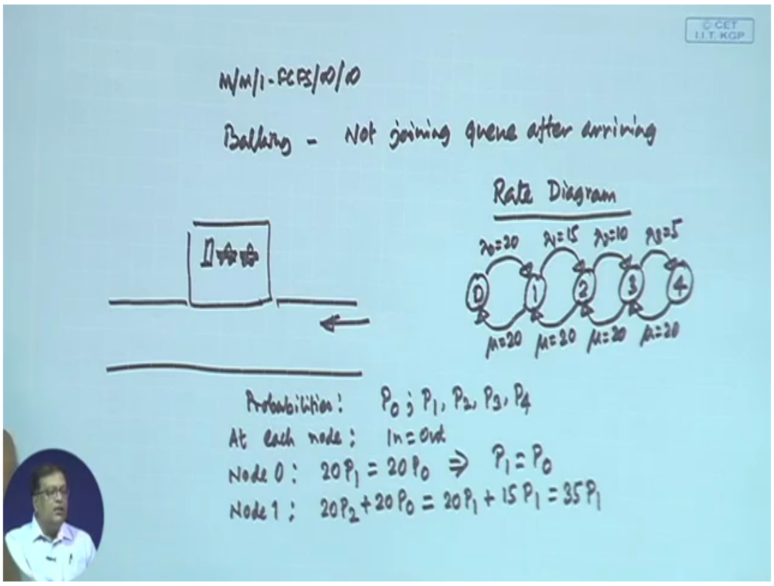
Course on Decision Modeling
Professor Biswajit Mahanty
Department of Industrial and System Engineering
Indian Institute of Technology Kharagpur
Lecture 19
Module 4
Finite Queue Space and Queuing Cost Models

Let us now begin once again to our next module that is on Finite Queue Space and Queuing Cost Models. Now the finite queue space and queuing cost models these are very important topics particularly in the context of queuing theory. Initially let us look at what really happens at finite queue space.

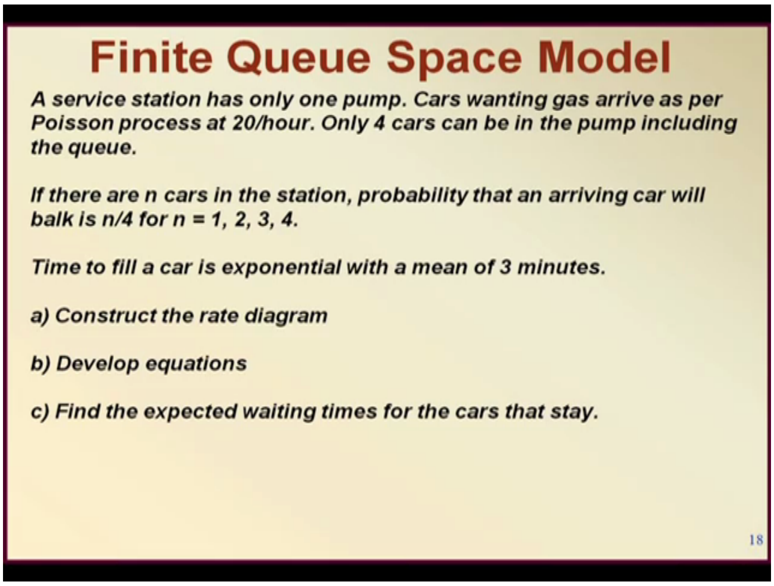
(Refer Slide Time: 0:56)



Please remember that all those things that we have said about M/M/1 and M/M/S system really when we talk about M/M1 really speaking this M/M/1 was something like M/M/1-FCFS/infinity/infinity whatever those last two infinites mean the last two infinites are meaning that there is an infinite queue space and there is an infinite input or calling population.

But what happens if they are really not infinite, see all this simple formula that we derived for our M/M/1 systems obviously the whole true for Exponential or Poisson distributions that is mark of processes also the queue space should be infinite and calling population is infinite. What happens if input space is not infinite? You recall we have discussed about a particular thing called balking what is really balking, balking means that not joining queue after arriving not joining queue after arriving the customer has come but it did not join the queue because because the may be the queue is long or may be the there is no place really to wait. Particularly in highways suppose there are something like a petrol pump and in that petrol pump there is not enough space that is available, right.

