Introduction to Evolutionary Dynamics

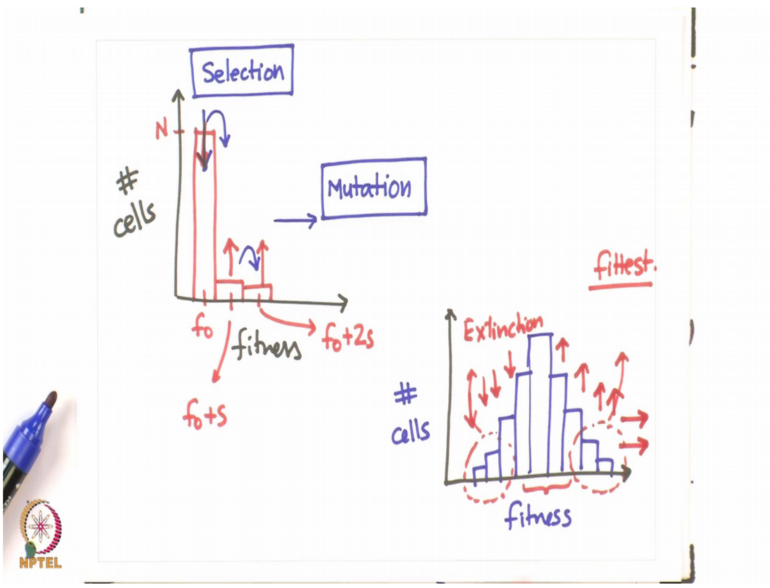
Prof. Supreet Saini
Department of Chemical Engineering
Indian Institute of Technology, Bombay

Lecture - 31
Evolutionary dynamics when mutations are rapid - 1

Hi everyone. Let us continue our discussion on the regime where mutations are rapid and which leads to the implication that there are more than 2 genotypes which coexist in an environment at the same given time. And this analysis is going to be slightly tricky as compared to mutations are regime, precisely because of the reason that we have to take into account coexistence of multiple genotypes in the environment at that particular instant.

But before we try and work towards development of the expression corresponding to the speed of revolution in such an environment let us try and intuitively understand why the representation, with which we stop the last lecture. And understand what is happening in this population as time moves forward starting from a population where each individual belongs to one type only.

(Refer Slide Time: 01:12)



So, let us think about this we have at equal to 0 we have all individuals at fitness f naught this is the picture at t equal to 0 at some future time a mutation event is going to take place. And this population one of the individuals is going to mutate and I will have an individual which is that f naught plus S.

As soon as this mutation gets established, its numbers are going to get going to increase and then selection at. And the numbers corresponding to this are going to decrease, but as the numbers of individuals which correspond to this particular genotype are increasing one of the individuals belonging to this particular population is going to mutate and give rise to an individual which belongs to a genotype whose fitness is given by if naught plus 2 S. And this keeps on progressing forward in time that each full of individuals at the highest sequence the leading edge as we call it of this distribution is going to keep giving rise to individuals of higher fitness.

That is going to keep on happening. So, what mutation is doing here is that mutation is creating variability mutation keeps having these newer genotypes arise in the population and keeps widening this distribution that we have. On the other hand, selection is trying to narrow this distribution selection is trying to read out the less pet individuals from among this distribution and that works towards removing individuals which correspond to this particular genotype for instance in the picture that we have just drawn.

And question that we should ask ourselves now is that suppose we started this particular experiment with all individuals corresponding to fitness f naught. Now if we wait a long enough time what does this distribution of individuals corresponding to particular fitness particular fitness bars on the x-axis look like, when we waited for a very long time and one way to think about this is that mutations when we have waited for a long time and observe this distribution.

Then individuals at the leading edge are going to be small in number by definition and that is because these particular genotypes have only just arisen in the population the mutation event that led to their coming into being have only just happened and they have not had time to flourish and spread through the population even though they are the ones which have the highest fitness in the population at that particular instant.

But, because they only just arrived in the population their numbers are going to be smaller than number despite them being fittest in the population the numbers will rise with time, but at this time their numbers the numbers associated with those particular genotypes are going to be small. So, these numbers are the leading edge is going to be small in number the lagging edge is also going to be smaller number because these are the genotypes which correspond to the least fitness in the environment at that particular time.

