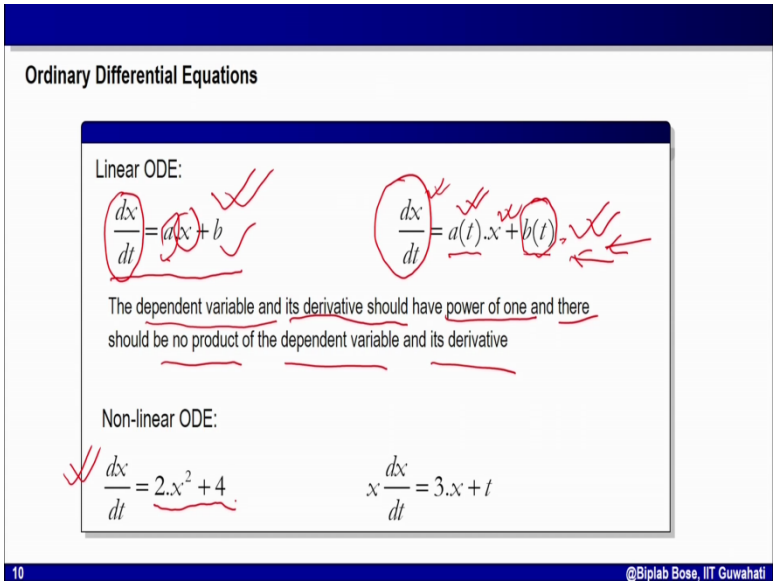
(Refer Slide Time: 22:48)



I believe most of you are quite well acquainted with calculus and in fact, for this course we require + 2 level school calculus, we will recapitulate some basic concepts of that and I will also advice you if you have forgotten some basic concept of calculus, please look into the basic those basic concepts and try to recapitulate those basic concepts.

By ordinary differential equation what do we mean, ODE ordinary differential equation is nothing but a derivative of a particular function. For example here, I have a function, $x=f(t)=\frac{t^2}{2}+C$. If you differentiate this one x with respect to t you get $\frac{dx}{dt}$ and by simple formula of differentiation you get $\frac{dx}{dt}=t$. So this $\frac{dx}{dt}=t$ is a differential equation and this is ordinary differential equation because here 'x' is changing only with respect to time, 'x' is the dependent variable because it is changing with time and it depends upon time whereas time is independent variable, so we have a dependent variable 'x' and we have a independent variable 't' and together we have got a ODE $\frac{dx}{dt}=t$ that is the derivative of the function $x=f(t)=\frac{t^2}{2}+C$.

(Refer Slide Time: 24:40)



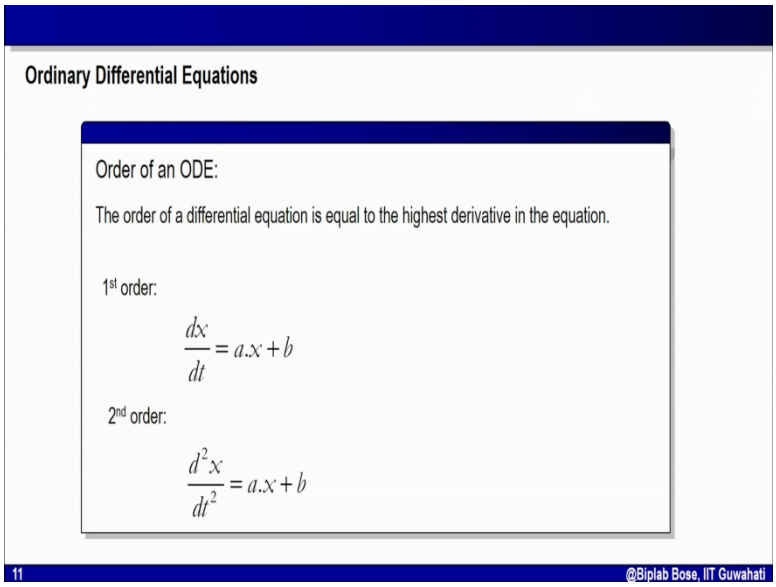
Now we will discuss about some basic difference between certain types of ODE. Let us start with the simplest one the linear ODE, a linear ODE is such that the dependent variable and if derivates have power one and there is no product of multiplication of its dependent variable and its derivate. Let us explain with this example, here this is a linear ODE, $\frac{dx}{dt} = a.x + b$.