(Refer Slide Time: 2:53)



So this is here is a problem only one pump is available cars that means it is a single server. Cars wanting gas arrive as per the Poisson process at 20 per hour. But only 4 cars can be in the pump including the queue, there is no other space and you just cannot wait on the highway, right you just cannot park your car on the highway outside the pump, right so you have to go away you have to wait for the next pump.

That is what really happens to the customers the next pump that is not what we really are interested here we are interested in what is happening in this particular pump. So the arrival process does not remain exactly Poisson why? Because the car has come but it has balked it did not join by seeing the length of the queue. So what exactly should be there? In fact I know sometimes there is another kind of balking there are two kinds of balking that is happening here.

First of all if there is the entire queue space is full then also the people go away but what really happens let us look at this. See this is the pump, this is the highway here is the pump and here is one car that is waiting, right. Now you are coming from here as you come here you think look here there is one car is already there should I join or should I go ahead and go to the next pump which I might get free.

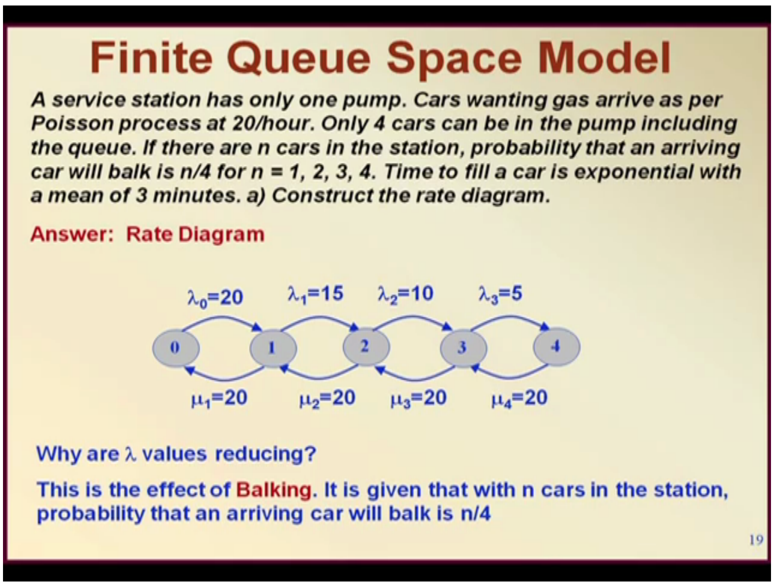
So you see even when there is may be only one car you know which is getting may be gas you may still think let me not join. So there is a chance of not joining even if there is one car obviously and if there are two cars already one car is getting fuel the other is waiting then you may think oh I have to wait long time, right so this chance of balking goes up.

So that is what look at the problem if there are n cars in the station probability that an arriving car will balk is n by 4, right. So n equal to 1, 2, 3, 4, right that means when there is only one car in the system the 25 percent cars are going away, when there are 2 cars in the system 50 percent of the cars are going away, when there are 3 cars in the system 75 percent cars are going away and when there are 4 cars in the system all cars are going away, why? Because there is no place to wait.

So really speaking it is a finite queue space model because at no point of time the pump will hold more than 4 cars, right they will be going away. So all the formula that we have derived earlier they are not applicable we have to redo the whole problem form the birth and death process, draw the rate diagram and move from there, right. So how how that rate diagram should look like, right rate diagram.

In this case the rate diagram should be first of all we have to know how many states how many states are there there are only 5 states that is 0, 1, 2, 3 and 4, right. So from the first 0 state one car can arrive from one, another car can arrive, another car can arrive and then another car can arrive and if there are service then all these cars will be going away. So you see that is the kind of rate diagram that we can look into.

(Refer Slide Time: 7:26)



So let us look at the rate diagram that we have here, in this rate diagram see there are first of all there are 5 states, n equal to 0, n equal to 1, n equal to 2, n equal to 3 and n equal to 4, right. And what are the values of lambdas and Mus, lambda 0 equal to 20, lambda 1 15, lambda 2 10 and lambda 3 5 and Mus are all 20. But why are lambda values reducing? Look at this balking effect the probability that an arriving car will balk is n by 4.

