J. David Logan
University of Nebraska Lincoln

A Primer on Population Genetics

August 11, 2009

Department of Mathematics
University of Nebraska Lincoln, Lincoln NE

Contents

Preface

This primer on population genetics is a supplement to the textbook J. D. Logan & W. R. Wolesensky, 2009. *Methods of Mathematical Biology*, John Wiley & Sons, New York. To cite this primer, please reference:

J. D. Logan, 2009. *Primer on Population Genetics*,
`http://www.math.unl.edu/~dlogan/MethMathBioBook`.

I am greatly indebted to books of N. Britton, J. Roughgarden, M. Nowak, and E. S. Allman & J. A. Rhodes. These notes show a strong influence from those sources, reflecting some of the same approaches and adopting similar philosophies. References are given at the end of the notes. I would like to thank Ben Nolting, a graduate student at the University of Nebraska, for many helpful discussions and suggestions. One of us (JDL) would like to thank the University of Nebraska for a Faculty Development Leave during Spring Semester 2009, which led to the opportunity to write up much of this material and put it an accessible form.

Suggestions, comments, and corrections are very welcome. Contact information is on the web site:
`http://www.math.unl.edu/ dlogan/MathMethBioBook.html`.

David Logan, Lincoln, Nebraska
August, 2009

A Primer on Population Genetics.
By J. David Logan

1
Selection

1.1 Introduction

Population genetics forms the mathematical basis for the key ideas in the evolution of species: **random variation** and **natural selection**. The subject of evolutionary biology began with Charles Darwin's observations and subsequent hypothesis that biological species change through time. He conjectured that some beneficial modification that may occur in a population would be preserved in future generations. For example, a change in a trait that increases an organism's fitness for its environment leads to survival. This strategy is called **survival of the fittest**. Nature, through natural selection, picks out the best suited traits.

The underlying reasons and biochemical foundations for natural selection were not known by Darwin. But, the work of Mendel, and those who followed, showed that genetics provides the underpinnings. For example, chromosomes in individuals carry genes that contain hereditary markers (DNA) that code for different traits. Through reproduction, where the parents' gametes (e.g., eggs and sperm) form to make the genetic material for their offspring, traits are passed on. Spontaneous, random mutations can occur, for example, through the copying of DNA. Therefore, genetic variation can naturally occur from one generation to another. The beneficial traits, or the ones that increase survival and reproductive success (fitness), are the ones preserved in future generations.

Mathematically, how do we model this complicated selection process? In its very simplest form, we keep track of gene frequencies from one generation to

A Primer on Population Genetics.
By J. David Logan

the next. The laws that govern this model are based upon both randomness and selection rules. In this primer we present the most elementary models.

A brief listing of several of the important, early contributions give a perspective of the time line involved and the researchers who played a key role in this early progression.

– 1859 C. Darwin. *The Origin of Species* first appeared.

– 1886 G. Mendel. His original paper on plant hybridization.

– 1908 G. H. Hardy and Weinberg (independently). Work on Mendelian proportions.

– 1918 R. A. Fisher. Correlations in Mendelian inheritance; R. B. Robbins, the mathematics of breeding.

– 1922 R. A. Fisher. Dominance

– 1924 J. B. S. Haldane. The mathematical theory of selection.

– 1926 and 1931. J. B. S. Haldane. The causes of evolution.

– 1930. R. A. Fisher. The theory of natural selection.

– 1937. S. Wright. The distribution of gene frequencies in a population.

– 1941. R. A. Fisher. The theory of gene substitution.

A further historical perspective up to 1977 can be found in Edwards (1977).

1.2 Mendelian Genetics

It is quite amazing that Mendel did his work on hybridization and breeding experiments in pea plants with only the abstract idea of a 'gene'. His generation knew nothing about its biochemical basis or DNA sequencing. It wasn't until the early 1900s that chromosomes, which carry the heritable traits, were identified under a microscope, and other researchers began to notice his work.

Mendel's reasoning was based on an observation—that offspring often exhibit the same traits of their parents, and even their grandparents. Although not totally predictable, it is an observation that we all have made in our own families.

The mathematical basis of Mendel's ideas is elementary probability and some assumptions about the process. In the simplest terms, using more modern terminology, we assume that a chromosome has a site, or **locus** (think of it as a location on chromosome defined by a DNA sequence), that can accommodate

one of two different heritable traits, a and b, called **alleles**. For example, a may carry the information for brown eyes, and b the information for blue eyes. In the chromosome there are other loci for other traits, but we are not concerned with those at this time. Also, several alleles can occur at a locus, but here we consider only two. Later we discuss the case of multiple alleles at a single locus. Refer to fig. 1.1 for the following more specific discussion. A chromosome in an individual's cell is composed of a pair of identical sister chromotids joined together at the centromere. Chromosomes come in homologous pairs where a particular locus occurs on all of the four chromotids. On each of the homologous chromosomes, the locus contains an a and a, or b and b. In one pair has a and a, and the other pair has b and b, we say the **genotype** is ab; if the pairs are $a-a$ and $a-a$, then the genotype is aa, and if the pairs are $b-b$ and $b-b$, the genotype is bb. The possibilities for the allele pairs (e.g., a at one locus and b at the other) are aa, ab, and bb. Thus, there are three possible genotypes. How these genotypes express themselves in an individual is deterimined by genotype and a dominance relation. For example, for eyes, if a is dominant over b, then both aa and ab give brown eyes, while bb presents itself as blue eyes. These two expressions, brown and blue, are called the **phenotypes**. We say aa and bb are homozygous genotypes, while ab is heterozygous. The genotype ab is the same as genotype ba.

How does information get to the next generation? During a complicated process of cell division, an individual carrying the two alleles creates gametes (e.g., eggs or sperm) that carry only one of the alleles. For example, an ab genotype produces gametes a and b. When this individual, say a male, mates with a female, say with genotype aa that produces gametes a and a, random union of the gametes is assumed to occur and produce zygotes for the next generation. These zygotes have possible genotypes ab and aa; here, the first the first element in the pair is from the male, and the second is from the female. Again, fig. 1.1 shows a crude diagram showing this process, or how alleles of an ab genotype eventually produces gametes and offspring. Other books and web articles have much more detailed information about this process and often present elaborate graphics.

If an bb male mates with ab female, then the male gametes are b and b, and the female gametes are a and b. The resulting zygotes have genotype bb, ab, thus producing two phenotypes, blue eyes and brown eyes. We can conveniently determine crosses between two genotypes using a **Punnett square**. For example, if an bb male is crossed with an ab female, then a Punnet square gives the information we seek.

	a	b
b	ba	bb
b	ba	bb

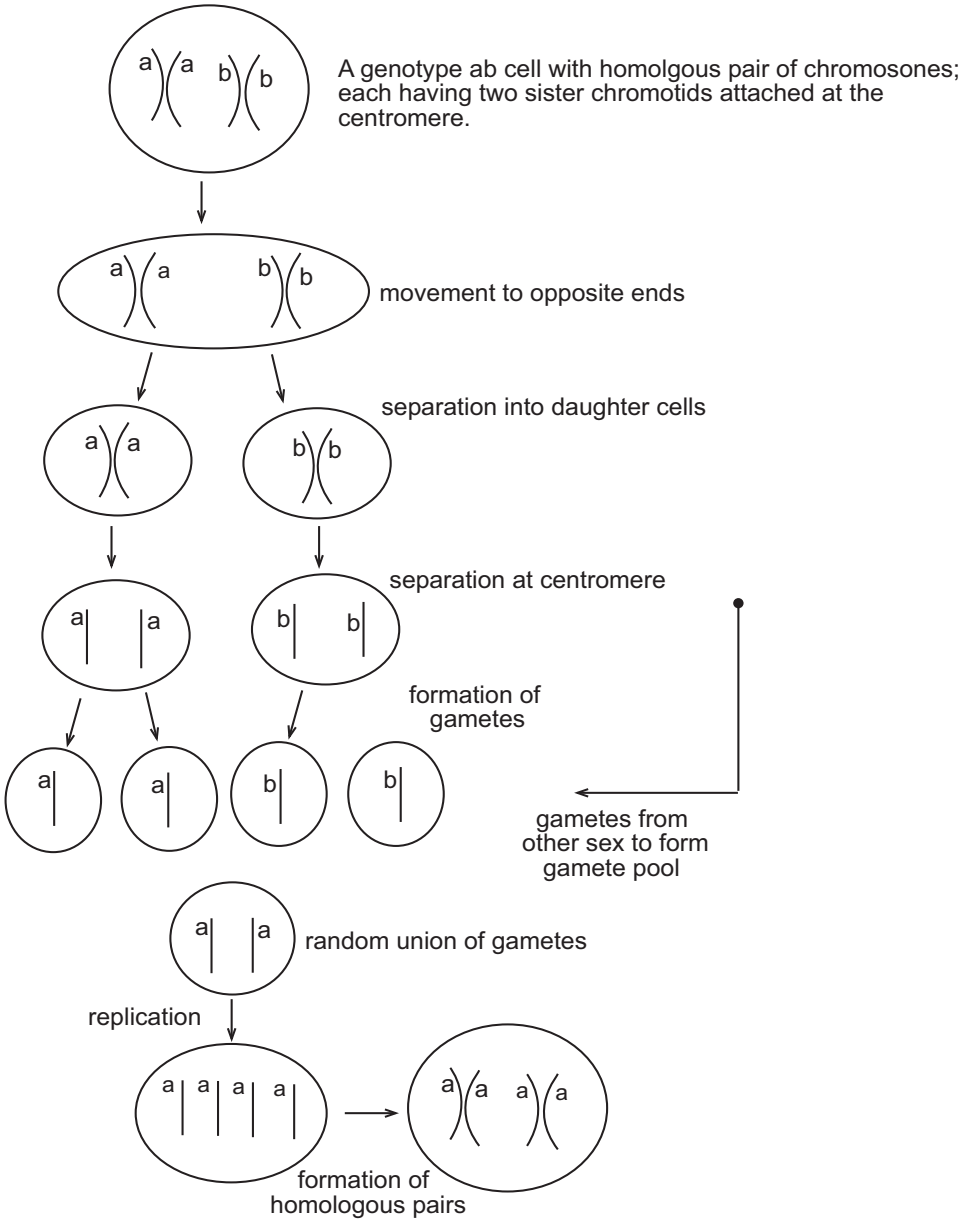

Figure 1.1 Diagram of the process of where a cell carrying genotype ab, forms gametes that combine with gametes from a mate to produce a gamete pool; random union of the gametes in the pool then form a new zygote.

