Project Number: MA-RYL-1001

Three-Allele Selection Model

A Major Qualifying Project

submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

by

Colbert Sesanker

Jesse Bowers

March 9, 2011

Approved

Professor Roger Y. Lui
Major Advisor

Abstract

In this MQP, we attempt to classify all the possible dynamics of a three-allele selection model. A complete classification of our model, a system of two nonlinear ODEs, first requires determining all possible cases of coexisting equilibria, then requires partitioning these cases into patterns based on which equilibria are asymptotically stable. Last, we decompose the patterns into all the possible dynamics based on the separatricies present in each pattern. This three step process completely classifies the behavior of our three-allele selection model.

Acknowledgments

We would like to thank Professor Lui for his extensive help, motivation and sense of humor. Also, we would like to thank Professor Lin Lin Su for her patience, suggestions and extensive help with the proofs, especially Lemmas 4.1-4.4.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Understanding the selection mechanisms for genes in a population is of both practical and theoretical value. Population genetics models have primary applications in epidemiology, ecology, human genetics and in animal and plant breeding; it also has connections in natural history, mathematics statistics and computing. Evolutionarily oriented population genetics models attempt to explain and predict the dynamics of various traits in a population. Our evolutionary model places emphasis on how different genes give individuals a survival and/or reproductive advantage over successive generations. A gene is composed of two alleles which determine the variation of the gene. Roughly an individual gets one allele from their mother and one from their father. Different combinations of alleles may yield different traits. These traits may give an individual a survival and/or reproductive advantage. For example, the peppered moth exists in both light and dark colors in the United Kingdom. After the industrial revolution, the trees became black with soot and within a few generations, the majority of the moths were dark. The dark colored moths had a survival advantage, because they had better camouflage and were less likely to be killed by predators. This is one of the simplest, and most often sited cases of a selective advantage [3]; the peppered moth example illustrates the types of simple traits our model addresses. A full account of evolutionarily forces beyond selection is very complex and requires the introduction of many parameters that address mutation, the mating system, epigenetic, environmental and other factors. Our goal is not to understand the evolutionary forces of nature, rather we endeavor to develop mathematical tools we can use to solve other problems.

1.1 Genes and Chromosomes

DNA is a double stranded helix composed of four bases, adenine (A), guanine (G), thymine (T) and cytosine (C). These bases, called nucleotides, are arranged sequentially in a double helix with pairings A-T and G-C; three bases, (e.g. AAA), code for one of 20 amino acids. A polypeptide chain is a folded structure composed of repeating units of amino acids, and a protein is composed of one or more polypeptide chains. The region of DNA that determines a polypeptide chain is referred to as a gene, its position is the gene locus, and variant sequences at this locus are called alleles. Chromosomes are tightly coiled, single threadlike complexes of DNA, bound by proteins, containing a linear arrangement of genes. The number of different chromosomes is characteristic of each species. Autosomal chromosomes refer to the 22 pairs of chromosomes not involved in the determination of sex. If the chromosome set is single, the organism is called haploid; if the chromosome set is doubled, the organism is called diploid. With the exception of bacteria, algae, mosses and fungi, most animals and higher plants are diploid. Every cell in an organism has the exact same chromosome set, with the exception of error and some special cells. An individual's genotype, is set of alleles at the loci under consideration; for a diploid organism, the genotype represents the allele present, at the same locus, on each of the doubled chromosomes.

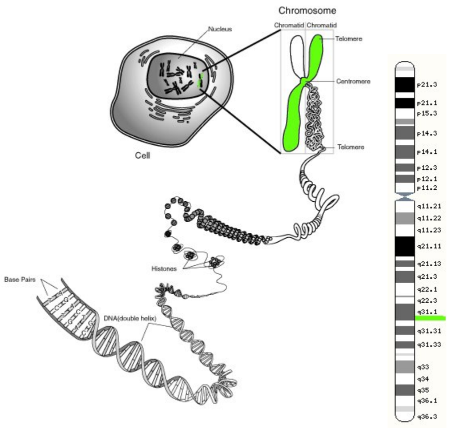


Figure 1.1: Chromosome in Diploid cell (left) Locus for Cystic Fibrosis (right)

1.2 Genotype

We express the genotype of a diploid individual by A_iA_j , where A_i represents allele i at the same locus on one of the two identical chromosomes ; it is clear that A_iA_j=A_jA_i. We use the word genotype synonymously with phenotype or the physical manifestation of a genotype. Also, in population genetics, it is common to use the term gene in place of allele; the proportion of an allele in a population may be referred to as gene frequency instead of allele frequency. A population refers to a collection of interbreeding organisms, or species. We consider a population with no distinct sexes where any individual can reproduce with any other regardless of age, genotype and ancestry. Heterozygous (homozygous) means an individual has two different (same) alleles at a particular locus. When we refer to heterozygous dominance (inferiority) we mean the genotype A_iA_j has greater (less) fitness than both A_iA_i and A_jA_j. Since we make no distinction between sexes, we only consider autosomal and not sex linked loci. In this idealization of diploid sexual reproduction, each individual gets one of two alleles from each of its two predecessors.

For example, the locus for ABO blood type in humans has six possible genotypes: AA, AO, BB, BO, AB and OO. It may appear that they are just three alleles that code for each of A, B and O but, as with many traits, they are actually many alleles that represent each of A, B and O.

Sickle cell anemia causes red blood cells to collapse into a half-moon shape resulting red-blood cell destruction and severely diminished oxygen carrying capacity; yet the disease is maintained at relatively high levels in some populations, because heterozygous carriers are resistant to malaria. Sickle cell anemia exemplifies the case of heterozygous dominance.

The locus for Cystic Fibrosis, caused by a mutation in the CFTR gene, is addressed by 7q31.2 (marked in green Figure 2.1) and is read chromosome 7, long arm (q), band 3, section 1 and sub-band 2 [4]. Again, they are multiple alleles at this locus that cause cystic fibrosis, but some are much more common than others; the most common mutation ΔF508 occurs in about 70% of cases worldwide [5]. The ΔF508 mutation is estimated to be at least 50,000 years old [6] and they are many theories that address how and why such a lethal gene could persist for so long; like sickle cell anemia, the theories claim the heterozygous carriers have a survival advantage against certain types of infections and diseases.

Why do lethal genes persist, and when does a fitter allele wipe out other alleles? Why do certain proportions of allele frequencies reach a stable equilibrium and how can we predict allele frequencies over time, given an initial proportion of allele frequencies. Many times the answers lie hidden in the mathematical structure of the selection mechanisms and remain inaccessible to ordinary reasoning. These are the types of questions our MQP attempts to answer. There has been much theoretical work done to answer these questions for two alleles, and various models are ubiquitous. Adding just one more allele, complicates matters significantly. The n-allele case is in its own catergory and requires a more theoretical approach.

1.3 Hardy-Weinberg Equilibrium

Our assumptions can be summarized with the principle of Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium states that the allele frequencies in a population will remain constant, unless acted upon by factors such as non-random mating and mutation, which we do not consider (see 1.4). Specifically, the eight assumptions of Hardy Weinberg equilibrium are: Diploid organism, sexual reproduction, non-overlapping generations, random mating, large population size (no genetic drift), equal allele frequencies in the sexes, no migration, no mutation and no selection. We consider all these assumptions except selection, so in our case Hardy-Weinberg equilibrium means that allele frequencies will change solely by the force of selection. We assume non-overlapping generations, but since our model is continuous, our assumption of Hardy-Weinberg equilibrium is an approximation.

The only assumption of Hardy Weinberg equilibrium we have not mentioned is genetic drift. Under the assumption of genetic drift, an offspring's alleles are sampled randomly from their parents; also, survival and reproduction are governed by chance. In small populations, the effects of genetic drift may be substantial, and certain alleles may disappear entirely solely by chance. An illustrative example is the loss of pigments in caved dwelling animals. In an isolated cave environment, the alleles for pigment have no survival advantage or disadvantage; allele frequencies may wander away from pigmented variants and stabilize towards lighter variants.

In a larger population this 'wandering away' of allele frequencies is much less pronounced, thus the assumption of a 'large population' is sometimes equated with 'no genetic drift'.

Again, the assumptions of our model are far too strict for practical application; practical application was not the intention. The real value of our endeavors is to develop techniques and insights that can be generalized to other more accurate population genetics models, and to expand the current literature.

Although our model may not have any practical application we do still envision a few experimental situations where it could apply. Consider the case of a large population of fruit flies in a large homogeneous environment with known frequencies of three alleles distributed equally among the sexes with relatively constant fitness coefficients. Within a few generations we should see the dynamics and equilibria predicted by our model.

1.4 Hardy-Weinberg Equilibrium Condition

Consider k alleles denoted by A_i for $i = 1, ..., k$; then P_{ij}, the unordered frequency of A_iA_j satisfies $P_{ij} + P_{ji} = 2P_{ij}$. The frequency of allele A_i in the population is $p_i = \sum_j^k P_{ij}$. The genotype A_iA_j $(i \neq j)$ can result from the unordered matings $A_iA_k \times A_lA_k$; all the possible unordered matings are: $A_iA_i \times A_jA_j$, $A_iA_k \times A_jA_j$, $A_iA_i \times A_lA_j$, $A_iA_j \times A_iA_j$ and $A_iA_k \times A_lA_j$ where $k \neq i$, $l \neq j$ and $(k, l) \neq (j, i)$. The probabilities that an offspring of these matings is of genotype A_iA_j are $1, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}$ and $\frac{1}{4}$ respectively. Therefore the total probability of A_iA_j occurring in the next generation is:

$$P'_{ij} = P_{ii}P_{jj} + \sum_{k \neq i} P_{ik}P_{jj} + \sum_{l \neq j} P_{ii}P_{lj} + \sum_{\substack{l \neq j, k \neq i \\ (k,l) \neq (j,i)}} P_{ik}P_{lj} \tag{1.1}$$

Notice, (1.1) is exactly $\sum_{k,l} P_{ik}P_{lj} = (\sum_k P_{ik})(\sum_l P_{il}) = p_ip_j$ Therefore we have the equation $P_{ij} = p_ip_j$ from the assumptions of Hardy-Weinberg equilibrium. A population that satisfies $P_{ij} = p_ip_j$ can be said to be in Hardy-Weinberg equilibrium [1].

Chapter 2

Mathematical Introduction

2.1 The Mathematical Model

Consider a gene on a single locus with J alleles, A_i, where $i \in K = \{1, 2, ..., J\}$. In this section, we find a differential equation that describes the proportion of an allele A_i in a population. This model was first introduced by Nagylaki and Crow in 1974 [2]. We follow R. Bürger in the derivation of this model. Suppose there is an individual with genotype A_iA_j with the former allele gained from his father and the latter allele from his mother. In this way, we distinguish between genotypes A_iA_j and A_jA_i while accepting that there is no practical difference between the two. Let $N = N(t)$ denote the population size, $n_{ij} = n_{ij}(t)$ denote the number of individuals with genotype A_iA_j in the population, and $p_{ij} := n_{ij}/N$ denote the proportion of genotype A_iA_j in the population. The symmetric viability matrix

$$R = \begin{bmatrix} r_{11} & r_{12} & ... & r_{1J} \\ r_{21} & r_{22} & ... & r_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ r_{J1} & r_{J2} & ... & r_{JJ} \end{bmatrix}$$

is such that for each r_{ij}:

$$\dot{n}_{ij} = r_{ij} n_{ij}$$

We assume everywhere that

Assumption 2.1. *$r_{ij} = r_{nm} \Leftrightarrow \{i, j\} = \{n, m\}$, and $r_{ii} < r_{jj} \Leftrightarrow i < j$*

We denote $n_i = n_i(t)$ as the total number of alleles A_i in the population, and $p_i := n_i/(2N)$ as proportion of allele A_i in the population. Note that:

$$\sum_{i=1}^{J} p_i = 1 \tag{2.1}$$

Now we derive an explicit form of n_i and $\dot{n}_i$ in turn:

$$n_i = \sum_{j=1}^{J} n_{ij} + \sum_{j=1}^{J} n_{ji}$$

$$\dot{n}_i = \sum_{j=1}^{J} r_{ij} n_{ij} + \sum_{j=1}^{J} r_{ji} n_{ji}$$

Now, we assume Hardy-Weinberge equilibrium: that is $P_{ij} = p_i p_j$. Then:

$$\dot{n}_i = N \sum_{j=1}^{J} r_{ij} P_{ij} + N \sum_{j=1}^{J} r_{ji} P_{ji} = 2N p_i \sum_{j=1}^{J} r_{ij} p_j$$

$$\dot{N} = \frac{\sum_{i=1}^{J} \dot{n}_i}{2} = N \sum_{i,j=1}^{J} r_{ij} p_i p_j$$

Now we have sufficient information to derive $\dot{p}_i$.