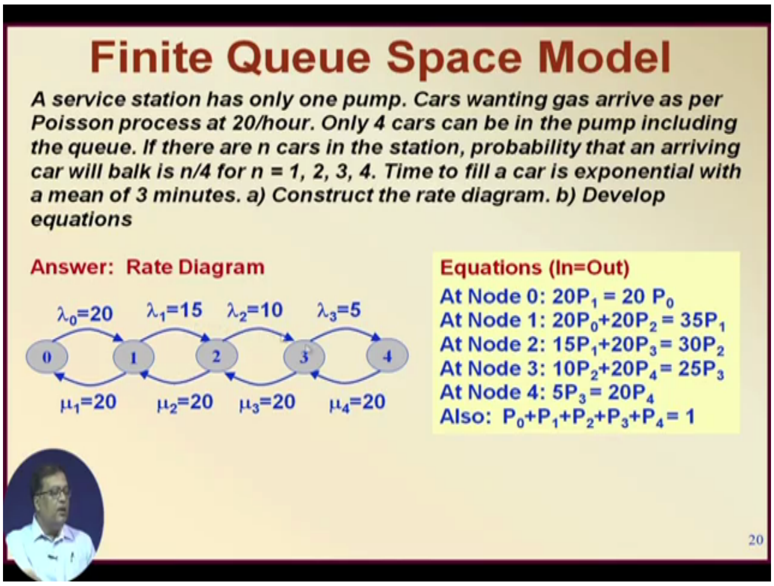
So when there are 0 cars in the system entire set of cars are coming that is at 20 per hour. But when n equal to 1 that means there are 1 car waiting then 25 percent cars are going away that is n by 4, in this case 1 by 4. So 1 by 4 of 20 is 5, so 5 cars go away that means 1there 5 cars will really join, alright. And there after 10 and finally 5 and then 0 that means there is no such system beyond this.

So in this let us look at lambda 0 equal to 20, lambda 1 equal to 15, lambda 2 equal to 10 and lambda 3 equal to 5 but Mus are all 20 so there is no need to differentiate them, right. So once we have this particular rate diagram we know we have to say and please remember the rate diagram really depicts the steady state situation. So if this is a steady state situation how do I get the probabilities, you see that what are some system probabilities the probabilities are there are how many different probabilities can you think off for this particular system there will be only 5. What are they? P0, P1, P2, P3 and P4 that means probability of 0 percents in the system probability of 1, probability of 2, probability of 3, probability of 4.

And how the equilibrium will be obtained already we know that In will be at every node In equal to Out. So at node 0 what happens at node 0 let us see at node 0 how much is In how much is In 20 P1, how much is Out 20 P0 alright. So because both the rates are 20, so that means there is an equilibrium and this will give what P1 equal to P0, so we could get one relation very easily by looking at the rate diagram.

What happens at node 1? How much is In how much is In is node 1 20 P2 20 P2 and 20 P0 and how much is Out? Out is 20 P1, right and 15 P1, alright. So that is equal to 35 P1.

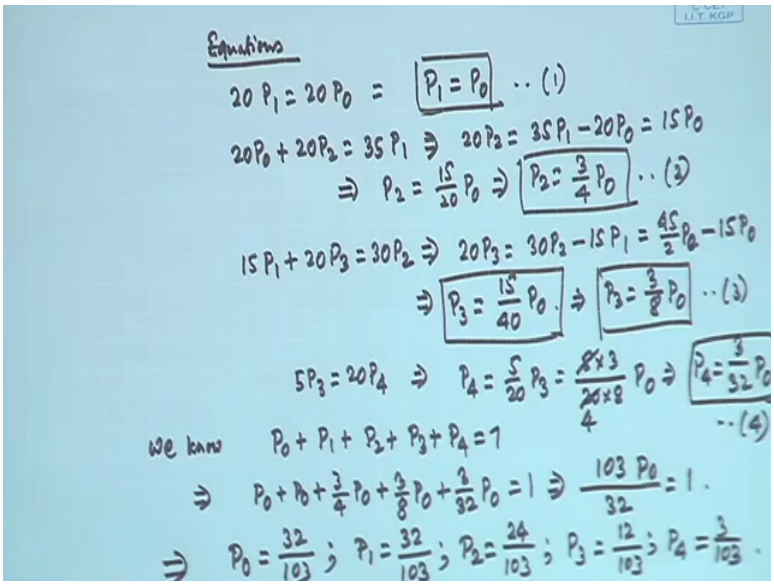
(Refer Slide Time: 11:52)



So like this if you do all these calculation you see all these node balance are to be obtained at different nodes, right and you can find at Node 0 20P1 equal to 20P0, at Node 1 20P0 20P0 so this is 20P0 that is n and 20P2 that is n equal to 35P1 why 35 because 15 and 20 and into P1. So I hope you can do all these node balances let us see one more that is at Node 4 at Node 4 what really happens 5P3 equal to 20P4, right 5P3 because 5P3 equal to 20P4 so we get all these equations and we should also remember the total process is P0 plus P1 plus P2 plus P3 plus P4 equal to 1. Now how do we find out so we get so we get so many equations 1, 2, 3, 4, 5, 6 equations and how many are knowns 5 see there are 5 equation and we have got 6 equations that means one of the equation must be redundant that is also interesting fact here that one equation is redundant.

