Measuring the Effect of Selection on Allele Frequency for Finite Populations

by

Nicholas Andrew Gellner

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

(Mathematics)

Queen’s University

August 2013
Acknowledgements

I would first like to thank my supervisor, Professor Peter Taylor, for his support and guidance throughout this school year. Without him this thesis would not exist.

Many thanks go out to my family and my girlfriend for their encouragement.

Thank you to Rian Dewji for your mathematical assistance as well.
Table of Contents

Acknowledgements ........................................................................................................................................... 1
Introduction .................................................................................................................................................. 3
Table of Notation ........................................................................................................................................ 4
Model ............................................................................................................................................................ 6
Analysis of Change in A-allele Frequency .................................................................................................... 9
  Technical Statements .................................................................................................................................. 12
Introduction to Reproductive Value ........................................................................................................... 13
  Outline ....................................................................................................................................................... 13
  Under Birth-Death ..................................................................................................................................... 13
  Under Death-Birth ..................................................................................................................................... 14
  What Changes? ......................................................................................................................................... 15
Analysis of Change in Reproductive Value A-allele Frequency .................................................................... 16
  Technical Statements with Reproductive Value ......................................................................................... 19
Application of Analysis ............................................................................................................................... 20
  Birth-Death Updating Mechanism with Fecundity Payoffs ................................................................. 20
  Death-Birth updating mechanism with Fecundity Payoffs ................................................................. 27
  Birth-Death Updating with Survival Payoffs ......................................................................................... 31
  Death-Birth Updating Mechanism with Survival Payoffs ................................................................. 35
Discussion .................................................................................................................................................. 39
References .................................................................................................................................................. 41
Introduction

Natural selection is a key mechanism of evolution whereby the frequency of biological traits changes in a population. The variation in these traits can be a function of mutation, as well as the genomes of individuals interacting with the environment. This variation allows the population to evolve over time. Scientists have always been interested in studying why certain organisms thrive in a particular ecosystem, how competition affects the balance of a population, and why certain genes/trait become dominant over a long period of time. Mathematics provides an enlightening way of studying the evolution of specific traits: by structuring the interactions within the population using graphs.

A graph is a mathematical structure used to model pairwise interactions between objects. The objects are represented using nodes, and the interactions between objects are represented using edges. A finite population can be modelled using a graph, where individuals occupy the nodes (vertices) in the graph. The edges connecting nodes together will represent which individuals interact within the population. They also dictate where potential offspring can be sent.

In this paper, we study evolutionary games on graphs. An edge connecting two vertices will mean that the occupants of those vertices play a game with each other. All edges will hold equal weight, meaning the game is played in the same way for each connected pair of nodes in the graph. The conditions of the game are the selection forces which will affect how the population evolves. We will study the long-term evolution of two types of alleles in the finite population (call them $A$ and $B$), with the selection forces of the game acting. Below is an example of what a two-allele graph for a size six population could look like. Individuals with genotype $A$ or $B$ occupy all the nodes in some combination, and we can easily see how individuals interact with one another from the edges.

![Graph Example](image)

Our goal will be to figure out what conditions allow the $A$ allele to be at a selective advantage over the $B$ allele. This will depend on how the game is defined. For the $A$ allele to be at a selective advantage, it must benefit from the introduction of the game. If the game favours the $A$ allele, then the frequency of these alleles will be higher over a long time period.

To solve this problem, we provide a general analysis studying the changes in allele frequency over time. Once this is complete, it will be important to see how the calculations are performed, given a finite population. The calculations are difficult, so we introduce the concept of reproductive value - the expected representation of a gene on some node in the future population - to make them easier. Using reproductive values in the analysis will make solving the original problem easier, and will give a better understanding of how the structure of the population plays an important role in the long-term.
dynamics. Before introducing the model, it is convenient to summarize the technical notation that will be used throughout this paper, and this is what we do next.

### Table of Notation

<table>
<thead>
<tr>
<th>Expression</th>
<th>Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>The size of the population (number of nodes in the graph)</td>
</tr>
<tr>
<td>$x_i$</td>
<td>Genotypic value of node $i$. This will be 1 for an $A$ individual and 0 for a $B$ individual</td>
</tr>
<tr>
<td>$y_i$</td>
<td>Genotypic value of an offspring born on node $i$</td>
</tr>
<tr>
<td>$S$</td>
<td>The state of a population, i.e. a specification of the genotypic value on each node</td>
</tr>
<tr>
<td>$\bar{x} = \frac{1}{N} \sum_i x_i$</td>
<td>The allele frequency: the average genotypic value of the nodes, which depends on the population state</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Parameter which indicates a component involving selection terms.</td>
</tr>
<tr>
<td>$b_i = b_i^N + b_i^\delta$</td>
<td>The fecundity of node $i$, which is the probability of contributing an offspring in the next time-step. It can be broken down into a neutral component $b_i^N$ and a selective component $b_i^\delta$, which will be further explained later. The equality is true to first order in $\delta$</td>
</tr>
<tr>
<td>$d_i = d_i^N + d_i^\delta$</td>
<td>The mortality of node $i$, which is the probability of dying in the next time-step. It can be broken down into its neutral component $d_i^N$ and its selective component $d_i^\delta$. The equality is true to first order in $\delta$</td>
</tr>
<tr>
<td>$\Delta_S \bar{x}$</td>
<td>The one-step change in allele frequency starting in some state $S$</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Mutation rate</td>
</tr>
<tr>
<td>$\Delta_S \bar{x}^{sel}$</td>
<td>The one-step change in allele frequency due to selection, starting in some state $S$</td>
</tr>
<tr>
<td>$\Delta_S \bar{x}^{mut,-,sel}$</td>
<td>The one-step change in allele frequency due to mutation and selection, starting in some state $S$</td>
</tr>
<tr>
<td>$\pi_S(\mu, \delta) = \pi_S^N + \pi_S^\delta$</td>
<td>The long-term frequency of state $S$, which can be broken down into its neutral part $\pi_S^N$ and its selective part $\pi_S^\delta$. The equality is true to first order in $\delta$</td>
</tr>
<tr>
<td>$w_i = w_i^N + w_i^\delta$</td>
<td>The fitness of node $i$, which is defined as $b_i - d_i$. It can be broken down into its neutral component $w_i^N$ and its selective component $w_i^\delta$.</td>
</tr>
</tbody>
</table>
\( \bar{b}_S = \bar{b}_S^N + \bar{b}_S^\delta \) and its selective component \( \bar{b}_S^\delta \). The equality is true to first order in \( \delta \).

\( \bar{d}_S = \bar{d}_S^N + \bar{d}_S^\delta \) Average mortality for a state \( S \), which can be separated into its neutral part \( \bar{d}_S^N \) and its selective part \( \bar{d}_S^\delta \). The equality is true to first order in \( \delta \).

\( v_i \) The reproductive value of node \( i \)

\( d_{ij} \) The probability that an offspring born on node \( i \) disperses to node \( j \), given that the occupant of node \( i \) gives birth

\( v_i^* = \sum_j d_{ij} v_j \) Average reproductive value of an offspring born on node \( i \)

\( x^* = \frac{1}{N} \sum_i v_i x_i \) Average reproductive value allele frequency

\( b_i^* = b_i v_i \) Reproductive value fecundity of the node \( i \), which can again be broken down into neutral and selective parts as before

\( d_i^* = d_i v_i \) Reproductive value mortality of the node \( i \)

\( w_i^* \) Reproductive value fitness of node \( i \), defined as \( b_i^* - d_i^* \)

\( \Delta_S x^* \) The one-step change in reproductive value allele frequency for some state \( S \)

\( \Delta_S x^*_{sel} \) The one-step change in reproductive value allele frequency due to selection, for some state \( S \)

\( \Delta_S x^*_{mut-sel} \) The one-step change in reproductive value allele frequency due to selection and mutation, for some state \( S \)

\( \text{deg}(j) \) The degree of a vertex \( j \) in a graph, which is the number of edges extending from \( j \)
Model

Before conducting any sort of analysis, it is imperative to know exactly how the evolution of a population works. Any finite population can be represented using a graph. The individuals that make up the population will occupy the nodes of the graph. If an edge connects two nodes $i$ and $j$ together, then $i$ and $j$ are said to be neighbors. If a game is introduced into the population, then node $i$ will play the game with its neighbors. An edge going from node $i$ to node $j$ also means that node $i$ can send offspring to node $j$, and node $j$ can send offspring to node $i$.

We will be considering populations with two possible genotypes, and we will refer to each individual as either an $A$ or a $B$. Individuals will die, and new individuals will be born, which will change the configuration of $A$’s and $B$’s on the graph. There will be a finite number of possible states $S$ for each graph. As an example, these are the possible $A,B$ configurations on the square graph:

```
A A A A
A A B B
A B A B
A B B B
```

The states consisting of only $A$ or only $B$ individuals are called pure states, and any state that contains at least one $A$ and one $B$ individual is called a mixed state. Notice that in the state with three $A$ and one $B$, we can permute the $B$ in four different ways. However, permuting the $B$ does not change the structure of the state, as all the nodes reside in the same class. For all nodes to be in the same class $C$, there must be a map $f$ from $C$ to itself such that $f(v) = w$ for any nodes $v,w$ in $C$. For a graph, the map will be a permutation $\sigma$ of the nodes in the class $C$, such that the pair of vertices $v,w$ form an edge if and only if the pair $(\sigma(v),\sigma(w))$ also form an edge. Intuitively, two nodes will be in the same class if they interact with other nodes in the same way. Every node in the square has two neighbors, and so all the nodes behave the same.

Once the structure of the population has been determined, we can define a game in any way we like. We analyze a game that is played once, at the beginning of each (discrete) time step. The game is played before any births or deaths happen. All $A$ individuals give a benefit $b$ to each of their neighbors, meanwhile incurring a personal cost $c$ for each benefit given. With this definition of the game, the $B$ individuals simply receive the benefits of the game, never having to incur a cost.

The benefits $b$ and costs $c$ can be viewed in a couple of different ways. We can treat $b$ and $c$ as fecundity payoffs. Any individual receives a fecundity benefit $b$ from each neighboring $A$ individual, while all $A$ individuals incur a fecundity cost of $c$ to themselves each time they give a benefit. These payoffs are added to a baseline fecundity of 1 for each individual, at the beginning of each time step. As an example, given a configuration of the square graph, the adjusted fecundities after the game is played are shown:
We can also view the benefits and costs as survival payoffs, where an individual receives a survival benefit from each A-neighbor, and the A individuals incur a personal survival cost of $c$ for each neighbor interaction. If an individual receives a survival benefit $b$, its mortality will be decreased by $b$. The survival cost $c$ incurred by an A-individual raises its mortality by $c$. The changes in mortality are added to a baseline of 1 for each individual. We convert survival payoffs to mortality payoffs in order to track fitness, which is fecundity minus mortality. As an example, given the same configuration of the square, the adjusted mortalities from the survival payoffs are as follows:

\[
\begin{array}{c|c|c|c}
A & A & 1+2b−2c & 1+b−2c \\
A & A & 1+b−2c & 1+2b \\
\end{array}
\]

Viewing the $b-c$ game in terms of fecundities or survivals, the gameplay can be inserted into a birth-death or death-birth updating mechanism. In a birth-death scenario, an individual is chosen to give birth, and the offspring replaces one of the neighbors. In a death-birth scenario, an individual is chosen to die, and once this happens the vacant node is replaced by an offspring of a neighbor. We first consider the evolution that occurs in one time-step, and extend it to a long-term average afterwards.

