

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high-quality educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](http://ocw.mit.edu).

**PROFESSOR:** OK, so let's get started. I want to talk mostly about countable state Markov chains today, which is the new topic we started on Wednesday. I want to talk just a little bit about the strong law proof that was in the third problem of the quiz. I'm not doing that because none of you understood anything about it. I'm doing it because all of you understood more about it than I thought you would. And in fact, I've always avoided saying too much about this proof because I thought everybody was tuning them out.

And for the class here, it looks like a lot of you have tried seriously to understand these things. So I thought I would explain that one part of the quiz in detail so that you'd see the parts you're missing, and so that [INAUDIBLE] all these other proofs that we have, talking about the strong law and the strong law for renewals, and putting them together and all of these things, all of them are essentially the same. And it's just a matter of figuring out how things fit together. So I wanted to talk about that because it's clear that most of you understand enough about it that it makes sense.

The situation in the quiz, which is very close to the usual queuing situation and little theorem type things, there's a sequence of  $Y$  sub  $i$ 's. They're IID. There's the service times for  $G/G$  infinity queue, and  $n$  of  $t$ , which is a renewal process for the arrivals to the process. We have arrivals coming in according to this renewal process, which means the [INAUDIBLE] arrival times,  $X$  sub  $i$ , were IID. And we want to put those two things together. And we want to find out what this limit is. If it's a limit, show that it's a limit. And hopefully show that it's a limit with probability 1.

And I think a large number of you basically realize that this argument consisted of a

bunch of steps. Some people with more detail than others. But the first step, which we do in all of these arguments, is to divide and multiply by  $n$  of  $t$  of  $\omega$ . So we're starting out, looking at some particular sample path, and start out by multiplying and dividing by  $n$  of  $t$  of  $\omega$ . The next thing is to claim that the limit of this times this is equal to the limit of this times the limit of this. Almost no one recognized that as a real problem, and that is a real problem.

It's probably the least obvious thing in this whole problem. I'm not saying you shouldn't have done that, because I've been doing that in all the proofs I've been giving you all along. It is sort of obvious that that works. And when you're constructing a proof, especially in a quiz when you don't have much time, things which are almost obvious or which look obvious, you should just go ahead and assume them and come back later when you have time to see whether that really makes sense. That's the way you do research also. You don't do research by painstakingly establishing every point in some linear path. What you do is you carelessly as you can and with as much insight as you can, you jump all the way to the end, you see where you're trying to go, you see how to get there, and then you come back and you try to figure out what each of the steps are.

So this is certainly a very reasonable way of solving this problem, because it looks like this limit should be equal to this limit times this limit. The next step in the argument is to claim that this sum, up to  $N$  of  $t$  of  $\omega$ , over  $N$  of  $t$  of  $\omega$ , as  $t$  approaches infinity-- the argument is that  $t$  approaches infinity-- this  $N$  of  $t$  of  $\omega$  goes through, one by one, a sequence, 1, 2, 3, 4, 5, and so forth. So this limit is equal to that limit. I've never been able to figure out whether that's obvious or not obvious. It is just on the borderline between what's obvious and not obvious, so I'm gonna prove it to you.

And then the next step is to see that  $N$  of  $t$  of  $\omega$  over  $t$  is equal to  $1/\bar{X}$  with probability 1. And this limit is equal to  $\bar{Y}$  with probability 1. The first argument, this equal to  $1/\bar{X}$  is because of the strong law with renewals. And this one over here is because of the strong law of large numbers. And most of you managed to get this. And the whole argument assumes that  $\bar{X}$  is less than infinity, and  $\bar{Y}$  bar

is less than infinity.

Now, how do you go back and actually see that this actually makes sense? And that's what I want to do next. And if you look at this, you can't do it by starting here and working your way down to there, because there's no way you're going to argue that this limit is equal to this product of limit unless you know something about this and you know something about this. So you really have to establish that these things have limits first before you can go back and establish this. In the same way, you have to know something about this before you can establish this.

So what you have to do way after you've managed to have the insight to jump the whole way through this thing, is to go back and argue each of the points, but argue them in reverse order. And that's very often the way you do research, and it's certainly the way you do quizzes. So let's see where those arguments were. Start out by letting  $A_1$  be the set of  $\omega$  for which this limit here,  $N$  of  $t$  of  $\omega$  over  $t$ , is equal to  $1/\bar{X}$ . By the strong law for renewal processes, the probability of  $A_1$  equals 1.

This is stating this in a little cleaner way, I think, than we stated the strong law of renewals originally, because we started out by tumbling together this statement and this statement. I think it's cleaner to say, start out, there's a set of  $\omega$  for which this limit exists. And what the strong law says is that that set of  $\omega$  has probability 1. And now we have some terminology for  $A_1$ , what it actually means. It is the set of  $\omega$  for which this works. You never have it working for all  $\omega$ , only for some  $\omega$ .

Then the next step, let  $A_2$  be the set of  $\omega$  for which the limit is  $n$  goes to infinity of  $1/n$  times the sum of  $Y$  sub  $i$  of  $\omega$ , is equal to  $\bar{Y}$ . By the strong law of large numbers, the probability of  $A_2$  is equal to 1. So now we've established there's an  $A_1$ , there's an  $A_2$ . Each of them have probability 1. On  $A_1$ , one limit exists. On  $A_2$ , the other limit exists. And we have two sets, both at probability 1. What's the probability of the intersection of them? It has to be 1 also.

So with that, you know that equation three is equal to equation four for  $\omega$  in the

sets,  $A_1$ ,  $A_2$ . And also, you know that the probability of  $A_1$ ,  $A_2$  is equal to 1. So we've established this part of the argument down here. Now we want to go up and establish this part of the argument, which as I said, I can't convince myself that it's necessary or not necessary. But since I can't convince myself, I thought, in trying to make up solutions for the quiz, I ought to actually write a proof of it. And I want to show you what the proof is so that, if it's not obvious to you, you'll know exactly how to do it. And if it is obvious, you can maybe sort out exactly why it's obvious.

So this is an epsilon delta kind of argument. We assume that  $\omega$  is in  $A_2$ . That's the set for which the strong law of large numbers holds. There exists some integer,  $n$ , which is a function of both epsilon and  $\omega$ . This is the funny thing about all of these strong law arguments. In almost all of them, you're dealing with individual sample paths. When you start saying something exists as a limit, you're not saying that it exists as a limit for the random variables. You're saying it exists as a limit for a set of sample paths. And therefore, this epsilon here that you're gonna choose, you need some integer there, such that this minus this is less than epsilon.