So you know you cannot use all the 6 equations you have to use only 5 equations but you must remember those 5 equations whatever they are the last equation has to be used, right that means we have to use the last equation and we have to use the another 4 equations out of these. So let us see that and try to calculate and find out what should be the value of the probabilities.

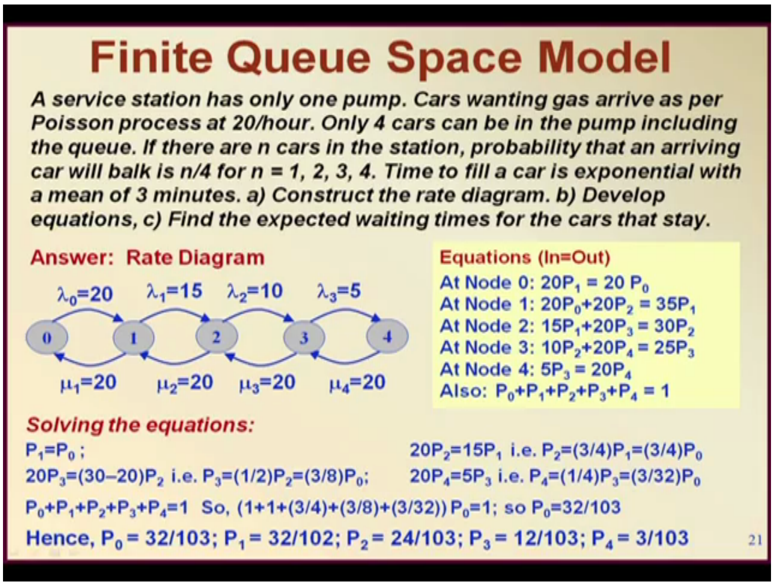
(Refer Slide Time: 13:50)



So these are the equations 20 P1 equal to 20 P0 so from here we can find out P1 equal to P0 one result. Second 20 P0 plus 20 P2 equal to 35 P1 is it alright? So this will give 20 P2 equal to 35 P1 minus 20 P0, but look at 1 P1 equal to P0, so 35 P1 is nothing but 35 P0 so this may be written as 15 P0, right. So this will give P2 equal to 15 by 20 P0 or P2 equal to 3 by 4 P0, right. So this may be our second equation.

Then and the third equation we have got is 15 P1 plus 20 P3 equal to 30 P2, right 15 P1 plus 20 P3 equal to 30 P2. So here we can write 20 P3 equal to 30 P2 minus 15 P1 but 30 P2 3 by 4 into P0, so equal to 90 by 4, right equal to 90 by 4 or 45 by 2 P2 minus sorry P0 minus P1 is P0 15 P0. So this will give this will give P3 equal to so this will be 15 by 2 so 15 by 40 P0, right. And finally 5 P3 that is P3 equal to 3 by 8 P0 and finally 5 P3 equal to 20 P4 where we get P4 equal to 5 by 20 P3 equal to 5 into 3 by 20 into 8 P0, so this will be 4 so P4 equal to 3 by 32 P0.

