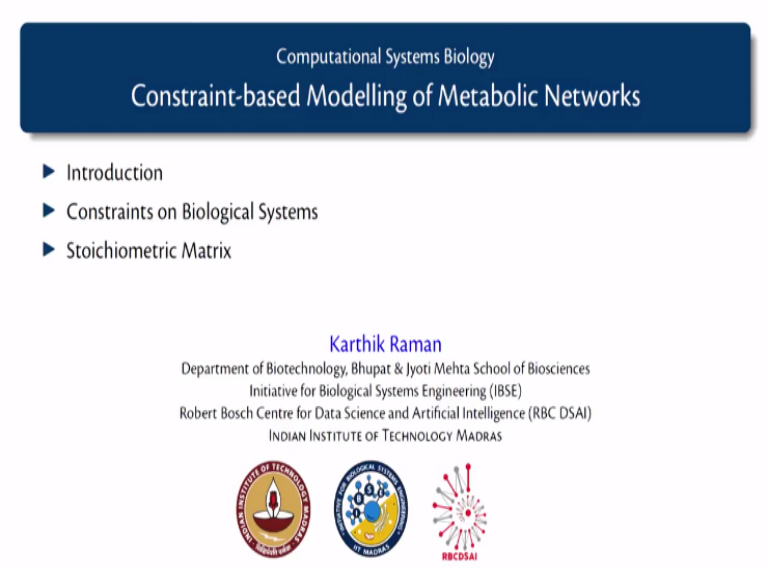
Computational Systems Biology
Karthik Raman
Department of Biotechnology
Indian Institute of Technology – Madras

Lecture - 56
Constraint-based Modelling of Metabolic Networks

(Refer Slide Time: 00:12)



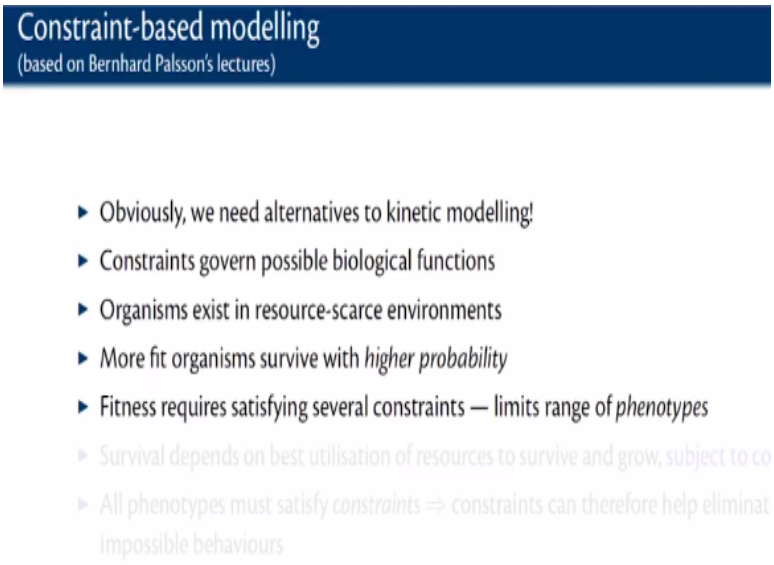
So beginning today, we will look at a long series of videos wherein we will study constraint-based modelling of metabolic networks. This happens to be one of the most important tools to study metabolic networks and you know a very favourite tool in systems biology because it can make predictions about large genome scale metabolic networks and so on.

So today I will introduce you the constraints-based modelling, what are all the constraints that prevail biological systems and how do we compute what is known as the stoichiometric matrix which is a nice representation of our metabolic network for further modelling. Welcome back. We will move over to the next module of the course where we are going to talk about constraint-based modelling.

We have seen a few other paradigms of modelling particularly network based modelling or network biology followed by studying dynamic models with special emphasis on parameter estimation and so on. Today, we will start looking at constraints-based modelling which happens

to be one of the most powerful techniques to study biological systems especially at a system scale.