The probability of those two parents producing an ab offspring is one-half, and the probability of bb is one-half. Note that the male gametes are listed in the leftmost column and the female gametes are listed along the top row. To repeat, by random union of the gametes, there is a 50–50 chance that an offspring is ab, and a 50–50 chance it will be bb.

The way we quantify all of this information is to track genotype frequencies and allele frequencies from one generation to the next. Suppose we have a pool of N individuals having the three possible genotypes aa, ab, bb . We denote the genotype frequency (or, proportion of the population) of aa, ab, and bb by x, y, and z, respectively. Thus, x is the number of aa's divided by the population N, and so on. It is clear that

$$x + y + z = 1.$$

If we know the genotype frequencies, we can calculate the allele frequencies simply. There are $2N$ alleles in the gene pool. Let p be the frequency (proportion of) of allele a and let q be the frequency of allele b. The number of a alleles is twice the number of aa genotypes plus one-half the number of ab genotypes. So,

$$p = \frac{\text{no. of } a \text{ alleles}}{2N} = \frac{2Nx + Ny}{2N} = x + \frac{1}{2}y. \tag{1.1}$$

Similarly,

$$q = \frac{\text{no. of } b \text{ alleles}}{2N} = \frac{2Nz + Ny}{2N} = z + \frac{1}{2}y. \tag{1.2}$$

Therefore, in a fixed generation, the genotype frequencies x, y, z uniquely determine the allele frequencies p and q. Clearly,

$$p + q = 1.$$

But, the opposite is not true. Knowing the allele frequencies p and q does not determine the genotype frequencies uniquely. It is very easy to invent a simple example with a population of $N = 2$; for example, aa and bb vs. ab and ab. Both have allele frequencies $p = q = 0.5$, but their genotype frequencies are $x = 0.5$, $y = 0$, $z = 0.5$ vs. $x = z = 0$, $y = 1$. Equations (1.1)–(1.2) *always* hold within a generation.

So much for one generation. How do allele frequencies change from one generation to the next? To fix the idea, let us begin with a population of adults of genotype frequencies x, y, z; we know the allele frequencies p and q from (1.1)–(1.2). The assumed progression of events is that the breeding adults produce a large pool of gametes, and the combination of the alleles, located in the gametes, takes place randomly, producing zygotes for the next generation. The zygotes are assumed to mature into adults, with no mortality or selection.

So, the frequencies of zygotes and adults are the same in this model (later we take into account selection). A diagram shows this schematically:

$$\text{adults} \ \rightarrow \text{gamete pool} \ \rightarrow \ \text{zygotes} \rightarrow \text{adults}.$$

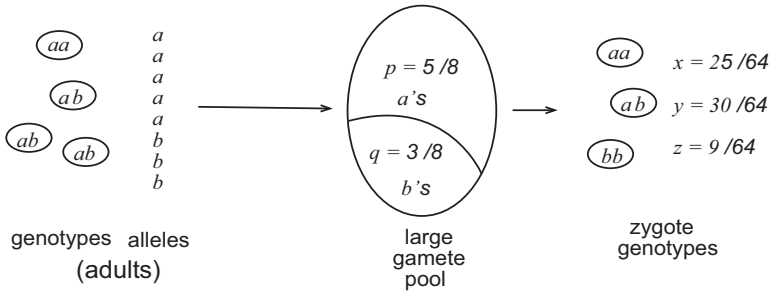

Figure 1.2

An example is shown pictorially in fig. 1.2. Let x', y', z' denote the genotype frequencies in this next generation (So, 'prime' is not a derivative, but rather means 'next generation'.). Frequency means probability, and so we can ask what is $x' = \Pr(aa)$, $y' = \Pr(ab)$, $z' = \Pr(bb)$. A zygote from the next generation is formed by randomly choosing two gametes from the large gamete pool. We can think of this formation as a binomial experiment; we first choose gamete 1, and then we choose gamete 2. So, we are filling two slots. Because the gamete pool is very large, we can consider these two choices as independent, even though we are choosing without replacement. The possible results of the two ordered choices are the four pairs of slots: $\underline{aa}$, $\underline{ab}$, $\underline{ba}$, and $\underline{bb}$. So, by independence, $x' = \Pr(aa) = \Pr(a)\Pr(a) = p \cdot p = p^2$. Similarly, $z' = \Pr(bb) = q^2$. Finally, the probability of the genotype ab is the probability is the probability of getting $\underline{ab}$ or $\underline{ba}$; hence,

$$\begin{aligned} y' &= \Pr(ab) = \Pr(\text{choosing } \underline{ab} \text{ or choosing } \underline{ba}) \\ &= \Pr(\text{choosing } \underline{ab}) + \Pr(\text{choosing } \underline{ab}) = \Pr(a)P(b) + \Pr(b)\Pr(a) \\ &= 2pq. \end{aligned}$$

Again, compare to a binomial experiment. We make two choices of a random variable X that gives the number of successes, say, picking an a. A success has probability p, and failure (picking a b) has probability q. The number of successes is binomial distributed. Thus, it has density

$$\Pr(X = k) = \binom{2}{k} p^k q^{2-k}, \quad k = 0, 1, 2.$$

These are the frequencies for the next generation. Clearly, $\Pr(X = 2) = p^2$, $\Pr(X = 1) = 2pq$, $\Pr(X = 0) = q^2$.

Now that we know the genotype frequencies for the next generation, we can compute the allele frequencies from (1.1)–(1.2). We have

$$\begin{aligned} p' &= x' + \frac{1}{2}y' = p^2 + \frac{1}{2}2pq = p, \\ q' &= z' + \frac{1}{2}y' = q^2 + \frac{1}{2}2pq = q. \end{aligned}$$

So the allele frequency did not change in the next generation. This fact will clearly be true if we iterate to the next generation, and so on. This is the famous **Hardy–Weinberg law** that we now state as a theorem. In stating the theorem we return to a time series notation for the allele frequencies; that is. let p_t and q_t be the allele frequencies in the tth generation. Generally, we refer to the zeroth generation as the *parent generation*, the first generation as the F_1, or first filial generation, F_2 the second filial generation, and so on.

Theorem 1.1

(Hardy–Weinberg) Let p_0 and q_0 be the frequencies of alleles a and b in the initial $t = 0$ population. Then, under the assumptions of random mating, random union of gametes, and allele frequencies independent of sex,

$$p_{t+1} = p_t, \quad q_{t+1} = q_t, \quad t = 0, 1, 2, ...,$$

or $p_t = p_0$, $q_t = q_0$, for all $t = 1, 2, 3, ...$ For generations *after* the initial generation, the genotype frequencies are

$$x_t = p_t^2, \quad y_t = 2p_tq_t, \quad z_t = q_t^2, \quad t = 1, 2, 3, ...$$

The ratios $p^2 : 2pq : q^2$ are called the **Hardy–Weinberg ratios**, and the system is said to be in **Hardy–Weinberg equilibrium**.

Example 1.2

. Consider a population with four individuals (see fig. HWfig) of genotype aa, ab, ab, ab. Thus $x_0 = 1/4$, $y_0 = 3/4$, $z_0 = 0$. Thus, $p_0 = 5/8$, $q_0 = 3/8$. The next generation has genotype frequencies $x_1 = 25/64$, $y_1 = 30/64$, $z_1 = 9/64$, and $p_1 = 5/8$, $q_1 = 3/8$. All future generations will have these same genotype and allele frequencies.

Observe that we have not tracked the population of genotypes or alleles, just their frequencies. Later, when we discuss fitness, which includes mortality

and fecundity, we also track population growth. Without knowing the vital statistics or life history parameters of a population it is impossible to know size, for example, of the gamete pool. We also point out that this model is the simplest type of accounting; we are keeping track of a genotypes from organisms that have a single life cycle, or non-overlapping generations. All the organisms, equally fit, reproduce once and then die, all at the same time. There is no evolution in such a system.

It is straightforward to extend Mendelian genetics to more complicated breeding systems. Consider the following example with two loci.

Example 1.3

. Mendel's experiments involved pea plants with different characteristics. We consider two loci with alleles D and d (tall or dwarf plants) and G and g (green pods or yellow pods). D and G are dominant over d and g. Consider a cross breed with a $DdGg$ and a $ddGg$. We assume the alleles at the loci sort independently, meaning that in gamete formation alleles for height and color separate, independently. We can draw a Punnett square to show the sixteen possible results of the cross. The leftmost column are the allele pairs of the $ddGg$ parent (parent 1), and the topmost row are the allele pairs of the $DdGg$ parent (parent 2).