$$\begin{aligned}
\dot{p}_i &= \frac{1}{2} \frac{N \dot{n}_i - n_i \dot{N}}{N^2} \\
&= \frac{N(2N p_i \sum_{j=1}^{J} r_{ij} p_j) - n_i (N \sum_{j,k=1}^{J} r_{jk} p_j p_k)}{2N^2} \\
&= p_i \sum_{j=1}^{J} r_{ij} p_j - \frac{n_i}{2N} \sum_{j,k=1}^{J} r_{jk} p_j p_k \\
&= p_i (\sum_{j=1}^{J} r_{ij} p_j - \sum_{j,k=1}^{J} r_{jk} p_j p_k)
\end{aligned}$$

We now seek to describe $\dot{p}_i$ in terms of two new variables r_i and r. r_i is the average survivability of allele A_i in the population, and r is the average survivability of the entire population. Explicit forms of these variables are below:

$$r_i := \frac{\sum_{j=1}^{J} r_{ij} n_{ij} + \sum_{j=1}^{J} r_{ji} n_{ji}}{n_i} = \frac{\sum_{j=1}^{J} r_{ij} P_{ij} + \sum_{j=1}^{J} r_{ji} P_{ji}}{2p_i} = \sum_{j=1}^{J} r_{ij} p_j$$

$$r := \sum_{i=1}^{J} r_i p_i = \sum_{i,j=1}^{J} r_{ij} p_i p_j$$

Substitutions of r_i and r into $\dot{p}_i$ gives us:

$$\dot{p}_i = p_i (r_i - r) \tag{2.2}$$

With (2.2) we can solve for the initial value problem that describes $(p_1(t), p_2(t), ..., p_J(t))$. Recall (2.1), an equation used to derive (2.2). This equation implies $p_J = 1 - \sum_{i=1}^{J-1} p_i$. Thus we can describe p_J as a dependent variable. This fact matched with equation (2.2) gives us the following initial value problem for any $(p_1^*, p_2^*, ..., p_{J-1}^*, t^*) \in \Re^J$:

$$\begin{array}{ll} \dot{p}_1 = p_1(r_1 - r) & p_1(t^*) = p_1^* \\ \dot{p}_2 = p_2(r_2 - r) & p_2(t^*) = p_2^* \\ \dots & \dots \\ \dot{p}_{J-1} = p_{J-1}(r_{J-1} - r) & p_{J-1}(t^*) = p_{J-1}^* \\ \multicolumn{2}{c}{p_J = 1 - \sum_{i=1}^{J-1} p_i} \end{array} \tag{2.3}$$

Consider the following fictitious example. Off the coast of North America, a study on a colony of wild boars was conducted. This study has shown that a gene ”alpha” had a significant impact on which boars lived to maturity. In this population there are two alleles which compose Alpha: B and C; that is to say $J = 2$. The study determined the following two facts. First, it found that AA was more favorable than BB which was in turn more favorable than AB. Specifically if we denote A by A_1, B by A_2, and consider time in millenia; we get the following survivability matrix:

$$R = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix}$$

Second, the study found that the proportion $A : B$ in the population is $1 : 2$, that is $(p_1(0), p_2(0)) = (1/3, 2/3)$. Given this data, we can project p_1 and p_2 over time by our model. We first consider the following average survivabilities:

$$\begin{array}{l} r_1 = \sum_{j=1}^{2} r_{1j} p_j = 2p_1 + (-1)p_2 = 2p_1 - p_2 \\ r_2 = \sum_{j=1}^{2} r_{2j} p_j = (-1)p_1 + (0)p_2 = -p_1 \\ r = \sum_{i=1}^{2} r_i p_i = (2p_1 - p_2)p_1 + (-p_1)p_2 = 2p_1(p_1 - p_2) \end{array}$$

This in turn gives us $\dot{p}_i$:

$$\begin{array}{l} \dot{p}_1 = p_1(r_1 - r) = p_1([2p_1 - p_2] - [2p_1(p_1 - p_2)]) = p_1(2p_1(1 - p_1) - p_2(1 - 2p_1)) \\ \dot{p}_2 = p_2(r_2 - r) = p_2([-p_1] - [2p_1(p_1 - p_2)]) = p_1p_2(2p_2 - (1 + 2p_1)) \end{array}$$

The solution cannot be easily expressed analytically. However, the solution can be approximated numerically. The solution such a numerical approximation is given below:

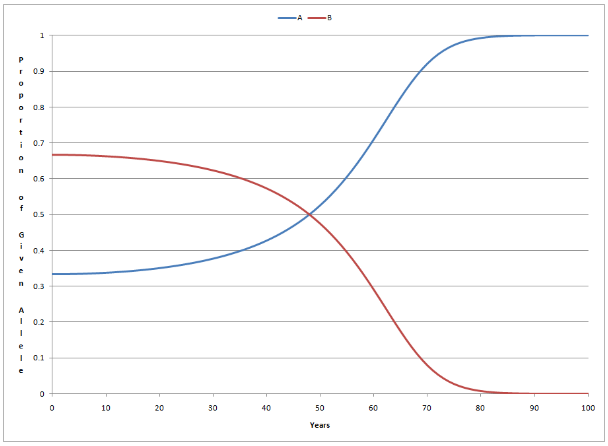


Figure 2.1: A graph of p_1 (blue) and p_2 (red) over time

Thus over the next hundred years, the population will rapidly eliminate allele B from the population in favor of allele A.

2.2 Mathematical Preliminaries

We now consider if a few basic assumptions of our model hold. We first consider if a solution to the initial value problem given by (2.3) is uniquely defined on $\Re$.

Proposition 2.2. *For any* $(p_1^*, p_2^*, ..., p_{J-1}^*) \in \Re^{J-1}$*, the initial value problem* (2.3) *has a unique solution on* $\Re$*.*

Proof. Consider (2.3). Take any $(p_1^*, p_2^*, ..., p_{J-1}^*) \in \Re^{J-1}$ as initial conditions to (2.3). Let $\dot{p}$ be a function mapping from $\Re^{J-1}$ to $\Re^{J-1}$ such that

$$\dot{p}(p_1, ..., p_{J-1}) := (p_1(r_1 - r), p_2(r_2 - r), ..., p_{J-1}(r_{J-1} - r))$$

Since $p_i(r_i - r)$ is a third degree polynomial of $J-1$ variables, $\dot{p}$ is Lipschitz continuous on any compact subset of its domain $\Re^{J-1}$. Thus by the Picard-Lindelöf theorem, a unique solution to (2.3) exists on $\Re$ □

Proposition 2.3. *The initial value problem* (2.3) *is invariant on* $\Phi := \{(p_1, ..., p_J) : [(\forall i)(p_i \geq 0)] \wedge [\sum_{i=1}^{J} p_i = 1]\}$.

Proof. Take any solution to (2.3) with initial condition $(p_1(t^*), ..., p_J(t^*)) \in \Phi$. In this paragraph we show that $\sum_{i=1}^{J} p_i(t) = 1$ for $t > t^*$. Let $q(t) = \sum_i^J p_i(t)$. Then $\dot{q} = \sum_{i=1}^{J} \dot{p}_i = \sum_{i=1}^{J} p_i(r_i - r) = \sum_{i=1}^{J} p_i r_i - r \sum_{i=1}^{J} p_i = r(1-q)$. Therefore, if $q(t^*) = 1$, uniqueness of solutions for the equation for q implies that $q \equiv 1$.

In this paragraph we show $p_i \geq 0$. We now complete the proof by contradiction. Assume that the solution in question leaves Φ. Suppose there exists t_1 and j such that $p_j(t_1) < 0$. Then for some $t_2 \in (t^*, t_1)$, $p_j(t_2) = 0$. However, the set $I := \{(p_1, ..., p_J) : p_j = 0, \ p_i \geq 0 \ \forall i \neq j\}$ is an invariant manifold with respect to system (2.3). Thus $p_j(t_2) = 0$ implies that $p_j(t) = 0$ for $t \geq t^*$ since the solution must remain on the invariant manifold I. This implies $p_j(t_1)$ cannot be negative.

□

The above Proposition also implies that the set Δ is invariant where

$$\Delta := \{(p_1, ..., p_{J-1}) : p_i \geq 0, \ \sum_{i=1}^{J-1} p_i \leq 1\}$$

Finally, we direct our attention to whether solutions to (2.3) converge to an equilibrium.

Proposition 2.4. (The Fundamental Theorem of Natural Selection [7]) *Let* $p = (p_1, p_2, ..., p_J) \in \Delta$ *be a solution to (2.3). Then* $\dot{r} > 0$ *when evaluated on any nonequilibrium solution* p*. If* p *is an equilibrium solution, then* $\dot{r} = 0$*.*

Proof. Take any solution $p := (p_1(t), ..., p_J(t))$ of (2.3). Then

$$\begin{aligned}
\dot{r} &= (\textstyle\sum_{i,j} r_{ij} p_i p_j)' \\
&= \textstyle\sum_{i,j} [r_{ij} \dot{p}_i p_j + r_{ij} p_i \dot{p}_j] \\
&= 2 \textstyle\sum_{i,j} r_{ij} \dot{p}_i p_j \\
&= 2 \textstyle\sum_i \dot{p}_i \left[\textstyle\sum_j^J r_{ij} p_j\right] \\
&= 2 \textstyle\sum_i \dot{p}_i r_i \\
&= 2 \textstyle\sum_i p_i (r_i - r) r_i \\
&= 2 \textstyle\sum_i p_i (r_i - r)(r_i - r + r) \\
&= 2 \textstyle\sum_i p_i (r_i - r)^2 + 2r \textstyle\sum_i p_i (r_i - r) \\
&= 2 \textstyle\sum_i p_i (r_i - r)^2 + 2r \textstyle\sum_i p_i r_i - 2r^2 \textstyle\sum_i p_i \\
&= 2 \textstyle\sum_i p_i (r_i - r)^2 + 2r^2 - 2r^2 \\
&= 2 \textstyle\sum_i p_i (r_i - r)^2 \\
&\geq 0.
\end{aligned}$$

Suppose, p is not an equilibrium. Then there exists some I such that $\dot{p}_I = p_I(r_I - r) \neq 0$. Then

$$\dot{r} = 2 \sum_i^J p_i (r_i - r)^2 \geq 2 p_I (r_I - r)^2 > 0$$

.

Suppose instead p is an equilibrium. Then for all i, $\dot{p}_i = 0$. Thus

$$\dot{r} = 2 \sum_i^J p_i (r_i - r)^2 = 2 \sum_i^J \dot{p}_i (r_i - r) = 0.$$

The proof of the Proposition in complete. □

Proposition 2.5. *All solutions to* (2.3) *converge to an equilibrium.*

Proof. Take any solution to (2.3) on Δ. Since Δ, $\dot{r} \geq 0$ on any trajectory, and r is bounded on Δ; r approaches a limit r^* as time approaches infinity. Thus $\dot{r}$ approaches 0 as time approaches infinity. Thus from the fundamental theorem of natural selection, the trajectory approaches a set of equilibria. There are finitely many equilibria, thus equilibria are isolated. Thus, trajectories approach a single equilibrium.

□

Chapter 3

Two Allele Model

3.1 The Nagylaki and Crow Model

We now assume that there are exactly two alleles (that $J = 2$). This problem has been well studied by previous minds. In this section we derive the differential equation describing the two allele model. We use the equations:

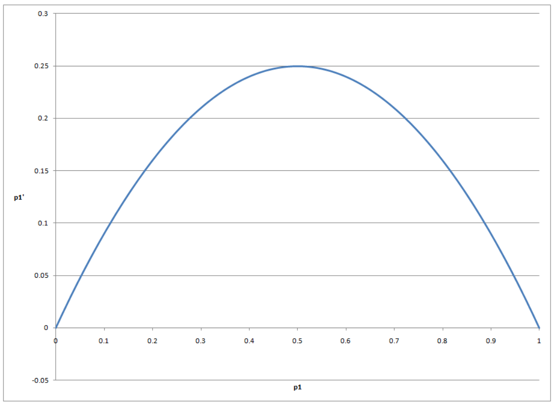
$$\dot{p}_1 = p_1(r_1 - r) = p_1\left(\sum_{j=1}^{2} r_{1j}p_j - \sum_{j,k=1}^{2} r_{j,k}p_jp_k\right), \qquad p_2 = 1 - p_1$$

To solve for the following representation of $\dot{p}_1$:

$$\dot{p}_1 = p_1(1-p_1)((r_{11} - 2r_{12} + r_{22})p_1 + (r_{12} - r_{22})) \tag{3.1}$$

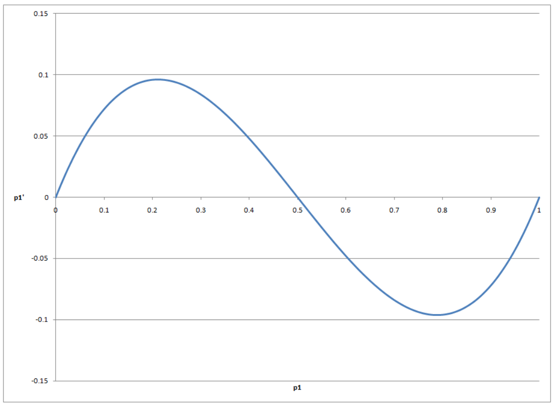
Different values of r_{ij} yield different behavior on the model. Without loss of generality we assume $r_{11} > r_{22}$.