Remember, the dependent variable here is 'x', its derivate is $\frac{dx}{dt}$, the power of 'x' is one here and there is no power higher than 1 for $\frac{dx}{dt}$ also and there is no product of 'x' and $\frac{dx}{dt}$. So that is why as these two conditions are met that the dependent variable and its derivative have power 1 and there is no product of the dependent variable and its derivative, this particular ODE, ordinary differential equation is a linear ODE.

I can have slight variation of this linear ODE also, let us see that. See in the previous example in this one 'a' is a constant, whereas in this equation the second one we have $\frac{dx}{dt} = a(t).x + b(t)$. So 'b' itself is the function of the dependent variable it depends upon 't' and 'a' itself is a dependent variable and depends upon 't' and changes with time.

But in this particular equation we have only written the derivate of ‘x’ and we have no product of $\frac{dx}{dt}$ and ‘x’. And the power of ‘x’ is maximum 1 and power of $\frac{dx}{dt}$ is also 1, so this equation is also linear ODE. Let us see some non-linear ODE, the concepts will be clear. The first one, $\frac{dx}{dt}=2.x^2+4$, so the power of ‘x’ on the dependent variable is 2 greater than 1 so it is non-linear. Look into this one, the second one; $x.\frac{dx}{dt}=3.x+t$, power of ‘x’ is 1, there is no issues but we have a product of derivate and the dependent variable of ‘x’, so it is a non linear ODE. Sometimes we differentiate between ODE based on its order, by order we mean the value it is equal to the highest derivate of the equation.

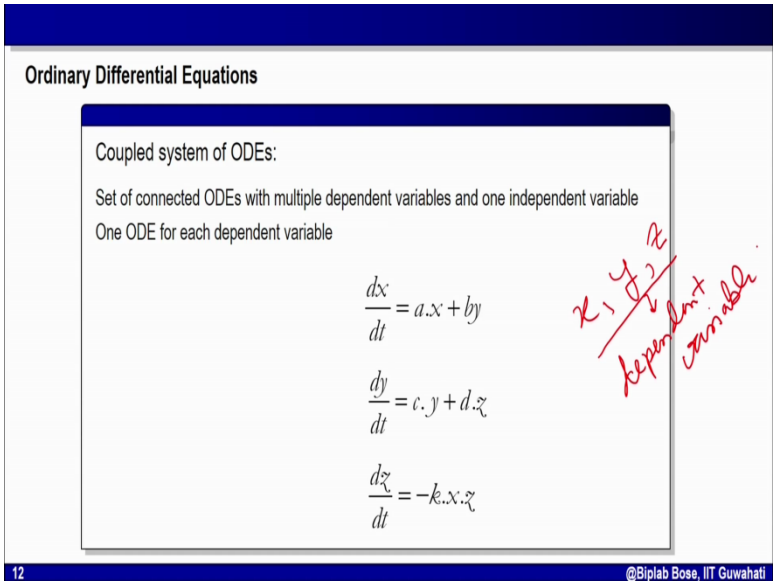
(Refer Slide Time: 27:43)



For example the first one, the first order equation. Here the derivate is $\frac{dx}{dt}$ whereas the second one is second order because we have $\frac{d^2x}{dt^2}$, so this is second order derivative. In most of the problem that we will model in this course, we will deal with first order ODE. We may have both, linear and non- linear, we will frequently meet situations where number of dependent variable is more than 1. In all the previous examples, we have only one dependent variable which is ‘x’, but you may have a situation and that is much more realistic and most of the biological problem you will have more than one dependent variable. You may have

another dependent variable ‘y’, you may have a third one ‘z’, so ‘x’, ‘y’ and ‘z’ all of them are dependent variable.