I think I'm going to have to give up on this. These things run out of batteries too quickly. So we have that this difference here must be less than epsilon if  $n$  is bigger than that  $m$  of epsilon  $\omega$ . That's simply what a limit means. That's the definition of a limit. The only way to define a limit sensibly is to say, for all epsilon greater than 0, no matter how small the epsilon is, you can always find an  $n$  big enough that this difference here is less than epsilon.

Then if  $\omega$  is also in  $A_1$ , the limit of  $N$  of  $t$  of  $\omega$  has to be equal to infinity. If you want to, you can just say, we proved in class that the limit of  $t$  of  $\omega$  is equal to infinity with probability 1, and introduce another set,  $A_3$ . You want to do a probability 1. But let's do it this way. And then there has to be a  $t$ , which is also a function of epsilon and  $\omega$ , such that  $N$  of  $t$  and  $\omega$  is greater than-- that's an integer, by the way. That's greater than or equal to  $m$  of epsilon of  $\omega$  for all  $t$ , which is greater than or equal to  $t$  of epsilon of  $\omega$ . That says that this difference here is less than  $\omega$ . And that's true for all epsilon greater than 0. And that says that, in fact, this limit has to exist. This limit over here is equal to  $\bar{Y}$

with probability 1.

So that's what we were trying to prove here, that this limit is the same as this limit. So we found out what this limit is. We found out that it exists with probability 1, namely on the set  $A_2$ . This is equal to this, not necessarily on  $A_2$ , but on  $A_1$  and  $A_2$ . So we got into there. Now how do we get the fact that this limit times this limit is equal to the limit of [? these. ?] Now we have a chance of proceeding, because we've actually shown that this limit exists on some set with probability 1, this limit exists on some set with probability 1. So we can look at that set and say, for  $\omega$  in that set, this limit exists and this limit exists. Those limits are non-0 and they're non-infinite. The important thing is that they're non-infinite. And we move on from there, and to do that carefully.

And again, I'm not suggesting that I expect any of you to do this on the quiz. I would have been amazed if you had. It took me quite a while to sort it out, because all these things are tangled together. Where am I? I want to be in the next slide. Ah, there we go. Finally, we can interchange the limit of a product of two functions-- say,  $f$  of  $t$ ,  $g$  of  $t$ -- with the product of the limits. Can we do that? If the two functions each have finite limits, as the functions of interests do for  $\omega$  in  $A_1$ ,  $A_2$ , then the answer is yes. And if you look at any book on analysis, I'm sure that theorem is somewhere in the first couple of chapters.

But anyway, if you're the kind of person like I am who would rather sort something out for yourself rather than look it up, there's a trick involved in doing it. It's this equality right here.  $f$  of  $t$  times  $g$  of  $t$  minus  $ab$ . What you want to do is somehow make that look like  $f$  of  $t$  minus  $a$ , which you have some control over, and  $g$  of  $t$  minus  $b$ . So the identity is this is equal to  $f$  of  $t$  minus  $a$  times  $g$  of  $t$  minus  $b$ -- we have control over that-- plus  $a$  times  $g$  of  $t$  minus  $b$ , plus  $b$  times  $f$  of  $t$  minus  $a$ . And you multiply and add all those things together and you see that that is just an identity. And therefore the magnitude of  $f$  of  $t$  times  $g$  of  $t$  minus  $ab$  is less than or equal to this.

And then you go through all the epsilon delta stuff again. For any epsilon greater

than 0, you choose a  $t$  such that this is less than or equal to  $\epsilon$  for  $t$  greater than the  $t$   $\epsilon$ . This is less than or equal to  $\epsilon$  for  $t$  greater than or equal to  $t$   $\epsilon$ . And then this difference here is less than or equal to  $\epsilon$  squared plus this. And with a little extra fiddling around, that shows you have that  $f$  of  $t$ ,  $g$  of  $t$  minus  $ab$  approaches 0 as  $t$  gets large.

So that's the whole thing. Now, let me reemphasize again, I did not expect you to do that. I did not expect you to know how to do analysis arguments like that, because analysis is not a prerequisite for the course. I do want to show you that the kinds of things we've been doing are not, in fact, impossible. If you trace them out from beginning to end and put in every little detail in them. If you have to go through these kinds of arguments again, you will in fact know how to make it precise and know how to put all those details in.

Let's go back to countable state Markov chains. As we've said, two states are in the same class as they communicate. It's the same definition as for finite state chains. They communicate if there's some path by which you can get from  $i$  to  $j$ , and there's some path from which you can get from  $j$  to  $i$ . And you can't get there in one step only, but you can get there in some finite number of steps. So that's the definition of two states communicating. The theorem that we sort of proved last time is that all states in the same class are recurrent or all are transient. That's the same as the theorem we have for finite state Markov chains. It's just a little hard to establish here.

The argument is that you assume that  $j$  is recurrent. If  $j$  is recurrent, then the sum has to be equal to infinity. How do you interpret that sum there? What is it?  $P_{jj}^n$ , super  $n$ , is the probability that you will be in state  $j$  at time  $n$  given that you're in state  $j$  at time 0. So what we're doing is we're starting out in time 0. This quantity here is the probability that we'll be in state  $j$  at time  $n$ . Since you either are or you're not, since this is also equal to the expected value of state  $j$  at time  $n$ , given state  $j$  at time 0. So when you add all these up, you're adding expectations. So this quantity here is simply the expected number of recurrences to state  $j$  from time 1 up to time infinity.

And that number of recurrences is equal to infinity. You remember we argued last time that the probability of one recurrence had to be equal to 1 if it was recurrent. If you got back to  $j$  once in finite time, you're going to get back again in a finite time again. You're going to get back again in finite time. It might take a very, very long time, but it's finite, and you have an infinite number of returns as time goes to infinity.

So that is the consequence of  $j$  being recurrent. For any  $i$  such that  $j$  and  $i$  communicate, there's some path at some length  $m$  such that the probability of going from state  $i$  to state  $j$  in  $m$  steps is greater than 0. That's by meaning of communicate. And there's some  $m$  and some  $p_{ji}$  of  $l$ . Oh, for [? some  $m$ . ?] And there's some way of getting back from  $j$  to  $i$ . So what you're doing is going from state  $i$  to state  $j$ , and there is some path for doing that. You're wobbling around, returning to state  $j$ , returning to state  $j$ , maybe returning to state  $i$  along the way. That's part of it. And eventually there's some path for going from  $j$  back to  $i$  again.