In one time-step, there is one birth and one death, in some order. The births and deaths will always occur after the game is played throughout the population. If an individual is chosen to give birth, it can only send offspring to one of its neighbors. Assuming we are in some state $S$, each node will have a probability of giving birth, $b_i$, and a probability of dying, $d_i$. These will of course depend on what the payoffs mean (fecundity or survival), as well as the updating mechanism (birth-death or death-birth). When a birthing event happens, there is a small chance of the offspring mutating (from an A to a B or from a B to an A), with probability $\mu$. Once a mutation occurs, the offspring will have genotype A with probability $p$ and genotype B with probability $1-p$. Now we can define four evolution protocols that we can analyze:

1) Birth-Death Updating with Fecundity Payoffs
2) Death-Birth Updating with Fecundity Payoffs
3) Birth-Death Updating with Survival Payoffs
4) Death-Birth Updating with Survival Payoffs

Under 1), the game is played, and the relative fecundities determine who gives birth. An individual will divide its fecundity by the total fecundity of the population to get its probability of giving birth. Once the offspring is born, it will occupy one of the neighboring nodes with equal probability.
Under 2), the game is played amongst the population, and then a random individual in the population is chosen to die. The relative fecundities from the game determine which neighbor sends its offspring to the vacated node.

Under 3), the game is played amongst the population, and then a random individual is selected to give birth. The relative mortalities then determine which of the neighboring individuals is chosen to die.

Lastly, under 4), the game is played amongst the population, and the relative mortalities determine who is chosen to die. After this, all neighbors send offspring to the vacant node with equal probability.

We can imagine that the order with which births and deaths occur, along with the structure of how the payoffs are calculated, will produce interesting dynamics in these four cases. Now that the model has been precisely defined, we can begin to answer the important question at hand: What are the necessary (first order) conditions in $b$ and $c$ for the $A$ allele to be at a selective advantage over the $B$ allele?

In working to first order in $b$ and $c$ (the same as working to first order in $\delta$), we will ignore terms of order two or higher in these parameters, and we will also ignore terms of order two or higher in $\mu$, the mutation parameter. We can then use the linear approximation $\frac{1}{1+\varepsilon} = 1 - \varepsilon$ when performing calculations, to get rid of quotients.

This question opens the gate for some interesting analysis, and this problem will touch on areas of mathematics such as dynamical systems, game theory, graph theory, and probability. But what is the question really asking? For $A$ to be at a selective advantage over the allele $B$, the $A$ allele must be better off when the game is played compared to when there is no game. When no game is played, the population is said to be evolving under neutrality. Once we introduce the game, selection forces are activated. We need to find out what the long-term frequency of individuals with $A$ alleles are in the neutral case, and then with selection activated. If the frequency of individuals with $A$ alleles is higher with selection activated than under neutrality, then the $A$ allele has benefited from the playing of the game. This implies that the $B$ allele is worse off than it was under neutrality. If the individuals with $A$ alleles are able to survive and reproduce better than the individuals with $B$ alleles, then the $A$ allele would be at a selective advantage over the $B$ allele.
Analysis of Change in A-allele Frequency

Before we begin, note that the analysis at times will bring up the use of some technical statements, and these are provided after the analysis is complete, for reference. Remember that we want to track the frequency of the A allele. We let \( x_i \) equal 1 if node \( i \) is occupied by an A-individual, and 0 if occupied by a B individual. Now take the population to be in some state \( S \). What would it mean to define the one-step change in A-allele frequency? There is exactly one birth and one death, in some order. If a node \( i \) is chosen to give birth, then the A-allele increases in frequency only if the offspring produced is an A individual. If that node \( i \) is chosen to die, then there is a loss of allele A only if the occupant of the node was an A individual. We sum over all nodes \( i \) in the population, and then divide by the size of the population to give us an equation for the average change in A-allele frequency:

\[
\Delta_s \bar{x} = \frac{1}{N} \sum_i \left( b_i y_i - d_i x_i \right)
\]

Remember \( y_i \) is defined as the genotype of the offspring born on node \( i \). We define \( y_i \) in terms of \( x_i \) as follows: Either there was mutation to provide the offspring or there wasn’t. If there was no mutation, then the offspring has the same genotypic value as the parent, which is \( x_i \). If there was a mutation, then the offspring has genotype A with probability \( p \). Thus we rewrite:

\[
y_i = (1 - \mu)x_i + \mu p = x_i - \mu(x_i - p)
\]

Then:

\[
\Delta_s \bar{x} = \frac{1}{N} \sum_i \left[ (b_i - d_i)x_i - \mu b_i(x_i - p) \right]
\]

The first term is the change due to selection, which does not have a mutation component, which we write as:

\[
\Delta_s \bar{x}^{sel} = \frac{1}{N} \sum_i w_i x_i, \text{ where } w_i = b_i - d_i
\]

The second term is also an effect of selection, but through mutation:

\[
\Delta_s \bar{x}^{mut-sel} = -\frac{\mu}{N} \sum_i b_i(x_i - p)
\]

The net change in allele frequency is of course the sum of these terms:

\[
\Delta_s \bar{x}^{tot} = \Delta_s \bar{x}^{sel} + \Delta_s \bar{x}^{mut-sel}
\]
What has just been presented was all for one step in time. We are interested in a long-term average of this change. Thus the change in each state $S$ must be weighted by its long-term frequency, which we call $\pi_S(\mu, \delta)$. Using angular brackets to account for this average:

$$\langle \Delta_S \vec{x}^{tot} \rangle = \langle \Delta_S \vec{x}^{sel} \rangle + \langle \Delta_S \vec{x}^{mut\text{-}sel} \rangle$$

At the mutation-selection equilibrium, the $A$-allele frequency must be stationary, so $\langle \Delta_S \vec{x}^{tot} \rangle = 0$.

Now it is time to break down these terms. As stated earlier, we assume the effects of selection are small enough so that we need only consider terms of first order in $\delta$, which we use to symbolize the selective component of an entity. Then (to first order in $\delta$):

$$w_i = w_i^N + w_i^\delta$$

$$\pi_S = \pi_S^N + \pi_S^\delta$$

This is just a simple separation of fitness and long-term frequency into their neutral and selective components. Note that the fitness does not depend on mutation, but the long-term frequencies do.

We begin with the selective term:

$$\langle \Delta_S \vec{x}^{sel} \rangle = \sum_S \frac{1}{N} \sum_i (w_i^N + w_i^\delta) x_i (\pi_S^N + \pi_S^\delta)$$

The neutral component of this is equal to zero:

$$\sum_S \frac{1}{N} \sum_i w_i^N x_i \pi_S^N = \frac{1}{N} \sum_i w_i^N \sum_S x_i \pi_S^N = \frac{1}{N} \sum_i w_i^N p = \bar{w}^N p = 0$$

The technical statements section will show why this is true. The first equality follows from L1, the second from L7 and the last from L2. The term purely based on selection,

$$\sum_S \frac{1}{N} \sum_i w_i^\delta x_i \pi_S^\delta$$

will be a second-order term in $\delta$, which is small enough to disregard by our previous assumptions. Thus (to first order in $\delta$) we are left with two terms:

$$\langle \Delta_S \vec{x}^{sel} \rangle = \sum_S \frac{1}{N} \sum_i w_i^\delta x_i \pi_S^N + \sum_S \frac{1}{N} \sum_i w_i^N x_i \pi_S^\delta$$

These two terms will account for the effect of selection on allele-frequency change. The first term represents the effect of selection on fitness, and the second term represents the effect of selection on state frequency.

Now we look at the selection-mutation term. From L6 on the technical statements page we need only average over the two pure states, all $A$ and all $B$. The mutation-selection term is of order $\mu$, and so are the long-term frequencies for the mixed states (you can only get to a mixed state through mutation
from a pure state). Since we ignore terms of order two in $\mu$, we need only look at the pure states, whose long-term frequencies will not depend on $\mu$:

$$\langle \Delta_S \vec{x}^{\text{mut-set}} \rangle = \mu \left( \frac{1}{N} \sum_i b_i (1-p) \pi_A + \mu \left( \frac{1}{N} \sum_i b_i (-p) \right) \pi_B = \mu [\bar{b}_A(1-p)\pi_A - \bar{b}_B p \pi_B] , \right.$$  

Where $\pi_A, \pi_B$ are frequencies for the pure states and $\bar{b}_A, \bar{b}_B$ are average fecundities for the pure states. Separating these into their neutral and selective parts:

$$\langle \Delta_S \vec{x}^{\text{mut-set}} \rangle = \mu \left[ (\bar{b}_A + \bar{b}_A^\delta)(1-p)\pi_A^N + (1-p)\pi_B^N + (\bar{b}_B^N + \bar{b}_B^\delta)(p)\pi_B^N \right] - (\bar{b}_B^N + \bar{b}_B^\delta)(p)\pi_B^N$$

At neutrality, by L7 we know that $\pi_A^N = p$ and thus $\pi_B^N = 1 - p$. Also, $\bar{b}_A^N = \bar{b}_B^N = \bar{b}$ by neutral state independence (L1). Then the neutral component of this expression is:

$$\mu [\bar{b}_A^N (1-p) \pi_A^N - \bar{b}_B^N p \pi_B^N] = \mu [\bar{b}(1-p)p - \bar{b}p(1-p)] = 0$$

We ignore the term of order two in $\delta$, and are left with:

$$\langle \Delta_S \vec{x}^{\text{mut-set}} \rangle = \mu \left[ (\bar{b}_A^\delta - \bar{b}_B^\delta) p(1-p) + \bar{b}((1-p)\pi_A^\delta - p \pi_B^\delta) \right]$$

Now we have the total one-step change, $\langle \Delta_S \vec{x}^{\text{tot}} \rangle$, as the sum of four terms $T1 + T2 + T3 + T4$, where:

Effect of selection through fitness: $T1 = \sum_S \frac{1}{N} \sum_i w_i^\delta x_i \pi_A^N$

Effect of selection through state frequency: $T2 = \sum_S \frac{1}{N} \sum_i w_i^N x_i \pi_S^\delta$

Mutation-selection effect through fecundity: $T3 = -\mu (\bar{b}_A^\delta - \bar{b}_B^\delta) p(1-p)$

Mutation-selection effect through state frequency: $T4 = -\mu \bar{b}((1-p)\pi_A^\delta - p \pi_B^\delta)$

As we previously pointed out, the change in allele frequency we have been dealing with is the average (over all states) one-step change and at equilibrium, this must be zero. To see when the $A$ allele is at a selective advantage, we compare the allele frequency at the neutral equilibrium to the allele frequency at the mutation-selection equilibrium. To zeroth order in $\mu$ by L4 we look at the pure states. Then the frequency (to zeroth order in $\mu$) of $A$ is just $\pi_A = \pi_A^N + \pi_A^\delta = p + \pi_A^\delta$. So the change in allele frequency between the neutral and selective equilibria is just $\pi_A^\delta$. Now since we consider only the pure states, $\pi_A + \pi_B = 1$ and $\pi_A^N + \pi_B^N = 1$. Thus we must have $\pi_B^\delta = -\pi_A^\delta$. We can say that $T4 = -\mu \bar{b} \pi_A^\delta$, and so the change in allele frequency between the neutral and selective equilibria has the opposite sign to $T4$. Since the total change must equal zero at equilibrium, we then deduce that the change in allele frequency between equilibria has the sign of $T1 + T2 + T3$.