If $r_{11} > r_{12} > r_{22}$, we say r_{12} is in the intermediate case. In the intermediate case, there are exactly two equilibriums on $[0, 1]$: $p_1 = 0$ and $p_1 = 1$. All solutions not starting on these equilibriums approach $p_1 = 1$ as time approaches infinity. This is often called the monostable case. Consider the graph of $(p_1, \dot{p}_1)$ for the intermediate case where $R = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.



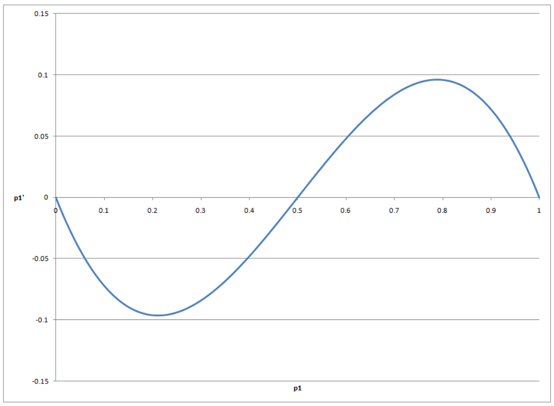
Intermediate Case

If $r_{12} > r_{11} > r_{22}$ we say that r_{12} is in the superior case. In the superior case, there are three equilibrium on $[0,1]$: $p_1 = 0$, $p_1 = 1$, and $p_1 = (r_{22} - r_{12})/(r_{11} - 2r_{12} + r_{22})$. All trajectories not starting on an equilibrium approach the intermediate equilibrium $p_1 = (r_{22} - r_{12})/(r_{11} - 2r_{12} + r_{22})$ as time approaches infinity. Consider the graph of $(p_1, \dot{p_1})$ for the superior case where $R = \begin{bmatrix} 0.001 & 1 \\ 1 & -0.001 \end{bmatrix}$:



Superior Case

If $r_{12} < r_{22} < r_{11}$ we say that r_{12} is in the inferior case. the inferior case, there are three equilibrium on $[0,1]$: $p_1 = 0$, $p_1 = 1$, and $p_1 = (r_{22} - r_{12})/(r_{11} - 2r_{12} + r_{22})$. Trajectories starting on $(0, (r_{22} - r_{12})/(r_{11} - 2r_{12} + r_{22}))$ approach $p_1 = 0$, and trajectories on $((r_{22} - r_{12})/(r_{11} - 2r_{12} + r_{22}), 1)$ approach $p_1 = 1$ as time approaches infinity. This is often called the bistable case. Consider the graph of $(p_1, \dot{p_1})$ for the inferior case where $R = \begin{bmatrix} -0.001 & -1 \\ -1 & 0.001 \end{bmatrix}$:



Inferior Case

Chapter 4

Three-Allele Model

4.1 Introduction

In this section we attempt to classify all the distinct behaviors of system (4.1) in a similar fashion to the two-allele case. Without loss of generality, we may assume the diagonal entries of the fitness matrix R are decreasing, by relabeling; in the three allele case this means $r_{11} < r_{22} < r_{33}$. Assuming Hardy Weinberg equilibrium holds, we consider $\dot{p}_i = p_i(r_i - r)$ for $i = 1, 2, 3$. Our model is

$$\begin{aligned} \dot{p}_1 &= p_1(r_1 - r) \\ \dot{p}_2 &= p_2(r_2 - r) \end{aligned} \tag{4.1}$$

There are three distinct behaviors in the two-allele case, completely determined by the existence and stability of the three possible equilibria. In the three-allele case, determining the existence and location of the stable equilibria is only the second step in determining the dynamics of (4.1). We must first break (4.1) into *cases*, group the cases into *patterns* and finally decompose the patterns based on the *separatricies* present in each pattern. A case is the set of all possible coexisting stable equilibria under the assumption (2.1). Note that the cases distinguish marginally stable from asymptotically stable equilibria. Cases that have the same asymptotically stable equilibria belong to the same pattern. A pattern is the set of all cases that share the same asymptotically stable equilibria. The patterns themselves are important because they describe which alleles may persist and which alleles may disappear in the long run. separatricies distinguish the dynamics present in each pattern, and precisely describe exactly which alleles persist and disappear in the long run. In the two-allele case, the dynamics of (3.1) are completely determined by the patterns (separatricies do not play a role) and not all patterns are possible. For example, in the two-allele case, there is no pattern where the origin is unstable and both the interior equilibrium and $(1, 0)$ are stable. Like the two-allele case, some patterns of the three-allele case are not possible, but unlike the two-allele case the separatricies must be analyzed to determine the dynamics. For example, in the three-allele case, there are only four cases where the equilibria $(0, 0)$, $(0, 1)$ and $(1, 0)$ are mutually asymptotically stable. It is clear from (2.3) that there are no periodic solutions in Δ, and that all solutions with initial conditions in Δ will exist in Δ for all time. Also, it can be shown that the eigenvalues of the Jacobian matrix evaluated at any equilibrium are real. Thus all trajectories in Δ must approach a stable equilibrium. Analyzing the existence and linearized stability of the equilibria is the first major step forward in classifying the dynamics of (4.1).

4.2 Linearized Stability

The Jacobian matrix, $\mathcal{J}$, for (4.1) is

$$\mathcal{J} = \begin{pmatrix} \mathcal{J}_{11} & \mathcal{J}_{12} \\ \mathcal{J}_{21} & \mathcal{J}_{22} \end{pmatrix} \tag{4.2}$$

Using Maple:

$$\begin{aligned}\mathcal{J}_{11} = {} & -\Delta_{23}p_2^2 - 3\Delta_{13}p_1^2 + 4p_1p_2(r_{13} + r_{23} - r_{12} - r_{33}) + 2p_1(\Delta_{13} + r_{33} - r_{13}) \\ & + p_2(r_{12} + 2r_{33} - 2r_{23} - r_{13}) + r_{13} - r_{33}\end{aligned}$$

$$\mathcal{J}_{12} = 2(r_{13} + r_{23} - r_{12} - r_{33})p_1^2 - 2\Delta_{23}p_1p_2 + (2r_{33} - 2r_{23} + r_{12} - r_{13})p_1.$$

$$\mathcal{J}_{21} = 2(r_{13} + r_{23} - r_{12} - r_{33})p_2^2 - 2\Delta_{13}p_1p_2 + (2r_{33} - 2r_{13} + r_{12} - r_{23})p_2$$

$$\begin{aligned}\mathcal{J}_{22} = {} & -3\Delta_{23}p_2^2 - \Delta_{13}p_1^2 + 4p_1p_2(r_{13} + r_{23} - r_{12} - r_{33}) + p_1(r_{12} + 2r_{33} - r_{23} - 2r_{13}) \\ & + 2p_2(\Delta_{23} + r_{33} - r_{23}) + r_{23} - r_{33}\end{aligned}$$

where $\Delta_{ij} = r_{ii} + r_{jj} - 2r_{ij}$ for $i \neq j,\ i, j = 1, 2, 3$. Note that $(0,0), (1,0)$ and $(0,1)$ are equilibria for (4.1). The Jacobian matrices at the three vertices are:

$$\mathcal{J}_{(0,0)} = \begin{pmatrix} r_{13} - r_{33} & 0 \\ 0 & r_{23} - r_{33} \end{pmatrix} \qquad \mathcal{J}_{(1,0)} = \begin{pmatrix} r_{13} - r_{11} & r_{13} - r_{12} \\ 0 & r_{12} - r_{11} \end{pmatrix} \tag{4.3}$$

$$\mathcal{J}_{(0,1)} = \begin{pmatrix} r_{12} - r_{22} & 0 \\ r_{23} - r_{12} & r_{23} - r_{22} \end{pmatrix}$$

It is clear from (4.3) that the eigenvalues corresponding to equilibria $(0,0)$, $(1,0)$ and $(0,1)$ are $\{(r_{13} - r_{33}), (r_{23} - r_{33})\}$, $\{(r_{13} - r_{11}), (r_{12} - r_{11})\}$, $\{(r_{12} - r_{22}), (r_{23} - r_{22})\}$, respectively.

4.3 Boundary Equilibria

The boundary equilibria and interior equilibrium are best expressed in the following notations:

$$\begin{array}{lll} \alpha_1 = r_{21} - r_{11} & \alpha_2 = r_{31} - r_{21} & \alpha_3 = r_{31} - r_{11} = \alpha_1 + \alpha_2 \\ \beta_1 = r_{22} - r_{12} & \beta_2 = r_{32} - r_{22} & \beta_3 = r_{32} - r_{21} = \beta_1 + \beta_2 \\ \gamma_1 = r_{23} - r_{13} & \gamma_2 = r_{33} - r_{23} & \gamma_3 = r_{33} - r_{13} = \gamma_1 + \gamma_2\,. \end{array}$$

Now , we investigate the existence of equilibria on the boundaries $p_i = 0, (i = 1, 2, 3)$ of Δ.

(i) On $p_2 = 0$, (4.1) reduces to

$$\dot{p}_1 = p_1(r_1 - r)$$

where $r_1 = r_{11}p_1 + r_{13}(1 - p_1)$, $r_3 = r_{31}p_1 + r_{33}(1 - p_1)$ and $r = r_1 p_1 + r_3(1 - p_1)$. Interior equilibrium p_1^* satisfies the equation $r_1 = r$ which is same as $r_1 = r_3$. Solving, the boundary equilibrium is

$$(p_1^*, p_2^*) = \left(\frac{r_{33} - r_{13}}{\Delta_{13}}, 0\right) = \left(\frac{\gamma_3}{\Delta_{13}}, 0\right) .$$

Therefore, $p_1^* \in (0, 1)$ if and only if $r_{13} < r_{33}$ or $r_{13} > r_{11}$. Substitution $p_1 = p_1^*, p_2 = 0$ into (4.2), the eigenvalues of $\mathcal{J}_{(p_1^*,0)}$ are:

$$\begin{aligned}
e_{21} &= -\frac{-r_{13}^2 + r_{13}r_{33} + r_{11}r_{13} - r_{11}r_{33}}{r_{11} + r_{33} - 2r_{13}} = \frac{(r_{11} - r_{13})(r_{33} - r_{13})}{r_{11} + r_{33} - 2r_{13}} = \frac{-\alpha_3\gamma_3}{\Delta_{13}} \\
e_{22} &= \frac{r_{13}^2 - r_{12}r_{13} + r_{12}r_{33} - r_{23}r_{13} + r_{23}r_{11} - r_{11}r_{33}}{r_{11} + r_{33} - 2r_{13}} \\
&= \frac{(r_{12} - r_{11})(r_{33} - r_{13}) + (r_{23} - r_{13})(r_{11} - r_{13})}{r_{11} + r_{33} - 2r_{13}} = \frac{\alpha_1\gamma_3 - \gamma_1\alpha_3}{-\alpha_3 + \gamma_3} \equiv \frac{U_2}{\Delta_{13}} .
\end{aligned}$$

Note the eigenvalue e_{21} corresponds to the eigenvector lying along the side $p_2 = 0$. On $\{p_2 = 0\}$ (4.1) reduces to the two-allele case. In general, $e_{i1}, i = 1, 2, 3$ corresponds to the eigenvector lying on invariant submanifold $S_i = \{p_i = 0\}$ of Δ.