So this sum here is greater than or equal. And now all I'm doing is summing up the paths which in  $m$  steps go from  $i$  to  $j$ , and those paths which in the final  $l$  steps go from  $j$  back to  $i$ . And they do whatever they want to in between. So I'm summing over the number of times they are in between. And this sum here is summing over  $p_{ijk}$ . And that sum is infinite, so this sum is infinite. So that shows that if  $j$  is recurrent, then  $i$  is recurrent also for any  $i$  in the same class. And you can do the same thing reversing  $i$  and  $j$ , obviously. And if that's true for all classes that are very recurrent, all law classes that are transient have to be in the same class also, because a state is either transient or it's recurrent.

If a state  $j$  is recurrent, then the recurrence time,  $T_{jj}$ . When you read this chapter or read my notes, I apologize because there's a huge confusion here. And the confusion comes from the fact that there's an extraordinary amount of notation here. We're dealing with all the notation of finite-state Markov chains. We're dealing with all the notation of renewal processes. And we're jumping back and forth between theorems for one and theorems for the other. And then we're inventing a lot of new notation. And I have to rewrite that section. But anyway, the results are all

correct as far as I know.

I mean, all of you can remember notation much better than I can. So if I can remember this notation, you can also. Let me put it that way. So I can't feel too sorry for you. I want to rewrite it because I'm feeling sorry for myself after every year I go through this and try to re-understand it again, and I find it very hard to do it. So I'm going to rewrite it and get rid of some of that notation. We've already seen that if you have a chain like this, which is simply the Markov chain corresponding to Bernoulli trials, if it's Bernoulli trials with  $p$  equals  $1/2$ , you move up a probability  $1/2$ , you move down with probability  $1/2$ . As we said, you eventually disperse. And as you disperse, the probability of being in any one of these states goes to 0.

And what that means is that the individual probabilities of the states is going to 0. You can also see, not so easily, that you're eventually going to return to each state with probability 1. And I'm sorry I didn't give that definition first. We gave it last time. If the expected value of the renewal time is less than infinity, then  $j$  is positive recurrent. If  $T_{jj}$ , the recurrence time, is a random variable but it has infinite expectation, then  $j$  is not recurrent. And finally, if none of those things happen,  $j$  is transient. So that we went through last time. And for  $p$  equals  $1/2$ , and in both of these situations, the probability of being in any state is going to 0. The expected time of returning is going to infinity. But with probability 1, you will return eventually.

So in both of these cases, these are both examples of no recurrence. Let's say more about positive recurrence and no recurrence. Suppose, first, that  $i$  and  $j$  are both recurrent and they both communicate with each other. In other words, there's a path from  $i$  to  $j$ , there's a path from  $j$  to  $i$ . And I want to look at the renewal process of returns to  $j$ . You've sorted out by now, I think, that recurrence means exactly what you think it means. A recurrence means, starting from a state  $j$ , there's a recurrence to  $j$  if eventually you come back to  $j$ . And this random variable, the recurrence of random variable, is the amount of time it takes you to get back to  $j$  once you've been in  $j$ . That's a random variable.

So let's look at the renewal process, starting in  $j$ , of returning to  $j$  eventually. This is

one of the things that makes this whole study awkward. We have renewal processes when we start at  $j$  and we bob back to  $j$  at various periods of time. If we start in  $i$  and we're interested in returns to  $j$ , then we have something called a delayed renewal process. All the theorems about renewals apply there. It's a little harder to see what's going on. It's in the end of chapter four. You should have read it, at least quickly. But we're going to avoid those theorems and instead go directly using the theorems of renewal processes. But there's still places where the transitions are awkward. So I can warn you about that.

But the renewal reward theorem, if I look at this renewal process, I get a renewal every time I return to state  $j$ . But in that renewal process of returns to state  $j$ , what I'm really interested in is returns to state  $i$ , because what I'm trying to do here is relate how often do you go to state  $i$  with how often do you go to state  $j$ ? So we have a little bit of a symmetry in it, because we're starting in state  $j$ , because that gives us a renewal process. But now we have this renewal reward process, where we give ourselves a reward of 1 every time we hit state  $i$ . And we have a renewal every time we hit state  $j$ . So how does that renewal process work?

Well, it's a renewal process just like every other one we've studied. It has this peculiar feature here that is a discrete time renewal process. And with discrete time renewal processes, as we've seen, you can save yourself a lot of aggravation by only looking at these discrete times. Namely, you only look at integer times. And now when you only look at integer times-- well, whether you look at integer times or not-- this is the fundamental theorem of renewal rewards.

If you look at the limit as  $t$  goes to infinity, there's the integral of the rewards you pick up. For this discrete case, this is just a summation of the rewards that you get. This summation here by the theorem is equal to the expected number of rewards within one renewal period. Namely, this is the expected number of recurrences of state  $i$  per state  $j$ . So in between each occurrence of state  $j$ , what's the expected number of  $i$ 's that I hit? And that's the number that it is.

We don't know what that number is, but we could calculate it if we wanted to. It's not

a limit of anything. Well, it's a sort of a limit, but not very much of a limit that's well defined. And the theorem says that this integral is equal to that expected value divided by the expected recurrence time of  $T_{jj}$ . Now, we argue that this is, in fact, the number of occurrences of state  $i$ . So it's in the limit. It's  $1$  over the expected value of the recurrence time to state  $i$ . So what this says is the  $1$  over the recurrence time to state  $i$  is equal to the expected number of recurrences to state  $i$  per state  $j$  divided by the expected time in state  $j$ .

If you think about that for a minute, it's something which is intuitively obvious. I mean, you look at this long sequences of things. You keep hitting  $j$ 's every once in a while. And then what you do, is you count all of the  $i$ 's that occur. So now looking at it this way, you're going to count the number of  $i$ 's that occur in between each  $j$ . You can't have a simultaneous  $i$  and  $j$ . The state is  $i$  or  $j$ .

So for each recurrence period, you count the number of  $i$ 's that occur. And what this is then saying, is the expected time between  $i$ 's is equal to the expected time between  $j$ 's divided by-- if I turn this equation upside down, the expected time between  $i$ 's is equal to the expected time between  $j$ 's divided by the expected number of  $i$ 's per  $j$ . What else would you expect? It has to be that way, right?

But this says that it indeed, is that way. Mathematics is sometimes confusing with countable-state chains as we've seen.