These are individuals which probably arouse sometime back once they must have been the better ones, but then they correspondingly their frequencies increase, but then fitness started acting mutations kept on generating newer individuals which has higher fitness corresponding to them. And hence the fitness the numbers came down and now they are on the verge of being extinction and something like this distribution will eventually meet that; that its numbers will keep on coming down and while they are coming down the numbers associated with these 2 and the subsequent fitness's will be higher and we can imagine that the lagging; lagging and individuals are also going to be small in number.

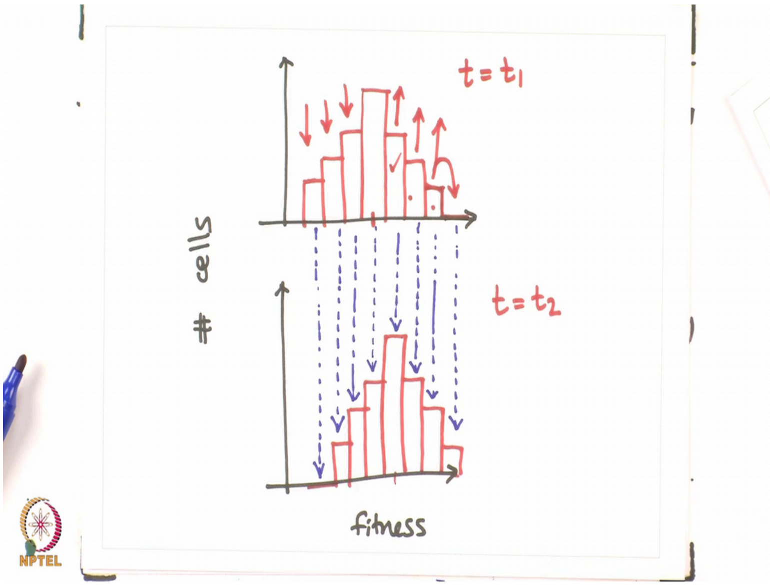
So, if it think about this in those terms intuitively what it tells us about steady state picture of the distribution if I take a snapshot of the population at any time t the distribution of individuals across the different frequency fitness bins would look something like this, but perhaps look something like this x axis again represents fitness and y axis represents number cells which correspond to a particular fitness and what happened is that these individuals; they are the fittest individuals in the population. So, they are the fittest, but still the numbers are small because they have only been in the population for a very short time at the time that the snapshot was taken it has not been very long. Since these mutations have come into existence in this environment hence the numbers are small.

Before these individuals came into being these were the fittest individuals. So, before these 2 came into being these were the fitter genotype and hence the numbers their frequencies for increasing the population. And hence they are represented by the highest frequency is among all distribution among the distribution of the fitness among individuals.

And then at the lagging end at the login end these are on the verge of extinction and if you compare these 2 graphs you can imagine that after we wait a while this plot would start to look something like this with the numbers associated with this particular genotype will decrease such that the number corresponding to this genotype is represented by this small bar here. So, at steady state it sort of makes intuitive sense to anticipate a distribution such as this and this distribution; however, is not steady in nature this is not steady in nature, because selection is continuously acting to remove these genotypes from the population.

Selection is also acting to increase the frequency of these genotypes in the population whereas; mutation is pushing this wave forward. So, at steady state this distribution acquires a moving wave kind of phenomena which is perhaps best represented by this graph. So, at equal to t equal to t 1 the distribution looks such as this, but then we have new mutations arising from this particular genotype and everybody is going to the right.

(Refer Slide Time: 08:21)



At the same time selection is acting to ensure that these genotypes are eliminated from the population and selection is also acting to make sure that these genotypes flourish in the population and the frequency is associated with them increase with time if that is the case then if I move forward in time to t equal to t 2; what would have happened because selection is acting to increase the frequencies of these genotypes their numbers would have gone up and time.

Now, in this snapshot this particular genotype was the mean genotype which was at the highest frequency that has changed now because, selection is acting to reduce its fitness selection is acting to increase the frequency associated with this genotype and its frequency has gone up. And now the maximum the genotype which is maximally represented in the population is now this particular genotype.