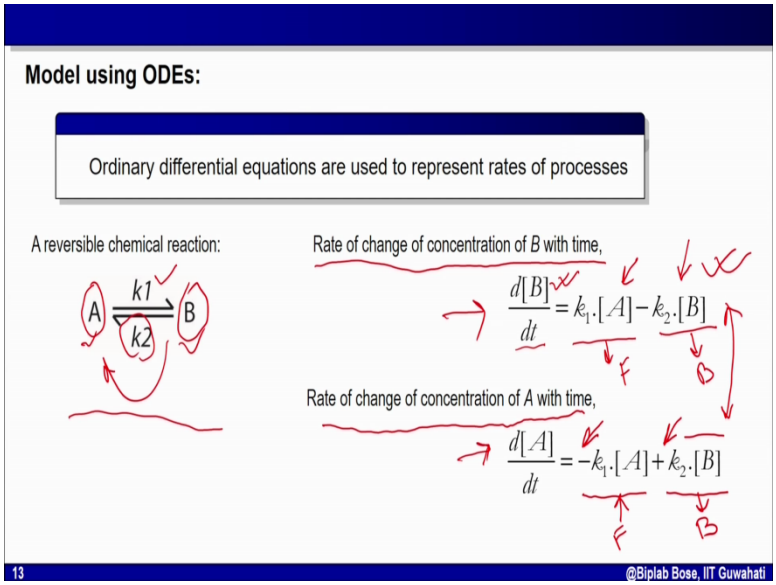
(Refer Slide Time: 28:55)



So they are all dependent on 't' and they are changing with time, so for this type of system we have to write separate derivative for each of these variables. So I will write one derivate for 'x' that mean $\frac{dx}{dt}$. I will write one separate variable for derivative for 'y' that means I will write $\frac{dy}{dt}$ and I will also make a separate variable for derivative for 'z'. And if you see the example, the first ODE is for $\frac{dx}{dt}$, this is my first ODE, second ODE is for $\frac{dy}{dt}$ and third one is for $\frac{dz}{dt}$; 3 dependent variable, 3 ODEs.

Now this system is called a set of ODEs because I have more than one ODE. One interesting thing here in this set of ODE is that notice that $\frac{dx}{dt}$ has 'y' term whereas $\frac{dy}{dt}$ has 'z' term. That means if I have to know the derivative of $\frac{dx}{dt}$, I have to know 'y' and that is controlled by these derivative again. Whereas dz dt also has 'x' that means I have to know about 'x' to know $\frac{dz}{dt}$. So 'x', 'y' and 'z' they are changing with time and their values depend upon each other. So these type of system or set of ODE is called Coupled system of ODEs. In most of the biological problem we will deal with we will have Coupled system of ODE.

(Refer Slide Time: 30:35)



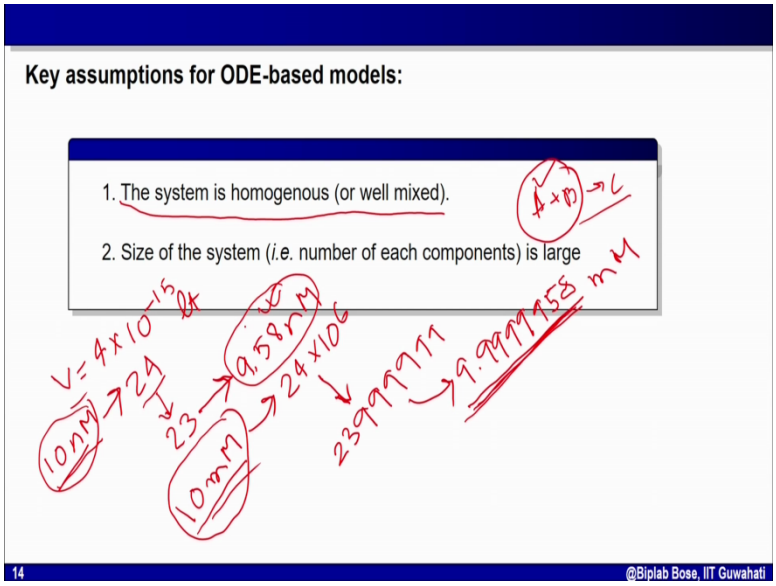
Now let us look into utilities of ODEs in our dynamical model. You have to remember the ordinary differential equations will be used to represent rate of processes that is the utility of ODE in our mathematical process. So take an example, suppose I have chemical reaction, A is getting converted to B and the rate constant for that is k_1, whereas it is as it is reversible, B is also becoming A and the rate constant is k_2 and I want to model this one. When I say I want to model this reversible process essentially I want to write mathematical equation that will represent the rate of change of B and rate of change of A. Let us see, so the first ODE that I have written that represents the rate of change of concentration of B with time.

So how I am representing the ODE, I have written $\frac{dB}{dt}$. Remember this square bracket represent molar concentration, so change in molar concentration of B with respect to time is given by $\frac{dB}{dt} = k_1.[A] - k_2.[B]$. So B is formed by the forward reaction, B is converted back to A by this reverse reaction or backward reaction. So $\frac{dB}{dt}$ the rate of change of concentration of B with time $¿k_1.[A] - k_2.[B]$.