(Refer Slide Time: 01:13)



So, I think 1 thing that you may be convinced with already is that we need alternatives to kinetic modelling. So why would we need alternatives to kinetic modelling essentially the key thing I want you to hit at right away is that kinetic models involve very large number of parameters. If you want to build a genome scale kinetic model people have by the way it is very tricky, very difficult simply.

Because you have to estimate something like 2000 parameters unless you have very respectable tight estimates for each of the parameters a priory you are not going to be able to estimate or build such models very effectively. But people are you know bravely attempting this and with good reason, but clearly you need a simpler method and you have already discussed at the very beginning it all depends on the questions you want to ask.

If all I want to know is whether post this gene deletion is my E. coli going to grew or not, I do not need to build a dynamic model most likely I do not need a dynamic model and we will see how to constraints-based models can actually give you answers to questions like that. So in constraints-based modelling was literally you know founded by Palsson and colleagues.

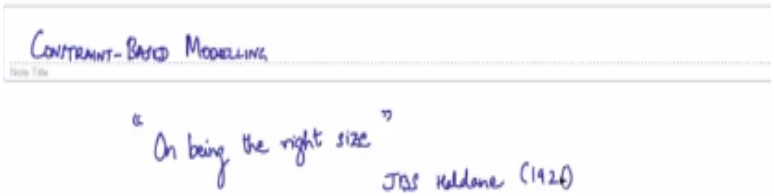
Palsson was at UCSD then and I think he is one of the you know major proponents of constraints-based modelling and their lab has to brought out a lot of work and there are many other fantastic scientists who have since contributed notably Costas Marana, Chaneton, Ruppin and few others, many others and there are other people like Nielsen and so on who made a lot of contributions to constraint-based modelling in terms of building models and using them to make predictions and so on.

So what is constraint-based modelling? What do we mean by constraints? **"Professor - student conversation starts"** I am just asking you a general question. I am not trying to motivate what you know preempt what I am going to cover today. What do we mean by constraint in general? Maybe some kind of limitations. So, some limitations, some I think limitation is a good word and it also makes a lot of sense in the context of biological system because you always talk about the growth limiting substrate something of that sort. **"Professor - student conversation ends"**

So there is always a constraint that every system has to function well. So constraints govern all possible biological functions and organisms exist in resource-scarce environments and fitter organisms survive with higher probability and fitness itself requires satisfying several constraints which limits the range of observable phenotypes. So at this point you know I want you to think about what kind of constraints do you need to satisfy to live or what organisms need to satisfy to survive.

"Professor - student conversation starts" What constraints must they satisfy? Can of think of some simple constraint? The nutrients, concentration, but think of something more macroscopic, water, oxygen. So essentially lot of nutrients and input and so on, but other than that there are fundamental constraints on even how big an organism can be. So this is very classic I say which I like to point to it at this point **"Professor - student conversation ends"**.

(Refer Slide Time: 04:53)



I really decorated biologist JBS Haldane who essentially wrote this essay in 1926 I think on being the right size where he essentially argues that you know a human being can only be so big and elephant can only be so big, whereas an insect does not need to have oxygen in its blood stream because it has got a such a small body to power it can take air by diffusion, take oxygen by diffusion. So there are constraints.

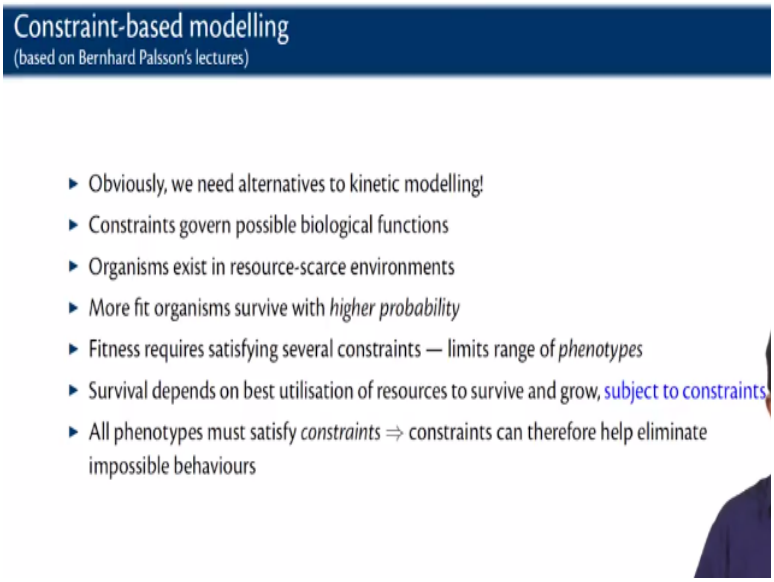
So if you want to have a bigger animal it has to have better machinery. **"Professor - student conversation starts"** So why do not we have a you know 20-foot-tall human? Could we have? We could but what will be the challenge? Bone won't be able to support. Yes, in fact that is actually what Haldane argues. The argue is that you know you cannot you know there is some dependency between the weight and organism can be and the cross section of its bones.