	DG	Dg	dG	dg
dG	$DdGG$	$DdGg$	$ddGG$	$ddGg$
dg	$DdGg$	$Ddgg$	$ddGg$	$ddgg$
dG	$DdGG$	$DdGg$	$ddGG$	$ddGg$
dg	$DdGg$	$Ddgg$	$ddGg$	$ddgg$

We assume that all results are equally probable. There are 6 different genotypes for the offspring: $DdGG$, $DdGg$, $Ddgg$, $ddGG$, $ddGg$, $ddgg$. Because D and G are dominant, there are four different phenotypes: tall-green, tall-yellow,dwarf-green, dwarf-yellow. We can count the number of phenotypes in each case and divide by 16 to get the probability of each occuring:

phenotype	probability
tall-green	$\frac{6}{16}$
tall-yellow	$\frac{2}{16}$
dwarf-green	$\frac{6}{16}$
dwarf-yellow	$\frac{2}{16}$

We can also compute the probabilities using the laws of probability and independent events. For example, to compute the probability of a dwarf-yellow plant we have:

$$\begin{aligned}
\Pr(ddgg) &= \Pr(dg \text{ from parent 1 and } dg \text{ from parent 2}) \\
&= \Pr(dg \text{ from parent 1}) \Pr(dg \text{ from parent 2}) \\
&= \Pr(d \text{ from parent 1}) \Pr(g \text{ from parent 1}) \\
&\quad \times \Pr(d \text{ from parent 2}) \Pr(g \text{ from parent 2}) \\
&= 1 \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{2}{16}
\end{aligned}$$

EXERCISES

1. Draw Punnett squares for a cross between an aa male and a bb female and a cross between an ab male and ab female.
2. Tay-Sachs is a disease that occurs in children of parents of Jewish Ashkenazi descent. Roughly 1 out of 31 carry the recessive allele b in the heterozygous state ab, where a is a dominant allele. Only the homozygous bb individuals get the disease, which is lethal in childhood. What is the probability that an offspring in this family of individuals will get the disease?
3. Suppose an $ab \times ab$ cross has four childrem. Use the binomialdistribution to find the probability that 3 of the 4 are aa. What is the expected number of children out of the 4 that will be aa? What is the variance. What is the expected number of ab's?
4. Draw a Punnett square for an $AaBb$ cross with an $AaBb$, where A and a are alleles at one locus, and B and b are alleles at another locus, and the loci assort independently. What are the possible genotypes? If A and B are dominant over a and b, respectively, how many phenotypes will there be?

1.3 Selection and Variability

Population genetics is about how the gene pool of a population changes from generation to generation and determining the mechanisms that cause the change. As many so aptly have pointed out, chemicals don't change over time, but biological organisms do. Changes that occur are determined by many factors. A genotype's fitness may be population dependent, frequency dependent, or even dependent upon environmental conditions; one of the genotypes, say, aa, could have some has some selective advantage over the other genotypes in that it contributes more than its Hardy-Weinberg share of alleles to the next generation.

In the simplest case we assume a life cycle that begins with the zygotes in generation t. The first place selection can enter the process is through mortality of the genotypes. Zygotes grow to become breeding adults, and the survivorship of the various genotypes through this growth process may be different. When breeding occurs, which is the second source of selection, some genotypes may be more successful than others in producing gametes, and some gametes may differentially form zygotes for the new generation at time $t+1$. Let's quantify this progression of events. Let l_x, l_y, and l_z denote the probabilities that the three genotypes, aa, ab, and bb, respectively, survive to breeding adults. These adults go through a complicated process of producing gametes that then combine to form the genotypes of the new $t+1$ generation. Generally, the overall gamete pool is very large and we cannot know or track the total numbers of gametes produced in the tth generation, or how these gametes unite. But, we can postulate the number of gametes produced by each genotype that are *actually incorporated* into the population at the $(t+1)$st generation. To proceed, we let $2m_x$ be the number of gametes shed by aa genotypes that are incorporated into the $(t+1)$st generation of zygotes, and similarly $2m_y$, and $2m_z$ for ab and bb genotypes. (For example, Suppose there are 100 aa's in the tth population and $l_x = 0.7$. Then there are 70 breeding adults. If the 70 breeding aa adults leave 300 gametes that are incorporated into the next generation, then $m_x = 35$.) The m's are fecundity factors representing the contributions to the next generation. All this information is in the table. We emphasize that the last column is the number *gametes* contributed to the next generation by the various genotypes; they are not genotypes, and we do not track the number of genotypes in the next generation. Recall that allele frequencies do not determine the genotypes.

zygotes (t)	No. genotypes (t)	breeding adults (t)	No. gametes $(t+1)$
aa	Np^2	l_xNp^2	$2m_xl_xNp^2$
ab	$2Npq$	$2l_yNpq$	$2m_y2l_yNpq$
bb	Nq^2	l_zNq^2	$2m_zl_zNq^2$

The total number of gametes in the $(t+1)$st generation is the sum of the last column, or

$$\begin{aligned} \text{total gametes} &= 2m_xl_xNp^2 + 2m_y2l_yNpq + 2m_zl_zNq^2 \\ &= 2(m_xl_xp^2 + 2m_yl_ypq + m_zl_zq^2)N. \end{aligned}$$

This equation leads to the total population N' of the $(t+1)$st generation in terms of the population N at the tth generation because the population is one-half the number of gametes. Thus,

$$N' = (m_xl_xp^2 + 2m_yl_ypq + m_zl_zq^2)N.$$

But we don't know the genotype frequencies; we track only allele frequencies. Now, the frequency of allele a is

$$\begin{aligned} p' &= \frac{\text{no. of gametes having } a \text{ type alleles}}{\text{total no. of gametes}} \\ &= \frac{\text{no. of gametes from } aa\text{'s} + \frac{1}{2}(\text{no. of gametes produced by } ab\text{'s})}{\text{total no. of gametes}} \\ &= \frac{2m_x l_x N p^2 + \frac{1}{2}\left(2m_y 2l_y N pq\right)}{2m_x l_x N p^2 + 2m_y 2l_y N pq + 2m_z l_z N q^2} \\ &= \frac{m_x l_x p^2 + m_y l_y pq}{m_x l_x p^2 + 2m_y l_y pq + m_z l_z q^2}. \end{aligned}$$

Next we introduce the **absolute fitnesses coefficients**

$$W_x = m_x l_x, \quad W_y = m_y l_y, \quad W_z = m_z l_z.$$

These coefficients are products of fecundities and survivorships of the three genotypes, and they measure the contribution of a given genotype the next generation. Although fitness can be measured in many ways, one certain measure is in terms of survivorships and fertility. (Some texts just introduce these coefficients without relating them to mortality and fertility rates, as we have done here.) Using these fitness definitions, we can write

$$p' = \frac{W_x p^2 + W_y pq}{W_x p^2 + 2W_y pq + W_z q^2}. \tag{1.3}$$

Usually we are not focused on population sizes, but only how the allele frequencies change; therefore, we can deal with **relative fitness coefficients**

$$w_x = \frac{W_x}{W_{\max}}, \quad w_y = \frac{W_y}{W_{\max}}, \quad w_z = \frac{W_z}{W_{\max}},$$

where $W_{\max} = \max(W_x, W_y, W_z)$. So, we are scaling the fitness coefficients by the largest of the three. Also, rather than use $W_{\max}$ for scaling, we can use any constant, e.g., W_z. This choice would scale the fitnesses by the fitness of the homozygous bb allele. From (1.3), whatever the scaling, the scaling constant, e.g., $W_{\max}$, will cancel out, and we get

Therefore we have

$$p' = \frac{(w_x p + w_y q)p}{\overline{w}}, \tag{1.4}$$

where

$$\overline{w} = w_x p^2 + 2w_y pq + w_z q^2 \tag{1.5}$$

is called the **mean fitness**. In a similar manner,

$$q' = \frac{(w_y p + w_z q)q}{\overline{w}}. \tag{1.6}$$

These are key selection equations due to Fisher, Haldane, and Wright that track the allele frequencies from one generation to the next. It is convenient to write these equations in terms of changes in p and q, for example, $\Delta p = p' - p$. Further, which simplifies the formulas for easier analysis, we introduce **weighted fitnesses** by

$$w_p = w_x p + w_y q, \quad w_q = w_y p + w_z q.$$

This makes the mean fitness

$$\overline{w} = p w_p + q w_q.$$

Then the change Δp can be written

$$\begin{aligned} \Delta p &= p' - p = p\frac{w_p - \overline{w}}{\overline{w}} \\ &= p\frac{w_p - (pw_p + qw_q)}{\overline{w}} \\ &= p\frac{w_p - ((1-q)w_p + qw_q)}{\overline{w}} \\ &= pq\frac{w_p - w_q}{\overline{w}} \\ &= pq\frac{p(w_x - w_y) + q(w_y - w_z)}{\overline{w}}. \end{aligned} \tag{1.7}$$

Equation (1.4) or (1.7) is the famous **Fisher–Haldane–Wright** (FHW) equation. We record the result as a theorem.

Theorem 1.4

(Fisher–Haldane–Wright) Given the assumptions above, the change in the frequency p of allele a from one to the next generation is given by

$$\Delta p = pq\frac{p(w_x - w_y) + q(w_y - w_z)}{\overline{w}}. \qquad \text{(FHW equation)}$$

This equation is shorthand for the difference equation

$$p_{t+1} = p_t + p_t q_t \frac{p_t(w_x - w_y) + q_t(w_y - w_z)}{\overline{w}}. \qquad \text{(FHW equation).}$$

Notice that the FHW equation reduces to the Hardy–Weinberg law if all the fitnesses are equal, that is, there is no selection.

