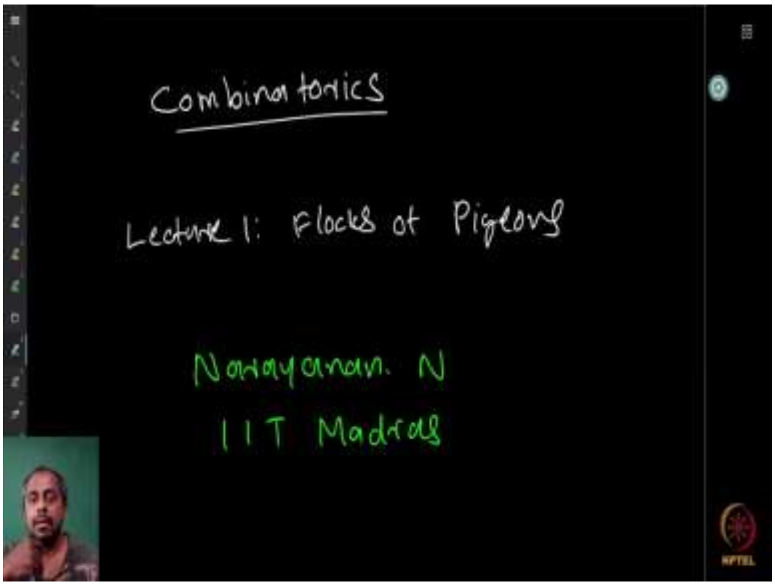
Combinatorics
Professor Doctor Narayanan N
Department of Mathematics
Indian Institute of Technology Madras
Lecture 01
Pigeonhole Principle

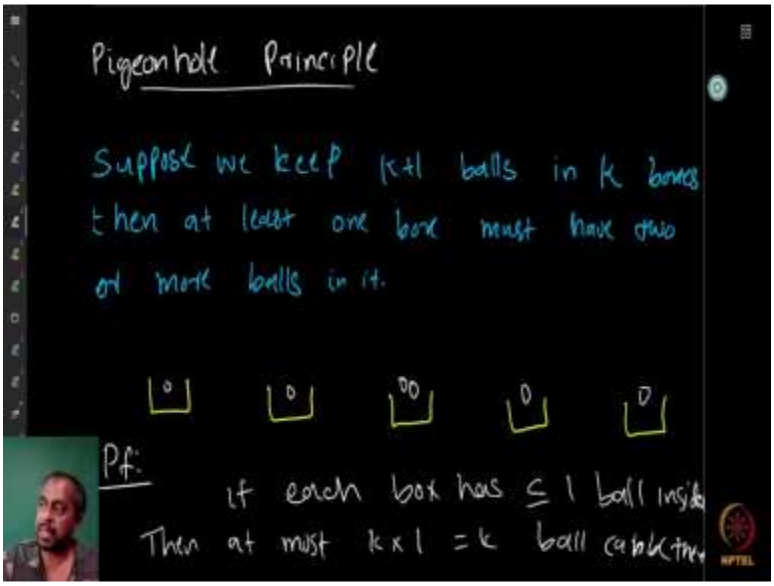
(Refer Slide Time: 0:15)



Our first lecture will be on one of the simplest techniques which is very often used in combinatorics. This is the method of pigeonhole or the pigeonhole principle. Now, it may be the simplest in terms of the principle, the statement and easy to understand. But that does not mean that it is by far the simplest to use.

In fact, we will see several examples where, you need to come up with some ingenious ways of thinking, so that we can put the question into the correct framework, where we can apply the pigeonhole principle. So, what is this principle?

(Refer Slide Time: 1:24)



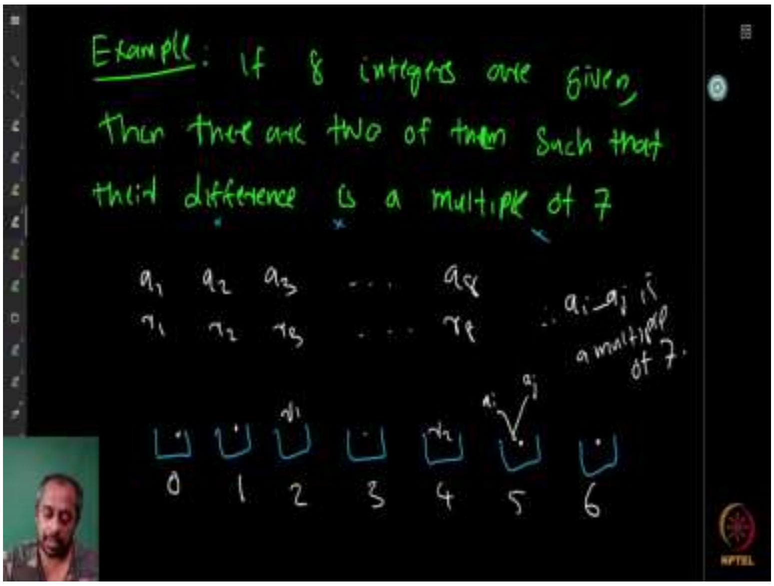
This is something that all of us know byheart I mean by experience and we might have already used it many times. So, this says the following. Suppose you have $k + 1$ pigeons, for some positive integer k. And then these guys sit into k pigeon holes or cages. Now, if $k + 1$ pigeons sit into k cages, then we can say that at least one of the cages must contain 2 or more pigeons inside.