As it turns out, by L2 we have $\bar{b}_A = \bar{b}_B = \bar{b} = 1/N$, and so $\bar{b}_A^\delta = \bar{b}_B^\delta = 0$. Looking above, this means $T3$ is zero. So we only need to calculate $T1$ and $T2$. We can calculate $T1$, because it uses the neutral state frequencies. This can be done using some linear algebra (as will be seen later). However, the calculation of $T2$ is much harder to perform because we need the $\pi_S^\delta$ terms. These values are much
harder to come by. We avoid calculating these by introducing the concept of reproductive value. The goal of using reproductive values will be to eliminate the \( T^2 \) term, which will allow us to solve the problem by just using \( T^1 \) and \( T^3 \). Much of the analysis will be the same, with only a few minor differences.

**Technical Statements**

**L1: Neutral State Independence.** In the neutral population, the two components of fitness at node \( i \), fecundity \( \bar{b}_i \) and mortality \( \bar{d}_i \), are independent of the population state \( S \). Thus the same is true of fitness \( \bar{w}_i \).

**L2: Average State Independence.** Even with selection acting, average (over the population) fecundity and mortality are independent of state. Thus \( \bar{b}_S = \bar{b} = \bar{d}_S = \bar{d} = 1/N \). It follows that average fitness is zero at every state: \( \bar{w}_S = \bar{w} = 0 \).

**L3:** To calculate \( \langle \Delta_S \bar{x}^{sel} \rangle \) we need only average over the mixed states \( S \).

**L4:** If \( S \) is a mixed state, \( \pi_S \) is of order \( \mu \).

**L5:** \( \langle \Delta_S \bar{x}^{sel} \rangle \) is of order \( \mu \).

**L6:** To calculate \( \langle \Delta_S \bar{x}^{mut-sel} \rangle \) we need only average over the pure states all \( A \), and all \( B \).

**L7: Neutral Allele Frequency.** At any node \( i \), for the neutral state distribution, the average allele frequency \( x_i \) is \( p \).

**Justification**

Property L1 is clear. The only thing distinguishing fitness effects among states is the distribution of \( A \)'s and \( B \)'s, and under neutrality these behave the same.

For Property L2, we allow selection to act and we take an average over all nodes. This follows from the model assumption that in one time-step, one individual reproduces, and that one offspring results in a single death.

Property L3 follows from L2. To show this, we show that \( \Delta_S \bar{x}^{sel} = \frac{1}{N} \sum_i w_i x_i \) is zero at both pure states. At the all \( B \) state, all the \( x_i \) are 0, so \( \Delta_S \bar{x}^{sel} = \frac{1}{N} \sum_i w_i x_i = 0 \). At the all \( A \) state, all the \( x_i \) are 1. So \( \Delta_S \bar{x}^{sel} = \frac{1}{N} \sum_i w_i x_i = \frac{1}{N} \sum_i w_i = 0 \) by L2.

Observe that if \( S \) is a mixed state, then \( \pi_S(0, \delta) = 0 \). This is because the mixed states cannot occur if there is no mutation. To first order in \( \mu \), \( \pi_S = A + \mu B \). But \( A = 0 \) since \( \pi_S(0, \delta) = 0 \). So \( \pi_S = \mu B \). This gives us L4.
Property L5 follows from L3 and L4.

Property L6 follows from L4, as $\langle \Delta_S x_{\text{mut}-\text{sel}} \rangle$ is already of order $\mu$.

To demonstrate L7, take a fixed node and take the gene at that node. Go back in ancestral time until the most recent mutation event. At that event, the allele that was produced was either A or B with probabilities $p$ and $1 - p$. Now in the genealogical tree from that event to the present, there was no selection, and that means that the paths followed by that gene were independent of their allelic value so that those paths are equally likely. Thus the allele that arrives is A or B with probabilities $p$ and $1 - p$.

**Introduction to Reproductive Value**

**Outline**

In order to make the previous analysis easier to work with, a change needs to be made. The way to go about this is to introduce reproductive values. This will eliminate the need to calculate the $\pi^i_2$ terms in the calculation of $T2$.

The reproductive value $v_i$ of node $i$ is the expected representation of the gene at node $i$ in the future population. This is measured in a neutral population with no mutation. To calculate, we use the fact that the $v_j$ should not change in one time-step. We set the expected change (expected gain minus expected loss) in reproductive value at node $i$ to be zero. This equation will be different depending on what update mechanism we use (birth-death or death-birth). Define $d_{ij}$ to be the probability that an offspring born on node $i$ disperses to node $j$, given that node $i$ is chosen to give birth.

**Under Birth-Death**

For a birth-death updating mechanism under neutrality, a random individual is chosen to give birth. Afterwards the neighbors of that individual all die with equal probability. For a node $i$, the gain in reproductive value will be $d_{ij}v_j$ for all nodes $j$ (note $d_{ii} = 0$). This is saying that if node $i$ is chosen to give birth, its gain in reproductive value will be $v_j$, $d_{ij}$ of the time. Then we sum over all nodes $j$. The loss in reproductive value would then be $d_{ji}v_i$, again for all nodes $j$. This gives us the equation:

$$\sum_j d_{ij}v_j - \sum_j d_{ji}v_i = 0$$

(1)
The first term is the expected increase through offspring, and the second is the expected decrease through mortality. Since $\sum_j d_{ij} = 1$, the first term can be interpreted as the average reproductive value of the offspring born on node $i$ and is denoted $v_i^* = \sum_j d_{ij} v_j$.

**Under Death-Birth**

For a death-birth updating mechanism we have a slightly different expected change in reproductive value. Under neutrality, a random individual is chosen to die, and the vacant node is replaced with offspring from any of the neighbors with equal probability. Then the loss in reproductive value for a node $i$ (given that it dies) will always be $v_i$. The gain in reproductive value can be figured out using the following logic:

Let $P(i \rightarrow), P(\rightarrow j)$ be the probabilities that $i$ gives birth and something disperses to $j$ respectively. Let $P(i \rightarrow j| i \rightarrow)$ be the probability that $i$ disperses to $j$, given that $i$ gives birth, and let $P(i \rightarrow j| j \rightarrow)$ be the probability that $i$ disperses to $j$, given something disperses to $j$. Then:

$$P(i \rightarrow j| i \rightarrow)P(i \rightarrow) = P(i \rightarrow j| j \rightarrow)P(\rightarrow j)$$

The gain in reproductive value will be exactly $P(i \rightarrow j| j \rightarrow)v_j$, summing over all nodes $j$. Using our notation,

$$\sum_j P(i \rightarrow j| j \rightarrow)v_j = \sum_j \frac{d_{ij}b_i}{d_j} v_j$$

So the equation for expected change in reproductive value for death-birth is:

$$\sum_j \frac{d_{ij}b_i}{d_j} v_j - v_i = 0 \quad (2)$$

We represent the gain in reproductive value of a node $i$ in a death-birth system in another way as well. Under a death-birth protocol, a node will gain whenever another node dies, more specifically one of the neighbors of the focal node. Once a node $j$ dies, all the neighboring nodes send offspring to the vacant spot with equal probability, which is exactly $\frac{1}{\deg(j)}$. Thus the focal node $i$ gains reproductive value $v_j$ with probability $\frac{1}{\deg(j)}$ and we sum over all neighboring nodes $j$. The second representation of the expected change equation for a death-birth protocol is as follows:

$$\sum_{\text{neighbors } j} \frac{1}{\deg(j)} v_j - v_i = 0 \quad (3)$$
Note that equations (2) and (3) are saying the same thing: Node \( i \) will send its offspring to node \( j \) with probability \( d_j \frac{1}{\deg(j)} \). Given that node \( i \) gives birth, it will then send its offspring to node \( j \) with probability \( d_{ij} = \frac{d_j}{\deg(j)b_i} \). Then substituting this expression for \( d_{ij} \) into the first formula:

\[
\sum_j d_{ij}b_i \frac{d_j}{d_f} v_j = \sum_{\text{neighbors} j} \frac{1}{\deg(j)} v_j
\]

We switch the sum from being over \( j \) to being over neighbors because \( d_{ij} = 0 \) if \( j \) is not a neighbor to \( i \).

**What Changes?**

The equations (1) and (2)/(3) for expected change in reproductive value we have derived will come into play later on, in order to justify why the new \( T'2 \) term should be zero. Note that reproductive value is determined up to a constant multiple, so we choose any value for them as long as the ratios remain consistent. Now we repeat the allele frequency analysis, with a couple of minor changes. Instead of counting the proportion of \( A \) alleles, we count the total reproductive value of \( A \) alleles. The star superscript will signal this weighting. The average reproductive value allele frequency will be \( x^* = \frac{1}{N} \sum_i v_i x_i \) and we’ll work with changes in this rather than changes in \( \bar{x} \). To calculate these changes, we define the reproductive value fecundity of the breeder on node \( i \) to be \( b_i^* = b_i v_i^* \), and its reproductive value mortality to be \( d_i^* = d_i v_i \). Its reproductive value fitness is then \( w_i^* = b_i^* - d_i^* \). With the notation slightly different this time, we proceed to conduct the same steps of the original analysis.
Analysis of Change in Reproductive Value A-allele Frequency

Here we track changes in reproductive value (RV) A-allele frequency. Note that there is a revised technical lemma section at the end of this analysis. Now, whenever a node $i$ gives birth, there is the possibility of gaining an $A$-individual but there is a gain in reproductive value as well, which is what $b_i^*$ signifies. Whenever node $i$ is chosen to die, it loses A-allele frequency as well as its own reproductive value. Intuitively, if a node has an offspring, it will gain the reproductive value of the node occupied by its offspring. If a node dies, it will lose its own reproductive value because it no longer occupies the node! From this, we realize that the overall expected change in RV allele frequency in one time step is (fix a state $S$):

$$
\Delta_S x^* = \frac{1}{N} \sum_i b_i^* y_i - d_i^* x_i
$$

From this equation we perform the new analysis exactly as before (some parts are left out that were included in the original analysis because they use the exact same reasoning):

$$
\Delta_S x^{* sel} = \frac{1}{N} \sum_i w_i^* x_i, \text{ and } \Delta_S x^{* mut-sel} = -\frac{\mu}{N} \sum_i b_i^*(x_i - p)
$$