At the same time the frequency associated with these 2 genotypes has also increased as we move forward in time from t 1 to t 2 what is also happened is that the leading edge at t equal to t 1 has produced a mutant which is carrying a beneficial mutation; hence the leading edge is. Now this been where there were no individuals at t equal to the one that has happened because of this leading edge the earlier; earlier leading edge producing a beneficial mutation.

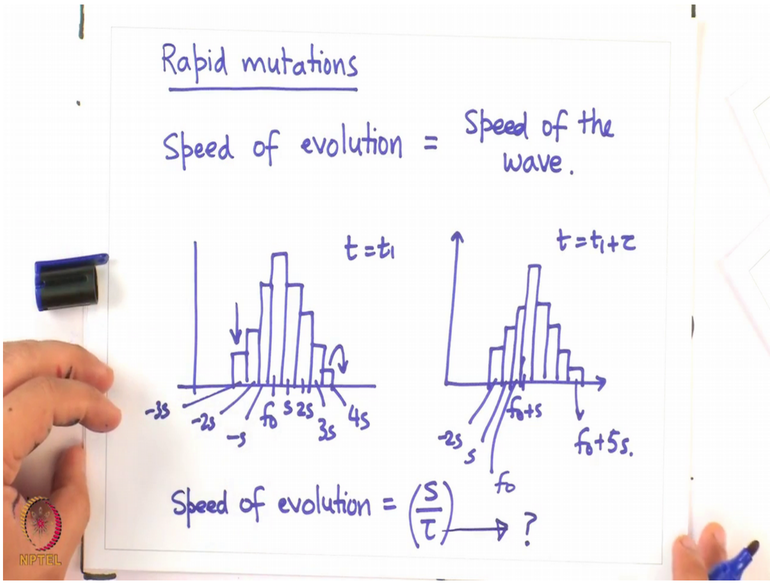
At the same time the lagging end has been acted upon by selection and that is no longer present at t equal to t 2. So, what this again tells us is that the steady state distribute of individuals in this population can be represented by this moving wave kind of a phenomena where this shape that is population has acquired this shape and the shape is this moving to the right with time leading to increase in fitness. Again the assumption that we have made here is that beneficial mutations are always available to the cell they do not try up. So, that to the mean fitness associated with the population keeps on increasing as we move forward in time.

But something that we did a couple of time couple of lectures earlier was that in the initial phase of evolution we saw that fitness increases linearly with time and then saturated. So, if you are only talking about that linear phase of evolution this result hold that beneficial mutations do not try up and it keeps increasing the mean fitness keeps increasing with time in a linear fashion.

So, if this way flight property of distribution of different genotypes in the pop in the environment and its importance to you what we will do next is try and derive the expression for speed of evolution in a setting like this in should be intuitively intuitive to note that speed of evolution the rate at which mean fitness of the environment is changing with time is actually just equal to the speed with which the wave is moving that we have just drawn.

So, essentially what we are interested in is that how fast does this way. So, that is something that will work towards now. So, again we are interested in the regime of rapid mutations and we want an expression for speed of evolution.

(Refer Slide Time: 12:03)



Which is also equal to speed of the wave that we just described one thing to note is that since this distribution is moving at a constant speed that is our hypothesis if this is the distribution at t equal to t 1 and then at t equal to t 1 plus tau what has happened is that every bar has gotten shifted to the right exactly.