This is something one can easily see, because, if every cage had at most 1 pigeon, there are k cages, where there is at most $k \times 1$ which is at most k pigeons inside. Therefore, we have a proof for same thing. So, we can say it in terms of terms of balls and bins, that you have $k + 1$ balls and k boxes or bins. So, here are the k bins, let us say and into this k bins suppose we are going to deposit k balls.

Now, what I want to prove is that at least one of them must have two or more balls. So, if I want to avoid that, then I will try to put at most 1 ball in each. If I have 5 bins and 6 balls, then I need to put at least two in one of them. So, this is easy to see. And this is basically what the pigeonhole principle is.

So, the proof was that if each box has less than or equal to 1 ball inside, then at most k into something less than or equal to 1 at most $k \times 1$ equal to k balls can be there. So, that is the proof. Now, this is supposedly a very simple looking principle. How do we use it? Can such simple statements have some real applications in mathematics? In fact, we are going to see many, many very nice applications of this in today's lecture.

(Refer Slide Time: 4:54)



We start with a very simple example. Suppose you have 8 integers and out of these 8 integers that are given to you there are two among them whose difference is divisible by 7. No matter which integers are given to you, you have to show that 2 among them are such that their difference is a multiple of 7. Now how do you prove something like this?

So now, if you want to use pigeonhole principle, the two things I need to figure out are what are the pigeons and what are the pigeonholes. Once you know what are the pigeons and what are the pigeonholes, we can very easily apply that principle because you know that there is some pigeonhole with these many and then use it to for whatever we want. But the question is that what are the boxes and balls.

So, how will you come up with an answer to this, what are the boxes and what are balls. And that is where the difficulty is. In applying the pigeonhole principle, always the difficulty is to figure out what are the boxes and what are the balls. Now, in this case, how will you proceed to solve this? So let me state something very general in the start of this course that, combinatorics is a topic where you need lots of practice, you need to do several exercises to be able to solve questions when you are asked for exam, for example or, like in later times.

Because unlike every other area in mathematics, there is less of a structural or general approach, but rather more of a situation where you get some idea to click and then you can use it. And you can use that idea to basically convert it into an already solved form or something like that or apply some general principle that you already know. But to be able to convert into that form, where you can apply this, you need to come up with some way.

And this usually is very very different for each of the problems depending on the flavor. That is what makes combinatorics one of the most difficult topics to work with. Now, the only solution to this is to solve as many questions as you can. So, when we finish the lecture, we will have some questions for you. Definitely you should try to solve it. But apart from that, you also look at the book, the books that we suggest have many exercises and go through the exercises and try to solve as many as you can.

This might be a little bit time consuming, but there is no other way. So yeah, this is my advice. Now, another advice I want to say is that, whenever a problem is presented to you, for a mathematics course, in fact in any course this is something that you should do. But in general, here, I want to say that once you hear the problem clearly, before the solution is discussed, you please pause, and try to think about this.

This is an advantage of an online course, where you can basically pause the video and then you can try to solve it on your own. Because on the class, you have very limited time, the teacher can give maybe a few minutes to think, but if he gives a lot of time for you to think, then of course, he cannot finish the topics. So, therefore he is much more restricted there. But on the other hand, for a lecture like this, where you have the option to stop the video and think about it, you should try to spend some time thinking.

And then once you think about it, you may not be able to solve it, that is okay. Once you think, several ways, and then you see the solution, that will definitely help you in understanding the applications tremendously. Okay, so with that, let us continue. So, how do we go about solving this question? First thing that you observe here is that, you need to show that, something is a multiple of 7.

So basically, what you are talking about is divisibility. Here, divisibility by 7. Now, when you talk about divisibility, what comes to mind is that if something is divisible, then the remainder is 0. So, then you get an idea that okay, we have to show the remainder is 0. Now, we want to show that remainder is 0 for the difference of two numbers. So, when is the difference of two numbers have remainder 0? Well, if they have the same remainder, when divided by the number, then the difference will have that remainder and will be gone. And then the difference will be multiple of the number that we are talking about.

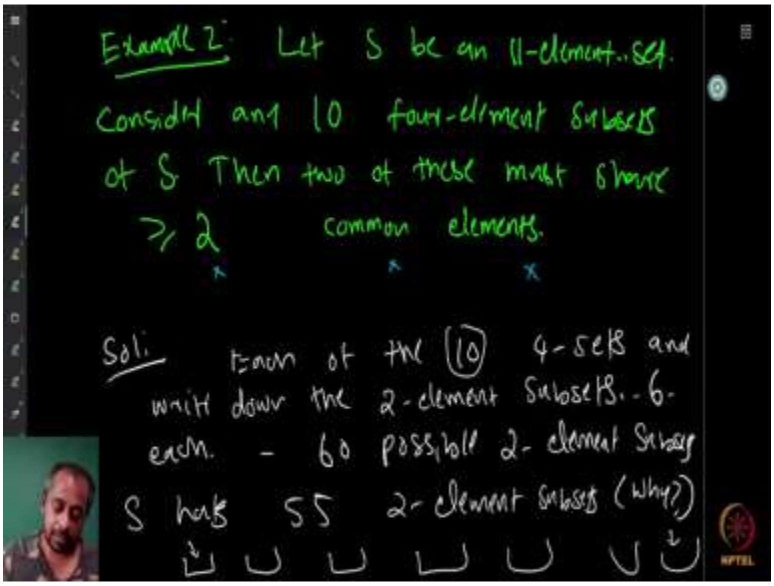
The difference will be 0 for the remainder, so remainder will be 0, and then we get the multiple. And that is what we are going to use here. We have, we are given 8 integers. And we have to show that difference is a multiple of 7 for some pair of them. So, what we are going to do is that we take our integers and divide all of them by 7, one by one. So, I take the numbers, let us say a_1, a_2, a_3,…, a_8. So, if I divide each of them by 7, then I get the remainders r_1, r_2, r_3, …, r_8.