(ii) The vertical side $p_1 = 0$. The boundary equilibrium is

$$(p_1^*, p_2^*) = \left(0, \frac{r_{33} - r_{23}}{\Delta_{23}}\right) = \left(0, \frac{\gamma_2}{\Delta_{23}}\right) .$$

Therefore, $p_2^* \in (0, 1)$ if and only if $r_{23} > r_{22}$ or $r_{23} < r_{33}$. The eigenvalues of $\mathcal{J}_{(0,p_2^*)}$ are:

$$\begin{aligned}
e_{11} &= -\frac{(-r_{23}^2 + r_{23}r_{33} + r_{22}r_{23} - r_{33}r_{22})}{r_{22} + r_{33} - 2r_{23}} = \frac{(r_{22} - r_{23})(r_{33} - r_{23})}{r_{22} + r_{33} - 2r_{23}} = \frac{-\beta_2\gamma_2}{\Delta_{23}} \\
e_{12} &= -\frac{(-r_{23}^2 + r_{12}r_{23} - r_{12}r_{33} + r_{13}r_{23} - r_{13}r_{22} + r_{33}r_{22})}{r_{22} + r_{33} - 2r_{23}} \\
&= \frac{(r_{12} - r_{22})(r_{33} - r_{23}) + (r_{13} - r_{23})(r_{22} - r_{23})}{r_{22} + r_{33} - 2r_{23}} = \frac{-\beta_1\gamma_2 + \gamma_1\beta_2}{-\beta_2 + \gamma_2} \equiv \frac{U_1}{\Delta_{23}} .
\end{aligned}$$

(iii) The hypotenuse $p_3 = 0$. The boundary equilibrium is

$$(p_1^*, p_2^*) = \left(\frac{r_{22} - r_{12}}{\Delta_{12}}, \frac{r_{11} - r_{12}}{\Delta_{12}}\right) = \left(\frac{\beta_1}{\Delta_{12}}, \frac{-\alpha_1}{\Delta_{12}}\right) .$$

Therefore, $p_1^*, p_2^* \in (0, 1)$ if and only if $r_{12} > r_{11}$ or $r_{12} < r_{22}$ Eigenvalues of $\mathcal{J}_{(p_1^*,p_2^*)}$ are:

$$
\begin{aligned}
e_{31} &= -\frac{(-r_{11}r_{22}+r_{11}r_{12}+r_{22}r_{12}-r_{12}^2)}{r_{11}+r_{22}-2r_{12}} = \frac{(r_{11}-r_{12})(r_{22}-r_{12})}{r_{11}+r_{22}-2r_{12}} = \frac{-\alpha_1\beta_1}{\Delta_{12}} \\
e_{32} &= -\frac{(r_{11}r_{22}-r_{23}r_{11}+r_{13}r_{12}-r_{13}r_{22}-r_{12}^2+r_{23}r_{12})}{r_{11}+r_{22}-2r_{12}} \\
&= \frac{-(r_{22}-r_{23})(r_{11}-r_{12})+(r_{13}-r_{12})(r_{22}-r_{12})}{r_{11}+r_{22}-2r_{12}} = \frac{-\beta_2\alpha_1+\alpha_2\beta_1}{-\alpha_1+\beta_1} \equiv \frac{U_3}{\Delta_{12}}\,.
\end{aligned}
$$

In summary, the eigenvalues of $\mathcal{J}$ at $P_i = 0$ are $\{e_{i1}, e_{i2}\}$. The eigenvalue e_{i1} corresponds to the eigenvector lying on the invariant submanifold $p_i = 0$. The dynamics on the invariant submanifolds $p_i = 0, i = 1, 2, 3$ are identical to the those described in (3.1). If the side $p_i = 0$ is in the inferior case, the unique boundary equilibrium, P_i, exists and is unstable on that side; if the side is in the superior case, then the boundary equilibrium also exists, but is stable on that side. The notation introduced in this section will be used throughout the rest of this paper for convenience.

4.4 Interior Equilibrium

An interior equilibrium satisfies $r_1 - r_2 = 0$, $r_2 - r_3 = 0$. This system can be written as

$$
A\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} \gamma_1-\alpha_1 & \gamma_1-\beta_1 \\ \gamma_2-\alpha_2 & \gamma_2-\beta_2 \end{bmatrix}\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}.
$$

$$
\det(A) = (\gamma_1-\alpha_1)(\gamma_2-\beta_2) - (\gamma_1-\beta_1)(\gamma_2-\alpha_2) = -(U_1+U_2+U_3) \neq 0.
$$

The unique solution to this system of equation is

$$
\begin{bmatrix} p_1^* \\ p_2^* \end{bmatrix} = \frac{1}{\det(A)}\begin{bmatrix} \gamma_1(\gamma_2-\beta_2)-\gamma_2(\gamma_1-\beta_1) \\ \gamma_2(\gamma_1-\alpha_1)-\gamma_1(\gamma_2-\alpha_2) \end{bmatrix} = \frac{1}{U_1+U_2+U_3}\begin{bmatrix} U_1 \\ U_2 \end{bmatrix}. \tag{4.4}
$$

Lemma 4.4.1 Necessary and sufficient conditions for the existence of an interior equilibrium are that $U_i, i = 1, 2, 3$ have the same sign where

$$
\begin{aligned}
U_1 &= \gamma_1\beta_2-\gamma_2\beta_1 = \gamma_1\beta_3-\gamma_3\beta_1 && (4.5)\\
U_2 &= \alpha_1\gamma_2-\alpha_2\gamma_1 = \alpha_1\gamma_3-\alpha_3\gamma_1 && (4.6)\\
U_3 &= \beta_1\alpha_2-\beta_2\alpha_1 = \beta_1\alpha_3-\beta_3\alpha_1 && (4.7)
\end{aligned}
$$

Proof. Note that $p_3^* = \frac{U_3}{U_1+U_2+U_3}$, thus if $U_1+U_2+U_3 > 0 \implies U_3 > 0$ and (4.4) $\implies U_1, U_2 > 0$. Following exactly the same reasoning, if $\det(A) < 0 \implies U_i < 0$ for $i = 1, 2, 3$. □

At an interior equilibrium, the Jacobian matrix of the vector field $(p_1(r_1 - r), p_2(r_2 - r))$ is

$$
\mathcal{J} = \begin{pmatrix} p_1(r_1-r)_{p_1} & p_1(r_1-r)_{p_2} \\ p_2(r_2-r)_{p_1} & p_2(r_2-r)_{p_2} \end{pmatrix}
$$

evaluated at an interior equilibrium, (p_1^*, p_2^*), $r_1 = r_2 = r_3$. Also, recall $p_3 = 1 - p_1 - p_2$. Therefore,

$$r_{p_1} = (r_1)_{p_1}p_1 + (r_2)_{p_1}p_2 + (r_3)_{p_1}p_3, \qquad r_{p_2} = (r_1)_{p_2}p_1 + (r_2)_{p_2}p_2 + (r_3)_{p_2}p_3\,.$$

From the above definitions, we have

$$(r_1)_{p_1} = -\alpha_3,\ (r_1)_{p_2} = -\alpha_2,\ (r_2)_{p_1} = -\beta_3,\ (r_2)_{p_2} = -\beta_2,\ (r_3)_{p_1} = -\gamma_3,\ (r_3)_{p_2} = -\gamma_2$$

and

$$\mathcal{J} = \begin{pmatrix} p_1^*[(\gamma_3-\alpha_3)(1-p_1^*) - (\gamma_3-\beta_3)p_2^*] & p_1^*[(\gamma_2-\alpha_2)(1-p_1^*) - (\gamma_2-\beta_2)p_2^*] \\ p_2^*[(\gamma_3-\beta_3)(1-p_2^*) - (\gamma_3-\alpha_3)p_1^*] & p_2^*[(\gamma_2-\beta_2)(1-p_2^*) - (\gamma_2-\alpha_2)p_1^*] \end{pmatrix}.$$

Therefore,

$$\begin{aligned} \det(\mathcal{J}) &= -p_1^*p_2^*p_3^*(\alpha_3\gamma_2 - \alpha_3\beta_2 + \gamma_3\beta_2 - \alpha_2\gamma_3 + \alpha_2\beta_3 - \gamma_2\beta_3) \\ &= -p_1^*p_2^*p_3^*(\alpha_1\gamma_2 - \alpha_1\beta_2 + \gamma_1\beta_2 - \alpha_2\gamma_1 + \alpha_2\beta_1 - \gamma_2\beta_1) \\ &= \ \ p_1^*p_2^*p_3^* \det(A) = -p_1^*p_2^*p_3^*(U_1 + U_2 + U_3) \end{aligned} \tag{4.8}$$

and

$$\operatorname{tr}(\mathcal{J}) = -\Delta_{13}(p_1^*)^2 - \Delta_{23}(p_2^*)^2 + (\Delta_{12} - \Delta_{23} - \Delta_{13})p_1^*p_2^* + \Delta_{13}p_1^* + \Delta_{23}p_2^*\,. \tag{4.9}$$

Under our assumptions if the interior equilibrium exists, it is unique. It is not necessary to obtain the eigenvalues explicitly, rather the determinant and trace of $\mathcal{J}$ carry all the necessary information to determine stability. If the inequalities in Lemma 4.4.1 are reversed so that $\det(A) < 0$, the interior equilibrium exists and is unstable. If the inequalities in Lemma 4.4.1 are positive, then the interior equilibrium is either stable or totally unstable depending on the sign of $\operatorname{tr}(J)$ given above.

4.5 Analysis of System (4.1)

In this section we find all possible cases of (4.1). Recall, Δ, defined in Section 2.2, is the triangle joining the vertices $(1,0),(0,1)$ and $(0,0)$. Also, recall we denote the three sides of Δ by $S_i = \{p_i = 0\}$ $(i = 1,2,3)$, and the boundary equilibrium on each of these sides by P_i $(i = 1,2,3)$, if it exists. The three possible types of equilibria are corner equilibria, boundary equilibria and interior equilibria. We do not search for patterns case by case as we did with 2 alleles, rather we implement a numerical simulation to find and test all possible cases, by randomly ordering the entries of R assuming (2.1), and checking the eigenvalues and existence of each of the seven possible equilibria. We conjectured all possible cases given in the following lemmas from a large sample $(100,000)$ of these simulations. The proofs that followed from these conjectures are found in Appendix B as they are technical and long.

Lemma 4.1. *Suppose none of the P_i's exists. Then the interior equilibrium does not exist.*

Lemma 4.2. *Suppose only one P_i exists. Then there are 8 cases:*

1. *If P_2 exists but not P_1 and P_3, then there are two subcases: (i) S_2 is in the inferior case, P_2 has two positive eigenvalues and interior equilibrium does not exi8st, (ii) S_2 is in the superior case, P_2 has two negative eigenvalues and interior equilibrium does not exist.*

2. *If P_1 exists but not P_2 and P_3, then there are three subcases: (i) S_1 is in the inferior case, P_1 has two positive eigenvalues and interior equilibrium does not exist, (ii) S_1 is in the superior case, P_1 has two negative eigenvalues and interior equilibrium exists with one positive and one negative eigenvalues, (iii) S_1 is in the superior case, P_1 has one negative and one positive eigenvalues and interior equilibrium does not exist.*

3. *If P_3 exists but not P_1 and P_2, then there are three subcases: (i) S_3 is in the superior case, P_3 has two negative eigenvalues and interior equilibrium does not exist, (ii) S_3 is in the inferior case, P_3 has two positive eigenvalues and interior equilibrium exists with one positive and one negative eigenvalue, (iii) S_3 is in the inferior case, P_3 has one positive and one negative eigenvalues and interior equilibrium does not exist.*

Lemma 4.3. *Suppose only two P_i's exist. Then there are 34 cases:*

Case A. Suppose P_2, P_3 exist but not P_1. Then there are 11 subcases:

1. *S_2, S_3 are in the superior case.*
   - *(a) P_2 or P_3 is stable with two negative eigenvalues and the other is unstable with one positive and one negative eigenvalues. Interior equilibrium does not exist.*
   - *(b) Both P_2 and P_3 are unstable with one positive and one negative eigenvalues. Interior equilibrium exists and is stable.*
   - *(c) Both P_2 and P_3 are stable with two negative eigenvalues. Interior equilibrium exists with one positive and one negative eigenvalue.*

2. *S_2, S_3 are in the inferior case.*
   - *(a) Both P_2 and P_3 are unstable, one has two positive eigenvalues and the other has one positive and one negative eigenvalues. Interior equilibrium does not exist.*
   - *(b) Both P_2 and P_3 are unstable with two positive eigenvalues. Interior equilibrium exists with one positive and one negative eigenvalues.*
   - *(c) Both P_2 and P_3 are unstable with one positive and one negative eigenvalues. Interior equilibrium exists with two positive eigenvalues.*

3. *S_2 is in the inferior case and S_3 is in the superior case. P_2 is unstable with two positive eigenvalues and P_3 is stable with two negative eigenvalues. Interior equilibrium does not exist.*

4. *S_2 is in the superior case and S_3 is in the inferior case.*
    (a) *P_2 is stable with two negative eigenvalues and P_3 is unstable one positive and one negative eigenvalues. Interior equilibrium does not exist.*

    (b) *P_2 is stable with two negative eigenvalues, P_3 is unstable with two positive eigenvalues. Interior equilibrium exists with one positive and one negative eigenvalue.*

Case B. Suppose P_1, P_3 exist but not P_2. Then there are 12 subcases. The results are identical to Case A except P_2 is changed to P_1 and P_1 is changed to P_2 in the statements of Case A. In addition, there is an extra case in part 4 which we call 4(c): P_1 is unstable with one positive and one negative eigenvalues, P_3 is unstable with two positive eigenvalues. Interior equilibrium does not exist.

Case C. Suppose P_1, P_2 exist but not P_3. Then there are 11 cases. The results are identical to Case A except P_1 is changed to P_3, P_2 is changed to P_1, and P_3 is changed to P_2 in the statements of Case A. In addition, statement 4(a) in Case A is changed to the following: P_1 is unstable with one positive and one negative eigenvalues, P_2 is unstable with two positive eigenvalues. Interior equilibrium does not exist.