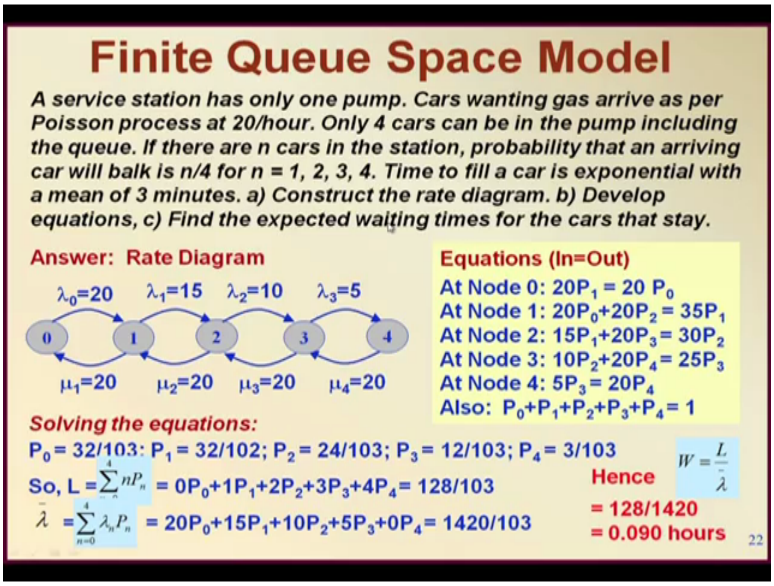
(Refer Slide Time: 17:20)



Let us look whether they are true or not, see this is what we have got solving P1 equal to P0 and then P2 equal to 3 by 4 P0 and then your P3 equal to 3 by 8 P0 and P4 equal to 3 by 32 P0. So we got all the correct results. So look at the calculations once again see therefore we know we know P0 plus P1 plus P2 plus P4 plus P4 equal to 1, right. So P0 plus P0 put all these results P0 plus P0 plus 3 by 4 P0 plus 3 by 8 P0 plus 3 by 32 P0 equal to 1. So if you add all of these so here you will get 32, here you will get 32, 32, 64 plus 24, 88 plus 12 that is 100 plus 3, So 103 P0 by 32 equal to 1.

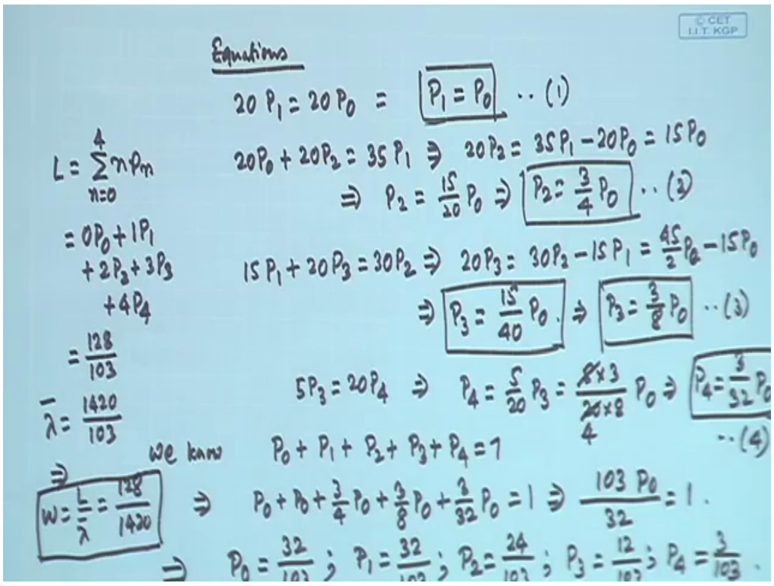
So this will give what is the value of P0? 32 by 103. What is the value of P1? 32 by 103. What is the value of P2? P2 will be 3 by 4 P0 so it will be 24 by 103. What will be the value of P3? 3 by 8 so 8 is 12 by 103 and finally P4 3 by 32 will be 3 by 103. So you see looks a little but difficult but not that difficult really you have to really write down the rate equations obtain that equations by In equal to Out and if you solve those equations which are nothing but some simple simultaneous equations and if you solve you can actually obtain all the probability values.

(Refer Slide Time: 20:06)



So once the probability values are obtained now you can do the rest from the first principles what are they you see already we have found out all the probabilities.

(Refer Slide Time: 20:16)