Now, when you divide a number by 7, what are the possible remainders? A remainder can be either 0 can be 1, 2, 3, 4, 5, or 6. These are the only possible remainders when you divide by 7. Now we have 8 numbers which we got as the remainders. And each of the remainder now can be put into one of the boxes labeled 0 to 6. Because they are the remainders when divided by 7, so therefore, it will be one of these numbers.

So, I am going to put these 8 numbers into these 7 boxes. Now, if I go and put it like this, then what do we know? Then, since there are 8 numbers, there will be some box in which two numbers will come together. That is, two of the remainders will be there. So that means that they must be the same because, I am going to put a remainder into a box only if it is that number, a particular number 5(say), if it is 5, the remainder is 5.

So, you take the corresponding numbers, let us say a_i and a_j for which the remainder is the same, so $a_i - a_j$ has remainder 0 when divided by 7, that is why it is divisible by 7. It is a multiple of 7. So that is all. So, once you figure out what are the pigeons and what are the pigeonholes, that is it, there is nothing else. So, this was the first example.

(Refer Slide Time: 14:29)



Now, let us look at another example. So here, the following question is, you are given a set with 11 elements, out of this set you are going to make 4 element subsets. So, you take any 4 of them, you will get a subset, you take some other 4, then you take another subset. So, you pick 4 element subsets and take any 10 of them whatever you like. So, pick any 10 of the 4 element subsets of the 11element set. Then, 2 of these 10 sets that we are talking about, must share at least two elements.

So, how do you show this? Okay. Again, if you like, please pause and think about it. Now, okay, so here is the solution! So, we are talking about two of these elements are going to have two or more elements in common. If I want to show there are two elements in common, I can say that the two element subsets have an intersection, that is also good enough. If I show that these two sets have a non-empty intersection and the two element subsets have non-empty intersection that is also fine.

So, to do this, what I take is that, I take each of the ten 4 element subsets and find consider the two element subsets of them. Now, if you take a 4 element sets, how many two element subsets you can find? This is the question that you need to figure out, we will see the general method later, but most probably you are already familiar with it. One can easily show that there will be 6 of them.

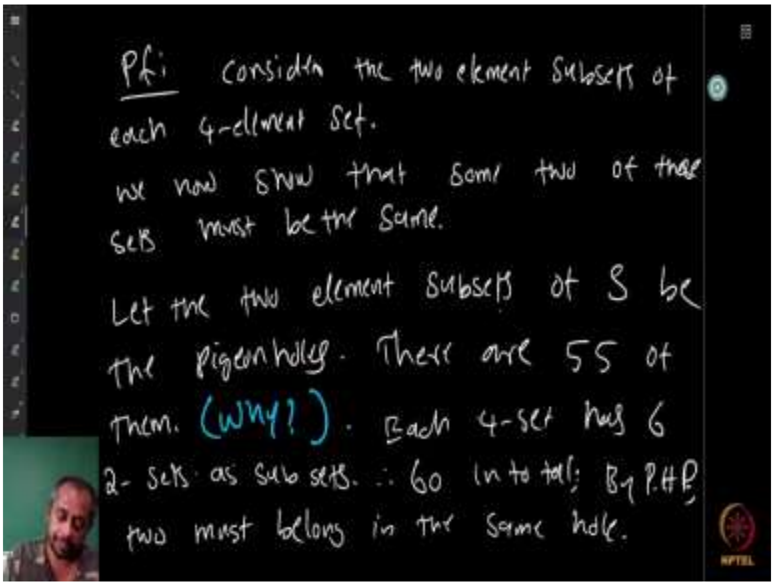
So, each 4 element set has 6 two element subsets. Now there are 10 sets, so there are how many possible subsets? There are 60. So, this gives 60 possible two element subsets. Now, if any two of them are the same, that means that those two sets had intersection of size at least 2. Right! So, we want to show that this is the case, there will be 2 with this property, 2 will be the same.

Now, how do you show this? Well, we started with an 11 elements subset, S is an 11-element set.

Now, look at all possible two element subsets of this. So, S has 55 2-element subsets. You know why? That is something that you need to figure out. Whenever I write 'why' that is the part that I have not given you a proof, that you have to prove it yourself, you have to think about it and do it. Okay. So, fifty five 2-element subsets are there. Now, once you know that there are only 55 possible 2-element subsets we can make pigeonholes, we make the 55 labeled boxes now, with each of these 2-element sets.