Lemma 4.4. *Suppose P_1, P_2, P_3 all exist. Then there are 26 cases:*

1. *S_1, S_2, S_3 are all in the inferior case. There are 4 subcases here.*
    (a) *P_1, P_2, P_3 are all unstable and each has one positive and one negative eigenvalues. Interior equilibrium exists and has two positive eigenvalues.*

    (b) *P_1, P_2, P_3 are all unstable. One of the P_i's has two positive eigenvalues and the rest have one positive and one negative eigenvalues. Interior equilibrium does not exist.*

2. *S_i is in the superior case and the other two sides are in the inferior case. There are 9 subcases here.*
    (a) *P_i is stable with two negative eigenvalues and the other two equilibria are unstable with two positive eigenvalues. Interior equilibrium exists with one positive and one negative eigenvalues.*

    (b) *P_i is stable with two negative eigenvalues and the other two equilibria are unstable, one has two positive eigenvalues where the dynamics on each of the three boundaries is identical to the two allele case; while the other has one positive and one negative eigenvalues. Interior equilibrium does not exist.*

3. *S_i is in the inferior case and the other two sides are in the superior case. There are 9 subcases here.*
    (a) *P_i is unstable with two positive eigenvalues and the other two equilibria are stable with two negative eigenvalues. Interior equilibrium exists with one positive and one negative eigenvalues.*

(b) *P_i is unstable with two positive eigenvalues, one of the other equilibria is unstable with one positive and one negative eigenvalues, and the third equilibrium is stable with two negative eigenvalues. Interior equilibrium does not exist.*

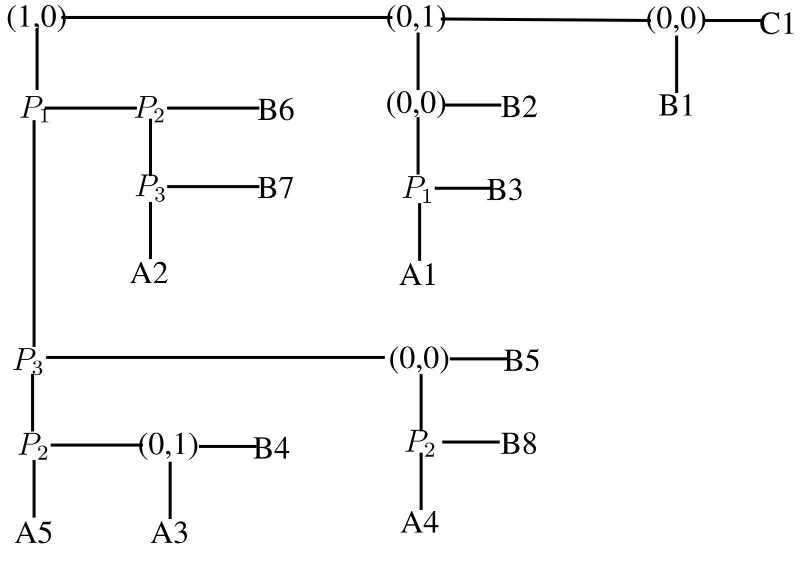
4. *All three edges are in the superior case. There are* 4 *subcases here.*

    (a) *P_1, P_2, P_3 are all unstable with one positive and one negative eigenvalues. Interior equilibrium exists with two negative eigenvalues.*

    (b) *One of the boundary equilibria is stable with two negative eigenvalues and the other two are unstable with one positive and one negative eigenvalues. Interior equilibrium does not exist.*

From the above Lemmas, we discovered that there are 69 cases which we can group into 14 patterns. These 69 cases and patterns are shown in Section 6.2, the tables in Appendix B.The 5 patterns with one stable equilibrium are denoted by A1-A5 and correspond to the five patterns in the first table. The 8 patterns with two stable equilibria are denoted by B1-B8 and correspond to the 8 patterns listed in the second table. The pattern with three stable equilibria is denoted by C1 which is table three. We summarize the 14 patterns for the three-allele case in the simple flowchart below. Start at the top left corner $(1, 0)$, go right if $(1, 0)$ is stable and go down if it is not. Repeat this process for each equilibrium, and you will arrive at one of 14 possible cases.

14 Patterns for Three-alleles



4.6 Dynamics of Solutions to System (4.1)

The 69 cases describe the existence and stability of the seven possible equilibria. The 69 cases can be broken down into 14 distinct patterns, uniquely determined by which equilibria exist and whether there are stable. From the tables in Appendix B, there are five patterns with one stable equilibrium, eight patterns with two stable equilibria and one pattern with three stable equilibria. The patterns yield some information on the long term genetic make up of the population, but to fully understand the dynamics of how the allele frequencies change with time, it is important to determine the locations of the separatracies for the 9 patterns with two or more stable equilibria. There are no separatricies for the case of one stable equilibrium; all solutions go to the stable equilibrium. The separatricies yield more information than just the pattern. A separatrix partitions Δ into invariant subregions, each with distinct asymptotic behavior. The separatricies are best analyzed by drawing out Δ for each case, and finding the intersection of the stable and unstable manifolds, which is the location of the separatrix. Decomposing each pattern into separatricies yields all the information about the possible dynamics of the three-allele case. The diagrams in (6.3), Appendix C, show the decomposition of each pattern into the separatricies present in that pattern.

From (6.3), the triangles in Appendix C, the following conclusions are apparent. The first set of conclusions describe which equilibria are stable. From the triangles it is clear that no two stable equilibria are adjacent. If there is one stable equilibrium, that equilibrium can be (1,0), any side equilibrium, or the interior equilibrium. If there are two stable equilibria, we have several cases. If both of the stable equilibria are corner equilibria then the stable equilibria are $(1, 0)$ and one other corner equilibrium. If the equilibria are composed of one corner equilibrium and one side equilibrium, then for any corner one has the pairing between the corner $p_i = 1$ and the side $p_i = 0$. If the equilibria are composed of two side equilibria, then any pairing is possible. These are the only possibilities for two stable equilibria. Note that here the interior equilibrium is never stable. If three equilibria are stable, then necessarily these equilibria are the corner equilibria.

The second set of conclusions tell of the separatrices. If there is one stable equilibrium, there is no separatrix. If there are two stable equilibria, then there is one separatix. The form of this separatrix can be found in the following way. Let e_1 and e_2 be the two stable equilibrium in question. Necessarily, these equilibria are on the boundary of Δ. Consider all of the equilibria between e_1 and e_2 clockwise. If there is a side equilibrium here, this is an endpoint of the separatrix. Otherwise, there is only one equilibrium to consider. In this case, this equilibrium is the end point of the separatrix. Consider all of the equilibria between e_1 and e_2 counter-clockwise in the same way, and get the second endpoint. If there is an internal equilibrium, the separatrix crosses the internal equilibrium. With this information, you have found the form of the separatrix. If there are three stable equilibria, there are either two or three separtrices. If there is an interior equilibria, there are three separatrices. Each of these separatrices connect the interior equilibrium with a side equilibrium. If there is no interior equilibrium, there are two separatrices. The side equilibrium with the positive eigenvalue corresponding to the internal eigenvector is the end point of two separtrices. The other end point of these separatrices are the other two side equilibria.

Chapter 5

Conclusion

We have succeeded in classifying our three allele selection model assuming (2.1). In [1], Bürger presents the following open problem: How many distinct patterns exist for k alleles? Counting patterns for the two-allele case is just an elementary exercise. Counting patterns for three-alleles, is rather involved, but can be accomplished in its entirety with the aid of simple numerical simulation and direct proof techniques. The three-allele case is special because a complete classification of of the dynamics is attainable by exhaustion. This is not the case for $k \geq 4$. For k alleles, the number of distinct orderings of the fitness coefficients r_{ij}, is

$$\mathrm{Or}(k) = \prod_{n=1}^{\binom{k}{2}} (n+k) \tag{5.1}$$

also it can be shown that for k alleles, there is at most one equilibrium on each of the $\binom{k}{1} + \binom{k}{2} + \binom{k}{3} + ... + \binom{k}{k}$ lower dimensional submanifolds, which means the maximum number of equilibira for k alleles is $2^k - 1$. We can use $\mathrm{Or}(k)$ the estimate the number of iterations necessary test all possible orderings of the r_{ij}'s for k alleles. Evaluating $\mathrm{Or}(4) = 151,200$ and $\mathrm{Or}(5) \approx 10,000,000,000$ makes it clear that $k = 4$ is the largest k amendable to practical numerical analysis. For $k \geq 5$ there is always the chance that we missed a case. Even for three-alleles there were a few rare cases that appeared, only a few times, after we doubled the standard amount of iterations. We have attempted to calculate the number of patterns, in Matlab, for $4 \leq k \leq 6$ by randomly ordering the entries of fitness matrix R, testing the signs and existence of the equilibria and finally we output this data to a text file we can analyze for patterns. Based on our simulations the following table shows the number of patterns, $P(k)$, for $k = 1, 2, ..., 6$ alleles:

k	$P(k)$
2	3
3	14
4	105
5	$\geq 1,500$
6	$\geq 20,000$

For $k \geq 5$ there are too many cases to find patterns using methods of exhaustion. Therefore we desire a way to cut down the number of cases so numerical simulation is feasible. Or even better, if possible, we wish to find an explicit form of $P(k)$. Aside from counting the patterns for k alleles, we also desire to explain interesting and unanticipated results. For example, for $k = 5$ they are 10 cases with 6 coexisting asymptotically stable equilibria. This is a surprise because we had found no cases for $k \leq 4$ where the number of coexisting stable equilibria exceeds the number of alleles. The final accomplishment, for k alleles, would be to count the number of distinct behaviors for each pattern. The distinct behaviors will be determined by how Δ is partitioned into $s = 1...k - 1$ dimensional submanifold separatricies.

Chapter 6

Appendix

6.1 Appendix A: Proofs of Lemmas 4.1 - 4.4

We only present proofs of selected five cases of the four lemmas since the rest are similar. In the following, we define $\bar{\alpha} = \alpha_1/\alpha_2$, $\bar{\beta} = \beta_1/\beta_2$, and $\bar{\gamma} = \gamma_1/\gamma_2$.

Proof of Lemma 4.1

Proof. From (4.4), interior equilibrium exists if and only if $U_i, i = 1, 2, 3$ are of the same sign. Suppose $U_i < 0$, $i = 1, 2, 3$. Then from (4.5), (4.6), and (4.7), we have $\bar{\beta} > 0, \bar{\gamma}/\bar{\beta} < 1$, $\bar{\beta}/\bar{\alpha} < 1$ and $\bar{\gamma}/\bar{\alpha} > 1$. If $\bar{\alpha} > 0$, then the last three inequalities contradict each other. If $\bar{\alpha} < 0$, then $\bar{\gamma} < 0$ which imply that $\gamma_1 > 0$ and $\alpha_2 > 0$. Thus $\gamma_1 + \alpha_2 = r_{23} - r_{12} > 0$. But this contradicts $\beta_1 + \beta_2 = r_{32} - r_{12} < 0$. The argument for the case $U_i > 0, i = 1, 2, 3$ is similar. The proof of the lemma is complete. ☐

Proof of Lemma 4.2 Part 1

Proof. The proof the non-existence of interior equilibrium is the same as in the proof of Lemma 4.1. To prove the rest of the lemma, since Δ_{13} is negative (positive) in the superior (inferior) case, from the middle term in (**??**), $-\gamma_3\alpha_3 > 0$, and it suffices to show that $U_2 = -\gamma_1\alpha_2 + \alpha_1\gamma_2 > 0$. Since $\alpha_1 < 0$ and $\gamma_2 < 0$, it suffices to show that $-\gamma_1\alpha_2 = (r_{23} - r_{13})(r_{21} - r_{31}) > 0$. Since S_1, S_3 are in the intermediate case, from (**??**), we have $r_{23} < r_{22} < r_{12}$ so that if $-\gamma_1\alpha_2 < 0$, we must have $\gamma_1 = r_{23} - r_{13} < 0$ and $\alpha_2 = r_{31} - r_{21} < 0$. These imply that $\gamma_3 = \gamma_1 + \gamma_2 < 0$ and $\alpha_3 = \alpha_1 + \alpha_2 < 0$, contradicting the above fact that $-\gamma_3\alpha_3 > 0$. The proof of the lemma is complete. ☐

Proof of Lemma 4.3 Cases B and C, Part 4

Proof. (Case B, Part 4) Suppose S_1 is in the superior case S_2 is in the intermediate case, and S_3 is in the inferior case. Then the signs of $\alpha_i, \beta_i, \gamma_i$'s are:

$$\alpha_1\,(-),\quad \alpha_2\,(?),\quad \alpha_3\,(-),\quad \beta_1\,(+),\quad \beta_2\,(+),\quad \beta_3\,(+),\quad \gamma_1\,(?),\quad \gamma_2,\,(-),\quad \gamma_3\,(-)$$

The signs of α_2 and γ_1 cannot be determined. Also, $\Delta_{23} < 0$, $\Delta_{12} > 0$ and the signs of the eigenvalues of the boundary equilibria are:

$$P_1 \left(-, \frac{U_1}{-}\right), \qquad P_3 \left(+, \frac{U_3}{+}\right)$$

From (4.5), (4.6), (4.7) and above, we have

$$\begin{aligned} U_1 &= \gamma_1\beta_2 - \gamma_2\beta_1 = \gamma_1\,(+) - (-)(+) \\ U_2 &= \alpha_1\gamma_2 - \alpha_2\gamma_1 = (-)(-) - \alpha_2\gamma_1 \\ U_3 &= \beta_1\alpha_2 - \beta_2\alpha_1 = (+)\,\alpha_2 - (+)(-) \end{aligned}$$

Suppose $\gamma_1 > 0$, then $U_2 > 0$ and $U_1 > 0$. If $U_3 < 0$, then the eigenvalues of P_1, P_3 are $P_1\ (-,-)$ and $P_3\ (+,-)$ and there is no interior equilibrium. If $U_3 > 0$, then the eigenvalues of P_1, P_3 are $P_1\ (-,-)$, $P_3\ (+,+)$ and interior equilibrium exists. From (**??**), the interior equilibrium is unstable with one positive and one negative eigenvalues. Now suppose $\gamma_1 < 0$, then if $\alpha_2 > 0$, we have $U_2 > 0$, $U_3 > 0$. If $U_1 > 0$, then the eigenvalues of the boundary equilibria are $P_1\ (-,-)$, $P_3\ (+,+)$ and interior equilibrium exists with one positive and one negative eigenvalues. This generates the same pattern as the case $\gamma_1 > 0$ and $U_i > 0, i = 1,2,3$ above. If $U_1 < 0$, then $P_1\ (-,+)$, $P_3\ (+,+)$ and there is no interior equilibrium. The case $\gamma_1 < 0$ and $\alpha_2 < 0$ is impossible since it would lead $r_{23} < r_{13} < r_{12}$ which contradicts our assumptions that S_1 is in the superior case ($r_{33} < r_{22} < r_{23}$) and S_3 is in the inferior case ($r_{12} < r_{22} < r_{11}$)) which imply that $r_{12} < r_{22} < r_{23}$. The proof of Case B Part 4 of Lemma 4.3 is complete.