Because it has a support that kind of a weight, but even simpler than that I would think of something else before even I go to bones. How will the blood flows? So can you have a heart that is powerful enough to pump blood 20 feet in a human being as it does it in giraffes, but can it do it in human beings? So those might be the constraints. So these are some sort of implicit constraint. **"Professor - student conversation ends"**

You do not really stop and think about these constraints, but clearly there are some constraints that you have to follow. If you step to the microscopic level, there are clearly other constraints

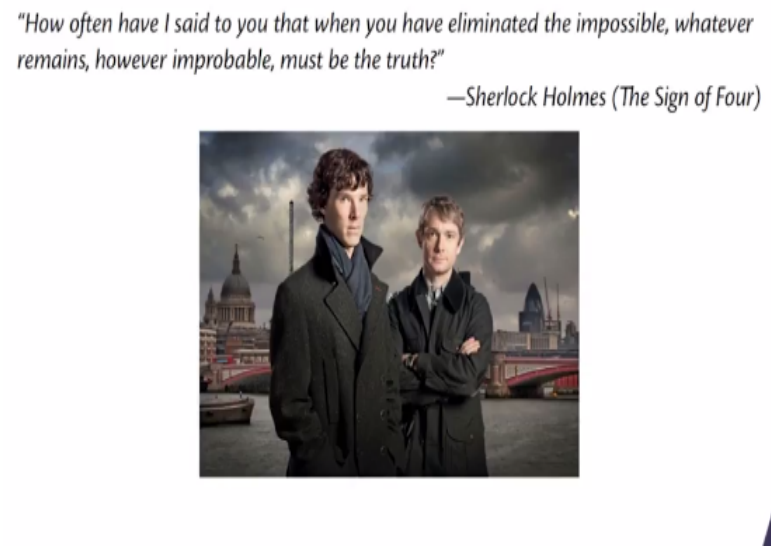
that one has to follow for example thermodynamics. No reaction in your body can violate thermodynamics. No reaction in your body can violate stoichiometry. So there are some such obvious constraints that exist in nature.

(Refer Slide Time: 07:20)



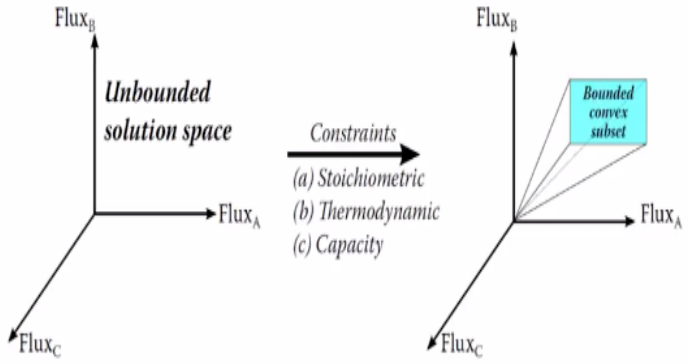
So fitness requires satisfying several constraints which limits the range of observable phenotypes and survival depends on the best utilization of resources to survive and grow, but always subject to constraints. You can never violate constraints. So the next idea is that since phenotypes must satisfy constraints, constraints can help eliminate impossible behaviours.