Similarly I have ODE, differential equation representing rate of change of concentration of A with time. So I have $\frac{dA}{dt} = -k_1.[A]$, remember by this process that is the forward process, forward reaction, but by this process A is getting used up that is why I have a minus sign

here, $+k_2.[B]$ which is the backward reaction and remember here I have positive sign because by this process this backward process, A is created. So actually both these ODE are exactly the same except the signs, the signs are inverted. So these two equations is essentially my mathematical model for this reversible reaction and the ODEs written here are representing rate of the processes; rate of formation of B, rate of formation of A.

(Refer Slide Time: 33:41)



When we are creating mathematical models using ODE, we have to make some basic assumptions and those two assumptions should always meet otherwise we cannot create ODE based model. The first criterion is that the system has to be homogenous or mixed.

Imagine a chemical reaction happening in a flask, A and B is reacting to make C.

$A + B \underset{\rightarrow}{yields} C$, if A and B are not well dispersed in the flask or not well mixed in the flask, then somewhere the concentration of A will be more and somewhere the concentration B will be more, so production of C will be somewhere more and somewhere it will be less. That means the process will depend upon the position of A and B molecule inside the flask, we do not want to do that. So we have to assume we have to make sure that A and B are very well mixed, so that I can imagine that the concentration of A and concentration of B are always same across the whole flask.

So this is one basic requirement if you have to make a ODE based model and you have to make sure that in reality it is true for your model. Second criteria is that the system has to be very large otherwise you cannot use ODE based model. Let me give you an example, suppose

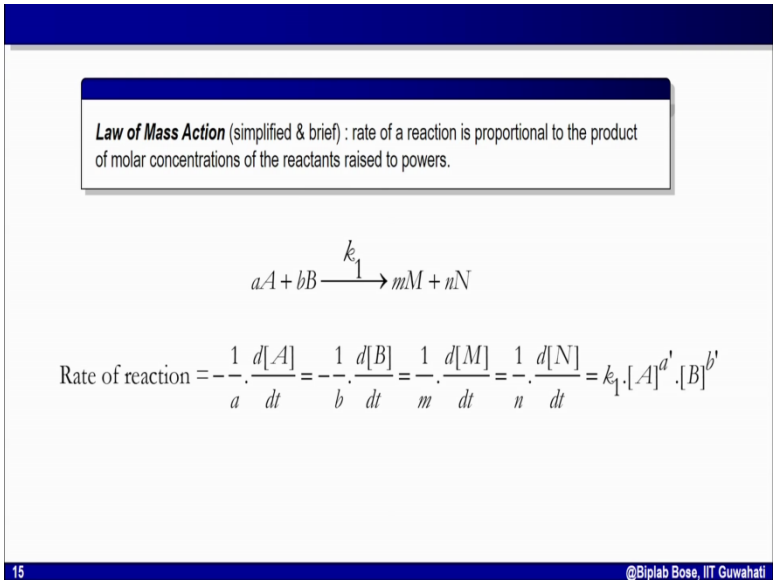
I take a cell of volume say 4×10^{-15} L. This cell is usually 1 micron in diameter so if you calculate the volume it will come in this range. And suppose I have a molecule A, which is also a concentration of 10 nanomolar. If you use Avogadro's number and use the volume of $4 \times 10^{-15} L$, 10 nanomolar essentially means 24 molecules.

So inside a cell of $4 \times 10^{-15} L$ volume, I have 24 molecules. Now suppose by this reaction $A + B \xrightarrow{yields} C$, one molecule of A is used up, so that becomes 23 molecules. Now if you convert this 23 molecule in concentration term it will become 9.58 nanomolar, so from 10 nanomolar it becomes abruptly 9.58 nanomolar. So there is abrupt change in the concentration just because one molecule has reacted to create C.

Now imagine in the same volume, if I have 10 millimolar of A that is equivalent to almost 24 into 10 to the power 6 molecules. Now if one molecule of A react that becomes 23999999 and this is equivalent to 9.9999958 or similar to that. So what I want to show here is that see if you have 10 millimolar this is also millimolar, change in one copy of the molecule will cause a slight change in the concentration. So remember here concentration change is not abrupt, if you have to model this system using a ordinary differential equation, the variable the dependent variable should not change abruptly, this should be changing continuously.