We shall be brief here. P_3 does not exist and the signs of the eigenvalues of P_1 and P_2 are $P_1\ (-, \frac{U_1}{-})$ and $P_2\ (+, \frac{U_2}{+})$. Also, $\Delta_{23} < 0$ and $\Delta_{13} > 0$. The signs of $\alpha_i, \beta_i, \gamma_1$'s are:

$$\alpha_1\,(-), \quad \alpha_2\,(-), \quad \alpha_3\,(-), \quad \beta_1\,(-), \quad \beta_2\,(+), \quad \beta_3\,(?), \quad \gamma_1\,(+), \quad \gamma_2\,(-), \quad \gamma_3\,(+)$$

From above formulas for U_i's, we have $U_2 > 0$, $U_3 > 0$ so that $P_2\ (+,+)$. Suppose $\beta_3 > 0$, then $U_1, > 0, U_2 > 0, U_3 > 0$ so that $P_1\ (-,-)$, $P_2\,(+,+)$ and interior equilibrium exists with one positive and one negative eigenvalues. Suppose $\beta_3 < 0$, then if $U_1 > 0$, we have $U_2 > 0, U_3 > 0$ and $P_1\ (-,-)$, $P_2\ (+,+)$. Interior equilibrium exists with one positive and one negative eigenvalues. The is the same pattern as before. If $U_1 < 0$, then $P_1\ (-,+), P_2(+,+)$ and interior equilibrium does not exist. The proof of Case C Part 4 of Lemma 4.3 is complete. □

Proof of Lemma 4.4

Proof. We only prove Parts 1 and 2. The proofs for the rest of the lemma are similar. Suppose all three sides are in the inferior case. Then the signs of $\alpha_i, \beta_i, \gamma_i$ are:

$$\alpha_1\,(-), \quad \alpha_2\,(?), \quad \alpha_3\,(-), \quad \beta_1\,(+), \quad \beta_2\,(-), \quad \beta_3\,(?), \quad \gamma_1\,(?), \quad \gamma_2\,(+), \quad \gamma_3\,(+).$$

The signs of $\alpha_2 = r_{31} - r_{21}, \beta_3 = r_{32} - r_{21}, \gamma_1 = r_{23} - r_{13}$ are indeterminate because the relations between r_{12}, r_{13}, r_{23} are unknown. From out assumptions, $\Delta_{12}\ (+), \Delta_{13}\ (+), \Delta_{23}\ (+)$ and the signs of the eigenvalues of $P_i's$ are

$$P_1 \left(+, \frac{U_1}{+}\right), \qquad P_2 \left(+, \frac{U_2}{+}\right), \qquad P_3 \left(+, \frac{U_3}{+}\right)$$

We shall use (4.5), (4.6), (4.7) to find the signs of the $U_i's$ under all possible relations between r_{12}, r_{13}, r_{23}.

(1.1) $r_{12} < r_{13} < r_{23}$. We have $\alpha_2\ (+), \beta_3\ (+), \gamma_1\ (+)$ and $U_1\ (-), U_2\ (-)$. If $U_3 < 0$, then (a) happens and if $U_3 > 0$, then (b) happens in which P_3 has two positive eigenvalues and P_1 and P_2 have one positive and one negative eigenvalues.

(1.2) $r_{13} < r_{12} < r_{23}$. We have $\alpha_2\ (-), \beta_3\ (+), \gamma_1\ (+)$ and $U_1\ (-), U_3\ (-)$. If $U_2 < 0$, then (a) happens and if $U_2 > 0$, then (b) happens in which P_2 has two positive eigenvalues and P_1 and P_3 have one positive and one negative eigenvalues.

(1.3) $r_{13} < r_{23} < r_{12}$. We have $\alpha_2\ (-), \beta_3\ (-), \gamma_1\ (+)$ and $U_1\ (-), U_3\ (-)$. If $U_2 < 0$, then (a) happens and if $U_2 > 0$, then (b) happens in which P_2 has two positive eigenvalues and P_1 and P_3 have one positive and one negative eigenvalues.

(1.4) $r_{23} < r_{13} < r_{12}$. We have $\alpha_2\ (-), \beta_3\ (-), \gamma_1\ (-)$ and $U_2\ (-), U_3\ (-)$. If $U_1 < 0$, then (a) happens and if $U_1 > 0$, then (b) happens in which P_1 has two positive eigenvalues and P_2 and P_3 have one positive and one negative eigenvalues.

(1.5) $r_{12} < r_{23} < r_{13}$. We have $\alpha_2\ (+), \beta_3\ (+), \gamma_1\ (-)$ and $U_1\ (-), U_2\ (-)$. If $U_3 < 0$, then (a) happens and if $U_3 > 0$, then (b) happens in which P_3 has two positive eigenvalues and P_1 and P_1 have one positive and one negative eigenvalues.

(1.6) $r_{23} < r_{12} < r_{13}$. We have $\alpha_2\ (+), \beta_3\ (-), \gamma_1\ (-)$ and $U_2\ (-), U_3\ (-)$. If $U_1 < 0$, then (a) happens and if $U_1 > 0$, then (b) happens in which P_1 has two positive eigenvalues and P_2 and P_3 have one positive and one negative eigenvalues.

Thus, the proof of Part 1 is complete. To prove Part 2, suppose $i = 1$, S_1 is in the superior case and S_2, S_3 are in the inferior case. The signs of $\alpha_i, \beta_i, \gamma_i$ are:

$$\alpha_1\ (-), \quad \alpha_2\ (?), \quad \alpha_3\ (-), \quad \beta_1\ (+), \quad \beta_2\ (+), \quad \beta_3\ (+), \quad \gamma_1\ (+), \quad \gamma_2\ (-), \quad \gamma_3\ (+)$$

We also have $\Delta_{12}\ (+), \Delta_{13}\ (+), \Delta_{23}\ (-)$. Note that $U_1 = \gamma_1\beta_2 - \gamma_2\beta_1 > 0$ and the signs of the eigenvalues of $P_i's$ are

$$P_1\ (-,-), \qquad P_2\ \left(+, \frac{U_2}{+}\right), \qquad P_3\ \left(+, \frac{U_3}{+}\right)$$

Suppose $\alpha_2 > 0$, then $U_3 > 0$ and $P_3\ (+,+)$. If $U_2 > 0$, then $P_2\ (+,+)$ and (a) happens. If $U_2 < 0$, then $P_2\ (+,-)$ and (b) happens. Suppose $\alpha_2 < 0$, then $U_2 > 0$ and $P_2\ (+,+)$. If $U_3 > 0$, then $P_3\ (+,+)$ and (a) happens and if $U_3 < 0$, then $P_3\ (+,-)$ and (b) happens. The proof of Part 2 is complete. □

6.2 Appendix B: Tables for Three-Allele Case

Table 6.1: Thirty cases 5 patterns with one stable equilibrium

Lemma	$(0,0)$	$(0,1)$	$(1,0)$	P_1	P_2	P_3	Interior
4.1	+ +	− +	− −	x x	x x	x x	x x
$4.2.1(i)$	− +	− +	− −	x x	+ +	x x	x x
$4.2.3(i)$	+ −	− +	− −	+ +	x x	x x	x x
$4.2.3(iii)$	+ +	+ +	− −	− +	x x	x x	x x
$4.3.B.4(c)$	+ +	+ −	− −	− +	x x	+ +	x x
$4.3.C.4(a)$	− +	+ +	− −	− +	+ +	x x	x x
$4.3.B.1(a)$	+ +	+ +	+ −	− −	x x	− +	x x
$4.3.C.1(a)$	+ +	+ +	− +	− −	− +	x x	x x
$4.4.3(b)$	+ +	+ −	− +	− −	− +	+ +	x x
$4.4.3(b)$	− +	+ +	+ −	− −	+ +	− +	x x
$4.4.4(b)$	+ +	+ +	+ +	− −	− +	− +	x x
$4.3.C.1(a)$	+ +	+ +	− +	− +	− −	x x	x x
$4.3.C.3$	+ −	− +	− +	+ +	− −	x x	x x
$4.4.3(b)$	+ +	+ −	− +	− +	− −	+ +	x x
$4.4.3(b)$	+ −	− +	+ +	+ +	− −	− +	x x
$4.4.4(b)$	+ +	+ +	+ +	− +	− −	− +	x x
$4.3.A.1(a)$	+ +	− +	+ +	x x	− −	− +	x x
$4.2.1(ii)$	+ +	− +	− +	x x	− −	x x	x x
$4.2.2(i)$	+ +	− +	+ −	x x	x x	− −	x x
$4.3.A.1(a)$	+ +	− +	+ +	x x	− +	− −	x x
$4.3.A.3$	− +	− +	+ −	x x	+ +	− −	x x
$4.3.B.1(a)$	+ +	+ +	+ −	− +	x x	− −	x x
$4.3.B.3$	+ −	− +	+ −	+ +	x x	− −	x x
$4.4.3(b)$	− +	+ +	+ −	− +	+ +	− −	x x
$4.4.3(b)$	+ −	− +	+ +	+ +	− +	− −	x x
$4.4.4(b)$	+ +	+ +	+ +	− +	− +	− −	x x
$4.3.A.1(b)$	+ +	− +	+ +	x x	− +	− +	− −
$4.3.B.1(b)$	+ +	+ +	+ −	− +	x x	− +	− −
$4.3.C.1(b)$	+ +	+ +	− +	− +	− +	x x	− −
$4.4.4(a)$	+ +	+ +	+ +	− +	− +	− +	− −