Again, the total change in reproductive value allele frequency is the sum of these two terms. We take a long-term average, so each state $S$ will be weighted by its long-term state frequency $\pi_S(\mu, \delta)$:

$$
\langle \Delta_S x^{* tot} \rangle = \langle \Delta_S x^{* sel} \rangle + \langle \Delta_S x^{* mut-sel} \rangle
$$

Once again, at the mutation-selection equilibrium, $\langle \Delta_S x^{* tot} \rangle = 0$. We assume that the effects of selection are small enough so we need not consider terms of order two in $\delta$. We write:

$$
w_i^* = w_i^{* N} + w_i^{* \delta}
$$

$$
\pi_S = \pi_S^{N} + \pi_S^{\delta}
$$

Beginning with the selection term:

$$
\langle \Delta_S x^{* sel} \rangle = \sum_S \sum_i \frac{1}{N} \sum_i (w_i^{* N} + w_i^{* \delta}) x_i (\pi_S^{N} + \pi_S^{\delta})
$$

Note that the neutral component of this is still zero:

$$
\sum_S \sum_i \frac{1}{N} w_i^{* N} x_i \pi_S^{N} = \frac{1}{N} \sum_i w_i^{* N} \sum_S x_i \pi_S^{N} = \frac{1}{N} \sum_i w_i^{* N} p = \bar{w}^{* N} p = 0
$$

The first equality comes from L1, the second from L7, and the third from L2.'
So (ignoring the term of order two in δ):

$$\langle \Delta Sx^{* \text{sel}} \rangle = \sum_{S} \frac{1}{N} \sum_{i} w_i x_i \pi_S^N + \sum_{S} \frac{1}{N} \sum_{i} w_i x_i \pi_S^\delta$$

These two terms provide the two components of the effect of selection on RV allele frequency change. We now analyze the second component responsible for RV allele frequency change:

$$\langle \Delta Sx^{* \text{mut-sel}} \rangle = \mu \left( \frac{1}{N} \sum_{i} b_i^*(1 - p)\pi_A + \frac{1}{N} \sum_{i} b_i^*(-p)\pi_B \right) = \mu [b_A^*(1 - p)\pi_A - b_B^*p\pi_B]$$

Separating these into their neutral and selective parts:

$$\langle \Delta Sx^{* \text{mut-sel}} \rangle = \mu \left[ (\bar{b}_A^N + \bar{b}_B^\delta)(1 - p)(\pi_A^N + \pi_B^\delta) - (\bar{b}_A^N + \bar{b}_B^\delta)(p)(\pi_A^N + \pi_B^\delta) \right]$$

At neutrality, by L7 we know that $\pi_A^N = p$ and $\pi_B^N = 1 - p$. Also, $\bar{b}_A^N = \bar{b}_B^N = \bar{b}$ by neutral state independence (L1). Then the neutral component of this expression is:

$$\mu [\bar{b}_A^N(1 - p)\pi_A^N - \bar{b}_B^Np\pi_B^N] = \mu [\bar{b}^*(1 - p)p - \bar{b}^*p(1 - p)] = 0$$

We ignore the term of order two in δ, and thus we are left with:

$$\langle \Delta Sx^{* \text{mut-sel}} \rangle = \mu \left( (\bar{b}_A^\delta - \bar{b}_B^\delta)p(1 - p) + \bar{b}^* \left( (1 - p)\pi_A^\delta - p\pi_B^\delta \right) \right)$$

This gives us the total one-step change $\langle \Delta Sx^{* \text{tot}} \rangle$ as the sum of four terms:

Effect of selection through fitness: $T_1^* = \sum_{S} \frac{1}{N} \sum_{i} w_i x_i \pi_S^N$

Effect of selection through state frequency: $T_2^* = \sum_{S} \frac{1}{N} \sum_{i} w_i x_i \pi_S^\delta$

Mutation-selection effect through fecundity: $T_3^* = -\mu (\bar{b}_A^\delta - \bar{b}_B^\delta)p(1 - p)$

Mutation-selection effect through state frequency: $T_4^* = -\mu \bar{b}^* \left( (1 - p)\pi_A^\delta - p\pi_B^\delta \right)$

With this new analysis, however, the $T_3^*$ term will not necessarily be zero anymore. The average fecundities with selection in the pure states will be different due to reproductive value weighting. However, we now show that for a birth-death or a death-birth updating mechanism, the $T_2^*$ term is zero. Remember this was the reason the original analysis was modified. To show that $T_2^* = 0$, it suffices to show that for a fixed node $i$, $w_i^* = 0$.

For birth-death, from (1) we have:

$$\sum_j v_j d_{ij} v_j - \sum_j d_{ji} v_i = 0, \text{ with } v_i^* = \sum_j d_{ij} v_j$$

By definition:
\[ w_i^{*N} = b_i^N v_i^* - d_i^N v_i \]

We can write \( d_i^N = \sum_j d_{ji} b_j^N \) since the mortality of node \( i \) is the sum of the mortality effects of all other nodes given that they give birth, and then:

\[ w_i^{*N} = b_i^N \sum_j d_{ji} v_i - \sum_j d_{ji} b_j^N v_i \]

In a birth-death updating game under neutrality, all nodes are equally likely to give birth, and so \( b_i^N = b_j^N \). Then we rewrite \( b_j^N \) as \( b_i^N \), pull it out of the sum, and then the expression for \( w_i^{*N} = 0 \).

For death-birth, from (2) we have:

\[ \sum_j \frac{d_{ij} b_i^N}{d_j^N} v_j - v_i = 0 \]

In a death-birth system under neutrality, everyone has the same mortality, so \( d_j^N = d_i^N \). So:

\[ \sum_j \frac{d_{ij} b_i^N}{d_j^N} v_j - v_i = \frac{b_i}{d_i} \sum_j d_{ij} v_j - v_i = \frac{b_i}{d_i} v_i^* - v_i = 0 \]

Now:

\[ w_i^{*N} = b_i v_i^* - d_i v_i = b_i v_i^* - d_i \frac{b_i}{d_i} v_i^* = 0, \text{ as required} \]

We presented equations (2) and (3) for expected change in reproductive value under a death-birth protocol, so that we could use (2) in the above calculation and (3) in a more intuitive light, which will be easier to use for larger graphs. For both the birth-death and death-birth updating mechanisms, we see that the \( T_2^* \) term is zero, which was the crucial point of the RV analysis. It follows that the change in allele frequency between the neutral and selective equilibria has the sign of \( T_i^* + T_3^* \). These can be calculated relatively easily. To zeroth order in \( \mu \), the equation we analyze to assess the change in allele frequency is this one in terms of the frequency of the all \( A \) state:

\[ T_4^* = -\mu \tilde{p}^* \pi_A^* = \sum_{s, N} \frac{1}{s N} \sum_i w_i^{*s} \pi_i^N - \mu (\tilde{b}_A^s - \tilde{b}_B^s) p (1 - p) \]

Now that the analysis with reproductive value is complete, we go through four examples with finite populations and perform the exact calculations of the \( T_i^* \). Performing these calculations will give great insight into how this analysis can be used, as well as how powerful and general it is. Before proceeding, note the changes in the technical statements from the non-reproductive value case to the case where reproductive values were introduced:
Technical Statements with Reproductive Value

The introduction of reproductive value does not affect the calculations of the state frequencies, as we use the original fecundities and mortalities to get the long term state frequencies. Lemmas L3, L4, L5, L6 remain unchanged from before as a result.

L1 still applies as well. In the neutral population, the two components of fitness are independent of state. This is because reproductive value is a property of a node, and whether an A or a B individual occupies the node does not affect the structure of the graph.

L2 still applies but in a different manner. We’ll call it L2'. Average fitness at every state will still be zero, because average fitness for each node will be zero (the $T^*_2$ term is zero). However, average fecundity and mortality are no longer state independent with selection acting. This is again simply due to the different weightings of the nodes in a heterogeneous population.

The logic for L7 is still the same as before, and so the only difference in the new technical statements section is in the application of L2'.
Application of Analysis

We consider four evolution protocols: birth-death with fecundity and survival payoffs, and death-birth with fecundity and survival payoffs. The analysis will be conducted on the star graph with $N = 3$, as shown below:

The node in the center is commonly referred to as the hub, and the nodes on either end are called leaves. There are six possible states/configurations (occupied by A’s and B’s) associated with this graph:

$$B-A-B \mid A-A-B \mid A-B-B \mid A-B-A \mid A-A-A \mid B-B-B$$

We see that there are four mixed states and two pure states. As a reminder, mixed states have both genotypes in the population, whereas pure states have the same genotype throughout the entire population.

The goal is to see, in this population, what the conditions are (first order in $b$ and $c$) for $A$ to be at a selective advantage over $B$. To figure this out, we need a direct calculation of $T_1^*, T_2^*, T_3^*$, and $T_4^*$. We will need to solve for fitnesses, reproductive values, and long-term frequencies at neutrality. The game and update mechanism will determine the fitnesses. The reproductive values will be calculated from our formulae for expected change in a birth-death or death-birth scenario. To calculate the long-term neutral state frequencies, we use a state transition matrix. The eigenvector associated with the eigenvalue 1 in the state transition matrix will reveal the relative state frequencies. After solving for all of these entities, the original question can be answered.

Birth-Death Updating Mechanism with Fecundity Payoffs

Under this model, the benefit-cost game is played among all nodes in the graph. All $A$ individuals give a benefit $b$ to each neighbor at cost $c$ for each neighbor interaction. These payoffs are added to a baseline fecundity of 1. Then, an individual is chosen to reproduce based on relative fecundities. Once this individual is chosen to give birth, the offspring will occupy one of the neighboring sites at random. For example:

Consider the 3-star graph, with an $A$ individual at the hub and $B$ individuals on the ends (the state known as B-A-B). Under this evolution protocol, the fecundity of the hub will be $1 - 2c$, and the fecundity of each $B$ leaf will be $1 + b$. The probabilities of reproduction for each individual are as follows:

$$B \frac{1 + b}{3 + 2b - 2c} = \frac{1}{3} (1 + b) \left(1 - \frac{2b}{3} + \frac{2c}{3}\right) = \frac{1}{3} \left(1 - \frac{2b}{3} + \frac{2c}{3} + b\right) = \frac{1}{3} + \frac{b}{9} + \frac{2c}{9}$$
Remember in working to first order in \( b \) and \( c \), we can use the linear approximation \( \frac{1}{1+\epsilon} = 1 - \epsilon \).