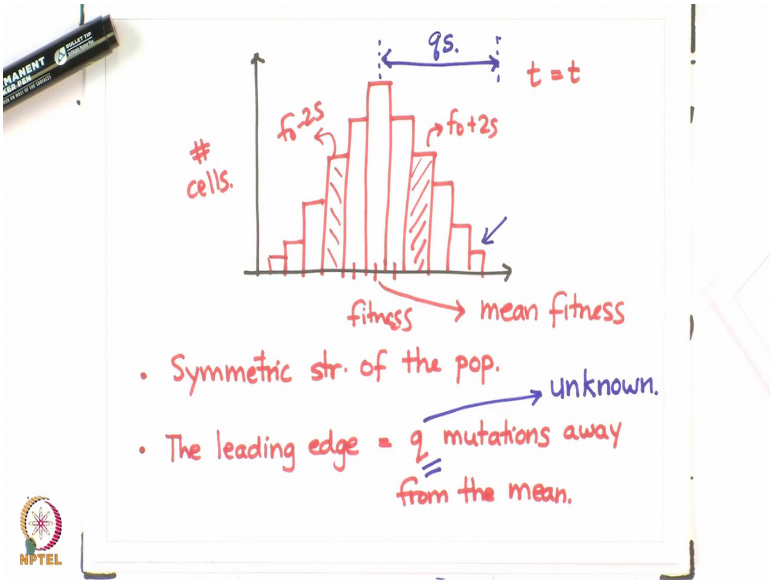
So, if this is the mean fitness f naught this is S this is f naught plus 2 S 3 S and 4 S now f naught have shifted to f naught plus S that is the maximally represented bar now and then you have does 2 S, 3 S, 4 S and 5 S and subsequently this is minus S this is minus 2 S this is minus 3 S. Now you have f naught f naught minus S minus 2 S and f naught. What is happen between in this 2 time events is that f naught minus 3 S has been acted upon by selection which has lead to its numbers corresponding to this particular genotyping reduced to 0. And mutation events in this particular population f naught plus 4 S have led to individuals belonging to a fitness of f naught plus 5 S getting established in the population.

So, one way to notice that in this tau time the wave has moved to the right exactly by S units, we can think about this in simple terms that every individual in this population it is almost equivalent to think that every individual in the population has jumped S units in fitness in this time tau. So, speed of evolution in a setting like this is this equal to S divided tau because the mean fitness has increased by S and we can say that because fitness of every individual in the population has increased by S.

Hence, the mean increase is also S and that change has happened in time tau which gives me the speed of evolution which is the rate of change of men fitness with time equal to S by tau. And what will do now- is try and develop an expression for this town we do not know what this town is and that is what will try and develop and quantify next

To do that we first understand how we first realize that we do not know the distribution we do not know the exact distribution of the genotypes at any particular time t even though we are drawing frequencies we do not exactly know the numbers associated with each fitness we have a intuitive hypothetical understanding of the distribution, but not in a quantitative sense.

(Refer Slide Time: 15:58)



So, what we will start with is a few assumptions we will assume a steady state distribution of population at any time t and in the distribution let me just draw this we are not sure how wide or how narrow this distribution is we do not exactly know anything about this distribution except for the fact that intituitly we feel that this should be the structure of the population at any time t. So, fitness number of cells if this is the case and this.

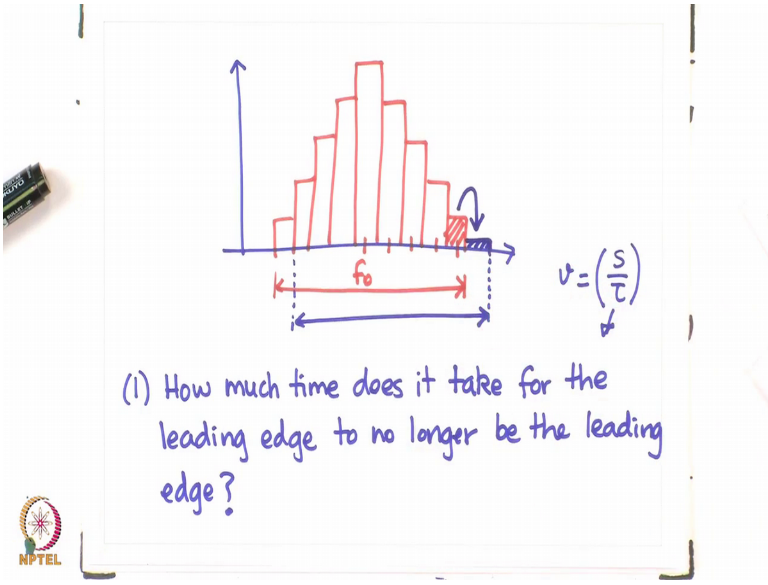
So, we will assume a symmetric structure of the population if we assume a symmetric structure of the population that automatically implies that the middle histogram here which is the highest in frequency represents the mean fitness because this is. So, because you are saying that the structure is symmetric about the mean then this is f naught plus S this is f naught minus S. And we are saying that the numbers of individuals belonging to each are equal in number. And hence, we have the same number of individuals which are at f naught plus S and the same number of individuals at f naught minus s. So, if we add these 2 together the mean fitness will be exactly cancelled out each other by each other the mean fitness will to be exactly f naught.