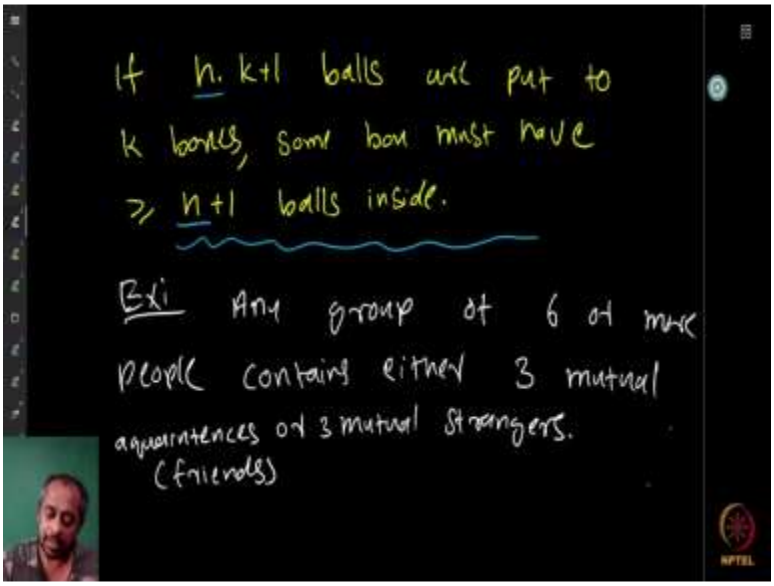
Now, we know that out of the 60 guys, each of them must be one of these two, I mean one of this 55 because there is no other possibility, they all come from the set S. All the elements come from the set S. So, I am going to put the 60 guys into the 55 boxes. 60 is more than 55. Therefore, one box must have at least two inside. And that says that these 2, 2 element subsets are the same and therefore they must have come from two of these, two different 4 element subsets that we are looking at out of the 10, now that is it.

(Refer Slide Time: 20:22)



So, we have proved that, this is indeed the case. So here is the proof. I have written here, consider the two element subsets of each of the 4-element set, we show that some of the two sets must be the same. And for this, we look at the 2 element subsets of S as the pigeonholes, there are 55 of them. Again, think about it why and each 4 element subsets has 6 two element subsets and again why? And therefore, 10×6 which is 60 in total, by pigeonhole principle two must be the same and that is the proof.

(Refer Slide Time: 21:03)



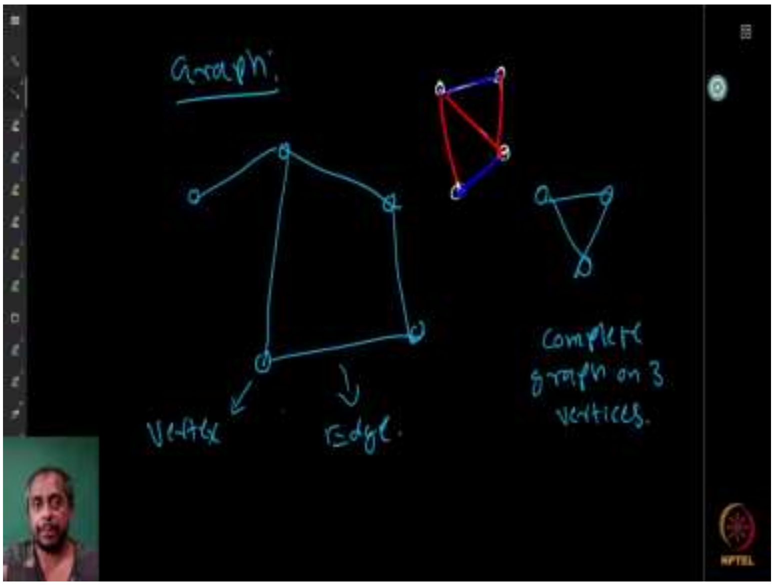
Now, the same pigeonhole principle we can write in a slightly more useable form which I will say as follows. Suppose, we have instead of $k + 1$ balls let us say we have $n.k + 1$ balls. Now, these are put into k boxes. That is we have n into k plus 1 balls and we are going to put them into k different boxes. Then one of the boxes must have greater than or equal to $n + 1$ balls inside. The same argument that we did earlier works.

Because, we know that if each box had at most n, then we have at most $n \times k$ balls inside. That is the proof. Now, we want to use it to prove a very simple but very, very important result. Look at this example, it may look like nothing but it is the start of a very big area in combinatorics, one of the most famous areas in combinatorics called Ramsey theory. And, all this can be started from this particular question that we are discussing. But let us look at what is it for the time being.

Any group of 6 or more people contains either 3 mutual acquaintances or 3 mutual strangers. So, you go to a bus stop let us say and you find 5 people sitting there and with you, there are 6 of them or maybe there are 6 of them or 7 of them already you join there are many of them. And now the claim is that there will be at least 3 guys who have met each other or at least 3 guys who have never met each other.

Acquaintance means that a person one knows slightly, but who is not a close friend. But in short, we will call it by friends. So, there are 3 mutual friends or 3 mutual strangers. And we want to prove this using pigeonhole principle. So how do you do this? Now, of course, we can use set theory to do this, but let us introduce a new concept just for proving this.

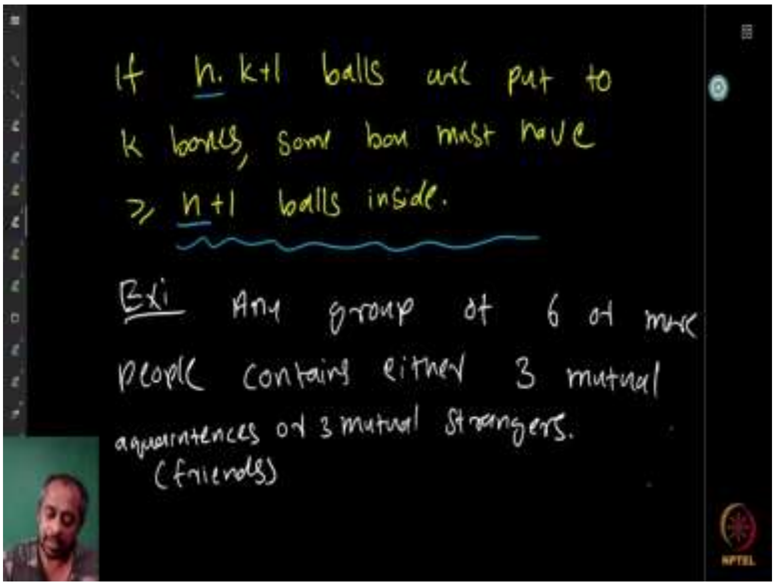
(Refer Slide Time: 23:58)