We also need to know the mortalities for each node, to help us obtain an expression for fitness \( w_i^* \). If the hub is chosen to give birth, then the offspring will occupy each leaf with probability \( \frac{1}{2} \). If a leaf is chosen to give birth then the offspring will automatically occupy the hub node. So the probabilities of mortality are as follows:

\[
B \left( 1+ \frac{1}{3} - \frac{b}{9} - \frac{4c}{9} \right) = \frac{1}{6} - \frac{b}{9} - \frac{2c}{9}
\]

\[
A \left( \frac{1}{3} + \frac{b}{9} + \frac{4c}{9} \right) = \frac{2}{3} + \frac{2b}{9} + \frac{4c}{9}
\]

\[
B \left( 1+ \frac{1}{3} - \frac{b}{9} - \frac{4c}{9} \right) = \frac{1}{6} - \frac{b}{9} - \frac{2c}{9}
\]

The next step is to find out what the reproductive values of the nodes are for this population. We immediately see that there are two classes of nodes in this graph. The two leaves are in one class and the hub is in another class. Thus both leaves will have the same reproductive value. Denote the reproductive value of a leaf by \( v_L \) and the reproductive value of the hub by \( v_H \). If the hub is chosen to give birth, it will disperse to either leaf with probability \( \frac{1}{2} \). The leaves will disperse to the hub each with probability 1, since the hub is the only place they can go. This gives us the values for \( d_{ij} \). Fix either of the leaves. By (1):

\[
\sum_j d_{ij} v_j - \sum_j d_{ji} v_i = 0
\]

\[
d_{L,H} v_H - d_{H,L} v_L = 0
\]

\[
v_H = \frac{v_L}{2}
\]

Let \( v_H = 1 \). Then \( v_L = 2 \). This means \( v_L^* = d_{L,H} v_H = 1 \). Fixing the hub, we have \( v_H^* = 2d_{L,H} v_H = 2 \).

Now we can complete a state fitness table:
<table>
<thead>
<tr>
<th>State</th>
<th>Fecundity</th>
<th>fecundity-RV</th>
<th>Mortality</th>
<th>mortality-RV</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>1/3+b/9+2c/9</td>
<td>1</td>
<td>1/6-b/9-2c/9</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>1/3-2b/9-4c/9</td>
<td>2</td>
<td>2/3+2b/9+4c/9</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1/3+b/9+2c/9</td>
<td>1</td>
<td>1/6-b/9-2c/9</td>
<td>2</td>
</tr>
</tbody>
</table>

| B     | 1/3-b/9-2c/9   | 1            | 1/6+b/9+c/18   | 2            |
| A     | 1/3+2b/9+c/9   | 2            | 2/3-2b/9-c/9   | 1            |
| B     | 1/3-b/9+c/9    | 1            | 1/6+b/9+c/18   | 2            |

| B     | 1/3-2b/9-c/9   | 1            | 1/6+2b/9+c/9   | 2            |
| A     | 1/3+4b/9+2c/9  | 2            | 2/3-4b/9-2c/9  | 1            |
| A     | 1/3-2b/9-c/9   | 1            | 1/6+2b/9+c/9   | 2            |

| B     | 1/3            | 1            | 1/6             | 2            |
| A     | 1/3            | 2            | 2/3             | 1            |
| B     | 1/3            | 1            | 1/6             | 2            |

| B     | 1/3-b/9+c/9    | 1            | 1/6+b/9-c/9     | 2            |
| A     | 1/3+2b/9-2c/9  | 2            | 2/3-2b/9+2c/9   | 1            |
| A     | 1/3-b/9+c/9    | 1            | 1/6+b/9-c/9     | 2            |

Finally, to perform the calculations from the reproductive value allele frequency analysis, we need the neutral long-term state frequencies. To do this we construct a one-step state transition matrix. Although we do not need the selection components displayed in the matrix, they are included to give insight into how the state transfers work. As an example:

Suppose we want to go from the state B-A-B to the state A-A-B. This can happen only by the hub node reproducing to an A individual. From the table, we see that the hub is chosen to reproduce with probability \( \frac{1}{3} - \frac{2b}{9} - \frac{4c}{9} \). An A individual will yield another A offspring with probability \( 1 - \mu + \mu p \). Thus the probability of transitioning from the state B-A-B to the state A-A-B is \( \left( \frac{1}{3} - \frac{2b}{9} - \frac{4c}{9} \right) \left( 1 - \mu + \mu p \right) \).

There is an important point to make here. In order to get to a mixed state from a pure state, a mutation is required. If there is mutation among mixed states, then we can ignore it. This is because the compounded mutation effect will be of order two. Anytime we begin and end in a mixed state, mutation can be ignored in the matrix. We compute all of the state transitions, and the one-step probability transition matrix looks as follows (where the \( i,j \) entry is the probability of transitioning from state \( j \) to state \( i \) in one time-step):
In order to attain the long-term state frequencies, we find the eigenvector associated with the eigenvalue 1, for the neutral state transition matrix. This equates to solving the equation \((M - I)x = 0\) for \(x\), where \(x\) is the eigenvector associated with eigenvalue 1. Again we use the neutral case, meaning \(b = c = 0\). Proceeding:

\[
\begin{bmatrix}
-1 & 0 & 0 & 0 & 0 & 2\mu p \\
1 & -1 & 1 & 0 & 0 & 0 \\
\frac{1}{3} & -\frac{2}{3} & 0 & 0 & \frac{1-p}{3} & 0 \\
0 & 1 & -1 & \frac{1}{3} & 0 & \frac{p}{3} \\
0 & 0 & 0 & -1 & 2\mu \frac{1-p}{3} & 0 \\
0 & \frac{1}{6} & 0 & \frac{2}{3} & -\mu + \mu p & 0 \\
\frac{2}{3} & 0 & \frac{1}{6} & 0 & 0 & -\mu p
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
x_5 \\
x_6
\end{bmatrix} = 0
\]

Using MAPLE to row reduce, we get:
We know that $x_5$ denotes the fixation probability of the pure state of $A$ individuals, and in a neutral population the average allele frequency at a node is $p$. So in order for $x_5$ to equal $p$, let $x_6 = 1-p$. Then the relative state frequencies are as follows:

\[
\begin{bmatrix}
1 & 0 & 0 & 0 & -2\frac{p}{3} & x_1 \\
0 & 1 & 0 & 0 & -10\frac{p}{3} & x_2 \\
0 & 0 & 1 & 0 & -10\frac{p}{3} & x_3 \\
0 & 0 & 0 & 1 & -2\frac{p}{3} & x_4 \\
0 & 0 & 0 & 0 & \frac{p}{p-1} & x_5 \\
0 & 0 & 0 & 0 & 0 & x_6 \\
\end{bmatrix}
= 0
\]

Now we calculate the terms responsible for the total change in allele frequency from the RV analysis.

\[
T_1^* = \sum_{s,N} \frac{1}{N} \sum_i w_i^s x_i \pi_s^N
\]

\[
= \frac{1}{3} \left( 2\mu \frac{p(1-p)}{3} \left( - \frac{4b}{9} - \frac{8c}{9} - \frac{2b}{9} - \frac{4c}{9} \right) + 10\mu \frac{p(1-p)}{3} \left( \frac{c}{3} - \frac{2c}{3} - \frac{c}{3} \right) \\
+ 10\mu \frac{p(1-p)}{3} \left( - \frac{b}{9} - \frac{2c}{9} - \frac{2b}{9} - \frac{c}{9} \right) \\
+ 2\mu \frac{p(1-p)}{3} \left( - \frac{4b}{27} - \frac{10b}{27} - \frac{8b}{27} - \frac{8c}{27} - \frac{20c}{27} - \frac{10c}{27} - \frac{4c}{27} \right) \right) \\
= \mu p(1-p) \left( - \frac{4b}{27} - \frac{10b}{27} - \frac{8b}{27} - \frac{8c}{27} - \frac{20c}{27} - \frac{10c}{27} - \frac{4c}{27} \right) = \mu p(1-p) \left( - \frac{22b}{27} - \frac{14c}{9} \right)
\]

$T_2^* = 0$

\[
T_3^* = -\mu p(1-p) \left( b_A^{*8} - b_B^{*8} \right) = -\mu p(1-p) \left( \frac{b}{9} + \frac{c}{9} - \frac{b}{9} + \frac{c}{9} + \frac{4b}{9} - \frac{4c}{9} \right) = -\mu p(1-p) \left( \frac{2b}{27} - \frac{2c}{27} \right)
\]
Now since the total one-step change must equal zero, we have \( T_1^* + T_2^* + T_3^* + T_4^* = 0 \), or \( T_4^* = -T_1^* - T_2^* - T_3^* \). But \( T_2^* = 0 \), so we have \( T_4^* = -T_1^* - T_3^* \):

\[
-\mu \pi_\delta b^* = \mu p (1-p) \left( \frac{22b}{27} + \frac{14c}{9} + \frac{2b}{27} - \frac{2c}{27} \right)
\]

\[
\pi_\delta = p (1-p) \left( -2b - \frac{10c}{3} \right)
\]

As long as \( \pi_\delta \) is positive, then the frequency of \( A \) alleles will be higher with the game played than at neutrality. To zeroth order in \( \mu \), \( A \) will be at a selective advantage over the allele \( B \). In order for this to occur, we need \(-2b - \frac{10c}{3} > 0 \).