The following MATLAB file plots the time series p_t using the FWH equation, with p_0 and the relative fitness coefficients given. (See fig. 1.3.)

```
function FWHequation
clear all
```

```
p=0.2; plist=2; numgenerations=200;
wx=1; wy=0.98; wz=0.8;
for t=1:numgenerations
q=1-p; wbar=wx*p.^2+2*wy*p .*q+wz*q.^2;
p=p+p .*q .*(p*(wx-wy)+q*(wy-wz)) ./ wbar;
plist=[plist,p];
end
time=0:numgenerations; plot(time, plist)
```

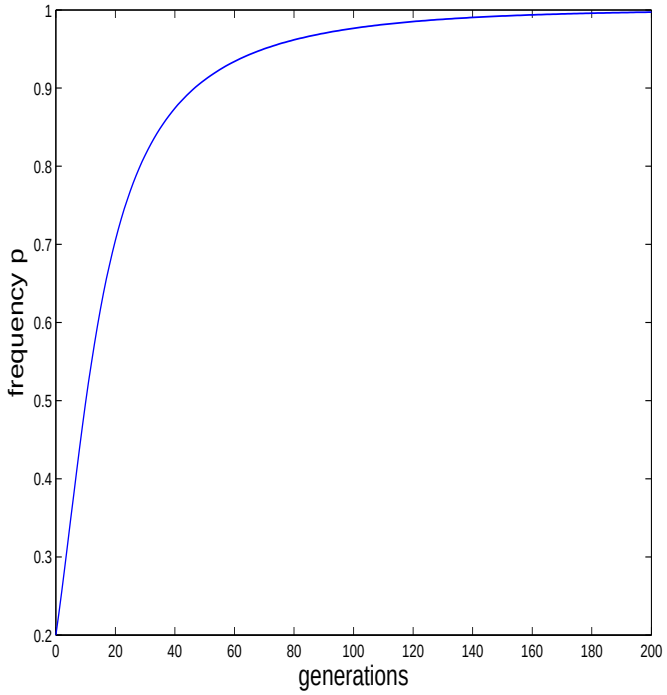

Figure 1.3 Time series plot of p_t. Note that p_t departs from its initial value very rapidly. In contrast, the approach to equilibrium is extremely slow. The rate that a selection process departs and arrives is an important issue we shall discuss later.

Example 1.5

If a is completely dominant and we select for it, then we mean

$$w_x = w_y = 1 + s, \quad w_z = 1,$$

where s is a positive **selection coefficient** measuring the strength of selection for allele a. In this case, we mean both genotypes aa and bb have the same fitness. Then the FHW equation becomes

$$p' = p + \frac{spq^2}{1 + s(p^2 + 2pq)}, \tag{1.8}$$

where $p = p_t$, $p = p_t$, and $p' = p_{t+1}$. We observe that $p' > p$, and so p_t, the frequency of allele a, is always increasing. We can use the preceding MATLAB program to plot the sequence for different values of s and different initial frequencies p_0. All the simulations show that the frequencies approach 1 as t gets large. As in fig. 1.3, the departure from p_0 is rapid, while the approach to equilibrium is slow.

1.3.1 Equilibria

The FHW equation is a difference equation, or discrete model of the form

$$p_{t+1} = f(p_t),$$

or in shorthand notation, $p' = f(p)$, where $f(p)$ represents the right side of the equation. A complete analysis of the FHW equation would include answering the following questions: Are there any equilibria, or constant solutions? Are those equilibria stable or unstable? If p approaches an equilibrium frequency, how fast is the approach? What is the time scale, or roughly the number of generations that it takes to get to equilibrium? To answer these question, we briefly review the key definitions and ideas. (This material is covered in Chapter 2 of Logan & Wolesensky 2009, and in most other mathematical biology texts.)

A discrete model

$$p_{t+1} = f(p_t)$$

has an equilibrium $p = p^*$ iff $p^* = f(p^*)$. In other words, the difference equation has a constant solution $p_t = p^*$. There is a convenient graphical interpretation of an equilibrium. On a $p'p$ coordinate system, we sketch plots of $p' = f(p)$ and the diagonal $p' = p$. Equilibria are values $p = p^*$ where the two graphs cross, or that satisfy $p = f(p)$. Often we sketch a cobweb diagram (e.g., see fig. 1.4 to determine how the time series p_t evolves. We say an equilibrium p^* is

locally asymptotically stable iff[1] $p_t \to p^*$ for all initial values p_0 that are sufficiently close to p^*. Analytically, an equilibrium is locally asymptotically iff

$$|f'(p^*)| < 1.$$

This means that the derivative (the slope of the tangent line) at an equilibrium

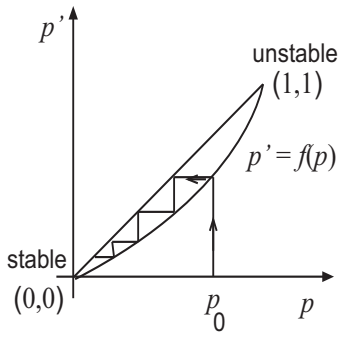

Figure 1.4 A cobweb diagram showing convergence to the asymptotically stable equilibrium $p = 0$. Beginning at the initial value p_0, one draws vertical and horizontal lines to the curve $p' = f(p)$ and to the diagonal $p' = p$, alternately. Note that $-1 < f'(0) < 1$, which implies stability, while $f'(1) > 1$, implying unstable at $p = 1$.

p^* cannot be too large; its absolute magnitude must not exceed 1. Thus, the graph of $f(p)$ cannot be too steep at the point p^* where it crosses the diagonal. An equilibrium p^* is not asymptotically stable iff

$$|f'(p^*)| > 1.$$

This means there is not an open interval I containing p^*, no matter how small, such that $p_t \to p^*$ for all p_0 in I. A weaker definition is that of **local stability**, which means that solutions beginning within a distance δ of an equilibrium p^* will remain within a distance ε of p^* for all $t > 0$; the equilibrium can be stable without being asymptotically stable—it just has to remain close. Asymptotic stable equilibria are stable, but not conversely. An equilibrium that is not locally stable is called **unstable**.

For equilibria of the FHW equation (1.7) we must have $\Delta p = 0$. Thus, $p^* = 0$ and $p^* = 1$ are always equilibria. To have a different equilibrium, not occurring at one of the extreme values, we need $p(w_x - w_y) + q(w_y - w_z) = 0$. This can only occur iff $w_x - w_y$ and $w_y - w_z$ have *opposite* signs. Therefore,

[1] We use the standard convention to abbreviate 'if, and only if' by 'iff'.

an internal equilibrium will occur iff $w_x < w_y$ and $w_y > w_z$, or $w_x > w_y$ and $w_y < w_z$. We formally record this result:

Theorem 1.6

The FHW equation (1.7) has equilibria $p^* = 0$ and $p^* = 1$; An internal equilibrium p^* (not equal to 0 or 1) of the FHW equation exists iff

$$w_x < w_y,\ w_y > w_z,\ \text{or}\ w_x > w_y,\ w_y < w_z, \tag{1.9}$$

and its value is

$$p^* = \frac{w_z - w_y}{w_x - 2w_y + w_z}. \tag{1.10}$$

If an internal equilibrium exists, it is unique.

Example 1.7

For an allele a that is completely dominant and there is selection for it (see Example ref) we have $w_x = w_y = 1 + s, \quad w_z = 1$. Therefore the condition (1.9) is not met, and therefore $p^* = 0, 1$ are the only equilibria. From the FHW equation in this case, equation (1.8), we see that $f(p) > p$, and so the curve $f(p)$ lies above the diagonal in the pp' plane. A cobweb shows that $p^* = 0$ is unstable and $p^* = 1$ is asymptotically stable. Often a graphical argument is easier to make than check the analytic stability criteria. Plots of both p and $f(p)$ can be sketched with the following MATLAB commands, where the user enters the formula for $f(p)$:

```
    p=0:0.01:1;
pprime=f(p);
plot(p,p,p,pprime)
```

One can sketch the cobweb from this plot and identify equilibria. There is software (e.g., on a TI-84 Plus) that plots a cobweb automatically.

Using the analytic stability criteria, we can make some general statements about the stability of 0 and 1. First we write the FHW equation (1.7) as

$$\begin{aligned} p' &= f(p) = p + pq\frac{p(w_x - w_y) + q(w_y - w_z)}{\overline{w}} \\ &= p\left(1 + (1-p)\frac{p(w_x - w_y) + (1-p)(w_y - w_z)}{\overline{w}}\right) \\ &= p(1 + (1-p)F(p)), \end{aligned}$$

where

$$F(p) = \frac{p(w_x - w_y) + (1-p)(w_y - w_z)}{\overline{w}}.$$

Taking the derivative using the product rule gives

$$f'(p) = p\frac{d}{dp}\left(1 + (1-p)F(p)\right) + \left(1 + (1-p)F(p)\right). \tag{1.11}$$

When we evaluate at $p = 0$, the first term vanishes and so

$$f'(0) = -(1 + F(0) = 1 + \frac{w_y - w_z}{w_z} = \frac{w_y}{w_z}.$$

Therefore, $p = 0$ is locally asymptotically stable if, and only if, $w_y < w_z$, or the fitness of the heterozygote ab is lower than the fitness of the homozygote bb. To evaluate the derivative (1.11) at $p = 1$, we note that the second term vanishes, but we still have to carry out the derivative in the first term. Notice that

$$\frac{d}{dp}\left(1 + (1-p)F(p)\right) = (1-p)F'(p) - F(p).$$

Evaluating at $p = 1$ gives $\frac{d}{dp}\left(1 + (1-p)F(p)\right)|_{p=1} = -F(1)$. Therefore, from (1.11),

$$f'(1) = -F(1) + 1 = -\frac{w_x - w_y}{w_x} + 1 = \frac{w_y}{w_x}.$$

Therefore, $p^* = 1$ is asymptotically stable if, and only if, $w_y < w_x$, or the fitness of the heterozygote ab is less than the heterozygote aa.