(Refer Slide Time: 07:50)



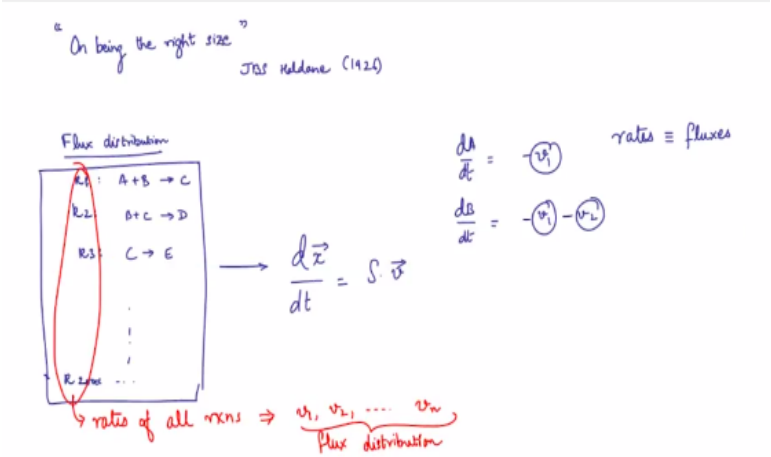
And this is where you know I borrow this classic slides from Palsson who asks you know how often have I said to you that when you eliminated the impossible, whatever remains, however improbable must be the truth. The famous fictitious character I think you must be familiar with. Sherlock Holmes. So you have a classic example on you know you basically start with the big pole of suspects and then you go about eliminating one by one to arrive at a smaller pole of suspects who have to be tested in a different way. We essentially do the same thing here.

(Refer Slide Time: 08:22)



We start with an unbounded solution space. So think as a space of possible solutions what do we mean by a flux distribution in the first place.

(Refer Slide Time: 08:25)



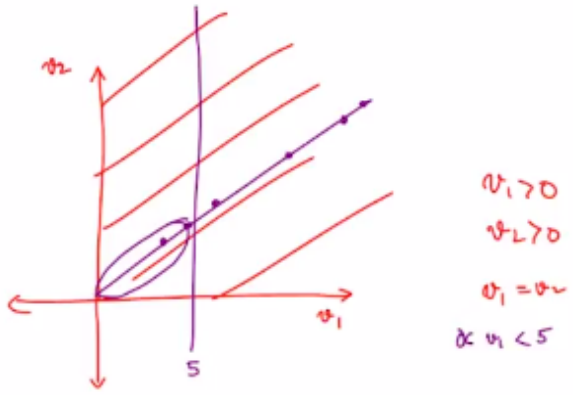
I want you to recall what a metabolic network is. A metabolic network for our purpose is a long list of reactions. As it is approximately this is what a metabolic network is if you recall we had an analogy with traffic networks and things like that and so on. So this is your metabolic network which can also be written in the form. You remember we did this on how do we write out each fluxes?

So you would say that dA/dt is going to be - vR1 or - v1 and dB/dt would be - v1 - v2 something of that sort. So now these are your rates which we also called fluxes. **"Professor - student conversation starts"** What is a flux in general chemical engineering. Quantity flow per unit area. Here we just round the unit area. We just talk about flow or rate of any given chemical reaction. So given this system we are tasked at finding rates of all reactions meaning v1, v2, vn. This is what we call a flux distribution. **"Professor - student conversation ends"**

So what are the fluxes of every single every individual reaction in a network? This essentially specifies a network for us. So think of this as a flux space. It is a 3 dimensional space so you have flux A, flux B and flux C. So just think of those 3 reactions v1, v2, v3. Now you apply constraints to eliminate improbable suspects or impossible, infeasible solutions, infeasible suspects or solutions and you come up with a much smaller subset.