OK, so the theorem then says for  $i$  and  $j$  recurrent, either both are positive-recurrent or both are null-recurrent. So this is adding to the theorem we had earlier. The theorem we had earlier says that all states within a class are either recurrent or they're transient. This now divides the ones that are recurrent into two subsets, those that are null-recurrent and those that are positive-recurrent. It says that for states within a class, either all of them are recurrent or all of them are-- all of them are positive-recurrent or all of them are null-recurrent.

And this theorem shows it because this theorem says there has to be an expected number of occurrences of state  $i$  between each occurrence of state  $j$ . Why is that? Because there has to be a path from  $i$  to  $j$ . And there has to be a path that doesn't

go through  $i$ . Because if you have a path from  $i$  that goes back to  $i$  and then off to  $j$ , there's also this path from  $i$  to  $j$ . So there's a path from  $i$  to  $j$ . There's a path from  $j$  to  $i$  that does not go through  $i$ . That has positive probability because paths are only defined over transitions with positive probability. So this quantity is always positive if you're talking about two states in the same class. So what this relationship says, along with the fact that it's a very nice and convenient relationship-- I almost put it in the quiz for finite state and Markov chains. And you can be happy I didn't, because proving it takes a little more agility than what one might expect at this point.

The theorem then says that if  $i$  and  $j$  are recurrent, either both are positive-recurrent or both are null-recurrent, what the overall theorem then says is that for every class of states, either all of them are transient, all of them are null-recurrent, or all of them are positive-recurrent. And that's sort of a convenient relationship. You can't have some states that you never get to. Or you only get to with an infinite recurrence time in a class and others that you keep coming back to all the time.

If there's a path from one to the other, then they have to work the same way. This sort of makes it obvious why that is.

And this is too sensitive. OK.

OK, so now we want to look at steady state for positive-recurrent chain. Do you remember that when we looked at finite state in Markov chains, we did all this classification stuff, and then we went into all this matrix stuff? And the outcome of the matrix stuff, the most important things, were that there is a steady state.

There's always a set of probabilities such that if you start the chain in those probabilities, the chain stays in those probabilities. There's always a set of  $\pi_i$ 's, which are probabilities. They all sum to 1. They're all non-negative. And each of them satisfy the relationship, the probability that you're in state  $j$  at time  $t$  is equal to the probability that you're in state  $i$  at time  $t - 1$  times the probability of going from state  $i$  to  $j$ . This is completely familiar from finite-state chains. And this is exactly the same for countable-state and Markov chains.

The only question is, it's now not at all sure that that equation has a solution anymore. And unfortunately, you can't use matrix theory to prove that it has a solution. So we have to find some other way of doing it. So we look in our toolbox, which we developed throughout the term. And there's only one obvious thing to try, and it's renewal theory. So we use renewal theory.

We then want to have one other definition, which you'll see throughout the rest of the term and every time you start reading about Markov chains. When you read about Markov chains in queuing kinds of situations, which are the kinds of things that occur all over the place, almost all of those Markov chains are countable-state Markov chains. And therefore, you need a convenient word to talk about a class of states where all of the states in that class communicate with each other. And irreducible is the definition that we use. An irreducible Markov chain is a Markov chain in which all pairs of states communicate with each other.

And before, when we were talking about finite-state Markov chains, if all states communicated with each other, then they were recurrent. You had a recurrent Markov chain, end of story.

Now we've seen that you can have a Markov chain where all the states communicate with each other. We just had these two examples-- these two examples here where they all communicate with each other. But depending on what  $p$  and  $q$  are, they're either transition transient, or they're positive-recurrent, or they're null-recurrent. The first one can't even be positive-recurrent, but it can be recurrent.

And the bottom one can also be positive-recurrent. So any Markov chain where all the states communicate with each other-- there's a path from everything to everything else-- which is the usual situation, is called an irreducible Markov chain. An irreducible can now be positive-recurrent, null-recurrent, or transient. All the states in an irreducible Markov chain have to be transient, or all of them have to be positive-recurrent, or all of them have to be null-recurrent. You can't share these qualities over an irreducible chain. That's what this last theorem just said.

OK, so if a steady state exists-- namely if the solution to those equations exist, and if the probability that  $X_{sub 0}$  equals  $i$  is equal to  $\pi_i$ . And incidentally, in the version that got handed out, that equation there was a little bit garbled. That one. Said the probability that  $X_{sub 0}$  was equal to  $\pi_i$ , which doesn't make any sense.

If a steady-state exists and you start out in steady-state-- namely, the starting state  $X_{sub 0}$  is in state  $i$  with probability  $\pi_{sub i}$  by for every  $i$ , this is the same trick we played for finite-state and Markov chains. As we go through this, I will try to explain what's the same and what's different. And this is completely the same. So there's nothing new here.

Then, this situation of being in steady-state persists from one unit of time to the next. Namely, if you start out in steady-state, then the probability that  $X_{sub 1}$  is equal to  $j$  is equal to the sum over  $i$  of  $\pi_{sub i}$ . That's the probability that  $X_{sub 0}$  is equal to  $i$ . Times  $P_{sub i j}$ , which by the steady-state equations, is equal to  $\pi_{sub j}$ .

So you start out in steady-state. After one transition, you're in steady-state again. You're in steady-state at time 1. Guess what, you're in state time 2. You're in steady-state again. And you stay in steady-state forever. So when you iterate, the probability that you're in state  $j$  at time  $X_{sub n}$  is equal to  $\pi_{sub j}$  also. This is assuming that you started out in steady-state. So again, we need some new notation here.

Let's let  $N_{sub j}$  of tilde be the number of visits to  $j$  in the period 0 to  $t$  starting in steady-state. Namely, if you start in state  $j$ , we get a renewal process to talk about the returns to state  $j$ .

If we start in steady-state, then this first return to state  $j$  is going to have a different set of probabilities than all subsequent returns to state  $j$ . So  $N_{sub j}$  of  $t$ , tilde is now not a renewal process, but a delayed renewal process. So we have to deal with it a little bit differently. But it's a very nice thing because for all  $t$ , the expected number of returns to state  $j$  over  $t$  transitions is equal to  $n$  times  $\pi_{sub j}$ .  $\pi_{sub j}$  is the probability that you will be in state  $j$  at any time  $n$ . And it stays the same for every  $n$ .

So if we look at the expected number of times we hit state  $j$ , it's exactly equal to  $n$  times  $\pi_j$ .

And again, here's this awkward thing about renewals and Markov. Yes?