So for this type of system if the concentration is higher at 10 millimolar than I can easily assume that these changes are happening smoothly, continuously, but if the concentration is very low suppose 10 nanomolar, then a simple change of one molecule will cause a abrupt change in concentration and I cannot use ODE for this type of systems. So the two key issues here if we remember, is that we have to remember that the system has to be homogenous or well mixed so that all molecules or all elements can interact with each other freely and the size of the system, number of molecules, number of people, number of bacteria that is all the elements involved in the process should be large in number.

(Refer Slide Time: 38:24)

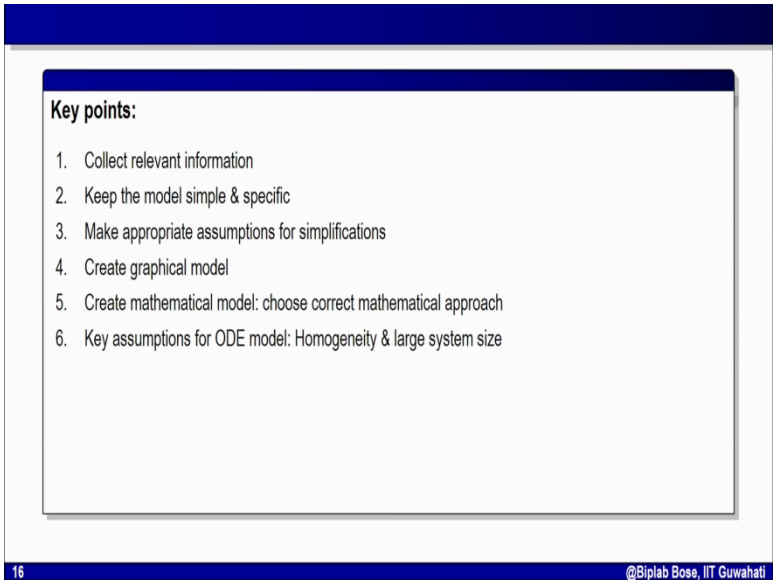


So let us briefly discuss one important issue in modelling bio-chemical reaction using ODE. Bio-chemical reaction can be very complex but in sometime we may use the law of mass action to simplify those processes. We all know law of mass action, if you do not remember please try to recapitulate it from any bio-chemistry book or physical chemistry book. In a sense, law of mass action says the rate of reaction is proportional to the product of molar concentration of reactants raised to power.

So take a example, A is reacting with B; the small 'a' and small 'b' are stoichiometric constant and giving rise to M and N with the rate constant k_1. So the rate of reaction should be proportional to the concentration of each of these reactant A and B raised to the power. So it is written as rate of reaction $¿k_1.[A]^{a'}.[B]^{b'}$; remember both these in square bracket are molar concentration. And the individual rate, the rate of change of A is given by this, the rate of change of B is also equal to this but its formation will be different, $\frac{dM}{dt}$ is rate of change of M and $\frac{dN}{dt}$ is the rate of change of N. And remember each of them is divided by a stoichiometric constant.

And notice the sign here, these are 'minus' because there will be decrease in A and there will be decrease in B whereas, these are 'plus' because M is increasing and N is increasing by this process. So whenever possible we will try to use this type of law of mass action equation but remember in all cases you cannot do that.

(Refer Slide Time: 40:12)



So let us jot down what we have learned in this module. If you have to create a mathematical model, first collect all relevant information that is critical, once you have collected all the information, try to simplify, keep the bare minimum essential thing then make some assumptions. And once you have made the assumptions to simplify the model, create a graphical model. In all cases you do not need it, but in many cases it is better to draw a graphical model because it will help us to write down the mathematical models.

Once you have done the graphical model, go for mathematical model. A mathematical model is nothing but a set of equations representing the processes. In this course we will write ordinary differential equation to represent each of the processes and remember a ODE represent rate of a particular process. Remember, when you have to make a ODE based model, I have certain key assumptions; the first one is system is homogenous that is well mixed and the system is very large. That means the number of molecules involve in the reaction are large in number. If I am modelling a cellular process involving multiple cells, so the number of cells will be large in number.

If I am modelling a process of a population of human being, the number of human being in that population should be large in number. So these two keys assumptions have to be met and then only we can use create ODE based dynamic models. Thanks for watching, in the next modules we will learn about creating simple models.