Table 6.2: Thirty-five cases 8 patterns with two stable equilibria

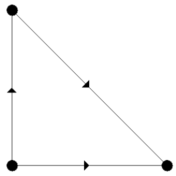
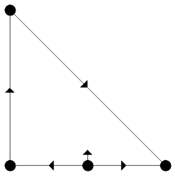
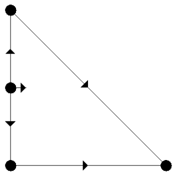
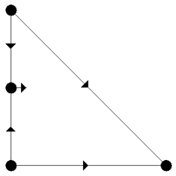
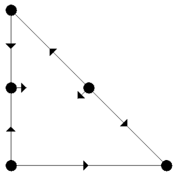
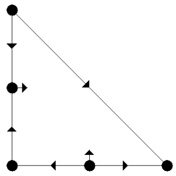
Lemma	$(0,0)$	$(0,1)$	$(1,0)$	P_1	P_2	P_3	Interior
$4.2.2(ii)$	$+ \quad +$	$- \quad -$	$- \quad -$	x x	x x	$+ \quad +$	$+ \quad -$
$4.2.2(iii)$	$+ \quad +$	$- \quad -$	$- \quad -$	x x	x x	$+ \quad -$	x x
$4.3.B.2(a)$	$+ \quad -$	$- \quad -$	$- \quad -$	$+ \quad -$	x x	$+ \quad +$	x x
$4.3.B.2(a)$	$+ \quad -$	$- \quad -$	$- \quad -$	$+ \quad +$	x x	$+ \quad -$	x x
$4.3.B.2(b)$	$+ \quad -$	$- \quad -$	$- \quad -$	$+ \quad +$	x x	$+ \quad +$	$+ \quad -$
$4.3.B.2(c)$	$+ \quad -$	$- \quad -$	$- \quad -$	$+ \quad -$	x x	$+ \quad -$	$+ \quad +$
$4.3.A.2(a)$	$- \quad +$	$- \quad -$	$- \quad -$	x x	$+ \quad -$	$+ \quad +$	x x
$4.3.A.2(a)$	$- \quad +$	$- \quad -$	$- \quad -$	x x	$+ \quad +$	$+ \quad -$	x x
$4.3.A.2(b)$	$- \quad +$	$- \quad -$	$- \quad -$	x x	$+ \quad +$	$+ \quad +$	$+ \quad -$
$4.3.A.2(c)$	$- \quad +$	$- \quad -$	$- \quad -$	x x	$+ \quad -$	$+ \quad -$	$+ \quad +$
$4.3.C.2(a)$	$- \quad -$	$- \quad +$	$- \quad -$	$+ \quad -$	$+ \quad +$	x x	x x
$4.3.C.2(a)$	$- \quad -$	$- \quad +$	$- \quad -$	$+ \quad +$	$+ \quad -$	x x	x x
$4.3.C.2(b)$	$- \quad -$	$- \quad +$	$- \quad -$	$+ \quad +$	$+ \quad +$	x x	$+ \quad -$
$4.3.C.2(c)$	$- \quad -$	$- \quad +$	$- \quad -$	$+ \quad -$	$+ \quad -$	x x	$+ \quad +$
$4.2.3(ii)$	$+ \quad +$	$+ \quad +$	$- \quad -$	$- \quad -$	x x	x x	$+ \quad -$
$4.3.B.4(a)$	$+ \quad +$	$+ \quad -$	$- \quad -$	$- \quad -$	x x	$+ \quad -$	x x
$4.3.B.4(b)$	$+ \quad +$	$+ \quad -$	$- \quad -$	$- \quad -$	x x	$+ \quad +$	$+ \quad -$
$4.4.2(a)$	$- \quad +$	$+ \quad -$	$- \quad -$	$- \quad -$	$+ \quad +$	$+ \quad +$	$+ \quad -$
$4.4.2(b)$	$- \quad +$	$+ \quad -$	$- \quad -$	$- \quad -$	$+ \quad +$	$+ \quad -$	x x
$4.4.2(b)$	$- \quad +$	$+ \quad -$	$- \quad -$	$- \quad -$	$+ \quad -$	$+ \quad +$	x x
$4.3.C.4(b)$	$- \quad +$	$+ \quad +$	$- \quad -$	$- \quad -$	$+ \quad +$	x x	$+ \quad -$
$4.3.A.4(a)$	$+ \quad +$	$- \quad -$	$- \quad +$	x x	$- \quad -$	$+ \quad -$	x x
$4.3.A.4(b)$	$+ \quad +$	$- \quad -$	$- \quad +$	x x	$- \quad -$	$+ \quad +$	$+ \quad -$
$4.4.2(a)$	$+ \quad -$	$- \quad -$	$- \quad +$	$+ \quad +$	$- \quad -$	$+ \quad +$	$+ \quad -$
$4.4.2(b)$	$+ \quad -$	$- \quad -$	$- \quad +$	$+ \quad +$	$- \quad -$	$+ \quad -$	x x
$4.4.2(b)$	$+ \quad -$	$- \quad -$	$- \quad +$	$+ \quad -$	$- \quad -$	$+ \quad +$	x x
$4.4.2(a)$	$- \quad -$	$- \quad +$	$+ \quad -$	$+ \quad +$	$+ \quad +$	$- \quad -$	$+ \quad -$
$4.4.2(b)$	$- \quad -$	$- \quad +$	$+ \quad -$	$+ \quad +$	$+ \quad -$	$- \quad -$	x x
$4.4.2(b)$	$- \quad -$	$- \quad +$	$+ \quad -$	$+ \quad -$	$+ \quad +$	$- \quad -$	x x
$4.3.C.1(c)$	$+ \quad +$	$+ \quad +$	$- \quad +$	$- \quad -$	$- \quad -$	x x	$+ \quad -$
$4.4.3(a)$	$+ \quad +$	$+ \quad -$	$- \quad +$	$- \quad -$	$- \quad -$	$+ \quad +$	$+ \quad -$
$4.3.B.1(c)$	$+ \quad +$	$+ \quad +$	$+ \quad -$	$- \quad -$	x x	$- \quad -$	$+ \quad -$
$4.4.3(a)$	$- \quad +$	$+ \quad +$	$+ \quad -$	$- \quad -$	$+ \quad +$	$- \quad -$	$+ \quad -$
$4.3.A.1(c)$	$+ \quad +$	$- \quad +$	$+ \quad +$	x x	$- \quad -$	$- \quad -$	$+ \quad -$
$4.4.3(a)$	$+ \quad -$	$- \quad +$	$+ \quad +$	$+ \quad +$	$- \quad -$	$- \quad -$	$+ \quad -$

Table 6.3: Four cases 1 pattern with three stable equilibria

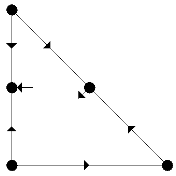
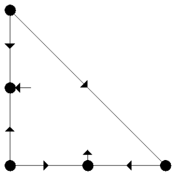
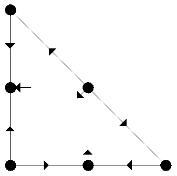
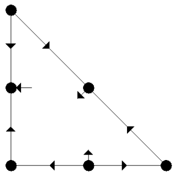
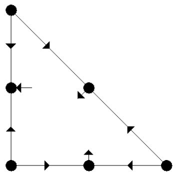
Lemma	$(0,0)$	$(0,1)$	$(1,0)$	P_1	P_2	P_3	Interior
$4.4.1(a)$	$-\ \ -$	$-\ \ -$	$-\ \ -$	$+\ \ -$	$+\ \ -$	$+\ \ -$	$+\ \ +$
$4.4.1(b)$	$-\ \ -$	$-\ \ -$	$-\ \ -$	$+\ \ +$	$+\ \ -$	$+\ \ -$	x x
$4.4.1(b)$	$-\ \ -$	$-\ \ -$	$-\ \ -$	$+\ \ -$	$+\ \ +$	$+\ \ -$	x x
$4.4.1(b)$	$-\ \ -$	$-\ \ -$	$-\ \ -$	$+\ \ -$	$+\ \ -$	$+\ \ +$	x x

6.3 Appendix C: Triangles for the 69 Cases

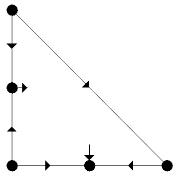
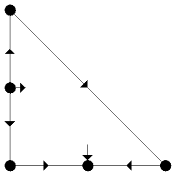
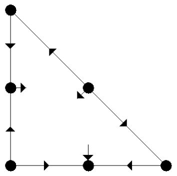
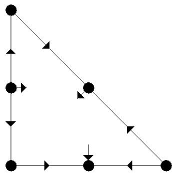
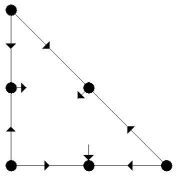
One Stable Equilibrium: Pattern 1