Similarly, this in this bunch of individual has a fitness f naught plus 2 S and this bunch of individuals has a fitness f naught minus 2 S, but because these 2 numbers are equal because of the symmetric rule that we have imposed here; that means, that the mean fitness of these 2 bars is f naught exactly the plus 2 S and minus 2 S cancel each other exactly and we get mean fitness equal to f naught. So, we take these bars in pair and what that what that will automatically lead us to is the fact that the mean fitness of the entire structure is just f naught then we assume that the leading edge is q mutations away from the mean.

What; that means, is that this particular this particular genotype which is the leading edge is from the mean at this particular time q mutations away; that means, in fitness terms each of these q mutations confers a fitness change of magnitude s. So, the fitness change between the fitness individual at the leading edge compared to the mean fitness is equal to q times S, it is important to know that this q is something that is yet unknown I do not know if q is 4 5 10 2 q could be anything, but at this point we do not have an idea of how much is q and since we are interested in finding the tau, because we need tau to be able to comment on the speed of revolution we will find that tau and q are quantities which are pretty closely connected to each other.

Alright, so, with these assumptions, now let us try and develop this further; how do you think about this we think about this in the following sense? So, when we just re draw the graph again.

(Refer Slide Time: 20:06)



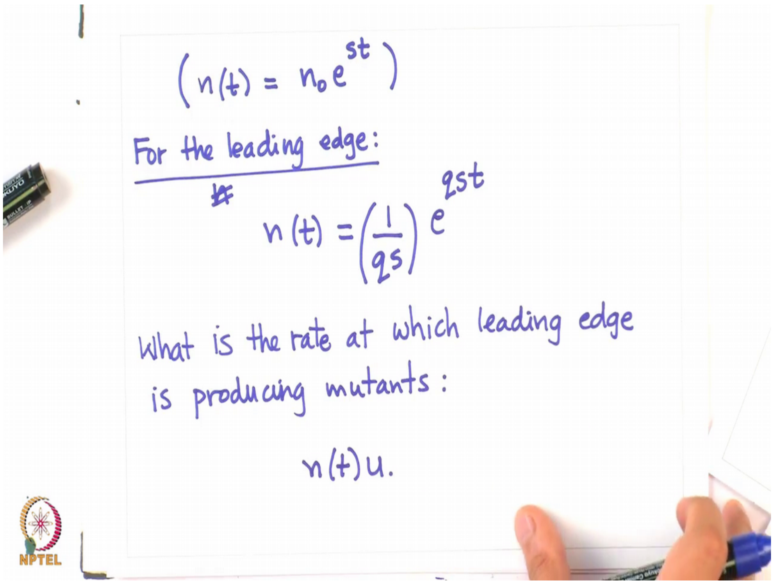
And let us fit individuals this is f naught first I am going to focus on the leading edge and let me call the time tau we already define tau time tau is the time by way in which this wave moves exactly S units in this fitness domain. So, from here to here right now the distribution is from here to here f naught plus S, 2 S, 3 S, 4 S. So, it moves from f naught minus 4 S to f naught for this is the range of the wave at this point at time t this wave move exactly by S units on the axis. So, at time t plus tau that wave will have a structure which is distributed across these fitness and hence we say that the fitness of every individual its almost of the fitness of every individual has increased by S and so, the speed of evolution is given by S by tau, but can we will comment on this tau what we are going to do is comment on the value of this tau in a couple of different ways.

The first way that we will comment on tau is that how much time does it take does it take for the leading edge to no longer be the leading edge which means at this at the time at this snapshot was taken this particular genotype represent the leading edge of the way, but eventually we know that another mutation is going to happened and a new genotype will arise and that genotype when established will form the leading edge of this wave and when that happens this particular genotype which is currently the leading wave will no longer be the leading wave.