Therefore we have established the following result.

Theorem 1.8

(Stability) The values $p^* = 0$ and $p^* = 1$ are equilibria for the FHW equation (1.7). (a) $p^* = 0$ is locally asymptotically stable iff $w_y < w_z$, and $p^* = 1$ is locally asymptotically stable iff $w_y < w_x$. (b) If $w_y > w_z$, the $p^* = 0$ is unstable, and if $w_y > w_x$,then $p^* = 1$ is unstable.

We can make a statement about the stability of $p^* = 0, 1$, when an internal equilibrium p^* exists. To this end, note that conditions (1.9) must hold. In comparing these conditions with the stability conditions in the preceding theorem, we easily obtain the following result.

Theorem 1.9

If an internal equilibrium p^* exists, then both $p^* = 0$ and $p^* = 1$ must be stable, or both $p^* = 0$ and $p^* = 1$ must be unstable.

The top row of fig. 1.5 shows the $p'p$ diagram in the two cases when an internal equilibrium does not exist. The bottom row of fig. (phasefig.eps) shows two cases when an internal equilibrium exists. The result below will confirm the two possibilities shown in this latter case. We examine the internal equilib-

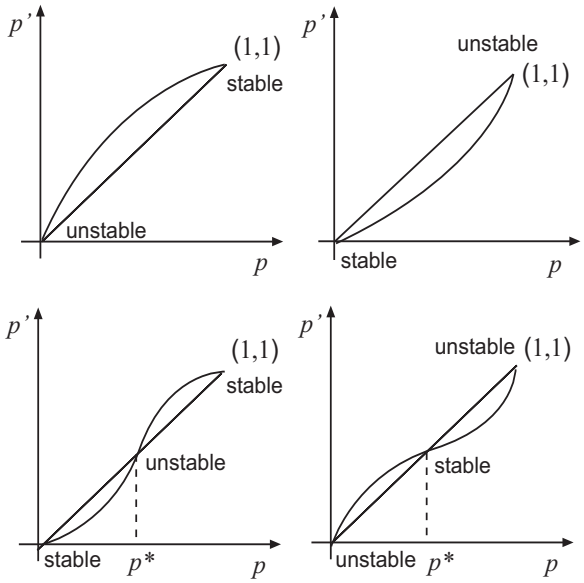

Figure 1.5 Four possible configurations of the $p'p$ phase plane. On the top row there is no internal equilibrium, and one equilibrium is stable and one unstable. On the bottom row there is an interior equilibrium that is either stable or unstable. When it is stable, both endpoints are unstable, and when it is unstable, both endpoints are stable.

rium p^* given by (1.10) in fig. (1.10). The derivation of the stability result is facilitated by scaling the fitness coefficients by the fitness of the heterozygote ab. Thus, we choose $w_y = 1$. Then we can write the FHW equation as

$$p' = f(p) = p + p(1-p)\frac{p(w_x - 1) + (1-p)(1-w_z)}{\overline{w}}.$$

We show the following fundamental result:

Theorem 1.10

The internal equilibrium p^*, when it exists, is asymptotically stable in the case

$p^* = 0, 1$ are both unstable, and unstable when $p^* = 0, 1$ are both asymptotically stable.

Proof

Although it seems to be a formidable task, we compute $f'(p)$ and evaluate it at the internal equilibrium. It is not as bad as it first appears. Let us write the the right side of the FHW equation as

$$f(p) = p + p(1-p)\frac{N(p)}{\overline{w}},$$

where N is the numerator $N(p) = p(w_x - 1) + (1-p)(1-w_z)$. Notice that $N(p^*) = 0$. By the product rule,

$$\begin{aligned} f'(p) &= 1 + (1-p)\frac{N}{\overline{w}} - p\frac{N}{\overline{w}} + p(1-p)\frac{\overline{w}\frac{dN'}{dp} - N(p)\frac{d\overline{w}}{dp}}{\overline{w}^2} \\ &= 1 + (1-p)\frac{N}{\overline{w}} - p\frac{N}{\overline{w}} + p(1-p)\frac{\overline{w}(w_x + w_z - 2) - N(p)\frac{d\overline{w}}{dp}}{\overline{w}^2}. \end{aligned}$$

Thus, evaluating at $p = p^*$, we get

$$f'(p^*) = 1 + p^*(1-p^*)\frac{w_x + w_z - 2}{\overline{w}},$$

where $\overline{w}$ is evaluated at p^*. There are two cases when the internal equilibrium exists:

$$w_x < 1,\ 1 > w_z,\ \text{or}\ w_x > 1,\ 1 < w_z.$$

The first is when the fitness of the heterozygote is larger than either homozygote, and the second is when the fitness of the heterozygote is less than the fitnesses of both homozygotes. In the second case, when $w_x > 1,\ 1 < w_z$, the numerator $w_x + w_z - 2 > 0$, and thus $f'(p^*) > 1$. Therefore, when both endpoints are asymptotically stable, the internal equilibrium is unstable.

Now we consider the harder case $w_x < 1,\ 1 > w_z$, where both both endpoints are unstable. In this case, $w_x + w_z - 2 < 0$, and $f'(p^*) < 1$. But we still need to show that $f'(p^*) > -1$. To this end, we obtain a bound on the negative fractional term

$$\frac{w_x + w_z - 2}{\overline{w}}.$$

This term will have its largest absolute value when the numerator is largest and the denominator is smallest. Both of these occur when the fitnesses w_x and w_z approach zero. In this limit, $w_x + w_z - 2 \to -2$ and $\overline{w} = w_x p^{*2} + 2p^*(1-p^*) + w_z(1-p^*)^2 \to 2p^*(1-p^*)$. Therefore,

$$p^*(1-p^*)\frac{w_x + w_z - 2}{\overline{w}} > p^*(1-p^*)\frac{-2}{2p^*(1-p^*)} > -1.$$

Therefore, $f'(p^*) > 0 > -1$. This shows p^* is asymptotically stable in the case both endpoints are unstable, which completes the proof. □

1.3.2 Allele Ratios

Sometimes it is easier to analyze the behavior of systems if we write the FHW equation in terms allele ratios $u = p/q$ or $v = q/p$. This is especially true in determining the asymptotic behavior (or, the speed of the approach to a stable equilibrium or away from an unstable equilibrium) of a system near $p^* = 0$ or $p^* = 1$. Why do we want to know this? Suppose at time $t = 0$ a deleterious allele, say, for a disease, enters the population. Clearly, determining how fast it spreads into the entire population is an important issue. Similarly, knowing how fast it goes to an equilibrium or dies out is an equally important issue.

To measure speed of approach or departure from an equilibrium, we use comparison functions. For example, suppose a function of t approaches 0 as $t \to \infty$. The function can approach 0 like e^{-at} $(a > 0)$, which is exponentially fast, or it can approach 0 like the power function $1/t^a$, which is algebraically fast. We learn in calculus that exponential functions decay faster than any power function. Similarly, exponential growth is much faster than algebraic growth; for example, e^{at} grows much faster than t^n for any n. If we can show that a function behaves like λ^t, as $t \to \infty$, where $0 < \lambda < 1$, then we have geometric decay; because $\lambda^t = (e^{\ln \lambda})^t = e^{(\ln \lambda)t}$, with $\ln \lambda < 0$, geometric decay is exponential decay.

Some examples illustrate how to proceed. First, we derive the following ratio forms of the FHW equation.