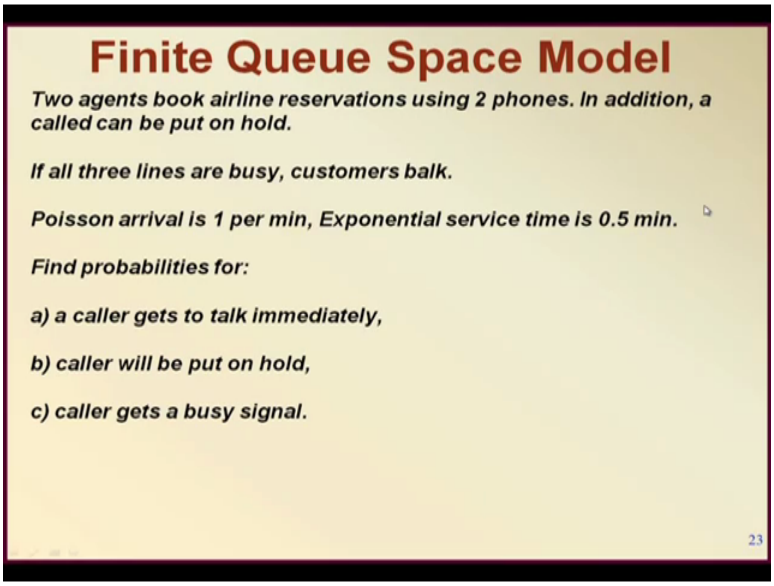


Now what is L, L is sum over nPn, n equal to 0 to 4 because in this case only upto 4, right. So that means 0P0 plus 1P1 plus 2P2 plus 3P3 plus 4P4, right. So if you add them up all then you get 128 by 103, so that is equal to L and then however we also have to remember that calculation of lambda that is because you see littles formula the lambda that they use is not really a lambda but it require to put what is known as lambda bar.

So what is lambda bar we have to really compute that is 20 P0 plus 15 P1 plus 10 P2 plus 5 P3 plus 0 P4, why? Because at different levels the lambdas are different. So you have to find out the weighted average for the lambda values also. So by so doing we have found the lambda bar equal to 1420 by 103 so these two results will give W equal to L by lambda bar equal to ratio of the two which is 128 by 1420, right. So you can find out the waiting time in this manner.

So look at the slide we get waiting time W L by lambda bar is 128 by 1420 is 0$090 hours, right. So that is how we solve what is known finite queue space models.

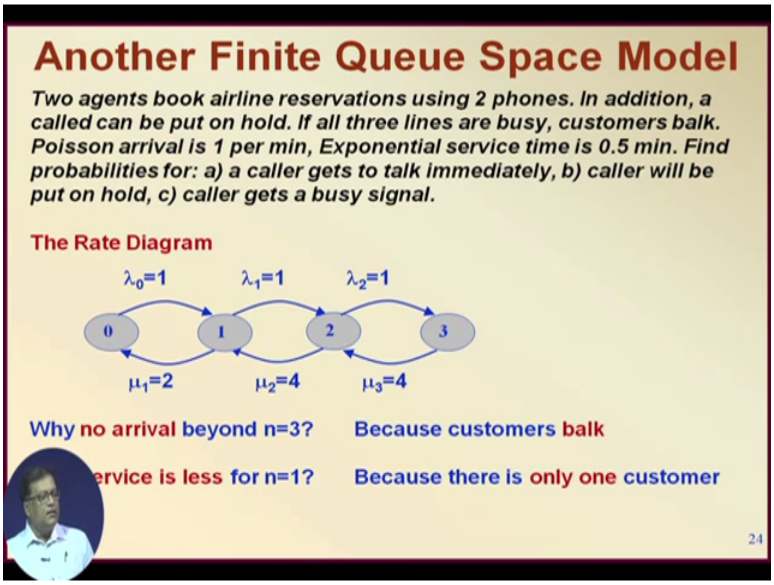
(Refer Slide Time: 22:20)



Now let us look at another problem quickly and try to understand this problem a little further. In this case there are two agents book airline reservations using 2 phones, right. In addiction, a called can be caller can be put on hold. If all three lines are busy, the customers balk. Poisson arrival is 1 per minute and Exponential service time is 0$5 minutes that means 2 per hour.

So find probabilities for a caller gets to talk immediately, caller will be put on hold and caller gets a busy signal. So what is happening there are 2 phones, right there are 2 phones a person but there is a single line and you know a call comes if only one call comes then anyone of the phones will be available, if two calls two people are calling then 2 phones will be called and if another person called then it can be used as a whole facility, right upto three but what happens if another person calls then nothing is possible the person has to go away that means the system state can be only upto that two people are busy they could be on the call and the third person could be put on hold, so that is all nothing else.