We can insert a check to see if the result we got is correct, by finding the expression for \( \pi_\delta \) using the one-step state transition matrix. The transition matrix can be separated into its neutral and selective components, and we try to isolate for \( \pi_\delta \). The transition matrix \( M \) can be represented in the following way:

\[
M = \begin{bmatrix} A + \delta A & \mu(B + \delta B) \\ C + \delta C & I - \mu D \end{bmatrix} \tag{4}
\]

All we have done so far is break up \( M \) where \( A, \delta A \) represent the entries for the mixed states, and the section \( I - \mu D \) represents the 2x2 block where we begin and end in a pure state. We want the selection component for the eigenvector associated with eigenvalue 1:

\[
\begin{bmatrix} A + \delta A & \mu(B + \delta B) \\ C + \delta C & I - \mu D \end{bmatrix} \begin{bmatrix} \mu(x + \delta x) \\ \mu(y + \delta y) \end{bmatrix} = \begin{bmatrix} \mu(x + \delta x) \\ \mu(y + \delta y) \end{bmatrix} \tag{5}
\]

Using this notation, \( x \) and \( y \) represent the neutral state frequencies for the mixed states and pure states respectively. \( \delta x \) and \( \delta y \) represent the selective components of the long-term state frequencies, and these are what we need to solve for. Expanding (5):

\[
\begin{bmatrix} A & \mu B \\ C & I - \mu D \end{bmatrix} \begin{bmatrix} \mu(\delta x) \\ \delta y \end{bmatrix} + \begin{bmatrix} \delta A & \mu(B) \\ \delta C & 0 \end{bmatrix} \begin{bmatrix} \mu x \\ y \end{bmatrix} = \begin{bmatrix} \mu(\delta x) \\ \delta y \end{bmatrix} \tag{6}
\]

Note that all the terms of order two in \( \delta \) go away and we have already evaluated the neutral case, so we ignore that equation. Then (6) simplifies to:

\[
\begin{bmatrix} I - A & -\mu B \\ -C & \mu D \end{bmatrix} \begin{bmatrix} \mu(\delta x) \\ \delta y \end{bmatrix} = \begin{bmatrix} \delta A & \mu(B) \\ \delta C & 0 \end{bmatrix} \begin{bmatrix} \mu x \\ y \end{bmatrix} \tag{7}
\]

Now we get a set of equations from (7):

\[
(I - A)\delta x - B(\delta y) = (\delta A)x + (\delta B)y
\]
\[-C(\delta x) + D(\delta y) = (\delta C)x\]

We can now solve the system for \(\delta x, \delta y\):

\[
\begin{bmatrix}
I - A & -B \\
-C & D
\end{bmatrix}
\begin{bmatrix}
\delta x \\
\delta y
\end{bmatrix}
= \begin{bmatrix}
(\delta A)x + (\delta B)y \\
(\delta C)x
\end{bmatrix}
\]

(8)

As a reminder,:

\[
M := \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 2\mu\frac{p}{3} \\
\frac{1}{3} - \frac{2b}{9} - \frac{4c}{9} & \frac{1}{2} - \frac{c}{6} & \frac{1}{3} - \frac{b}{9} - \frac{2c}{9} & 0 & \mu(1 - p) \left(\frac{1}{3} + \frac{2b}{9} - \frac{2c}{9}\right) & 0 \\
0 & \frac{1}{3} + \frac{c}{3} & \frac{1}{2} + \frac{c}{6} & \frac{1}{3} + \frac{4b}{9} + \frac{2c}{9} & 0 & \mu\frac{p}{3} \\
0 & 0 & 0 & 0 & \mu(1 - p) \left(\frac{2}{3} - \frac{2b}{9} + \frac{2c}{9}\right) & 0 \\
0 & \frac{1}{6} - \frac{c}{6} & 0 & \frac{2}{3} - \frac{4b}{9} - \frac{2c}{9} & 1 - \mu + \mu p & 0 \\
\frac{2}{3} + \frac{2b}{9} + \frac{4c}{9} & 0 & \frac{1}{6} + \frac{b}{9} + \frac{c}{18} & 0 & 0 & 1 - \mu p
\end{bmatrix}
\]

Substituting values from the matrix \(M\) into (8):

\[
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & -\frac{2p}{3} \\
-\frac{1}{3} & \frac{1}{2} & -\frac{1}{3} & 0 & \frac{p - 1}{3} & 0 \\
0 & -\frac{1}{3} & \frac{1}{2} & -\frac{1}{3} & 0 & -\frac{p}{3} \\
0 & 0 & 0 & 1 & \frac{2(p - 1)}{3} & 0 \\
0 & -\frac{1}{6} & 0 & -\frac{2}{3} & 1 - p & 0 \\
-\frac{2}{3} & 0 & -\frac{1}{6} & 0 & 0 & p
\end{bmatrix}
\begin{bmatrix}
\delta x \\
\delta y
\end{bmatrix}
= \begin{bmatrix}
0 \\
p(1 - p)(-8b - 49c) \\
p(1 - p)(49c + 8b) \\
p(1 - p)(2c - 2b) \\
p(1 - p)(-19c - 8b) \\
p(1 - p)(14b + 13c)
\end{bmatrix}
\]

Using MAPLE to row-reduce the system, we arrive at:
27

From this calculation, we arrive at:

\[ \pi_A^\delta + \frac{p\pi_B^\delta}{p-1} = -\frac{2p(5c+3b)}{3} \]

From the RV allele frequency analysis, we found that to zeroth order in \( \mu \), \( \pi_B^\delta = -\pi_A^\delta \). Then:

\[ \pi_A^\delta = p(1-p)\left(-2b - \frac{10c}{3}\right) \]

which is the exact same answer we obtained from the reproductive value allele-frequency calculation.

**Death-Birth updating mechanism with Fecundity Payoffs**

We now analyze a different type of interaction within the population. Under this updating protocol, an individual is chosen to die at random, and then the neighbors compete for the vacant spot based on relative fecundities.

One will notice that in the three-star population, much of the births are determined without selection playing a part. This is simply due to the graph structure, whereby if a leaf is chosen to die the hub automatically gets to fill the vacant spot with an offspring. Nevertheless we arrive at some interesting dynamics.

If a random individual is chosen to die, then the probability of mortality for each node is 1/3, regardless of which allele occupies the node.

We now look at the birthing probabilities for the nodes in the A-A-B state. The benefit-cost game is played, and so the A-leaf has fecundity \( 1 + b - c \), the hub has fecundity \( 1 + b - 2c \), and the B-leaf has fecundity \( 1 + b \). A leaf can only give birth when the hub is chosen to die, at which point it must compete with the other leaf for the vacant spot. We can see that the hub will give birth 2/3 of the time in this population, when either leaf is chosen to die. Mathematically, the probabilities of giving birth are:
\[
\frac{1+b-c}{3(2+2b-c)} = \frac{1}{6} (1 + b - c) \left(1 - b + \frac{c}{2}\right) = \frac{1}{6} \left(1 - b + \frac{c}{2} + b - c\right) = \frac{1}{6} - \frac{c}{12}
\]
\[
A = \frac{2}{3}
\]
\[
\frac{1+b}{3(2+2b-c)} = \frac{1}{6} (1 + b) \left(1 - b + \frac{c}{2}\right) = \frac{1}{6} \left(1 - b + \frac{c}{2} + b\right) = \frac{1}{6} + \frac{c}{12}
\]

The reproductive values for the nodes can be calculated from (2):

\[
\sum_j \frac{d_j b_j}{d_i} v_j - v_i = 0
\]

Fix either of the leaves. Then:

\[
\frac{d_{iH} b_L}{d_H} v_H - v_L = 0
\]

Under a death-birth mechanism, the leaves give birth only when the hub dies. Then they split the offspring on the vacant node, so that \(b_L = 1/6\).

\[
\frac{v_H}{2} = v_L
\]

Letting \(v_H = 2\), \(v_L = 1\). \(v_L^* = d_{LH} v_H = 2\). Then \(v_H^* = 2d_{LH} v_L = 1\).

Now we complete the state fitness table:
<table>
<thead>
<tr>
<th>State</th>
<th>Fecundity</th>
<th>fecundity-RV</th>
<th>Mortality</th>
<th>mortality-RV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( b_i )</td>
<td>( v_i^* )</td>
<td>( d_i )</td>
<td>( v_i )</td>
</tr>
<tr>
<td>B</td>
<td>1/6</td>
<td>2</td>
<td>1/3</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>2/3</td>
<td>1</td>
<td>1/3</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>1/6</td>
<td>2</td>
<td>1/3</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>1/6-c/12</td>
<td>2</td>
<td>1/3</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>2/3</td>
<td>1</td>
<td>1/3</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>1/6+c/12</td>
<td>2</td>
<td>1/3</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>1/6</td>
<td>2</td>
<td>1/3</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>2/3</td>
<td>1</td>
<td>1/3</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>1/6</td>
<td>2</td>
<td>1/3</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1/6</td>
<td>2</td>
<td>1/3</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>1/6</td>
<td>2</td>
<td>1/3</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>2/3</td>
<td>1</td>
<td>1/3</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>1/6</td>
<td>2</td>
<td>1/3</td>
<td>1</td>
</tr>
</tbody>
</table>

All we are missing to complete the allele frequency calculations is the long-term state frequency eigenvector for the neutral population. Once again, we find the one-step probability transition matrix for the new update mechanism: As an example:

To go from the state A-A-A to A-A-B, we first need a leaf to die. This happens with probability \( 2/3 \). Then we need the hub, an A, to mutate to a B. This happens with probability \( \mu(1 - p) \). Thus the probability of going from A-A-A to A-A-B is \( \frac{2\mu(1-p)}{3} \). Following this logic, we arrive at the one-step transition matrix \( M \):
Once again, we need the eigenvector associated with eigenvalue 1 for this matrix, with \( b = c = 0 \) to give us neutrality. Proceeding:

\[
(M - I)y = 0 : 
\]

\[
\begin{bmatrix}
-1 & 0 & 0 & 0 & 0 & \frac{\mu p}{3} \\
\frac{2}{3} & -\frac{1}{3} & \frac{1}{6} & 0 & \frac{2\mu(1-p)}{3} & 0 \\
0 & \frac{1}{6} & -\frac{1}{2} & \frac{2}{3} & 0 & \frac{2\mu p}{3} \\
0 & 0 & 0 & -1 & \frac{\mu(1-p)}{3} & 0 \\
0 & \frac{1}{3} & 0 & \frac{1}{3} & -\mu + \mu p & 0 \\
\frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & -\mu p \\
\end{bmatrix} \begin{bmatrix}
y_1 \\
y_2 \\
y_3 \\
y_4 \\
y_5 \\
y_6 \\
\end{bmatrix} = 0
\]

Row reducing in MAPLE, this simplifies to:

\[
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & \frac{\mu p}{3} \\
0 & 1 & 0 & 0 & \frac{8\mu p}{3} & 8\mu p \\
0 & 0 & 1 & 0 & \frac{8\mu p}{3} & \frac{8\mu p}{3} \\
0 & 0 & 0 & 1 & \frac{\mu p}{3} & \mu p \\
0 & 0 & 0 & 0 & \frac{p}{p-1} & p \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix} \begin{bmatrix}
y_1 \\
y_2 \\
y_3 \\
y_4 \\
y_5 \\
y_6 \\
\end{bmatrix} = 0
\]
Now we have all we need to calculate the total change in allele frequency from the RV analysis:

\[
T_1^* = \sum_i \frac{1}{N} \sum x_i \frac{1}{N} = \frac{1}{3} (\frac{c}{6} - \frac{c}{6}) bp(1-p) = \frac{cp(1-p)}{27} = -8 \mu \frac{cp(1-p)}{27}
\]

\[
T_2^* = 0
\]

\[
T_3^* = -\mu p(1-p) \left( b_A^* - b_B^* \right) = 0
\]

\[
T_4^* = -\mu p_A^* b^* = 8 \mu \frac{cp(1-p)}{27} = -T_1^*.
\]

This gives us \( p_{A}^* = -\frac{2cp(1-p)}{3} \)

In order for \( A \) to be at a selective advantage over \( B \), we must have \( c < 0 \).