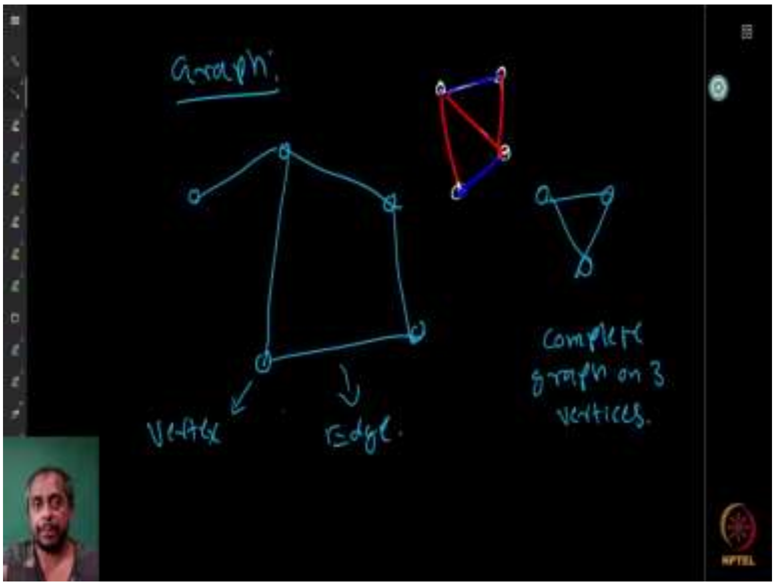
This is a concept of graphs. What is a graph? A graph is a set, so you have a base set, called the set of vertices which we are going to denote as points or dots on the plane. So here we have these dots which are called the vertices of the graph. And then, you have two element subsets, this forms relation, a binary relation among the elements of V. So, among the vertices, you have a relation which is called edge relation.

So whenever two vertices are in a relation then I will say there is an edge and I represent edge by a line segment joining the corresponding vertices. And for the time being, this is a graph for us. So, we have vertices and two element subsets. And now, the advantage is that, this graph can represent many things. For example, vertices can represent people or it can represent terminals in a communication network or components in an electronic circuit or many, many things like this.

And then the edges can be like for example, the channels of communication or you can say that like whenever we are friends then I put a line between them, you know, there is a relation. Now we can even go further, we can say that, I can represent different relations like knowing and not knowing can be a relation. Now, here, for knowing relation I will put let us say blue edge and not knowing relation I put a red edge.



So now, I can have colored edges. If the two guys know each other, then I put a red edge or whatever. And, if they do not know each other, I put a red edge between them. So, these are called edge coloring. We have a graph and then color the edges and the color can represent properties or relations. Now, there are many types of graphs we will not go into details at this time, we might look more of this in later part of the course.

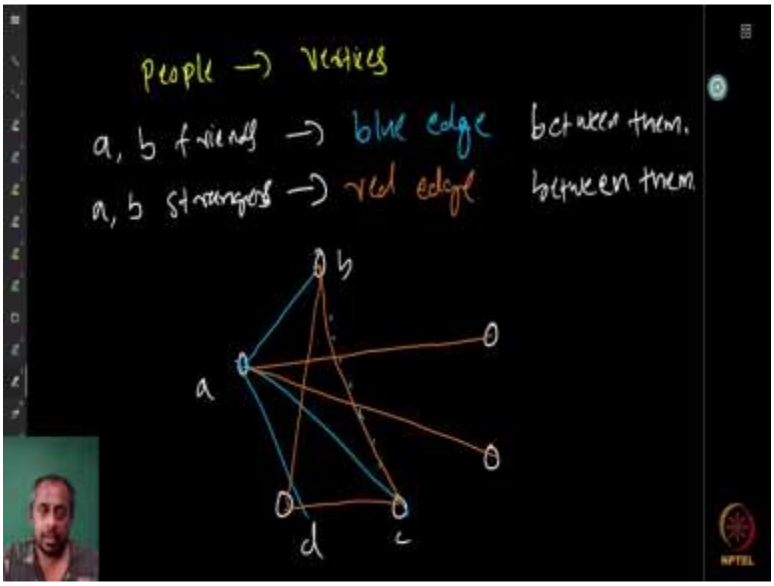


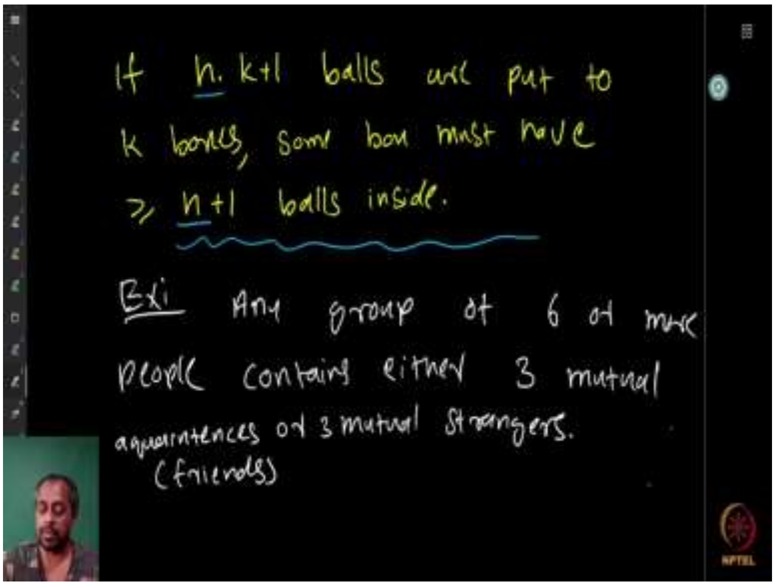
But for the time being if every possible two element subsets are present for a vertex set, then this is called a complete graph. So here is a complete graph on 3 vertices, there are 3 possible edges there, all edges are present, so it is a complete graph. So, with this in mind, we are going to design or model our problem into in terms of graph question. What was our question?