So you will note that from 1 infinite subset you have gone to another infinite subset, but clearly this is quite restricted. So the idea is simple. What we do is

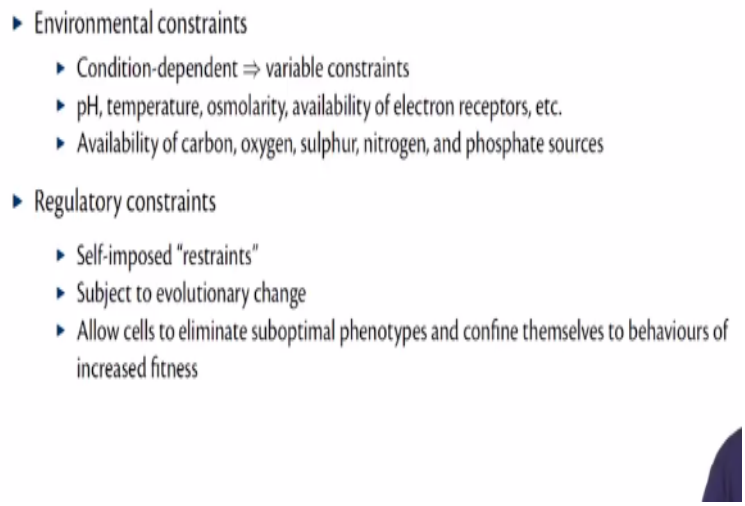
(Refer Slide Time: 12:20)



You may actually be very familiar with something of this sort I will come back to it in a moment in a few slides. You start with so now any point in this space in fact I should draw it this way. Any point in this space is a possible solution. For a system with 2 reactions, but now if I say v1 > 0, v2 > 0 it becomes this region. Now if I say v1 = v2 it becomes this very line. So every solution on this is a possibility.

Now if I know something more if I know you know 0 < v1 < 5. At that point say that all the solutions are on this part let us say, this is 5. So you can these constraints basically help eliminate impossible behaviours, impossible flux solutions. What kind of constraints to be put? How do we impose these constraints? Let us look at it.

(Refer Slide Time: 13:46)



So we can think of different types of constraints some classic examples would be environmental constraints something everybody thinks about first of what are the nutrients available. What is the availability of my carbon, oxygen, sulphur, nitrogen, and phosphate sources. Of course pH, temperature, osmolarity, and availability of NADH, NADPH and so on and many of these are condition dependent constraints.

So they are actually variable they are not fixed. Use 1 set of constraints the other really interesting set of constraints are there are some self-imposed restraints by the cell which is almost counter intuitive. So you expect that the cell will do something. The conditions are actually feasible for the cell to do something, but it will actually does not why? **"Professor - student conversation starts"** Can you think of an example? Surface area, volume constraint. It has a restraint.

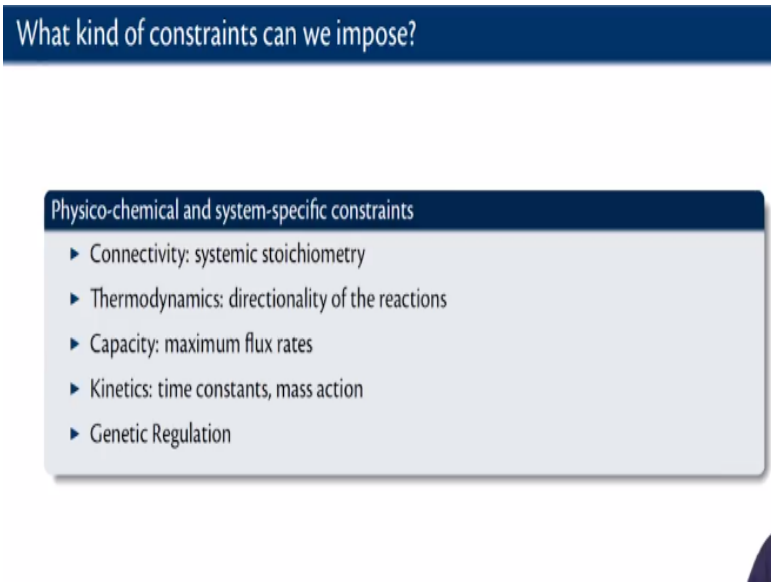
Well, so that could be one, because it is not something you think this I can potentially survive with a larger volume or something where it has to break down into 2. Yes, what else (()) (15:05) but that is not a self-imposed restrained actually. What will do a self-imposed restraint? A cell can go, but it may choose not to for example. I think very familiar with 2 examples of this sort. Why does a cell choose not to use a substrate? Catabolite Repression.

