So we have just seen that a clever trick based on the frequency interpretation of the transitions between successive states, like here, allows us to write a simple set of equations which can be solved recursively, given here, giving pi i plus 1 as a function of pi of i.

More specifically, we have pi i plus 1 equals pi of i times p of i. Divide by q of i plus 1. And this is true for i equal 0 up to m. And to start the recursion, we need to find pi of 0. And this can be done using this normalization condition-- which leads to pi of 0 times 1 plus p0 over q1 plus et cetera equals 1.

Let's illustrate the details of this procedure on a special case. Let's assume that all the p's are the same and all the q's are the same. So this is a special case in which we are interested. So at each point in time, if we are somewhere in the middle, you have probability p of moving up, and probability q of moving down.

Define rho to be the ratio of p over q. Rho can be interpreted as the frequency of going up versus the frequency of going down. If it's a service system, you can think of it as a measure of how loaded the system is. If p equals q, that means that if you are at this state-- you are equally likely to move left or right. So the chain does not have a tendency to move in that direction or in that direction.

If rho is bigger than 1, so that p is bigger than q, it means that whenever we are at some state in the middle, we are more likely to move right, as opposed to moving left. Which means that the chain has a tendency to move in that direction.

And if you think of this as a number of customers in queue, it means your system has the tendency to become loaded and to build up a queue. So rho being bigger than 1 corresponds to a heavy load, where queues build up. Rho less than one corresponds to the system where queues have the tendency to drain down. The system is going to move in that direction.

Now let us write down these equations for that special case. We end up with that, which is pi i times rho, by definition of rho. Once you look at this equation, you realize that pi of 1 is pi of 0 times rho. And pi of 2 is pi of 1 times rho equals pi of 0 times rho square. And so on and so forth.

And you find that you can express pi of i as pi of 0 times rho at the power i for any possible i between 0 and m. And now if we use the normalization condition, we get that pi of 0 times 1 plus rho plus rho squared plus rho at the power m is equal to 1.

Let's now complete the calculations for two special cases. If rho is equal to 1, that means p equals q. Then pi i equals pi of 0 for all i. It means that all the steady state probabilities are equal. This special case is called a symmetric random walk. So you start at the state at a point in time. Either you stay in place, or you have an equal probability of going left or right. There is no bias in either direction.

You might think that in such a process, you will tend to get stuck either near one end or the other end. It turns out that no, in the long run, the symmetric random walk is equally likely to be at any of those states.

And for the special case-- this equation here-- is simply that pi of 0 times 1 plus m equals one. That means that pi of 0 equals 1 over 1 plus m. Which is consistent with the fact that all steady-state probabilities are the same. They are all equally likely. They are end states. And so each one of them, pi i is pi of 0, which is 1 over 1 plus m. The Markov chain is equally likely to be in any of these m plus 1 states in the long run.

Suppose now instead of p equals q, that m is very, very large, a very large number. Let's take m going to infinity. And suppose that the system is on the stable side. That means that p is less than q, which means that there's a tendency for customers to be served faster than they arrive.

In other words, the chain is drifting toward that direction. So that means that rho is less than 1 and what it means is that this infinite series, when m goes to infinity, is the geometric series. And this series is going to be 1 over 1 minus rho. That is, this infinite series is 1 over 1 minus rho. And since pi of 0 is 1 over this infinite series, we end up having pi 0 equals 1 minus rho.

And since we have pi of i equals pi 0 times rho at the power i, we end up having that pi of i equals pi of 0, which is 1 minus rho times rho at the power i, for i equal-- this pi i can be seen as coming from the probability distribution.

They tell us that if we observe that chain at time-- let's say one billion-- and ask-- where is the state of the Markov chain? The answer will be the chain is in state zero, that is, the system is empty with a probability 1 minus rho, or there is one customer in the system. And that happens with probability 1 minus rho times rho. And so on.

So the distribution can be drawn like that. You have here i corresponding to a state and if you put pi of i here, 0 here, then 1, 2, 3-- then pi of 0 is 1 minus rho here. pi of 1 will be rho times 1 minus rho and pi of 2 and so forth.

So if you look at this distribution here, it's pretty much a geometric distribution, except that it has shifted so that it starts at 0 instead of starting at 1. So it's a shifted geometric. This model is the first and simplest model that one encounters when studying queuing theory.

So a final note-- the PMF that we have here has an expected value. And the expectation is given here-- e of x of m is-- let me rewrite it here-- it's rho over 1 minus rho.

And this formula-- which is interesting to anyone who tries to analyze a system of this kind-- tells you the following-- that as long as rho is less than 1, then the expected number of customers in the system is finite. But if rho, this little rho, becomes very close to 1, then you're going to have 1 over something that is very close to 0. And that number will be very, very big.

So when rho becomes very close to 1, that means the load factor is something like-- let's say 0.99-- you expect to have a very large number of customers in the system at any given time.