Any group of 6 or more people contains either 3 mutual friends or 3 mutual strangers. Suppose there is 100 people. If you show that, out of this 100 if you take any 6, even within that 6, you have 3 friends or 3 strangers, then the entire group also have 3 friends or 3 strangers. So, we do not really need to look at the larger numbers, we can just work with 6, so that is one advantage.

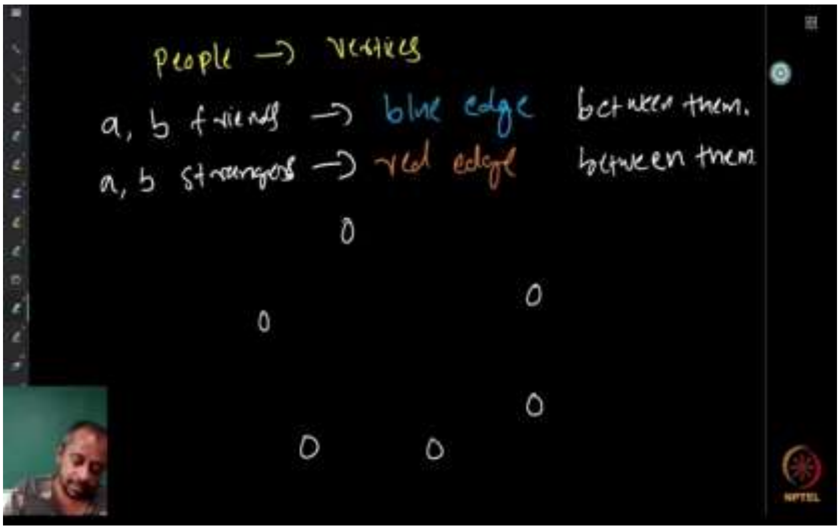
So, we are going to represent now, a graph with the 6 vertices where each vertex represent a person and you have all possible edges and the edges are going to represent if two guys are going to be friend or not.

(Refer Slide Time: 27:59)

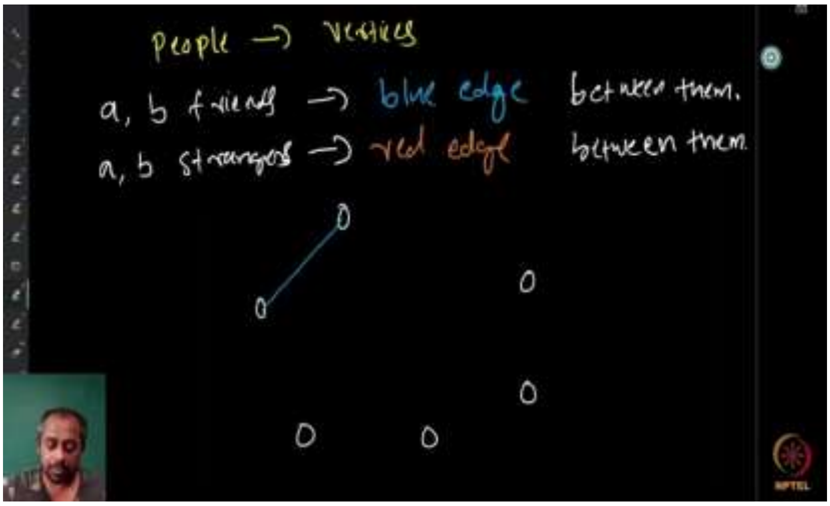




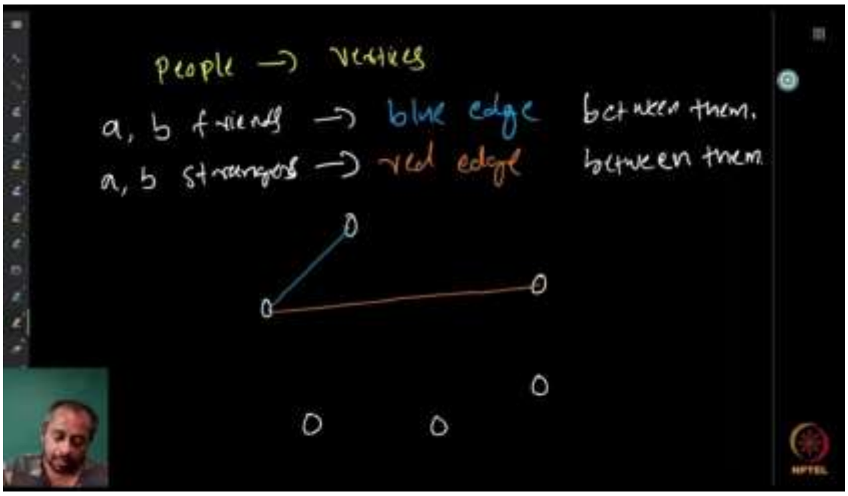
So, if I say, if there is a blue edge, then I say those two guys are friends, if a and b are friends, then there is an edge connecting a and b with blue color. So, I take the 6 different people.