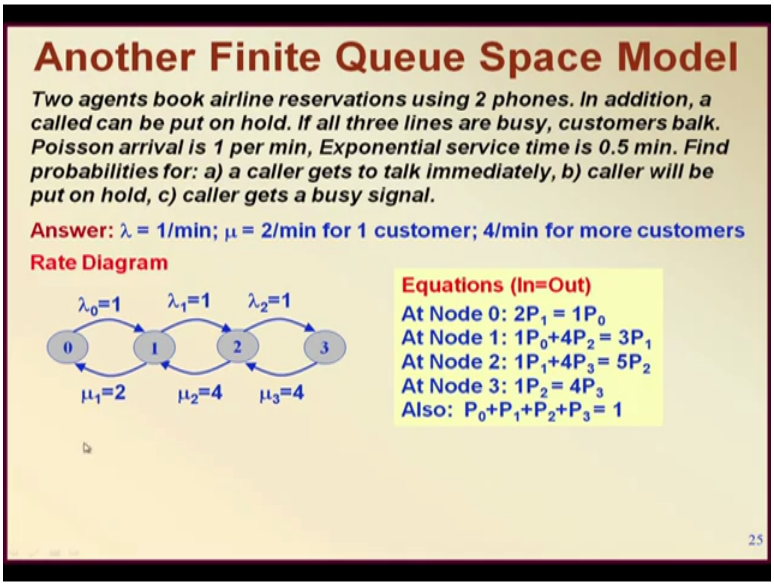
(Refer Slide Time: 23:54)



So look how many states in this case look at the rate diagram the rate diagram will therefore show that there will be 0, 1, 2 and 3, right there will be only 4 different states and out of those 4 different states what are the lambdas and Mu values, the arrivals could be 1 for each case but Mu values if there is a state only upto 1 then Mu value is 2, why Mu value is 2? Because service time is 0$5 minutes, right. So that means it will be 2 per minute but when you know more than 1 because there are 2 phones at that time the service time could be 4 is not or because service time will increase under those situations.

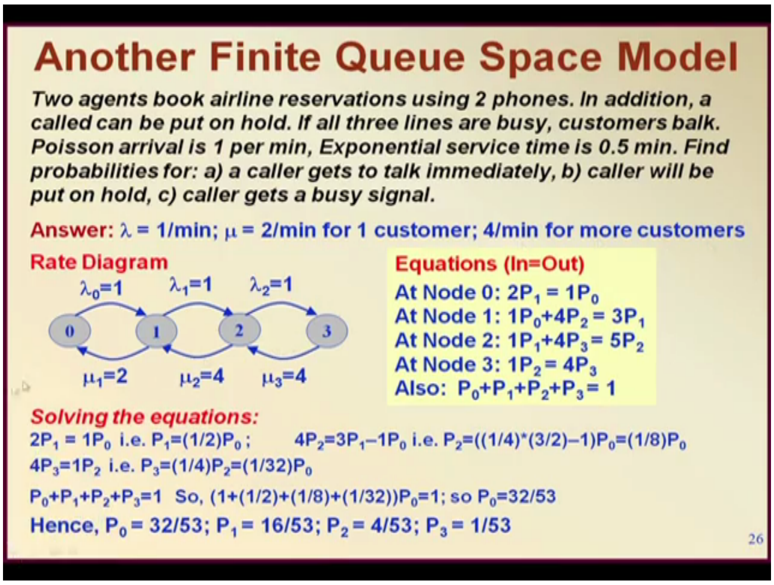
So what happens why no arrival beyond n equal to 3? Because customers balk. Why service is less for n equal to 1? Because there is only one customer, so this must be remembered.

(Refer Slide Time: 25:07)



So what we do again by similarly way we find out the equations how do we find the equations look here at Node 0 2P2 that is In sorry 2P1 equal to 1P0 that goes out. At 1 what is In 1P0 plus 4P2 equal to 2P1 plus 1P1 equal to 3P1. Similarly at Node 2, 1P1 plus 4P3 equal to 5P2 and at Node 3, we have what is In is 1P2, what is Out 4P3 and P0 plus P1 plus P2 plus P3 equal to 1. So these are all the equations that we need to write.