**AUDIENCE:** So is that sort of like an ensemble average--

**PROFESSOR:** Yes.

**AUDIENCE:** Or is the time average [INAUDIBLE]?

**PROFESSOR:** Well, it's an ensemble average and it's a time average. But the thing we're working with here is the fact there's a time-- is the fact that it's an ensemble average, yes. But it's convenient because it's an exact ensemble average.

Usually, with renewal processes, things are ugly until you start getting into the limit zone. Here, everything is nice and clean all the time. So we start out in steady-state and we get this beautiful result. It's starting in steady-state. The expected number of visit to state  $j$  by time  $n$ -- oh, this is interesting. That  $t$  there should be  $n$  obviously.

Well, since we have  $t$ 's everywhere else, that  $n$  there should probably be  $t$  also. So you can fix it whichever way you want.  $n$ 's and  $t$ 's are the same. I mean, for the purposes of this lecture, let all  $t$ 's be  $n$ 's and let all  $n$ 's be  $t$ 's.

This works for some things. This starts in steady state, stays in steady state. It doesn't work for renewals because it's a delayed renewal process, so you can't talk about a renewal process starting in state  $j$ , because you don't know that it starts in state  $j$ . So sometimes we want to deal with this. Sometimes we want to deal with this. This is the number of returns to  $t$  starting in state  $j$ . This is the number of returns to state  $j$  over  $0$  to  $t$  if we start in steady-state.

Here's a useful hack, which you can use a lot of the time. Look at what  $N_{ij}(t)$  is. It's the number of times you hit state  $j$  starting in state  $i$ . So let's look at it as you go for while, you hit state  $j$  for the first time. After hitting state  $j$  for the first time, you then go through a number of repetitions of state  $j$ . But after that first time you hit state  $j$ , you have a renewal process starting then. In other words, you have a

delayed renewal process up to the first renewal. After that, you have all the statistics of a renewal process.

So the idea then is  $N_{ij}(t)$  is 1. Counts 1 for the first visit to  $j$ , if there are any. Plus,  $N_{ij}(t) - 1$  for all the subsequent recurrences from  $j$  to  $j$ .

Thus, when you look at the expected values of this, the expected value of  $N_{ij}(t)$  is less than or equal to 1 for this first recurrence, for this first visit, plus the expected value of  $N_{jj}(t)$  of some number smaller than  $t$ . But  $N_{jj}(t)$  grows with  $t$ . It's a number of visits over some interval. And as the interval gets bigger and bigger, the number of visits can't shrink. So you just put the  $t$  there to make it an upper bound.

And then, when you look at starting in steady-state, what you get is the sum overall starting states  $\pi_i$  of the expected value of  $N_{ij}(t)$ . And this is less than or equal to 1 plus the expected value of  $N_{jj}(t)$  also. So this says you can always get from  $\tilde{N}(t)$  to  $N_{jj}(t)$ , by just giving up this term 1 here as an upper bound.

If you don't like that proof-- and it's not really a proof. If you try to make it a proof, it gets kind of ugly. It's part of the proof of theorem 4 in the text, which is even more ugly. Because it's mathematically clean with equations, but you don't get any idea of why it's true from looking at it. This you know why it's true from looking at it, but you're not quite sure that it satisfies the equations that you would like.

I am trying to move you from being totally dependent on equations to being more dependent on ideas like this, where you can see what's going on. But I'm also urging you, after you see what's going on, to have a way to put the equations in to see that you're absolutely right with it.

OK, now, we come to the major theorem of countable-state and Markov chains. It's sort of the crucial thing that everything else is based on. I mean, everything beyond what we've already done.

For any irreducible Markov chain-- in other words, for any Markov chain where all the states communicate with each other, the steady-state equations have a solution if and only if the states are positive-recurrent.

Now, remember, either all the states are positive-recurrent or none of them are. So there's nothing confusing there.

If all the states are positive-recurrent, then there is a steady-state solution. There is a solution to those equations. And if the set of states are transient, or null-recurrent, then there isn't a solution to all those equations.

If a solution exists, then the probability, the steady-state probability is state  $i$  is  $1$  over the mean recurrence time to state  $i$ . This is a relationship that we established by using renewal theory for finite-state and Markov chains. We're just coming back to it here.

One thing which is important here is that  $\pi_i$  is greater than 0 for all  $i$ . This is a property we had for finite-state Markov chains also. But it's a good deal more surprising here. When you have a countable number of states, saying that every one of them has a positive probability is-- I don't think it's entirely intuitive. If you think about it for a long time, it's sort of intuitive. But it's the kind of intuitive thing that really pushes your intuition into understanding what's going on. So let's give a Pf of this, of the only if part.

And I will warn you about reading the proof in the notes. It's ugly because it just goes through a bunch of logical relationships and equations. You have no idea of where it's going or why. And finally, at the end it says, QED.

I went through it. It's correct. But damned if I know why. And so, anyway, that has to be rewritten. But, anyway here's the Pf.

Start out by assuming that the steady-state equations exist. We want to show positive-recurrence. Pick any  $j$  and any  $t$ . Pick any state and any time.  $\pi_j$  is equal to the expected value of  $N_j$  of  $t$ . That we chose for any Markov chain at all. If you start out in steady-state, you stay in steady-state. So under the

assumption that we're in steady-state-- under the assumption that we start out in steady-state, we stay in steady-state. This  $\pi_j$  has to be the expected value of the number of recurrences to state  $j$  over  $t$  time units.

And what we showed on the last slide-- you must have realized I was doing this for some reason. This is less than or equal to  $1 + t$  times the expected recurrence time of state  $j$ .

So  $\pi_j$  is less than or equal to  $1/t$  times this expected recurrence time for state  $j$ . And if we go to the limit as  $t$  goes to infinity, this  $1/t$  dribbles away to nothingness. So this is less than or equal to the limit of expected value of  $N_{jj}(t)/t$  over  $t$ . What is that?

That's the expected number, long-term rate of visits to state  $j$ . It's what we've shown as equal to  $1/\tau_j$ .

Now, if the sum of the  $\pi_j$ 's is equal to 1, remember what happens when you sum a countable set of numbers. If all of them are 0, then no matter how many of them you sum, you have 0. And when you go to the limit, you still have 0. So when you sum a set of countable set of non-negative numbers, you have to have a limit. Because it's non-decreasing. And that sum is equal to 1.

Then somewhere along the line, you've got to find the positive probability. One of the  $\pi_j$ 's has to be positive. I mean, this is almost an amusing proof because you work so hard to prove that one of them is positive. And then, almost for free, you get the fact that all of them have to be positive.