**Birth-Death Updating with Survival Payoffs**

The third type of game we present is of a slightly different tone than the previous ones. In this evolution protocol, a random individual in the population is chosen to give birth, but the deaths are now determined by the game, using survivals instead of fecundities. Since the deaths are now determined by the game, we will use mortalities to determine the probabilities of transitioning among states. An \( A \) individual giving a benefit \( b \) to a neighbour will reduce that neighbours’ mortality by \( b \). The \( A \) individual incurs a cost \( c \) to itself, which actually increases its mortality by \( c \). We start with a baseline mortality of 1 for each individual. As an example, we look at the mortalities for individuals in the A-A-B state. A leaf will be randomly chosen to give birth 2/3 of the time, and so the hub will die with probability 2/3. If the hub is chosen to give birth, then the game determines which leaf dies:

\[
A \text{ } \frac{1-b+c}{3(2-2b+c)} = \frac{1}{6} (1-b+c) \left( 1 + b - \frac{c}{2} \right) = \frac{1}{6} \left( 1 + b - \frac{c}{2} - b + c \right) = \frac{1}{6} + \frac{c}{12}
\]
Note that even though the game is defined using survival payoffs, when performing calculations we use mortalities. This is for simplicity, as a survival benefit is the same as subtracting from mortality, and a survival cost is the same as adding to mortality. The reproductive values are the same as in the birth-death with fecundity game, because we use neutrality to figure them out. Now we provide the state fitness table:

<table>
<thead>
<tr>
<th>State</th>
<th>Fecundity</th>
<th>fecundity-RV</th>
<th>Mortality</th>
<th>mortality-RV</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>1/3</td>
<td>1</td>
<td>1/6</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>1/3</td>
<td>2</td>
<td>2/3</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1/3</td>
<td>1</td>
<td>1/6</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>1/3</td>
<td>1</td>
<td>1/6+c/12</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>1/3</td>
<td>2</td>
<td>2/3</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1/3</td>
<td>1</td>
<td>1/6-c/12</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>1/3</td>
<td>1</td>
<td>1/6</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>1/3</td>
<td>2</td>
<td>2/3</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>1/3</td>
<td>1</td>
<td>1/6</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>1/3</td>
<td>1</td>
<td>1/6</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>1/3</td>
<td>1</td>
<td>1/6</td>
<td>2</td>
</tr>
</tbody>
</table>

Now we find the neutral state frequencies for the long term, and so the one-step probability transition matrix is needed. As an example, say we wanted to go from state A-A-B to itself. Then either the A-leaf must produce an A if chosen, the B-leaf must produce an A if chosen, or the hub gives birth and goes to the appropriate leaf depending on the offspring. Because we ignore mutation in the mixed states, then we only care about the probability of the A-leaf being chosen and producing an A, or the hub reproducing to an A and sending it to the leaf which is already occupied by an A. The A-leaf is

\[
A \frac{1-b+2c}{3-3b+3c} = \frac{2}{3}
\]

\[
B \frac{1-b}{3(2-2b+c)} = \frac{1}{6} \left(1 - b\right) \left(1 + b - \frac{c}{2}\right) = \frac{1}{6} \left(1 + b - \frac{c}{2} - b\right) = \frac{1}{6} - \frac{c}{12}
\]
chosen to give birth 1/3 of the time. If it is chosen, it produces an A individual with probability $1 - \mu + \mu p$. If the hub is chosen to give birth, which occurs 1/3 of the time, then the leaves die according to their mortalities. We would need the A-leaf to die, which will happen with probability \( \frac{1-b+c}{2-2b+c} = \frac{1}{2} (1 - b + c) \left( 1 + b - \frac{c}{2} \right) = \frac{1}{2} \left( 1 + b - \frac{c}{2} - b + c \right) = \frac{1}{2} + \frac{c}{4} \) of the time. Thus the total probability of transitioning from the A-A-B state to itself is:

\[
\frac{1}{3} \left( 1 - \mu + \mu p \right) + \frac{1}{3} \left( \frac{1}{2} + \frac{c}{4} \right) = \frac{1}{2} + \frac{c}{12} \text{ (ignoring mutation among mixed states)}
\]

To get the rest of the matrix we use a similar analysis, and the end result looks like this:

\[
M := \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & \frac{2\mu p}{3} \\
\frac{1}{3} & \frac{1}{2} + \frac{c}{12} & \frac{1}{3} & 0 & \frac{\mu (1-p)}{3} & 0 \\
0 & \frac{1}{3} & \frac{1}{2} - \frac{c}{12} & \frac{1}{3} & 0 & \frac{\mu p}{3} \\
0 & 0 & 0 & \frac{2\mu (1-p)}{3} & 0 & 0 \\
0 & \frac{1}{6} - \frac{c}{12} & 0 & \frac{2}{3} & 1 - \mu + \mu p & 0 \\
\frac{2}{3} & 0 & \frac{1}{6} + \frac{c}{12} & 0 & 0 & 1 - \mu p
\end{bmatrix}
\]

Since we are interested in neutral frequencies, we set \( b = c = 0 \), and solve \((M - I)z = 0\), which equates to:

\[
\begin{bmatrix}
-1 & 0 & 0 & 0 & 0 & \frac{2\mu p}{3} \\
1 & -1 & 1 & 0 & \frac{\mu (1-p)}{3} & 0 \\
\frac{1}{3} & -\frac{2}{3} & 0 & \frac{\mu p}{3} & 0 \\
0 & 0 & 0 & -1 & \frac{2\mu (1-p)}{3} & 0 \\
0 & 1 & 0 & \frac{2}{3} & -\mu + \mu p & 0 \\
\frac{2}{3} & 0 & \frac{1}{6} & 0 & 0 & -\mu p
\end{bmatrix} \begin{bmatrix}
z_1 \\
z_2 \\
z_3 \\
z_4 \\
z_5 \\
z_6
\end{bmatrix} = 0
\]

We then row reduce this in MAPLE, and we get:
\[
\begin{bmatrix}
1 & 0 & 0 & 0 & -\frac{2\mu p}{3} \\
0 & 1 & 0 & 0 & -\frac{10\mu p}{3} \\
0 & 0 & 1 & 0 & -\frac{10\mu p}{3} \\
0 & 0 & 0 & 1 & -\frac{2\mu p}{3} \\
0 & 0 & 0 & 0 & -\frac{3p}{p-1} \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
z_1 \\
z_2 \\
z_3 \\
z_4 \\
z_5 \\
z_6
\end{bmatrix}
= 0
\]

And finally, the neutral state frequencies are:

\[
\begin{bmatrix}
z_1 \\
z_2 \\
z_3 \\
z_4 \\
z_5 \\
z_6
\end{bmatrix}
= \begin{bmatrix}
\frac{2\mu p(1-p)}{3} \\
\frac{10\mu p(1-p)}{3} \\
\frac{10\mu p(1-p)}{3} \\
\frac{2\mu p(1-p)}{3} \\
\frac{3}{p} \\
\frac{1-p}{1-p}
\end{bmatrix}
\]

And now we analyze the total change in allele frequency since we have all the necessary components:

\[T_1^* = \sum_S \frac{1}{N} \sum_i w_i^s x_i \pi_S^N = \frac{1}{3} \left(-\frac{c}{6} - \frac{c}{6}\right) \left(10\mu \frac{p(1-p)}{3}\right) = -10c\mu \frac{p(1-p)}{27}\]

\[T_2^* = 0\]

\[T_3^* = -\mu p(1-p) \left(b_A^* - b_B^*\right) = 0\]

\[T_4^* = -\mu \pi_A^* b^*\]

The total change in allele frequency has to equal zero, so \(T_4^* = -T_1^*\).

\[T_4^* = -\mu \pi_A^* b^* = 10c\mu \frac{p(1-p)}{27} = -T_1^*\]

This yields \(\pi_A^* = -\frac{5cp(1-p)}{6}\)

So in order for \(A\) to be at a selective advantage over \(B\), we must have \(c < 0\).
Death-Birth Updating Mechanism with Survival Payoffs

The last type of evolution protocol we look at operates as follows: An individual in the population is chosen to die, based on relative mortalities from the benefit-cost game. Once the node has been vacated, then all of the neighbors have an equal chance of providing an offspring to fill the node. As in the last evolution protocol, a benefit $b$ will decrease one’s baseline mortality and a cost incurred by giving this benefit will increase $A$’s baseline mortality. Because we are analyzing the 3-star graph, if the hub is chosen to die, then the leaves will each have a $\frac{1}{3}$ chance of providing the offspring to the vacant hub. Now we have a look at the probabilities of dying, which will then determine the probabilities of giving birth, looking at the state B-A-B:

\[
B: \frac{1-b}{3-2b+2c} = \frac{1}{3} (1 - b) \left(1 + \frac{2b}{3} - \frac{2c}{3}\right) = \frac{1}{3} \left(1 + \frac{2b}{3} - \frac{2c}{3} - b\right) = \frac{1}{3} \left(\frac{b}{9} - \frac{2c}{9}\right)
\]

\[
A: \frac{1+2c}{3-2b+2c} = \frac{1}{3} (1 + 2c) \left(1 + \frac{2b}{3} - \frac{2c}{3}\right) = \frac{1}{3} \left(1 + \frac{2b}{3} - \frac{2c}{3} + 2c\right) = \frac{1}{3} \left(\frac{2b}{9} + \frac{4c}{9}\right)
\]

\[
B: \frac{1-b}{3-2b+2c} = \frac{1}{3} (1 - b) \left(1 + \frac{2b}{3} - \frac{2c}{3}\right) = \frac{1}{3} \left(1 + \frac{2b}{3} - \frac{2c}{3} - b\right) = \frac{1}{3} \left(\frac{b}{9} - \frac{2c}{9}\right)
\]

From here, we figure out the probabilities of each individual in this population giving birth. If either of the leaves die based on their mortality, then the hub automatically gives birth, and if the hub dies, then the leaves have equal probability of giving birth:

\[
B: \frac{1}{2} \left(\frac{1}{3} + \frac{2b}{9} + \frac{4c}{9}\right) = \frac{1}{6} + \frac{b}{9} + \frac{2c}{9}
\]

\[
A: 2 \left(\frac{1}{3} - \frac{b}{9} - \frac{2c}{9}\right) = \frac{2}{3} - \frac{2b}{9} - \frac{4c}{9}
\]

\[
B: \frac{1}{2} \left(\frac{1}{3} + \frac{2b}{9} + \frac{4c}{9}\right) = \frac{1}{6} + \frac{b}{9} + \frac{2c}{9}
\]