(Refer Slide Time: 25:54)



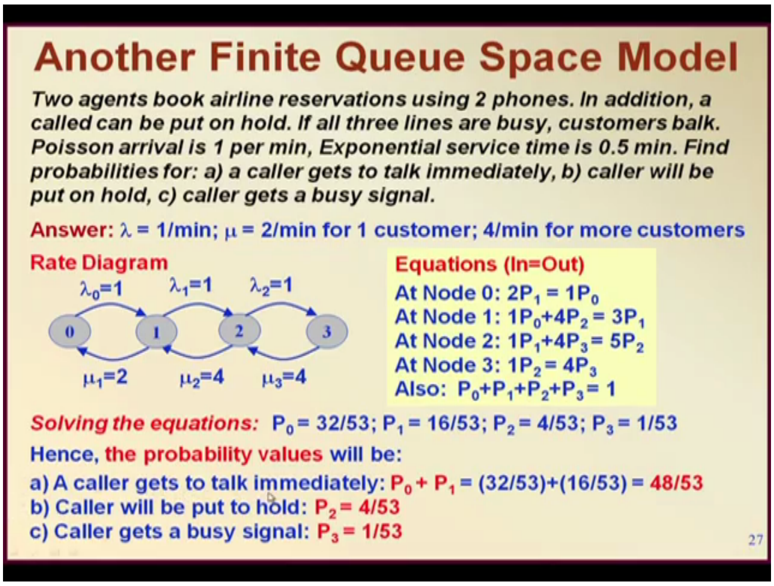
And once we write all those equations then we can sole like we solved in our previous case obviously using out of 4 may be 3 equations and the final equations and then after finding each probability in terms of P0 when you solve them all then you may find P0 is 32 by 53, P1

16 by 53, P2 4 by 53 and P3 1 by 53. Now once we find out all the probabilities now let us try to get the answers.

What is the probability that a caller gets to talk immediately? If these are my 4 probabilities when a caller gets to talk immediately either the system is free or there is only one is it not. The caller will gets to talk immediately if at least one of the 2 phones are available that means that probability should be P0 plus P1. What is that probability? 32 by 53 plus 16 by 53 equal to 48 by 53. What is the probability that caller will be put on hold? That means both the phones are busy but the hold facility is available, when that can happen when we have the probability of P2 means there are only two things two people in the system.

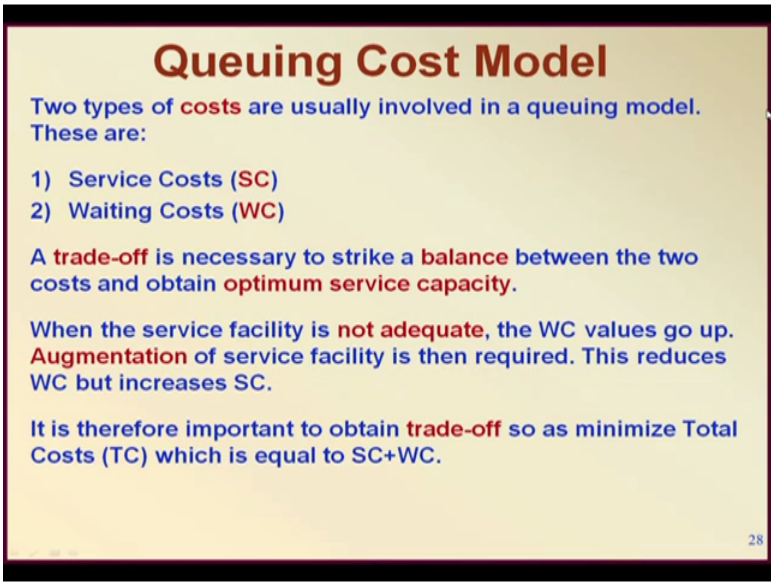
So that probability will be P2 but it is 4 by 53 and when the caller will get a busy signal that is P3 because probability more than this because there cannot be any other system state because rest of the customers balk that is 1 by 53.

(Refer Slide Time: 27:43)



And the same is shown here in the result caller gets to talk immediately P0 plus P1 that is 48 by 53, caller will be put to hold P2 is 4 by 53, caller gets a busy signal 1 by 53, right. So these are the details of the finite queue space model.

(Refer Slide Time: 28:08)



Next one which we need to discuss further that is called the Queuing Cost Model, right. The queuing cost models essentially talks about the there are two types of cost that are involved in a queuing model, what are they? Service cost, cost of service and the other is an waiting cost because the customers are waiting. So trade-off is necessary, why? Because if more customers are waiting waiting cost is more what we tend to do we tend to give better supply, a better service.

As we try to give better service the waiting cost will go down, people wait less and less but service cost will go up. So you see total cost which is service cost plus waiting cost for the station is going to go up if the service cost is too high or waiting cost is too high that is where a trade-off is necessary. And the queuing cost models really tries to achieve this kind of trade-off let us go into let us discuss this in our next class, thank you very much.