So some  $\pi_j$  is greater than 0. If  $\pi_j$  is less than or equal to this, thus the limit as  $t$  approaches infinity of the expected value of  $N_{jj}(t)/t$  is greater than 0 for that  $j$ , which says  $j$  has to be positive-recurrent. Which says all the states have to be positive-recurrent because we've already shown that. So all the states are positive-recurrent. Then you still have to show that this inequality here is equality, and you've got to do that by playing around with summing up these things. Something has been left out, we have to sum those up over  $j$ . And that's another mess. I'm not

going to do it here in class. But just sort of see why this happened. Yeah?

**AUDIENCE:** [INAUDIBLE]. Why do you have to show the equality?

**PROFESSOR:** Why do I have to show the equality? Because if I want to show that all of the  $\pi_j$ 's are positive, how do I show that?

All I've done is started out with an arbitrary-- oh, I've started out with an arbitrary  $j$  and an arbitrary  $t$ . Because I got the fact that this was positive-recurrent by arguing that at least one of the  $\pi_j$ 's had to be positive. From this I can argue that they're all positive-recurrent, which tells me that this number is greater than 0. But that doesn't show me that this number is greater than 0.

But it is. I mean, it's all right. It all works out. But not quite in such a simple way as you would hope.

OK, so now let's go back to what we called birth-death chains, but look at a slightly more general version of them. These are things that you-- I mean, queuing theory is built on these things. Everything in queuing theory. Or not everything, but all the things that come from a Poisson kind of background. All of these somehow look at the birth-death chains.

And the way a birth-death chain works is you have arbitrary self-loops. You have positive probabilities going from each state to the next state up. You have positive probabilities going from the higher state to the lower state. All transitions are limited from--  $i$  can only go to  $i + 1$ , or  $i$ , or  $i - 1$ . You can't make big jumps. You can only make jumps of one step. And other than that, it's completely general.

OK, now we go through an interesting argument. We look at an arbitrary state  $i$ . And for this arbitrary state  $i$ , like  $i$  equals 2, we look at the number of transitions that go from 2 to 3. And the number transitions that go from 3 to 2 for any old sample path whatsoever. And for any sample path, the number of transitions that go up-- if we start down there, before you can come back, you've got to go up. So if you're on that side, you have one more up transition than you have down transition. If you're on that side, you have the same number of up transitions and down transitions. So

that as you look over a longer and longer time, the number of up transitions is effectively the same as the number of down transitions.

If you have a steady-state,  $\pi_i$  is the fraction of time you're in state  $i$ .  $\pi_i p_i$  is the fraction of time you're going from state  $i$  to state  $i+1$ . And  $\pi_{i+1} q_{i+1}$  is the fraction of time you're going from state  $i+1$  down to state  $i$ . What we've just argued by the fact that sample path averages and ensemble averages have to be equal is that  $\pi_i p_i$  is equal to  $\pi_{i+1} q_{i+1}$ .

In the next slide, I will talk about whether to believe that or not. For the moment, let's say we believe it.

And from this equation, we see that the steady-state probability of  $i+1$  is equal to the steady-state probability of  $i$  times  $p_i$  over  $q_{i+1}$ . It says that the steady-state probability of each  $\pi_i$  is determined by the steady-state probability of the state underneath it. So you just go up. You can calculate the steady-state of each, the probability of each if you know the probability of the state below it.

So if you recurse on this,  $\pi_{i+1}$  is equal to  $\pi_i$  times this ratio is equal to  $\pi_{i-1}$  times this ratio times  $p_{i-1}$  over  $q_i$  is equal to  $\pi_{i-2}$  times this triple of things. It tells you that what you want to do is define  $r_i$  as the difference of these two probabilities, namely  $r_i$ , for any state  $i$ , is the ratio of that probability to that probability. And this equation then turns into  $\pi_{i+1} = \pi_i r_i$ . If you put all those things together-- if you just paste them one after the other, the way I was suggesting-- what you get is  $\pi_i$  is equal to  $\pi_0$  times this product of terms. The product of terms looks a little ugly. Why don't I care about that very much? Well, because usually, when you have a chain like this, all the  $P$ s are the same and all the  $Q$ s are the same-- or all the  $P$ s are the same for some point beyond someplace, they're different before that. There's always some structure to make life easy for you. Oh, that's my computer. It's telling me what time it is. I'm sorry.

OK. So  $\pi_i$  is this. We then have to calculate  $\pi_0$ .  $\pi_0$  is then 1 divided

by the sum of all the probabilities is  $\pi_0$  times all those other things. It's 1 plus the sum here. And now if you don't believe what I did here, and I don't blame you for being a little bit skeptical. If you don't believe this, then you look at this and you say, OK, I can now go back and look at the steady state equations themselves and I can plug this into the steady state equations themselves. And you will immediately see that this solution satisfies the steady state equations.

OK. Oh, damn. Excuse my language. OK. So we have our birth-death chain with all these transitions here. We have our solution to it. Note that the solution is only a function of these rows. It's only a function of the ratio of  $p_i$  to  $Q_i + 1$ . It doesn't depend on those self loops at all. Isn't that peculiar? Completely independent of what those self loops are. Well, you'll see later that it's not totally independent of it, but it's essentially independent of it.

And you think about that for a while and suddenly it's not that confusing because those equations have come from looking at up transitions and down transitions. By looking at an up transition and a down transition at one place here, it tells you something about the fraction of time you're over there and the fraction of time you're down there if you know what these steady state probabilities are. So if you think about it for a bit, you realize that these steady state probabilities cannot depend that strongly on what those self loops are. So this all sort of makes sense.

The next thing is the expression for  $\pi_0$ -- namely this thing here-- is a product of these terms. It converges and therefore the chain is positive recurrent because there is a solution to the steady state equation. It converges if the row sub  $i$ 's are asymptotically less than 1. So for example, if the row sub  $i$ 's-- beyond  $i$  equals 100-- are bounded by, say, 0.9, then these terms have to go to 0 rapidly after  $i$  equals 100 and this product has to converge. I say essentially here of all these particular cases where the row sub  $i$ 's are very close to 1, and they're converging very slowly to 1 and who knows. But for most of the things we do, these row sub  $i$ 's are strictly less than 1 as you move up. And it says that you have to have steady state probabilities.

So for most birth-death chains, it's almost immediate to establish whether it's

recurrent, positive recurrent, or not positive recurrent. And we'll talk more about that when we get into Markov processes, but that's enough of it for now.