Now, suppose I pick up a particular person. I can talk about him being friend with the other people. For example, like let us say that, these two guys are friends. Either they are friends or they are not friends, so we assume that one of them is the case. So let us say that they are friends.



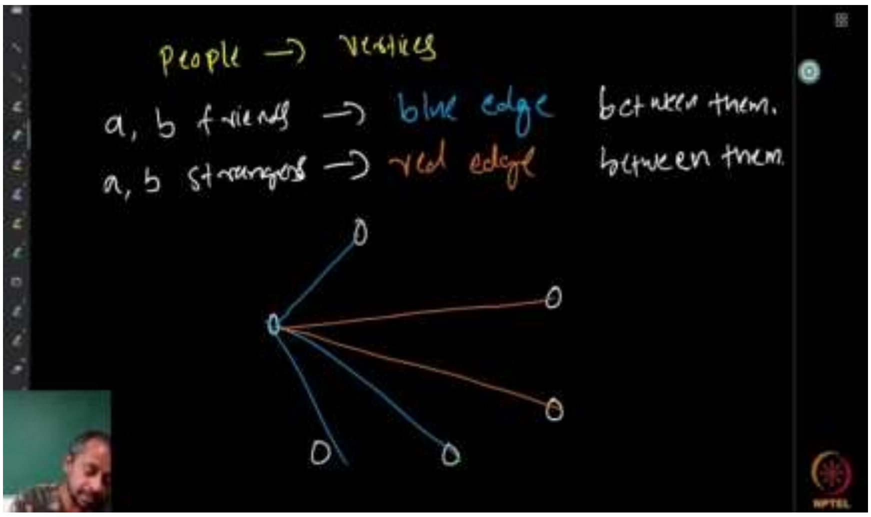
Now, I take this guy and this guy, maybe they are not friends, then they are going to be strangers. So, I am going to put a red edge between them.



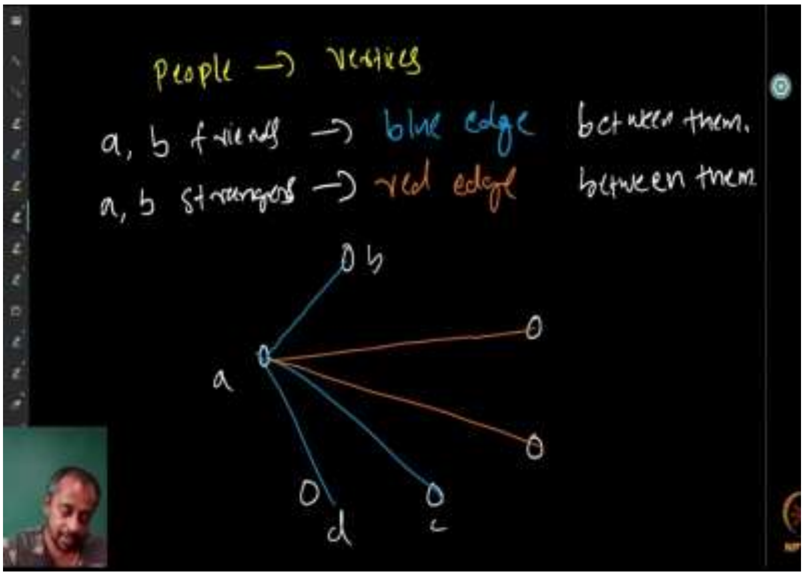
Similarly, we can talk about this particular, first guy being friends with remaining 3 guys also. Now, by our pigeonhole principle, the second version that we looked at, if $n.k + 1$ balls are put into k boxes, some box must have greater than or equal to n plus 1 balls inside.

Now, here we have 5 possible edges going out of this first guy. Now out of these 5 possible edges, we are going to color all of them with two different colors. Just two colors, either blue or red. If they are friends blue edge, and if they are not friends red edge. Now applying the earlier principle that we talked about, we know that since there are only two boxes and there we are going to put 5 inside, at least one of them must contain 3 or more inside.

So, which means that at least for some 3 vertices, the color is the same. So maybe they are, friends or not. That does not matter, since both are identical, symmetric cases. So, the 3 of them must be of the same color. Let us say that 3 of them are not friends. Maybe, all of them are friends, that is okay. But we just want at least 3. So now, we know that at least 3 we have the same color.