That will be a classic example. There is a lactose and glucose available to the cell. The cell should be growing, but it will just use a should be growing at a higher rate using both glucose and lactose, but the cell which choose to use only glucose because there is a catabolite repression on the utilization of lactose. Once glucose is exhausted you have a dioxide shift and you know the lactose enzymes are then expressed, the genes are expressed and then the lactose metabolism kicks in. **"Professor - student conversation ends".**

So the cell essentially chooses not to use lactose. So another very similar example as cell chooses not to grow. Sporulation. The conditions seem to be reasonable for growth but the cell is trying to eliminate a suboptimal phenotype which could be reasonable growth and there is some harsh conditions and confine itself to a behaviour of increased fitness. It really depends upon what you define fitness.

But if you talk about fitness as the ability to survive and you know live on for more generations may be the cell is achieving that by sporulating. So the spore is more stable and then the condition become more favourable the cell can you know again start growing and so on. So there are such self-imposed restraints as well, but we will focus on simpler constraints.

(Refer Slide Time: 17:23)



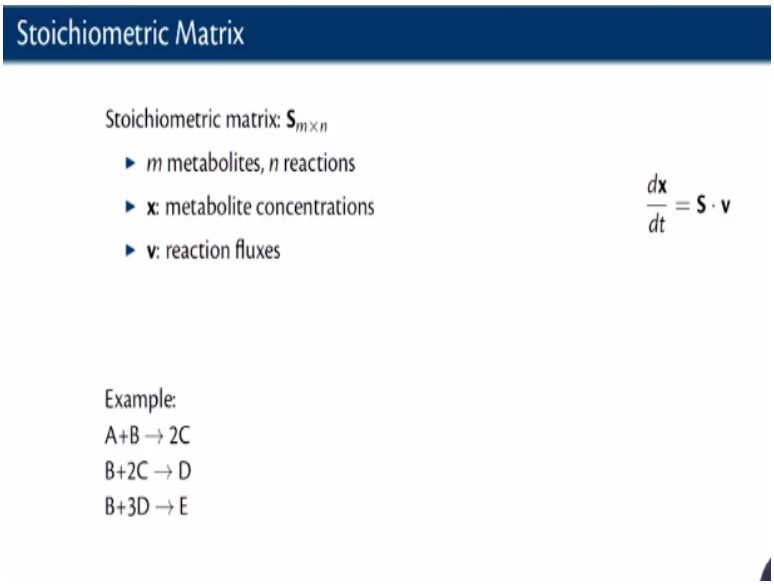
Essentially connectivity coming from systems stoichiometry first step. **"Professor - student conversation starts"** What is stoichiometry? Reaction in terms of ratios. So there are clear

dependencies. So, 1 mole of glucose will produce exactly 1 mole of fructose, not a molecule more, not a molecule less. One mole of glucose will produce exactly 2 moles of pyruvate via glycolysis.

So this becomes a nice constraint and of course there is thermodynamics. **"Professor - student conversation ends"** There are some reactions that can only go in a particular direction. Some reactions would be reversible. They can go in either direction, but you know there are such constraints then there are maximum flux rates. So every enzyme has a particular turn over number.