Comment on methodology. We could check the renewal results carefully, because what we're doing here is assuming something rather peculiar about time averages and ensemble averages. And sometimes you have to worry about those things, but here, we don't have to worry about it because we have this major theorem which tells us if steady state probabilities exist-- and they exist because they satisfy these equations-- then you have positive recurrence. So it says the methodology to use is not to get involved in any deep theory, but just to see if these equations are satisfied. Again, good mathematicians are lazy-- good engineers are even lazier. That's my motto of the day.

And finally, birth-death chains are going to be particularly useful in queuing where the births are arrivals and the deaths are departures.

OK. Now we come to reversibility. I'm glad we're coming to that towards the end of the lecture because reversibility is something which I don't think any of you guys even-- and I think this is a pretty smart class-- but I've never seen anybody who understands reversibility the first time they think about it. It's a very peculiar concept and the results coming from it are peculiar, and we will have to live with it for a while. But let's start out with the easy things-- just with a definition of what a Markov chain is.

This top equation here says that the probability of a whole bunch of states--  $X_{n+k}$  down to  $X_{n+1}$  given the stated time,  $n$ , down to the stated time 0. Because of the Markov condition, that has to be equal to the probability of these terms just given  $X_n$ . Namely, if you know what  $X_n$  is, for the future, you don't have to know what any of those previous states are. You get that directly from where we started with the Markov chains-- the probability of  $X_{n+1}$ , given all this stuff, and then you just add the other things onto it.

Now, if you define  $A_+$  as any event which is defined in terms of  $X_{n+1}$ ,  $X_{n+2}$ , and so forth up, and if you define  $A_-$  as anything which is a

function of  $X_{n-1}$ ,  $X_{n-2}$ , down to  $X_0$ , then what this equations says is that the probability of any  $A_{+}$  given  $X_n$  and  $A_{-}$  is equal to the probability of  $A_{+}$  given  $X_n$ . And this hasn't gotten hard yet. If you think this is hard, just wait.

If we now multiply this by the probability of  $A_{-}$  given  $X_n$  and what I'm trying to get at is, how do you reason about the probabilities of earlier states given the present state? We're used to proceeding in time. We're used to looking at the past for telling what the future is. And every once and a while, you want to look at the future and predict what the past had to be. It's probably more important to talk about the future given the past, because sometimes you don't know what the future is. But mathematically, you have to sort that out.

So if we multiply this equation by the probability of  $A_{-}$ , given  $X_n$ , we don't know what that is, but it exists. It's a defined conditional probability. Then what we get is the probability of  $A_{+}$  and  $A_{-}$ , given  $X_n$ , is equal to the probability of  $A_{+}$ , given  $X_n$ , times the probability of  $A_{-}$ , given  $X_n$ . So that the probability of the future and the past, given what's happening now, is equal to the probability of the future, given what's happening now, times the probability the past, given what's happening now. Which may be a more interesting way of looking at past and future and present than this totally asymmetric way here. This is a nice, symmetric way of looking at it. And as soon as you see that this has to be true, then you can turn around and write this the opposite way, and you see that the probability of  $A_{-}$ , given  $X_n$ , and  $A_{+}$  is equal to the probability of  $A_{-}$  given  $X_n$ . Which says that the probability of the past, given  $X_n$  and the future, is equal to the probability of the past just given  $X_n$ .

You can go from past the future or you can go from future to past. And incidentally, if you people have trouble trying to think of the past and the future as being symmetric animals-- and I do too-- everything we do with time can also be done on a line going from left to right, or it can be done on a line going from bottom up. Going from bottom up, it's hard to say that this is symmetric to this. If you look at it on a line going from left to right, it's kind of easy to see that this is symmetric

between left to right and right to left. So every time you get confused about these arguments, put them on a line and argue right to left and left to right instead of earlier and later. Because mathematically, it's the same thing, but it's easier to see these symmetries.

And now, if you think of  $A$  minus as being  $X_{n-1}$ , and you think of  $A$  plus as being  $X_{n+1}$ ,  $X_{n+2}$  and so forth up, what this equation says is that the probability of the last state in the past, given the state now and everything in the future, is equal to the probability of the last state in the past given  $X_n$ .

Now, this isn't reversibility. I'm not saying that these are special process. This is true for any Markov chain in the world. These relationships are always true. This is one reason why many people view this as the real Markov condition, as opposed to any of these other things. They say that three events have a Markov condition between them if there's one of them which is in between the other. Where you can say that the probability of the left one, given the middle one, times the right one, given the middle one, is equal to the probability of the left and the right given the middle. It says that the past and the future, given the present, are independent of each other. It says that as soon as you know what the present is, everything down there is independent of everything up there. That's a pretty powerful condition. And you'll see that we can do an awful lot with it, so it's going to be important.

OK. So let's go on with that. By Bayes rule-- and incidentally, this is why Bayes got into so much trouble with the other statisticians in the world. Because the other statisticians in the world really got emotionally upset at the idea of talking about the past given the future. That was almost an attack on their religion as well as all the mathematics they knew and everything else they knew. It was really hitting them below the belt, so to speak. So they didn't like this. But now, we've recognized that Bayes' Law is just the consequence of the axioms of probability, and there's nothing strange about it. You write down these conditional probabilities and that's sitting there, facing you.

But what it says here is that the probability of the state at time  $n-1$ , given the

state of time  $n$ , is equal to the probability of the state of time  $n$ , given  $n$  minus 1, times the probability of  $X_{n-1}$  divided by the probability of  $X_n$ . In other words, you put this over in this side, and it says the probability of  $X_n$  times the probability of  $X_{n-1}$  given  $X_n$  is that probability up there. It says that the probability of  $A$  given  $B$  times the probability of  $B$  is equal to the probability of  $A$  times the probability of  $B$  given  $A$ . And that's just the definition of a conditional probability, nothing more.

OK. If the forward chain is in a steady state, then the probability that  $X_{n-1}$  equals  $j$ , given  $X_n$  equals  $i$ , is  $p_{ji}$  times  $\pi_j$  divided by  $\pi_i$ . These probabilities become just probabilities which depend on  $i$  but not on  $n$ . Now what's going on here is when you look at this equation, it looks peculiar because normally with a Markov chain, we start out at time 0 with some assumed probability distribution. And as soon as you start out with some assumed probability distribution at time 0 and you start talking about the past condition on the future, it gets very sticky. Because when you talk about the past condition on the future, you can only go back to time equals 0, and you know what's happening down there because you have some established probabilities at 0.