We know that that time when the leading edge is defined by this particular genotype is exactly tau from when this picture was taken. So, the first thing we are going to try and do is estimate that how much time does it take for a beneficial mutation from this genotype to generate this particular genotype of a higher fitness and then this mutation to get established that is the first thing that we are going to start with.

So, how do we think about this now because the leading edge is the fittest individual in the population it is reasonable to expect that its numbers are going to grow exponentially because that instant there is nobody more fitter than that particular individual in the population.

(Refer Slide Time: 23:39)

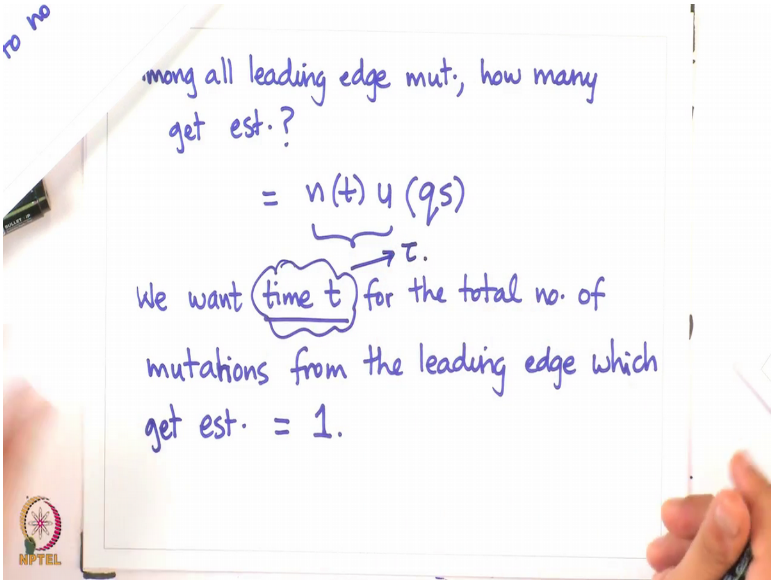


So, for the leading edge we can write that n t is equal to n naught e to the power S t this is a general formula that holds for any exponential growth for the particular case that we are interested in we can write it for the leading edge we can write that n t this is the number of individuals belonging to leading edge is equal to n naught n now what is n naught n naught is number of individuals of a leading edge at the start which is equal to one upon qs because remember the advantage that the leading edge holds advantage that the leading edge holds over the mean is exactly equal to qs.

If the advantage that the leading edge holds over the mean is equal to q s; that means, to become established it needs us one upon qs number of individual in the population. So, when the leading edge has one upon qs number of individuals presented that is when it becomes established after that its growth is exponential and again the growth factor is just qs into t that is the rate at which the leading edge is growing what is the rate t which its producing mutations.

So, leading edge is numbers where is like this what is the rate at which leading edge is producing mutants and that is just equal to n t times u that is the rate at which mutations are happening at the leading edge and next is among all these mutations that are happening at the leading edge how many mutations are actually getting fixed or established.

(Refer Slide Time: 26:09)



So, among all leading edge mutations how many get established and that is just going to nt times u this is the total number of mutations that are happening at the leading edge we know that the probability that these mutations will get established is just equal to q times s right, because the probability that a mutation which has an advantage s over the mean is the probability that the mutation carrying an advantage s over the mean survives is equal to s.

So, the probability that this mutation which has the advantage qs over the mean survives is also equals to q s; now what we want here is that. So, this is the total number of mutations which are getting established we want what we are interested in is that we want the time it takes t for the total number of mutations total number of mutations from the leading edge which get established equal to 1.

We are interested in solving this equation for the time such that the total number of mutations that get established is equal to one and that is. So, because if the total number of mutations which have established from the leading edge is equal to one what; that means, is that there is a beneficial mutation which happened from the leading edge and that mutation has gotten established.

What; that means, is that the leading edge is no longer the leading edge and has now been relegated to the second bar on the distribution that we have here and that time it takes for the leading edge to produce a mutation which gets established is also equal to tau this is going to give me a value of tau. So, this is slightly involved.

And we will we will continue this discussion in our next lecture, and arrive at a value of tau in a couple of different ways. And then eventually, we will solve for tau which will give us the speed of evolution for a case like this we will pick those up in the next lecture on words.

Thank you.