Theorem 1.11

In terms of the ratios u and v of allele frequencies, the FHW equation may be written

$$\Delta u = u' - u = u\frac{(w_x - w_y)u + w_y - w_z}{uw_y + w_z}, \tag{1.12}$$

$$\Delta v = v' - v = v\frac{(w_z - w_y)v + w_y - w_x}{vw_y + w_x}, \tag{1.13}$$

Proof

We use the two equations (1.4) and (1.6), which we write in the form

$$p' = \frac{pw_p}{\overline{w}}, \quad q' = \frac{qw_q}{\overline{w}}.$$

We prove (1.12) and leave (1.13) as an exercise, which is similar. Therefore,

$$\begin{aligned}\Delta u &= u' - u = \frac{p'}{q'} - u \\ &= u\left(\frac{w_p}{w_q} - 1\right) = u\frac{w_p - w_q}{w_q} \\ &= u\frac{pw_x + qw_y - pw_y - qw_z}{pw_y + qw_z}.\end{aligned}$$

Dividing the numerator and denominator by q gives equation (1.12). □

Example 1.12

Consider the spread of an allele a into a population where where the selection coefficients are $w_x = 1 + 2s$, $w_y = 1 + s$, $w_z = 1$. The FHW equation is

$$p_{t+1} = p_t + \frac{sp_t(1 - p_t)}{1 + sp_t}. \tag{1.14}$$

It is an easy exercise to see that $p^* = 0$ and $p^* = 1$ are the only equilibria, with 0 being unstable and 1 being asymptotically stable. Let us first deal with the rate of departure from the unstable equilibrium $p^* = 0$. Assuming p_0 is small and near 0, we can assume the frequency p_t is near 0 for for t near 0. We can try to approximate the right side of (1.14) for small p_t. One tool to accomplish this is to use the geometric series

$$\frac{1}{1 + z} = 1 - z + z^2 - z^3 + \cdots,$$

which is valid for $|z| < 1$. Applying this to the fraction on the right side of (1.14), we get

$$\frac{1}{1 + sp_t} = 1 - sp_t + (sp_t)^2 - (sp_t)^3 + \cdots.$$

Thus,

$$\begin{aligned}p_{t+1} &= p_t + sp_t(1 - p_t)(1 - sp_t + (sp_t)^2 - (sp_t)^3 + \cdots) \\ &= (1 + s)p_t - s(1 + s)p_t^2 + \mathrm{O}(p_t^3).\end{aligned}$$

Therefore, to leading order,

$$p_{t+1} = (1 + s)p_t.$$

This is the geometric growth equation and it has the simple solution

$$p_t = p_0(1 + s)^t.$$

Thus p_t grows geometrically since $1+s > 1$. So, the departure from equilibrium is exponentially fast. The allele a enters the population quickly.

Next we examine convergence to 1. It is hard to expand the right side of (1.14) for p_t near 1. It is better to try a ratio u_t or v_t and determine how the ratio converges; that will tell us how p_t converges. Notice, as $p_t \to 1$, we have $u_t \to \infty$ and $v_t \to 0$. Either will work here, and sometimes one is easier to deal with than the other. Here, let's use the v ratio. From (1.13) the FWH equation is

$$\begin{aligned} v_{t+1} &= v_t - v_t \frac{s(1+v_t)}{(1+s)v_t + (1+2s)} \\ &= v_t - v_t s(1+v_t) \frac{1}{1+2s} \frac{1}{1+\frac{1+s}{1+2s}v_t}, \end{aligned}$$

where we have prepared the equation to use the geometric series. Letting

$$\lambda = \frac{1+s}{1+2s} < 1,$$

we have

$$\begin{aligned} v_{t+1} &= v_t - v_t(1+v_t)\frac{s}{1+2s}\left(1 - \lambda v_t + \mathrm{O}(v_t^2)\right) \\ &= v_t\left(1 - \frac{s}{1+2s} + \mathrm{O}(v_t)\right) \\ &= \frac{1+s}{1+2s} v_t = \lambda v_t, \end{aligned}$$

to leading order. Therefore, to leading order,

$$v_t = v_0 \lambda^t, \quad v_0 = \frac{q_0}{p_0}.$$

Therefore, v_t decays to 0 exponentially. This means, to leading order,

$$\frac{q_t}{1-q_t} = v_0 \lambda^t.$$

Solving for q_t gives

$$\begin{aligned} q_t &= \frac{v_0\lambda^t}{1+v_0\lambda^t} = v_0\lambda^t\left(1 - v_0\lambda^t + \mathrm{O}(\lambda^t)^2\right) \\ &= v_0\lambda^t, \end{aligned}$$

to leading order. Thus,

$$p_t - 1 = v_0\lambda^t,$$

as $t \to \infty$. This means $p_t \to 1$ exponentially fast.

1.3.3 The Course of Evolution

A question that naturally arises is: does selection, as we have modeled it, push evolution is any preferred direction? For example, does it automatically optimize the fitness of the population? Does it make the population grow more rapidly? Do the fittest adapt more quickly to change? These are the key questions in Darwin's view of evolution, that selection uses genetic variation to pick those organisms that will adapt better to changes in their environment.

How do we measure adaptation? Well, we have measured fitness of various genetic traits in the population, and how their frequencies in the gene pool evolve over time. Fitness is a good place to start. Therefore, we ask, how does the average fitness of a population change over time? Does selection cause it to increase?

The mean fitness of a population is

$$\overline{w} = \overline{w}(p) = w_x p^2 + 2pqw_y + w_z q^2, \quad q = 1 - p.$$

S. Wright's view in the 1930s was that $\overline{w}$, viewed as a function of the allele frequency p, was a landscape, or **adaptive topography**. A way to measure a system's view of where to go next is to see how it changes from its current state p to the future, or how it responds to increases in p. Let us calculate what happens when p changes to $p + \Delta p$.

The FHW equation (1.4) gives

$$\begin{aligned} \Delta p &= \frac{pw_x + qw_y}{\overline{w}} - p \\ &= \frac{pw_x + qw_y - p\overline{w}}{\overline{w}}. \end{aligned}$$

Substituting for $\overline{w}$ and carrying out a lot of algebra gives

$$\Delta p = \frac{pq}{\overline{w}} \left(p(w_x - w_y) - q(w_z - w_y) \right).$$

Now, taking the derivative of $\overline{w}(p)$ with respect to p yields, after simplification,

$$\frac{d\overline{w}}{dp} = 2 \left[p(w_x - w_y) - q(w_z - w_y) \right].$$

Combining these last two expressions gives a fundamental result in population genetics:

Theorem 1.13

Under the assumptions of natural selection,

$$\Delta p = \frac{pq}{2\overline{w}} \frac{d\overline{w}}{dp}. \tag{1.15}$$

Equation (1.15) relates changes in allele frequency to the slope of the mean fitness curve, or adaptive topography. Clearly, $\overline{w}$ increases as p increases, and $\overline{w}$ decreases as p decreases. At a local minimum or maximum, $\Delta p = 0$ and there is an equilibrium at that value of p. Figure 1.6 shows two of the cases, where the heterozygote is most fit (left) with $w_x = 1 - s$, $w_y = 1$, $w_z = 1 - 2s$, and where the heterogygote is least fit (right) with $w_x = 1 + s$, $w_y = 1$, $w_z = 1 + s$. It is obvious that we *cannot* conclude that $\overline{w}$ always increases, or that selection always acts to increase the mean fitness. However, we do observe that we can approximate (1.15) by

$$\Delta p = \frac{pq}{2\overline{w}} \frac{\Delta \overline{w}}{\Delta p},$$

or

$$\Delta \overline{w} = \frac{2\overline{w}(\Delta p)^2}{pq} \geq 0.$$

Therefore, selection does act at each value of p to insure that the *change* in the mean fitness is always nonnegative.

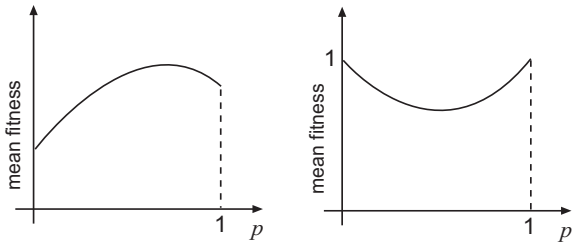

Figure 1.6

Example 1.14

In the case of selection against a recessive allele, with $w_x = 1 + s$, $w_y = 1 + s$, $w_z = 1$, the stable equilibrium is at $p = 1$ and the fitness diagram is shown in fig. 1.7.

EXERCISES

1. Derive Equation (1.13).
2. Find the following selection models, find the FHW equation and its equilibria, and determine the stability of the equilibria; draw a generic cobweb indicating the behavior of the solution.

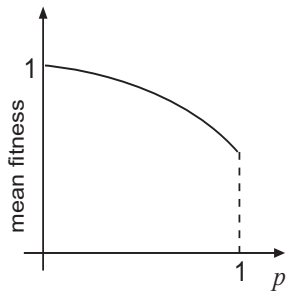

Figure 1.7

a) Selection for a recessive allele, where a is recessive and advantageous: $w_x = 1 + s$, $w_y = w_z = 1$.

b) Selection for an advantageous allele a that is not completely dominant: $w_x = 1 + 2s$, $w_y = 1 + s$, $w_z = 1$.

c) Selection for a deleterious allele a where: $w_x = 1$, $w_y = 1 + s$, $w_z = 1 + 2s$.

d) Selection where $w_x = 1$, $w_y = 1 + s$, $w_z = 1 - r$, where $0 < r \leq 1$, $s > 0$. Find the FHW equation and determine the equilibria. Draw a generic cobweb and discuss stability.

3. In the previous exercise assume $w_z = 0.2$ and the system has the equilibrium $p^* = 0.8$, $q^* = 0.2$. Find the selection coefficients r and s, and find the relative fitness of the aa genotype.

4. Consider selection favoring the homozygote with $w_x = 1 - s$, $w_y = 1$, $w_z = 1 - s$. Sketch time series p_t vs. t for 450 generations, taking $s = 0.025$, 0.05, 0.1, 0.25, 0.5, 1, with $p_0 = 0.95$. Repeat for $p_0 = 0.05$. Hint: Use the MATLAB code *FWHequation,* inserting the commands in a 'for $s = [0.025$, 0.05, 0.1, 0.25, $0.5, 1]$ ...end' loop. Within a loop, place 'hold on' after the plot command, and after all the loops are completed, type 'hold off'.

5. Repeat the last problem with selection against the heterozygote with relative fitnesses $w_x = 1$, $w_y = 1 - s$, $w_z = 1$.

6. For selection with fitnesses $w_x = 0$, $w_y = w_z = 1$, find the FHW equation for the allele frequency p_t and solve the equation. (Hint: let $y_t = 1/p_t$.)