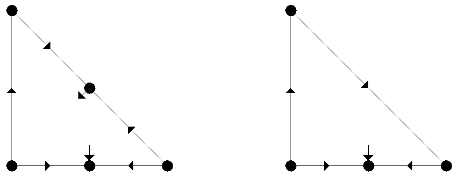







One Stable Equilibrium: Pattern 2

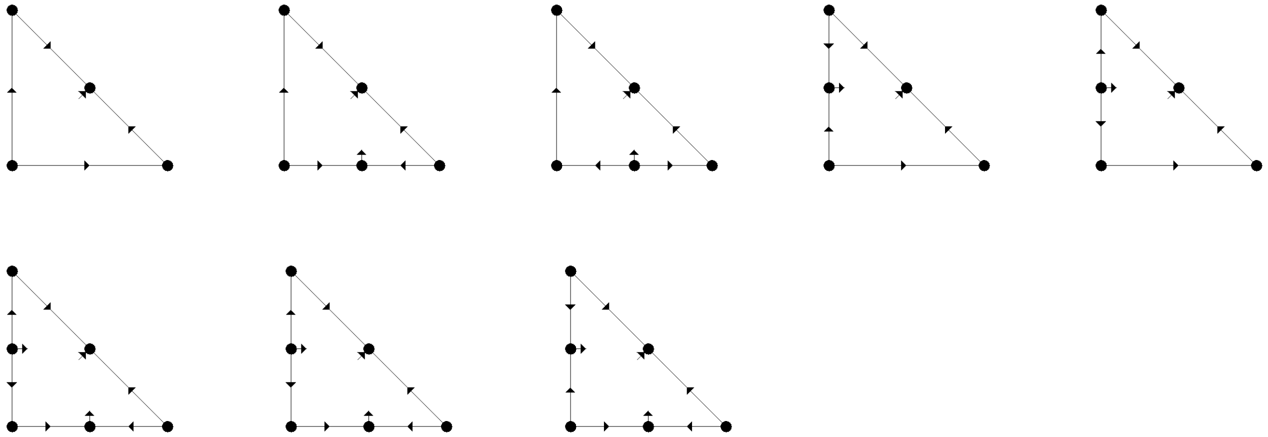






One Stable Equilibrium: Pattern 3

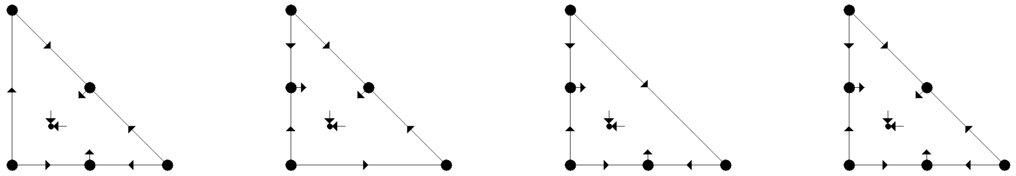








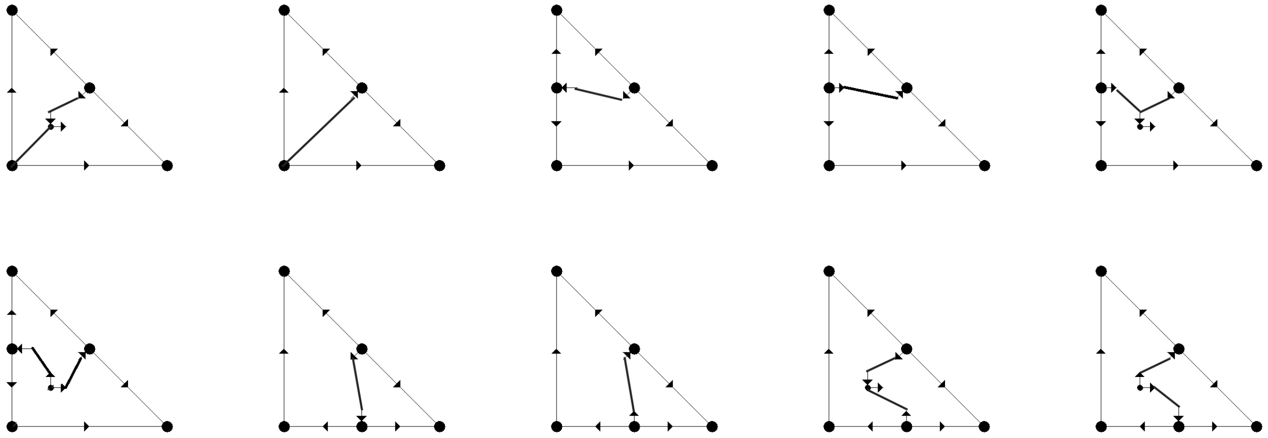
One Stable Equilibrium: Pattern 4



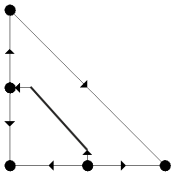
One Stable Equilibrium: Pattern 5

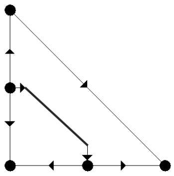


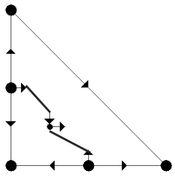
Two Stable Equilibria: Pattern 1

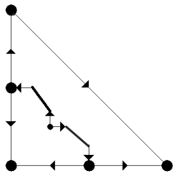


Two Stable Equilibria: Pattern 2

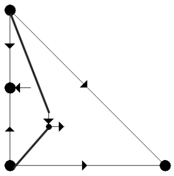


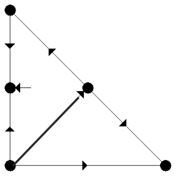


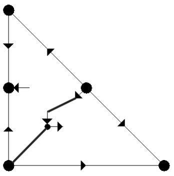


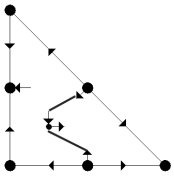


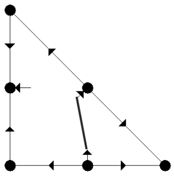
Two stable equilibria: Pattern 3

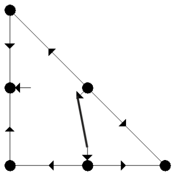


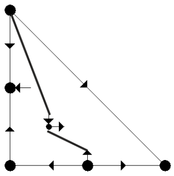




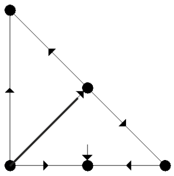


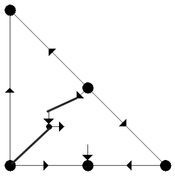


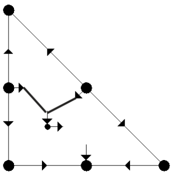


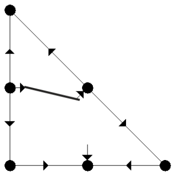


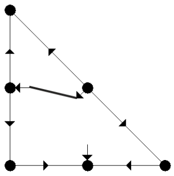
Two Stable Equilibria: Pattern 4



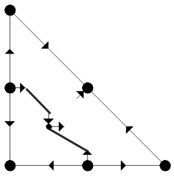


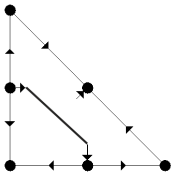


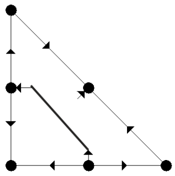




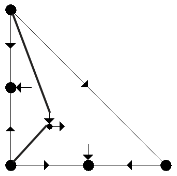
Two Stable Equilibria: Pattern 5

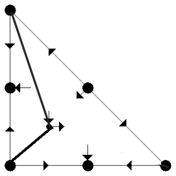




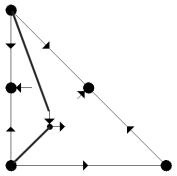


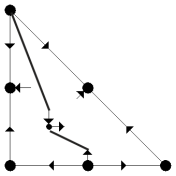
Two Stable Equilibria: Pattern 6



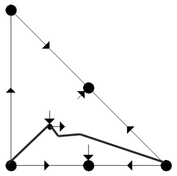


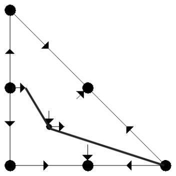
Two Stable Equilibria: Pattern 7



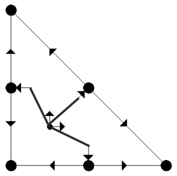


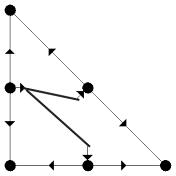
Two Stable Equilibria: Pattern 8

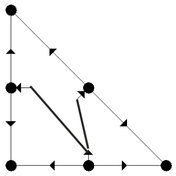


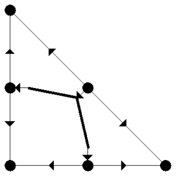


Three Stable Equilibria: Pattern 1









6.4 Appendix D: Traveling Wave Solutions

Initially, we started with a model that assumes the selection coefficient $r_{ij} = (r_{ii}+r_{jj})/2$. In this model we are interested in analyzing the fitness of individual alleles denoted by $s_k = r_{kk}/2$ All cases of dominance are considered, and we analyze the linear stability about each of the corner equilibria. Under these assumptions, they are no boundary or interior equilibria. We found the preceding model far too oversimplified and revised our model to allow the allele frequencies to diffuse in space using the equations:

$$\begin{cases} p_{1,t} = d_1 p_{1,xx} + p_1(r_1 - r) \\ p_{2,t} = d_2 p_{2,xx} + p_2(r_2 - r) \end{cases} \tag{6.1}$$

Where $p_{i,xx}$ denotes the second derivative with respect to space and we take $d_1 = d_2 = 1$. We assume a traveling wave solution in the form $p_i(x,t) = p_i(x - ct) = p(z)$ for $i = 1, 2$ and reduce the system of PDEs into as system of four nonlinear ODEs:

$$\begin{cases} \dot{p_1} = q_1 \\ \dot{q_1} = -cq_1 - p_1(A - Ap_1 - ap_2) \\ \dot{p_2} = q_2 \\ \dot{q_2} = -cq_2 - p_2(a - Ap_1 - ap_2) \end{cases} \tag{6.2}$$

Where $r_1 - r = A - Ap_1 - ap_2$ and $r_2 - r = a - Ap_1 - ap_2$ with $A = (r_{11} - r_{13})/2$ and $a = (r_{11} - r_{13})/2$. We explored the possibility of traveling waves from $(1, 0, 0, 0)$ to $(0, 0, 1, 0)$ along the intersection of the stable and unstable manifolds for $(p_1, q_1, p_2, q_2) \in \mathbb{R}^4$. Motivated by Tang and Fife [8] we attempt to prove the existence of an invariant region that bounds the heteroclinic trajectory from the unstable manifold about $(0, 0, 1, 0)$ to the stable manifold about $(1, 0, 0, 0)$ forward and backward in time. This profile represents the elimination of the fittest allele in favor of the second fittest allele; the conditions that allow for this traveling wave are highly nontrivial, and would be of great theoretical and practical interest to a biologist. We attempted to construct this set using the eigenvalues about the equilibria, but found placing bounds on the derivatives yielded better results. The idea behind a traveling wave solution is illustrated in the following section.

6.4.1 Two-Allele Traveling Wave Solutions

Consider a population of diploid individuals inhabiting a one-dimensional habitat which we assume to be the real line. We consider the case of no dominance ($r_{12} = s_1 + s_2$) for two alleles, A_1, A_2, constrained to a single locus and assume $s_1 > s_2$. Let $p(x,t)$ and $p_2(x,t) = 1 - p(x,t)$ be the density of allele A_1 and A_2 respectively at position x at time t. Under the above assumptions, $p(x,t)$ satisfies the following PDE:

$$p_t(x,t) = dp_{xx}(x,t) + p(x,t)(r_1(x,t) - r(x,t)) \tag{6.3}$$

where d is the diffusion constant. We can scale d to 1 if we rescale x to $y = xk$ and take k as $1/\sqrt{d}$. If we assume a traveling wave solution in the form $p(y,t) = w(y - ct) = w(z)$, $c \in \mathbb{R}$, we can write(6.3) as the following ODE:

$$-cw'(z) = w''(z) + w(z)(r_1(z) - r(z)) \tag{6.4}$$

where $w_{yy}(z) = w''(z)$, $w_t(z) = -cw(z)$, $w(z)(r_1(z) - r(z)) = \alpha w(z)(1 - w(z))$ and $\alpha = s_1 - s_2$¿0 since $s_1 > s_2$. If we let $q = w'$ and $q' = w''$, we can reduce(6.4) to the following first order system :

$$\begin{cases} w' = q \\ q' = -cq - \alpha w(1 - w) \end{cases} \tag{6.5}$$

Steady states for the system are $(0, 0)$ and $(1, 0)$. Notice w' is zero on $q = 0$ while q' is strictly negative on $q = 0$ as $\alpha w(1 - w)$ is a strictly positive quantity. Also notice q' is zero on $\frac{-\alpha}{c} w(1 - w)$ while on $\frac{-\alpha}{c} w(1 - w)$, w' is negative as $\frac{-\alpha}{c} w(1 - w)$ is a strictly negative quantity. We compute the Jacobian as

$$\mathcal{J} = \begin{pmatrix} 0 & 1 \\ & \\ 2\alpha w - \alpha & -c \end{pmatrix} \tag{6.6}$$

At $(0, 0)$, the determinant of(6.6) is the positive parameter, α , and the trace is the negative quantity, $-c$, which means both eigenvalues are negative real numbers or complex conjugates. The characteristic equation is $\lambda^2 + c\lambda + \alpha = 0$ thus if $c^2 \geq 4\alpha$ both eigenvalues are real and $(0, 0)$ is stable. At $(1, 0)$, the determinant of(6.6) is the negative quantity $-\alpha$ and the trace is the negative quantity $-c$ which means one eigenvalue is a positive real number and the other eigenvalue a negative real number, thus $(1, 0)$ is a saddle. At $(1, 0)$, the eigenvector corresponding to the positive eigenvalue, λ , lies under the line $q = 0$ as eigenvectors are in the form $y = \lambda x$. As $w \in [0, 1]$, q' is strictly negative on $q = 0$ and trajectories starting at $(1, 0)$ are bounded below by $q = -\frac{\alpha}{c} w(1 - w)$ all solutions are trapped in the closed region, Ω, bounded by the lines $q = 0, w = 0, w = 1$ and $-\frac{\alpha}{c} w(1 - w)$. Thus, solutions starting near $(1, 0)$ enter Ω along the eigenvector corresponding to the positive eigenvalue and approach $(0, 0)$ along the corner of Ω. As there are no interior equilibria, there exists no periodic behavior in Ω from the Poincaré-Bendixson Theorem. We have shown there exists a heteroclinic trajectory connecting $(1, 0)$ to $(0, 0)$ under the condition that $c^2 \geq 4\alpha$. As q is strictly negative, the graph of w against z starts at $w = 1$ and asymptotically decreases to zero as $z \to \infty$. As we increase t this curve just moves to the right at speed c. The heteroclinic trajectory is shown in the following figure:

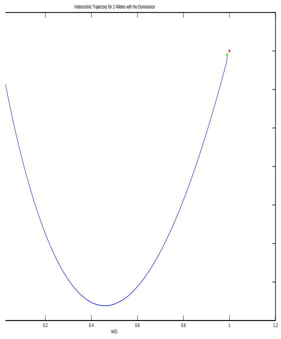


Figure 6.1: Heteroclinic Trajectory for 2 Alleles with No Dominance, where $s_1 = .9$ and $s_2 = .5$

6.4.2 Allele Frequency Diffusion

We were able to establish good bounds for the trajectory in forward time, but found that in backward time it diverged out the region. We continued to search for the traveling wave using numerical methods, but abandoned this approach because 'landing on' the stable manifold numerically is near impossible given even the tiniest numerical error.
On closer examination, we discovered the diffusion of gene frequencies is not interchangeable with the diffusion allele densities, though the work of Aronson and Weinberger [9] makes it clear when this interchange is possible. Aronson and Weinberger showed the a solution of the diffusion equation:

$$u_t = du_{xx} + f(u) \tag{6.7}$$

in the same form as those in(6.7), approximates a solution of PDE system:

$$\begin{cases} \rho_{1,t}(x,t) = \rho_{1,xx} - \tau_1\rho_1 + r(\rho_1 + \frac{\rho_2}{2})^2/\rho \\ \rho_{2,t}(x,t) = \rho_{2,xx} - \tau_2\rho_1 + 2r(\rho_1 + \frac{\rho_2}{2})(\rho_3 + \frac{\rho_2}{2})/\rho \\ \rho_{3,t}(x,t) = \rho_{2,xx} - \tau_3\rho_3 + r(\rho_3 + \frac{\rho_2}{2})^2/\rho \end{cases} \tag{6.8}$$

when $u = (\rho_3 + \frac{\rho_2}{2})/(\rho_1 + \rho_2 + \rho_3)$, where ρ_1, ρ_2 and ρ_3 denote the linear density of genotypes aa, aA and AA for two alleles respectively; the total genotype density $\rho = \rho_1 + \rho_2 + \rho_3$. This approximation holds when the initial conditions, birthrate, r, and deathrate, τ, satisfy certain constraints. The proof is rather involved an we refer the reader to [9] for details.

We concluded it would be best to give a complete classification of our first model, allowing the r_{ij}'s to vary continuously with the only assumption being $r_{11} > r_{22} > r_{33}$. Our revised assumptions allow for both internal and boundary equilibria. Two-allele population genetics models are ubiquitous, and the mechanisms of gene selection have been thoroughly investigated. Three-allele population genetics models are far less well understood, and a complete classification of even the simplest model is a formidable, but rewarding task.

Bibliography

[1] R. Bürger. *The mathematical theory of selection, recombination, and mutation.* Wiley, 2000.

[2] T. Nagylaki. *Introduction to Theoretical Population Genetics.* Springer, N.Y. 1992.

[3] Ken Miller. The Peppered Moth - An Update. http://www.millerandlevine.com/km/evol/Moths/moths.html. Brown University, August 1999.

[4] Ensembl Project. http://www.uswest.ensembl.org/index.html. European Bioinformatics Institute, 1999.

[5] Vinay Kumar. *Robbins Basic Pathology.* WB Saunders Co, 2007.

[6] C. Wiuf. (2001). Do ΔF508 heterozygotes have a selective advantage?. Genet.Res. **78** (1): 41-7 PMID 1155613.

[7] R.A. Fisher. *The Genetical Theory of Natural Selection.* Clarendon Press, Oxford, 1930.

[8] M.M. Tang and P.C. Fife (1980). Propagating Fronts for Competing Species Equations with Diffusion. Archive for Rational Mechanics and Analysis. **73**, 69-77.

[9] D.G. Aronson and H.F Weinberger (1975). Nonlinear Diffusion in Population Genetics, Combustion, and Nerve Pulse Propagation. Partial Differential Equations and Related Topics, Lecture Notes in Mathematics. **446**, 5-49.