So what it says is in this equation here, it says that the Markov chain, defined by this rule-- I guess I ought to back and look at the previous slide for that. This is saying the probability of the state at time  $n-1$ , conditional on the entire future, is equal to the probability of  $X_{n-1}$  just given  $X_n$ . This is the Markov condition, but it's the Markov condition turned around. Usually we talk about the next state given the previous state and everything before that. Here, we're talking about the previous state given everything after that. So this really is the Markov condition on what we might view as a backward chain. But to be a Markov chain, these transition probabilities have to be independent of  $n$ . The transition probabilities are not going to be independent of  $n$  if you have these arbitrary probabilities at time 0 lousing everything up.

So you can get around this in two ways. One way to get around it is to say let's restrict our attention to positive recurrent processes which are starting out in a steady state. And if we start out in a steady state, then these probabilities here--

looking at these probabilities here-- if you go from here down to here, you'll find out that this does not depend on  $n$ . And if you have an initial state which is something other than steady state, then these will depend on it. Let me put this down in the next chain up.

The probability of  $X_{n-1}$  given  $X_n$  is going to be independent of  $n$  if this is independent of  $n$ , which it is, because we have a homogeneous Markov chain. And this is independent of  $n$  and this is independent of  $n$ . Now, this will just be the probability of  $p_i$  if  $X_{n-1}$  is equal to  $i$ . And this will be  $p_i$  if probability of  $X_n$  is equal to  $i$ . So this and this will be independent of  $n$  if in fact we start out in a steady state. In other words, it won't be.

So what we're doing here is we normally think of a Markov chain starting out at time 0 because how else can you get it started? And we think of it in forward time, and then we say, well, we want to make it homogeneous, because we want to make it always do the same thing in the future otherwise it doesn't really look much like the a Markov chain. So what we're saying is that this backward chain-- we have backward probabilities defined now-- the backward probabilities are homogeneous if the forward probabilities start in a steady state. You could probably make a similar statement but say the forward probabilities are homogeneous if the backward probabilities start in a steady state. But I don't know when you're going to start. You're going to have to start it sometime in the future, and that gets too philosophical to understand.

OK. If we think of the chain as starting in a steady state at time minus infinity, these are also the equations of the homogeneous Markov chain. We can start at time minus infinity wherever we want to-- it doesn't make any difference-- because by the time we get to state 0, we will be in steady state, and the whole range of where we want to look at things will be in steady state.

OK. So aside from this issue about starting at 0 and steady state and things like that, what we've really shown here is that you can look at a Markov chain either going forward or going backward. Or look at it going rightward or going leftward.

And that's really pretty important. OK. That still doesn't say anything about it being reversible.

What reversibility is-- it comes from looking at this equation here. This says what the transition probabilities are, going backwards, and this is the transition probabilities going forward. These are the steady state probabilities. And if we define  $P^*$  of  $ji$  as a backward transition probabilities-- namely, the probability that at this time or in stage  $A$ -- given that in this next time, which to us, is the previous time, we're in state  $i$ , is the probability of going in a backward direction from  $j$  to  $i$ . This gets into whether this is  $P^*$  of  $ij$  or  $P^*$  of  $ji$ . But I did check it carefully, so it has to be right.

So anyway, when you substitute this in for this, the conditions that you get is  $\pi_i$  sub  $i$  times  $P^*$  of  $ij$  is equal to  $\pi_j$  times  $P$  of  $ji$ . These are the same equations that we had for a birth-death chain. But now, we're not talking about birth-death chains. Now we're talking about any old chain. Yeah?

**AUDIENCE:** Doesn't this only make sense for positive recurring chains?

**PROFESSOR:** Yes. Sorry. I should keep emphasizing that, because it only makes sense when you can define the steady state probabilities. Yes. The steady state probabilities are necessary in order to even define this  $P^*$  of  $ji$ . But once you have that steady state condition, and once you know what the steady state probabilities are, then you can calculate backward probabilities, you can calculate forward probabilities, and this is a very simple relationship that they satisfy. It makes sense because this is a normal form. You look at state transition probabilities and you look at the probability of being in one state and then the probability of going to the next state. And the question is the next state back there or is it over there? And if it's a star, then it means it's back there.

And then we define a chain as being reversible if  $P^*$  of  $ij$  is equal to  $P$  sub  $ij$ , for all  $i$  and all  $j$ . And what that means is that all birth-death chains are reversible. And now let me show you what that means.

If we look at arrivals and departures for a birth-death change, sometimes you go in

a self loop, so you don't go up and you don't go down. Other times you either go down or you go up. We have arrivals coming in. Arrivals correspond to upper transitions, departures correspond to downward transitions so that when you look at it in a normal way, you start out at time 0. You're in state 0. You have an arrival. Nothing happens for a while. You have another arrival, but this time, you have a departure, you have another departure, and you wind up in state 0 again.

As far states are concerned, you go from state 0 to state 1. You stay in state 1. In other words, this is the difference between arrivals and departures. This is what the state is. You stay in state 1. Then you go up and you get another arrival, you get a departure, and then you get a departure, according to this chain. Now, let's look at it coming in this way. When we look at it coming backwards in time, what happens? We're going along here, we're in state 0, suddenly we move up. If we want to view this as a backward moving Markov chain, this corresponds to an arrival of something. This corresponds to another arrival. This corresponds to a departure. We go along here with nothing else happening, we get another departure, and there we are back in a steady state again.

And for any birth-death chain, we can do this. Because any birth-death chain we can look at as an arrival and departure process. We have arrivals, we have departures-- we might have states that go negative. That would be rather awkward, but we can have that. But now, if we know that these steady state probabilities govern the probability of arrivals and the probabilities of departures, if we know that we have reversibility, then these events here have the same probability as these events here. It means that when we look at things going from right back to left, it means that the things that we viewed as departures here look like arrivals.

And what we're going to do next time is use that to prove Burke's Theorem, which says using this idea that if you look at the process of departures in a birth-death chain where arrivals are all with probability  $P$  and departures are all with probability  $Q$ , then you get this nice set of probabilities for arrivals and departures. Arrivals are independent of everything else-- same probability at every unit of time. Departures are the same way.

But when you're looking from left to right, you can only get departures when your state is greater than 0. When you're coming in looking this way, these things that looked like departures before are looking like arrivals. These arrivals form a Bernoulli Process, and the Bernoulli Process says that given the future, the probability of a departure, at any instant of time, is independent of everything in the future.

Now, that is not intuitive. If you think it's intuitive, go back and think again. Because it's not. But anyway, I'm going to stop here. You have this to mull over. We will try to sort it out a little more next time. This is something that's going to take your cooperation to sort it out, also.