7. For selection with fitnesses $w_x = 1$, $w_y = w_z = 1 + s$, show that the u-ratio equation is

$$u' = u + \frac{su}{1 + (1 + s)u},$$

and that approximately, for large u,

$$u' = u + \frac{s}{1+s}.$$

Solve this difference equation and obtain the result that $q_n \to 0$ algebraically as $n \to \infty$.

8. Consider a selection process with fitness coefficients $w_x = 1 + s$, $w_y = 1$, $w_z = 1$. Show that $p = 0$ is an unstable equilibrium. Show that the rate of departure from $p = 0$ cannot be calculated using a geometric series expansion on the FHW equation for small p. However, show that the departure is algebraic using the FHW equation for the ratio $v = q/p$. Show that the approach to $p = 1$ is exponential.

9.

10. Consider a selection process with fitness coefficients $w_x = 1 + s$, $w_y = 1$, $w_z = 1$. For different values of the selection coefficients, plot the evolution of the allele a over time. Compare these results with with mutation alone.

1.4 Mutations and Selection

There is some ambiguity about the a definition of a mutation. If a mutation is defined as a change in the genetic code, then recombination (for example, chromosomal cross-over) is a mutation. But many reserve the term mutation for actual changes in allele frequency, in which case recombination is not considered a mutation.

One can think of meiosis as introducing two non-mutational mechanisms for producing novelty by rearranging existing genetic information. The first is crossover, and second is the random alignment of chromosomes in meiosis I (for two pairs of homologous chromosomes, it is random which pair of non-homologous chromosomes end up in cells together after the first division).

In some models of mutation, neither of these processes are are assumed to occur. Rather, mutation occurs as DNA replicates to produce gametes, for example, by a base substitution in copying a sequence; a base substitution can be a transition where a purine (A or G) is interchanged with the other purine, or a pyrimidine (T or C) is interchanged. A transversion is a replacement of a purine (A or G) by a pyrimidine (T or C), or vice-versa[2] An excellent, elementary discussion of modeling molecular evolution can be found in Allman & Rhodes (2004), Chapter 4.

[2] .

1.4.1 Mutation

We focus here on a simple model. We assume **recurrent mutations** through the generations with *no selection*; this is the Hardy–Weinberg case. We assume for this model that alleles mutate to other alleles already in the population, and new alleles are not produced nor migrate into the population ('novel mutations'). We consider alleles at a locus which a or b can occupy, and we assume during a single generation a is replaced by b with probability u, and b is replaced by a with probability v. Generally, u and v are very small probabilities, perhaps 1 in a million, or on the order 10^{-6}. If p_t and q_t denote, as usual, the frequency of a and b, respectively, at time t, then

$$\begin{aligned} p_{t+1} &= \text{alleles } a \text{ that did not mutate } + \text{ allele } b \text{ that mutated to a} \\ &= (1-u)p_t + vq_t. \end{aligned} \tag{1.16}$$

This is the governing equation for recurrent mutation. It is easy to solve for p_t. Note that, upon replacing q_t by $1 - p_t$, we get

$$p_{t+1} = (1 - u - v)p_t + v,$$

which is the geometric decay equation ($1 - u - v < 0$) with a source term v. The equilibrium solution is clearly

$$p^* = \frac{v}{u+v}.$$

The general solution is the sum of the general solution to the homogeneous equation and a particular solution; thus,

$$p_t = C(1 - u - v)^t + p^*.$$

To determine the the arbitrary constant C we set $t = 0$ to get

$$C = p_0 - p^*.$$

Therefore

$$p_t = (p_0 - p^*)(1 - u - v)^t + p^*.$$

Notice that, because $1-u-v$ is very close to 1, the convergence to equilibrium p^* is not very fast. For example, if $u = v = 0.000001$, then $1-u-v = 0.999998$. Clearly, 0.999998^t converges to zero extremely slowly. Therefore, recurrent mutations take many generations to have an effect. For example, for $0.999998^t = \frac{1}{2}$, we get

$$t = -\frac{\ln 2}{\ln 0.999998} \simeq 346,573 \text{ generations.}$$

In comparison, selection usually acts on a much shorter time scale, especially strong selection, as measured by the selection coefficient s. In the case selection

is weak (small s with relative fitness coefficients close to one) the time scale for change can be more comparable to mutation time scale. A MATLAB code that implements the preceding recurrent mutation over 10^5 generations is given below.

```
    function mutation
u=0.000001; v=0.000001; generations=100000; p0=0.05; pstar=v/(u+v);
t=0:generations;
p=(p0-pstar)*(1-u-v).^ t +pstar;
plot(t,p)
```

1.4.2 Mutation with Selection

Now we include selection along with mutation. We recall that selection can occur as zygotes mature to adults (survivorship), and it can occur from the breeding adult stage to the production of gametes in the gamete pool for the next generation (fecundity or fertility). Where in this process does mutation occur? The key in the simple model we examine is that the only selection taken into account is differential survivorship from the zygote to breeding adult stage, and this selection is followed by a random mutation, with no fertility. Therefore, the fitness coefficients w are the relative probabilities of different genotypes surviving from zygote to breeding. We are assuming that fertility is independent of genotype, or all individuals produce the same number of gametes, and different combinations of gametes have the same probability of producing healthy zygotes. So really the w's are 'relative survivorships', not fitnesses. Interpreted differently, with the assumptions above, relative survivorships and relative fitness are the same. In a more general and realistic model, fertility would be included and mutation would occur in a complicated mix with differential gamete production.

In an individual's life, we might conclude that selection acts on the mutations that occur. but, from the perspective of keeping track of alleles, selection happens first to determine which individuals produce how many gametes, and then mutation occurs in the reproduction process. In summary, 'we mutate the selected values', or

$$p_{\text{mut}} = \frac{pw_p}{\overline{w}} \text{ and } q_{\text{mut}} = \frac{qw_q}{\overline{w}}$$

to get

$$p' = (1-u)\frac{pw_p}{\overline{w}} + v\frac{qw_q}{\overline{w}}.$$

This is the fundamental equation.

Example 1.15

Consider the case of mutation along with selection against a deleterious allele a, where

$$w_x = 1 - s, \quad w_y = w_z = 1, \quad s \ll 1, \quad u, v \ll s.$$

We determine the equilibrium value p^*. For equilibrium we require p^* satisfy the equation

$$p = (1-u)\frac{pw_p}{\overline{w}} + v\frac{qw_q}{\overline{w}},$$

or

$$(w_p - \overline{w})p - upw_p + vqw_q = 0.$$

Substituting values for the selection coefficients and simplifying gives

$$sp^3 + usp^2 - sp^2 - up - vp + v = 0. \tag{1.17}$$

This is a cubic equation for the equilibrium value, and cannot readily be solved exactly. So we will seek a leading order solution. For definiteness, assume

$$u, v \sim \mathrm{O}(\varepsilon^2), \quad s \sim \mathrm{O}(\varepsilon), \quad \text{as } \varepsilon \to 0.$$

So, ε is a small parameter, and we can determine the order of the coefficients in the cubic (1.17). The question is to determine the leading order approximation of the solution p. This problem is one of singular perturbation theory, a full discussion of which is contained in Logan (2008). Here, however, we proceed directly without the general theory. We assume $p \sim \mathrm{O}(\varepsilon^a)$, for some $a \geq 0$. To determine the value of a we look at the order of the six terms in the cubic (1.17). The orders are, from first to last,

$$\varepsilon^{1+3a},\ \varepsilon^{3+2a},\ \varepsilon^{1+2a},\ \varepsilon^{2+a},\ \varepsilon^{2+a},\ \varepsilon^2.$$

We want to make a simplification, so we first look for the dominant balance between two terms in the equation. This means we want two terms that are the same order with the remaining four terms small in comparison. There are $\binom{6}{2}$ possible choices, or fifteen. All but one of these choices will lead to a contradiction. For example, if the first two balance and the remaining four are small, then $1+3a = 3+2a$, giving $a < 0$, which is a contradiction. If the first and third form the dominant balance, then $1 + 3a = 1 + 2a$, giving $a = 0$ and $p \sim \mathrm{O}(1)$ This means the first and third terms are order ε. The remaining four terms are smaller order, but the leading order behavior implies $sp^3 \approx sp^2$, which means $p = 1$, again giving a contradiction because $p = 1$ is not a root of (1.17). Other cases are similar, giving a contradiction, but one. In this one case we try a balance between the third and last terms. We get $1 + 2a = 2$, or $a = \frac{1}{2}$. This means $p \sim \mathrm{O}(\sqrt{\varepsilon})$, and the third and last terms are order ε^2. The remaining

four terms are order $\varepsilon^{5/2}$, ε^4, $\varepsilon^{5/2}$, $\varepsilon^{5/2}$, which are all of smaller order than ε^2. This is the dominant balance. Therefore, approximately, $sp^2 \approx v$, or

$$p = p^* \approx \sqrt{\frac{v}{s}}.$$

This is the leading order approximate solution to equation (1.17), which gives the equilibrium.

1.4.3 Weak Selection Approximation

In many evolutionary changes, the selection coefficient s is on the order of 10^{-3}, or 1000 years for changes to occur. Therefore we can regard s as a small parameter in the FHW model and make approximations to leading order in s. This is the weak selection approximation.