We want to obtain the reproductive values for each node in this type of game. They will be the same as in the death-birth game with fecundity, since we work under neutrality to calculate reproductive values. Now the state fitness table can be provided:
<table>
<thead>
<tr>
<th>State</th>
<th>Fecundity</th>
<th>fecundity-RV</th>
<th>Mortality</th>
<th>mortality-RV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( b_i )</td>
<td>( v_i^j )</td>
<td>( d_i )</td>
<td>( v_i )</td>
</tr>
<tr>
<td>B</td>
<td>( 1/6 + b/9 + 2c/9 )</td>
<td>2</td>
<td>( 1/3 - b/9 - 2c/9 )</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>( 2/3 - 2b/9 - 4c/9 )</td>
<td>1</td>
<td>( 1/3 + 2b/9 + 4c/9 )</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>( 1/6 + b/9 + 2c/9 )</td>
<td>2</td>
<td>( 1/3 - b/9 - 2c/9 )</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>( 1/6 + c/6 )</td>
<td>2</td>
<td>( 1/3 )</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>( 2/3 - c/3 )</td>
<td>1</td>
<td>( 1/3 + c/3 )</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>( 1/6 + c/6 )</td>
<td>2</td>
<td>( 1/3 - c/3 )</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>( 1/6 - b/9 - c/18 )</td>
<td>2</td>
<td>( 1/3 + b/9 + 2c/9 )</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>( 2/3 - 2b/9 + 4c/9 )</td>
<td>1</td>
<td>( 1/3 - 2b/9 - c/9 )</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>( 1/6 - b/9 - c/18 )</td>
<td>2</td>
<td>( 1/3 + b/9 - c/9 )</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>( 1/6 - 2b/9 - c/9 )</td>
<td>2</td>
<td>( 1/3 + 2b/9 + c/9 )</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>( 2/3 - 4b/9 + 2c/9 )</td>
<td>1</td>
<td>( 1/3 - 4b/9 - 2c/9 )</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>( 1/6 - 2b/9 - c/9 )</td>
<td>2</td>
<td>( 1/3 + 2b/9 + c/9 )</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>( 1/6 )</td>
<td>2</td>
<td>( 1/3 )</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>( 2/3 )</td>
<td>1</td>
<td>( 1/3 )</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>( 1/6 )</td>
<td>2</td>
<td>( 1/3 )</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>( 1/6 - b/9 + c/9 )</td>
<td>2</td>
<td>( 1/3 + b/9 - c/9 )</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>( 2/3 - 2b/9 + 2c/9 )</td>
<td>1</td>
<td>( 1/3 - 2b/9 + 2c/9 )</td>
<td>2</td>
</tr>
<tr>
<td>A</td>
<td>( 1/6 - b/9 + c/9 )</td>
<td>2</td>
<td>( 1/3 + b/9 - c/9 )</td>
<td>1</td>
</tr>
</tbody>
</table>

The next step is to find out what the neutral long-term state frequencies are. To provide an example, we find the probability of transitioning from state B-A-B to state A-A-B is. We see that the only for this to happen is to have one of the leaves die and have the A provide an A offspring to the vacant node. From the table we see that the leaves will die with probability \( 2 \left( \frac{1}{3} - \frac{b}{9} - \frac{2c}{9} \right) \), and if this happens the hub will provide an A offspring with probability \( 1 - \mu + \mu p \). Once again ignoring mutation among mixed states, the probability of going from B-A-B to A-A-B in one time step is \( \frac{2}{3} \left( \frac{2b}{9} - \frac{4c}{9} \right) \). We continue to completely fill in the one-step transition matrix:
To obtain the eigenvector associated with eigenvalue 1, we solve 

\[(M - I)k = 0 \text{ (with } b = c = 0)\].

This equates to:

\[
\begin{bmatrix}
-1 & 0 & 0 & 0 & 0 & \mu p \\
\frac{2}{3} & 1 & 1 & 0 & 2\mu(1 - p) & 0 \\
0 & \frac{1}{2} & 1 & \frac{2}{3} & 0 & 2\mu p \\
0 & 0 & 0 & -1 & \mu(1 - p) & 0 \\
0 & \frac{1}{3} & 0 & \frac{1}{3} & -\mu + \mu p & 0 \\
\frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & -\mu p
\end{bmatrix}
\]

Using MAPLE to row-reduce, we arrive at:

\[
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & \frac{\mu p}{3} \\
0 & 1 & 0 & 0 & 0 & 8\mu p \\
0 & 0 & 1 & 0 & 0 & 3\mu p \\
0 & 0 & 0 & 1 & 0 & 3\mu p \\
0 & 0 & 0 & 0 & 1 & \frac{p}{p - 1} \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

This gives us the neutral state frequencies in the long-term:
Now we conduct the calculation of the $T_1^*$:

$$T_1^* = \sum_{s \in S} \frac{1}{N} \sum_i w_i^x x_i \pi_s^N$$

$$= \frac{1}{3} \left( \begin{array}{c}
- \frac{2b}{9} - \frac{4c}{9} - \frac{4b}{9} - \frac{8c}{9} \\
\mu \frac{p(1-p)}{3} \end{array} \right) + \left( \begin{array}{c}
\frac{c}{3} - \frac{2c}{3} \\
8\mu \frac{p(1-p)}{3} \end{array} \right)$$

$$= \mu p(1-p) \left( -\frac{6b}{81} - \frac{12c}{81} - \frac{24b}{81} - \frac{24c}{81} - \frac{12b}{81} - \frac{6c}{81} \right)$$

$$= \mu p(1-p) \left( -\frac{14b}{27} - \frac{90c}{81} \right)$$

$T_2^* = 0$

$$T_3^* = -\mu p(1-p) \left( \frac{b_A^x}{b_A^*} - \frac{b_B^x}{b_B^*} \right)$$

$$= -\mu p(1-p) \left( \frac{1}{3} \left( -\frac{2b}{9} + \frac{2c}{9} + \frac{2c}{9} + \frac{2b}{9} + \frac{2c}{9} \right) \right)$$

$$= -\mu p(1-p) \left( -\frac{2b}{27} + \frac{2c}{27} \right)$$

$T_4^* = -\mu r_A^x b^x$

The total change in allele frequency has to equal zero at equilibrium, so $T_4^* = -T_1^* - T_3^*$.

$$T_4^* = -\mu r_A^x b^x = \mu p(1-p) \left( \frac{14b}{27} + \frac{90c}{81} \right) + \mu p(1-p) \left( -\frac{2b}{27} + \frac{2c}{27} \right) = \mu p(1-p) \left( \frac{12b}{27} + \frac{32c}{27} \right)$$

This gives us $\pi_A^* = p(1-p) \left( -b - \frac{8c}{3} \right)$.

In order for $A$ to be at a selective advantage over $B$, we must have $-b - \frac{8c}{3} > 0$. 

38
Discussion

The allele frequency analysis outlined what factors influence how a finite population evolves over time. Working through the calculations with a small population was a good way to see how the analysis could be applied. The state transition matrix will get quite large with even a small increase in population size, as the number of possible configurations of A and B individuals increases very rapidly as population sizes increase. This is another reason why a small population size was chosen.

It was nice to uncover a simple first order expression for when the A allele is at an advantage upon introducing the game, for the four different evolution protocols. We showed how to get the same expression using only the transition matrix (in the birth-death with fecundity payoff scenario), however the allele frequency analysis seemed to give a lot more insight into the intricacy of the problem. We saw how changing the structure of the game affected the conditions for the A allele to be at an advantage.

The idea of using reproductive values proved to be quite useful. To obtain expressions for the $\pi^d$ terms is very cumbersome. By tweaking the original allele frequency analysis to track reproductive value allele frequency, we eliminated the need to obtain the $\pi^d$ terms. The state transition matrix is easier to work with if we only need to find the long term frequencies under neutrality. Certainly the allele frequency analysis without reproductive values could have been used to perform the calculations. However, for larger populations, using reproductive values is more feasible.

For populations where all nodes have the same reproductive value, the two analyses are exactly the same (by making all the reproductive values equal to 1, $b_i = b_i^*$ and $d_i = d_i^*$). This will happen when all nodes are in the same class, i.e. when the graph is regular. In a regular graph, all nodes have the same number of neighbors, and so all nodes behave the same way.

The conditions for A to be at a selective advantage are included below to remind the reader:

<table>
<thead>
<tr>
<th>Evolution Protocol</th>
<th>First Order Condition for A to be at a Selective Advantage over B</th>
</tr>
</thead>
<tbody>
<tr>
<td>i) B-D with Fecundity Payoff</td>
<td>$-2b - \frac{10c}{3} &gt; 0$</td>
</tr>
<tr>
<td>ii) D-B with Fecundity Payoff</td>
<td>$c &lt; 0$</td>
</tr>
<tr>
<td>iii) B-D with Survival Payoff</td>
<td>$c &lt; 0$</td>
</tr>
<tr>
<td>iv) D-B with Survival Payoff</td>
<td>$-b - \frac{8c}{3} &gt; 0$</td>
</tr>
</tbody>
</table>

Of the four scenarios analyzed in the application section, it seemed as though i) was very similar to iv). In both of these scenarios the first event that occurred (birth or death) was determined by the game. Thus, the benefits and costs were seen frequently throughout the state-transition matrices for these scenarios, as they had a large impact on the dynamics. Assuming that the incurred cost $c$ is always
positive, under i) and iv), the benefit b must be spiteful \((b < 0)\), for A to be at a selective advantage. This is to be expected because if A gave positive benefits while incurring a positive personal cost, while B is simply receiving benefits, then the B individuals would definitely be better off.

There are also similarities between the evolution protocols ii) and iii). This was due to the fact that the first event (birth or death) happened at random. The randomness of the first event tended to dictate what happened in the second event because of the small population size. As a result, the benefits and costs rarely appeared in the transition matrices for these evolution protocols. We even arrived at the same first order condition in \(b\) and \(c\) for A to be at a selective advantage. If the costs incurred by A are positive, then the A individuals can never reach a selective advantage, no matter how altruistic \((b > 0)\) or spiteful \((b < 0)\) the benefits are.

Under ii) and iii), we notice that the \(T^*_3\) term is zero. The reason for this is because \(\bar{b}^*_A = 0\). The \(\bar{b}^*_B\) component of \(T^*_3\) will always be zero because in a population of all B, there are no selection terms present, as the benefits and costs come from the presence of A individuals. We always have \(\bar{b}^*_A = 0\) under iii), because birth happens at random, and so the game plays no part. Under ii), certain graph structures will give us \(\bar{b}^*_A = 0\). No matter what node is chosen to die, if all the neighbors are equally likely to send offspring to the vacant node, then there will be no selection component to the average fecundity term. We see that in a regular graph, everyone will have the same fecundity in an all A population under ii). Thus \(b_i = 1/N\), and the \(T^*_3\) term would be zero. In a N-star, where there is one hub node connected to \(N - 1\) other nodes (which are only connected to the hub), we see that if the hub dies, the rest of the nodes will have an equal chance at attaining the vacant spot because they all have the same fecundity. Of course, if one of the \(N - 1\) leaves dies, then only the hub can replace it. In these graphs, the selection components do not play a factor, and so \(\bar{b}^*_A = 0\).

In a future endeavor it would be interesting to see if the birth-death with fecundity payoff scenario ever gives the same result as the death-birth with survival payoff scenario (in terms of selective advantage), and when the birth-death with survival payoff protocol yields the same result as the death-birth with fecundity payoff protocol. One could analyze finite or infinite populations, as well as graphs where all nodes have the same reproductive value. The analysis conducted in this paper would provide a good framework for this continuation.
References