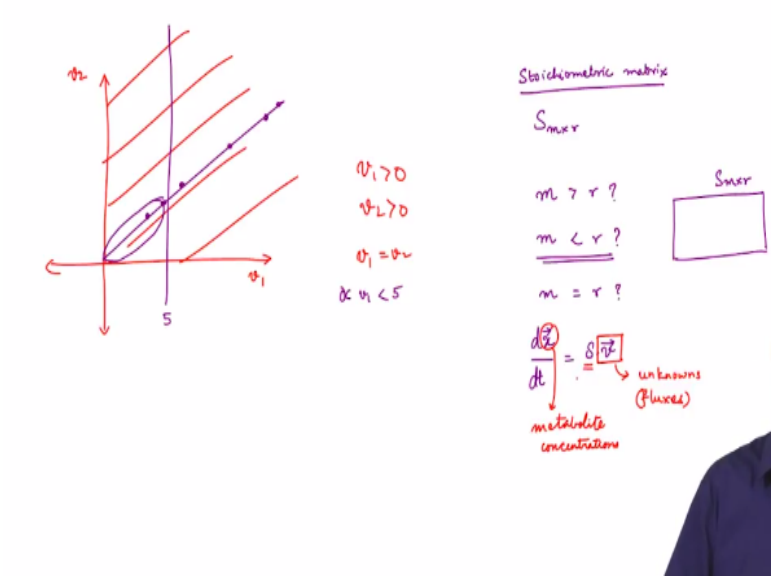
It may not be able to turn over more than x number of molecules per second and you know the genetic regulation, thus you need to worry about time constants, mass action. There are many other constraints that you may be able to think up off, but most importantly I would say are connectivity, thermodynamics, and capacity. So what would this translate into practice we will see shortly.

(Refer Slide Time: 18:48)



So let us take a very simple example. What is the size of this stoichiometric matrix for the system? 5 x 3. You have 5 metabolites and 3 reactions and you have a 5 x 3 system. So let us just go back to the concept of this stoichiometric matrix.

(Refer Slide Time: 19:32)

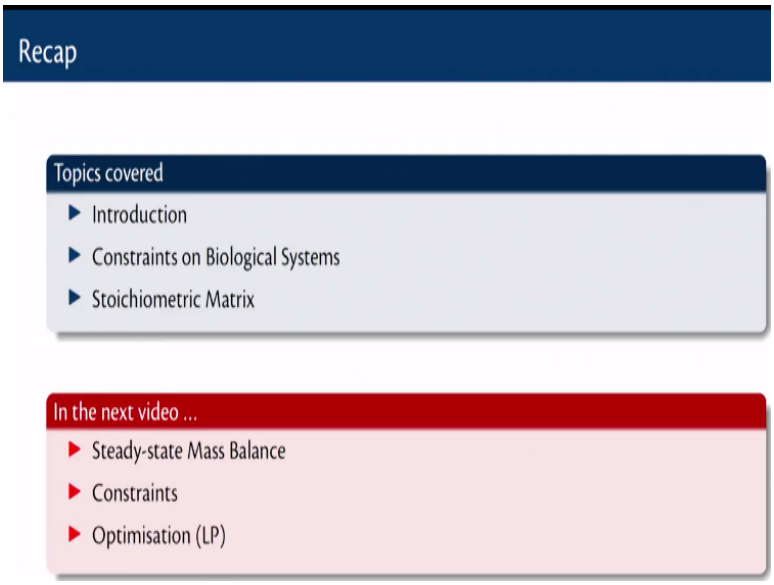


What do you think do you usually have more metabolites than reaction or more reactions than metabolites or roughly the same number of metabolites and reactions? So are you going to have a square stoichiometric matrix or a fat stoichiometric matrix or a tall matrix? **"Professor - student conversation starts"** Tall matrix. Tall? That would be this. Why? But the same metabolite can also participate in more than 1 reaction.

So you would see that ATP participates in 100 reactions and may be you know glucose participates in like 30-40 reactions and there are other metabolites that will participates 1 or 2 reactions. **"Professor - student conversation ends"** You might see a sort of you know power law as well normally. So you would normally observe that the same metabolites mix and match amongst different reactants obviously some metabolites keep changing so immediately you have m < r which means you have fewer rows than columns and you have a fat stoichiometric matrix.

We will see what implications this has. So for now you remember that you can set up the equations like dx/dt = s * v. So what are the vs again? Fluxes. What is x? Concentrations. Metabolite concentrations. So, rate of change of metabolite concentrations link via the stoichiometric matrix.

(Refer Slide Time: 22:03)



So I hope in today's video you got an introduction to this interesting field of constraints based modelling followed by a brief insight into the kind of constraints that exist in biological systems and how one goes about building this stoichiometric matrix. In the next video, we will start looking at flux balance analysis and we will start with the steady state mass balance followed by how one adds constraints to the mass balance problem and then how one sets up the optimization problem is typically a linear programming LP problem.