To be specific, let us assume in a given model that the relative fitness coefficients are

$$w_x = 1 + hs, \quad w_y = 1 + ks, \quad w_z = 1, \quad s \ll 1,$$

where g and h are order 1 constants. Then the mean fitness is

$$\begin{aligned} \overline{w} &= (1 + hs)p^2 + 2(1 + ks)pq + q^2 \\ &= p^2 + 2pq + q^2 + \mathrm{O}(s) = 1 + \mathrm{O}(s). \end{aligned}$$

Then the FHW equation is

$$\begin{aligned} \Delta p &= p' - p = pq\frac{p(w_x - w_y) + q(w_y - w_z)}{\overline{w}} \\ &= pq\frac{[(h-k)p + kq]s}{1 + \mathrm{O}(s)} \\ &= pq[(h-k)p + kq]s + \mathrm{O}(s^2). \end{aligned}$$

Because of weak selection we expect very small changes in p over very small changes in time. Therefore, it is reasonable to use small time steps Δt rather than unit time steps in the difference equation. Thus $\Delta p = p_{t+\Delta t} - p_t$, and we observe that the last equation is a difference approximation (the Euler approximation) to the differential equation

$$\frac{dp}{dt} = pq[(h-k)p + kq]s, \tag{1.18}$$

where we have ignored the $\mathrm{O}(s^2)$ terms. This equation can be solved by separation of variables to get

$$\int_{p_0}^{p} \frac{d\rho}{\rho(1-\rho)[(h-k)\rho + k(1-\rho)]} = s(t - t_0).$$

The integral can actually be resolved using a partial fraction expansion of the integrand.

Example 1.16

Consider the case when $h = 2$ and $k = 1$. Then (1.18) becomes

$$\frac{dp}{dt} = sp(1-p),$$

which is the logistic equation. This equation can be solved exactly.

Example 1.17

In the case of an advantageous, dominant allele a, we have $h = k = 1$. The the differential equation (1.18) is

$$\frac{dp}{dt} = sp(1-p)^2.$$

If, for example, in 350 generations there is a change in p from 0.01 to 0.09, then

$$\int_{0.01}^{0.09} \frac{d\rho}{\rho(1-\rho)^2} = 350s.$$

Calculating the integral numerically using a calculator gives 15.78. Therefore the selection coefficient is

$$s = 0.045.$$

EXERCISES

1. For $s = 0.045$, solve the exact FHW difference equation for the problem in Example 1.17, and then use Euler's method to numerically solve the approximating differential equation with a small step size $\Delta t = 0.0001$. Compare the two solutions and the values of p at $t = 350$.

2. Obtain the weak approximation in the case the heterozygote being dominant over the two homozygotes:

   $$w_x = 1, \quad w_y = 1 + ks, \quad w_z = 1, \quad s \ll 1.$$

   Draw the phase line for the differential equation approximation (1.18) and discuss the dynamics with regard to equilibria and their stability.

3. In mutation and selection, making the same order assumptions as in Example 1.15, show that if the allele a is dominant, then its equilibrium value is, to leading order,

   $$p^* = \frac{v}{s}.$$

1.5 Density Dependent Selection

So far our discussion of selection has ignored the effects of population size. We have only tracked allele frequencies, using the relative fitness coefficients, w_x, w_y, w_z, which were assumed to be constant. Yet we know from elementary studies in population dynamics that a populations growth can be limited by population size, or density dependent growth. For example, in the logistic model for population growth, the intraspecific competition for resources limits the growth rate as the population increases. Now we want to build this type of competition into the natural selection model, where the fitness coefficients for the various alleles, or traits, depend on the total population. This concept is **density-dependent selection**.

To this end, we will use absolute fitness coefficients

$$W_x = W_x(N), \quad W_y = W_y(N), \quad W_z = W_z(N),$$

where $N = N_t$ is the total population at time t. This means that the population in the next generation is now, in terms of the definitions in Section 1.2, that the survivorships and fecundities are now functions of the total population. Therefore, the total population at the next generation is

$$N' = (W_x(N)p^2 + 2W_y(N)pq + W_z(N)q^2)N,$$

where the average absolute fitness is

$$\overline{W}(N,p) = W_x(N)p^2 + 2W_y(N)pq + W_z(N)q^2.$$

Further, the allele frequency equation becomes

$$p' = \frac{W_x(N)p^2 + W_y(N)pq}{\overline{W}(N,p)}.$$

Written out in difference equation form, the governing equations are

$$p_{t+1} = \frac{W_x(N_t)p_t^2 + W_y(N_t)p_tq_t}{\overline{W}(N_t,p_t)}, \tag{1.19}$$

$$N_{t+1} = \overline{W}(N_t,p_t)N_t, \tag{1.20}$$

which is a two-dimensional system of nonlinear difference equations.

We can study (1.19)–(1.20) in the usual way (for example, see Chapter 3 of Logan & Wolesensky, 2009). The equilibria are solutions to

$$p^* = \frac{W_x(N^*)p^{*2} + W_y(N^*)p^*q^*}{\overline{W}(N^*,p^*)},$$

$$N^* = \overline{W}(N^*,p^*)N^*,$$

or

$$\overline{W}(N^*, p^*) = 1, \quad W_x(N^*)p^* + W_y(N^*)q^* = 1.$$

To make further analytical progress we have to make some assumptions about the dependence of the fitness coefficients on the total population size. We expect that fitness of all the genotypes decrease with increasing population. One assumption is to take logistic-type dependence

$$W_x(N) = 1 + r_x - \frac{r_x}{K_x}N, \quad W_y(N) = 1 + r_y - \frac{r_y}{K_y}N, \quad W_z(N) = 1 + r_z - \frac{r_z}{K_z}N,$$

where the r's are the growth rates and the K's are carrying capacities for the different genotypes. Still, an analytic approach is untenable. Now, we can input numerical values for the growth and carrying capacities and study the evolution of the system (1.19)–(1.20) in the pN phase plane. The general case involves a complicate mix of both genotypes. But a system where there are tradeoffs among growth rates and carrying capacities has more interest and possibilities. This leads back to the classic ecological question of r selection vs. K selection.

Example 1.18

The following MATLAB m-file plots an orbit of (1.19)–(1.20) in the pN phase plane (N is scaled by the initial population) for given initial values and given growth and carrying capacities. In the case shown the heterozygote has a smaller growth rate than both homozygotes, but a larger carrying capacity. The plot, in fig. 1.8 shows a small initial dip in the population and then a gradual growth up to an equilibrium $p = 0.227$, $N = 1.38$.

```
function densitydepsel
rx=.1; ry=.09; rz=.1; kx=100; ky=150; kz=100;
p=0.2; q=1-p; pevolve=p; Ninit=125; N=Ninit; Nevolve=N;
generations = 200;
for t=1:generations
wx=1+rx-(rx/kx)*N; wy=1+ry-(ry/ky)*N; wz=1+rz-(rz/kz)*N;
w=wx*p.^2+2*wy*p.*q+wz*q.^2 ;
NN=w.*N; PP=(wx*p+wy*q).*p./w ;
N=NN; p=PP;
Nevolve=[Nevolve,N]; pevolve=[pevolve,p];
end
time=0:generations; plot(pevolve,Nevolve/Ninit,'LineWidth',1.0),
xlabel('allele frequency p','FontSize',14); ylabel('scaled population','FontSize',14);
```

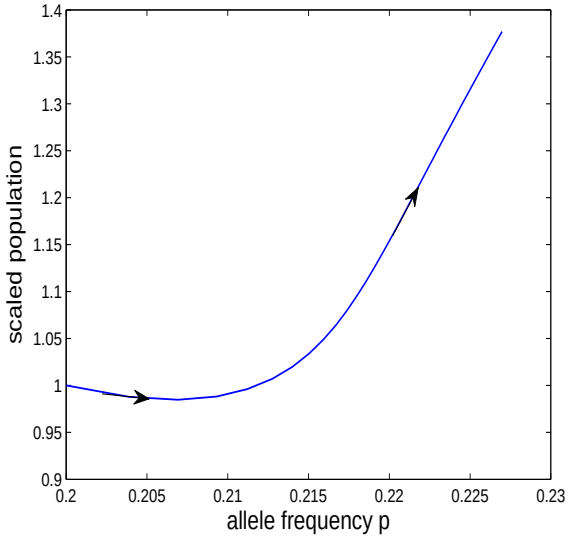

Figure 1.8 Numerical solution of the system (1.19)–(1.20) for fitnesses of logistic form. On the vertical axis the population is scaled by the initial population, in this case 125.

References

1. E. S. Allman & J. A. Rhodes. 2004. *Mathematical Models in Biology*, Cambridge University Press, Cambridge.

2. N. F. Britton, 2003. *Essential Mathematical Biology*, Springer-Verlag, London.

3. A. W. F. Edwards, 1977. *Foundations of Mathematical Genetics*, Cambridge University Press, Cambridge.

4. J. Hofbauer & K. Sigmund, 1998. *Evolutionary Games and Replicator Dynamics*, Cambridge University Press, Cambridge.

5. J. D. Logan, 2008. *Applied Mathematics*, 3rd ed., Wiley–Interscience, New York.

6. J. D. Logan & W. Wolesensky, 2009. *Methods of Mathematical Biology*, John Wiley & Sons, New York.

7. M. A. Nowak, 2006. *Evolutionary Dynamics*, Belknap/Harvard Univ. Press, Cambridge, MA.

8. J. Roughgarden, 1979. *Theory of Population Genetics and Evolutionary Biology*, Macmillan Pub. Co., Inc. New York.

9. J. Roughgarden, 1998. *Primer of Ecological Theory*, Prentice-Hall, Upper Saddle River, NJ.