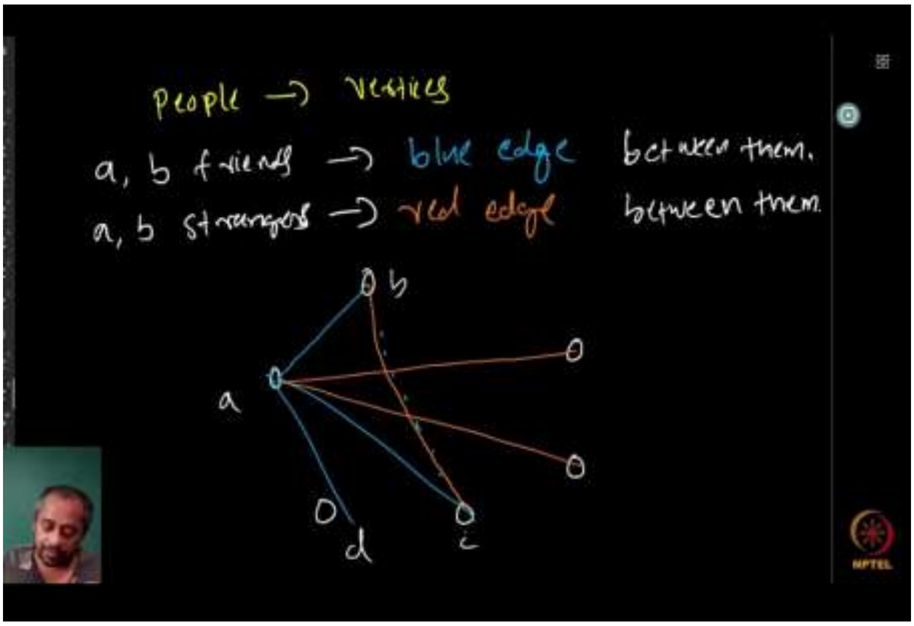
Now, what do we know? So, we have applied pigeonhole principle. First let us put some names to the vertices, say a, b, c, d. Now b, c and d are all friends with a.



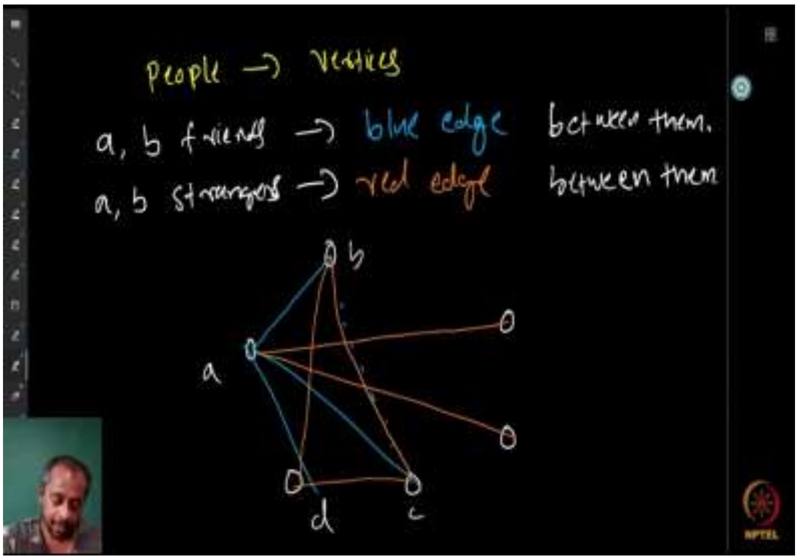
Now, the question is that what can you say about the guys b and c?

Here, b and c are friends then what does it say? It says that there is a blue triangle which says that all the 3 are mutual friends, which is what we wanted to prove to begin with.

If there is a blue edge, then I am already done. So therefore, I can assume that, (b, c) is a red edge.



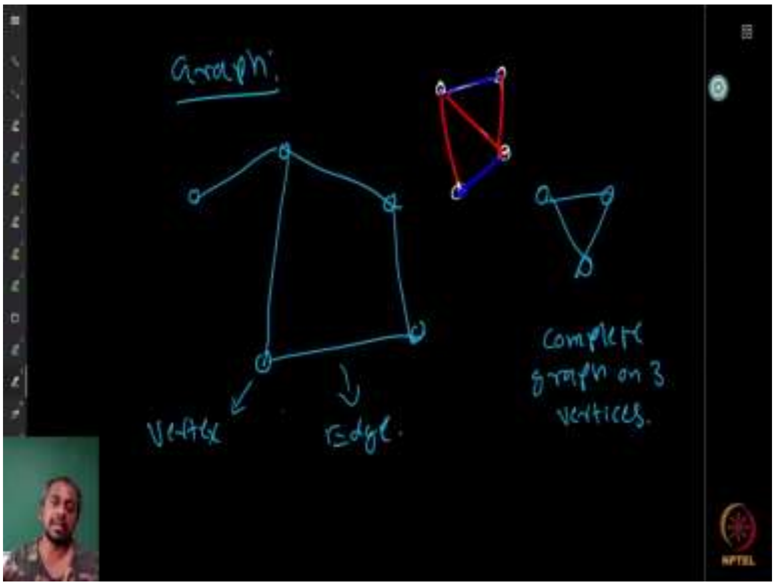
Now, similarly, if there is a blue edge between c and d then also I am done because I get a blue triangle. So, this must be red.



Now, between b and d, if it is blue, again I get a blue triangle. So, it must be red. But then I get a red triangle between b, c and d. So, if I do not have a blue triangle, I have a red triangle.

But maybe, this did not happen, maybe instead edge (b, d) was blue, but then, again I have a blue triangle namely abd. So blue triangle means that there are 3 mutual friends, red triangle means that there are 3 mutual strangers, so one of them must be always there. So, this is what this result says. You have any 6 or more people, there will be either 3 mutual friends or 3 mutual strangers.

(Refer Slide Time: 32:57)



Now this says something very fundamental. We will discuss a little more about this later. But it says that, this is something very much like the pigeonhole principle that if you are talking about very, very large numbers. 6 is not very, very large, but compared to 3, yes it is. If we are talking about very very large numbers, then no matter how chaotic you try to make it there will be still some order.

And that is what this really boils out. So, what we are saying is that, no matter how we are going to color with the two colors for the time being, or some fixed number of colors (we will see later) you can still find very structured subgraphs inside, smaller sized but still with a lot of structure and properties that we are looking for.