Genetic theory of intra-specific competition

John Richard Sedcole
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Genetics Commons

Recommended Citation
https://lib.dr.iastate.edu/rtd/6622

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University Microfilms International
300 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND
SEDCOLE, JOHN RICHARD
GENETIC THEORY OF INTRA-SPECIFIC COMPETITION.
IOWA STATE UNIVERSITY, PH.D., 1979
Genetic theory of intra-specific competition

by

John Richard Sedcole

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Departments: Statistics
          Agronomy
Co-Majors: Statistics
           Plant Breeding

Approved:

In Charge of Major Work

For the Major Departments

For the Graduate College

Iowa State University
Ames, Iowa

1979
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAVEAT</td>
<td>iv</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1. Bulk-population Breeding</td>
<td>2</td>
</tr>
<tr>
<td>1.2. Selection Under Competition</td>
<td>7</td>
</tr>
<tr>
<td>1.3. Outline of the Study</td>
<td>7</td>
</tr>
<tr>
<td>2. LITERATURE REVIEW</td>
<td>10</td>
</tr>
<tr>
<td>2.1. Model Due to Sakai (1955)</td>
<td>10</td>
</tr>
<tr>
<td>2.2. Model Due to Schutz, Brim and Usanis (1968)</td>
<td>16</td>
</tr>
<tr>
<td>2.3. Model Due to Allard and Adams (1969)</td>
<td>20</td>
</tr>
<tr>
<td>2.4. Model Due to Cockerham and Burrows (1971)</td>
<td>23</td>
</tr>
<tr>
<td>2.5. Model Due to Mather (1969)</td>
<td>27</td>
</tr>
<tr>
<td>2.6. Model Due to Schutz and Usanis (1969)</td>
<td>30</td>
</tr>
<tr>
<td>2.7. Model Due to Cockerham et al. (1972)</td>
<td>33</td>
</tr>
<tr>
<td>2.8. Model Due to Nei (1971)</td>
<td>38</td>
</tr>
<tr>
<td>2.9. Truncation Selection</td>
<td>55</td>
</tr>
<tr>
<td>2.10. Selection with Reference to Biological Groups</td>
<td>62</td>
</tr>
<tr>
<td>3. COMPETITION IN GROUPS GREATER THAN SIZE TWO</td>
<td>73</td>
</tr>
<tr>
<td>3.1. Generalization of the Pairwise Model</td>
<td>73</td>
</tr>
<tr>
<td>3.2. Triplet Diploid Model with Constant Fitness</td>
<td>81</td>
</tr>
<tr>
<td>3.3. Constant Fitness Model for Groups of Size n</td>
<td>96</td>
</tr>
<tr>
<td>3.4. Haploid Genotypes in Competing Groups with Constant Fitness</td>
<td>105</td>
</tr>
<tr>
<td>3.5. Variable Fitness Competition Models</td>
<td>115a</td>
</tr>
</tbody>
</table>
CAVEAT

Before the advent of the digital computer the applied scientist would collect data and construct a model. If the model did not fit the data the scientists discarded it and began again. Today, if the model does not fit the data, the scientist often discards the data and publishes the model. . . . There is a growing tendency to take a set of differential equations so chosen that the computer can solve them and then to believe the results which emerge. . . .

F. N. David, 1975

Biometrics 31:253
1. INTRODUCTION

In this study we are concerned about the genetic interrelationships between organisms competing against each other. While most of the study is developed in terms applicable to plant breeding, certain aspects of the development are relevant also to natural populations of organisms like animals which also are subject to mutual intra-specific competition.

The particular relevance to plant breeding is two-fold. From a population genetics view, the studies on random-mating populations under intra-specific competition have models directly related to the problem of bulk-population breeding; the technique introduced and studied by Suneson (1949) and later dubbed "Evolutionary plant-breeding" (Suneson, 1956).

The second aspect concerns the problem facing plant breeders where plants, selected on the basis of their performance in competition against one set of genotypes, are then assessed against a background consisting of a different set of genotypes. It can be appreciated that this interaction mimics a genotype by environment interaction, itself the subject of much study. (For a review of this topic, see Freeman, 1973.) In certain instances negative gains from selection have been recorded (Wiebe, Petr and Stevens, 1963).
1.1. Bulk-population Breeding

To discuss the first aspect in more detail we turn to the first published studies in bulk-population breeding by Harlan and Martini (1938), who measured the changes occurring in a population synthesized equipropor­tionately from eleven barley varieties, sown in ten statewide locations, and studied over four to twelve years.

Barley, an autogenous plant in domesticated varieties, may outcross occasionally, but this was not held to be of significance in this study. The marked changes that occurred in the population, changes that differed from location to location, were ascribed to differences in competitive ability between the varieties. Generally it was shown that acknowledged good varieties for particular locations became predominant in the mixture grown there, while acknowledged poor varieties made up less and less of the bulk over the years the experiment was run.

Allard (1960) discussed these results, pointing out that while the proportion of good competitors would be expected to increase and that of poor competitors decrease, that of intermediate types, however, would tend to increase so long as the proportion of superior competitors remained relatively low. When this no longer is the case the proportion of intermediate types would decrease. These three types of response were, in fact, observed.

Following the study of Harlan and Martini (1938), Suneson and Wiebe (1942), and Suneson (1949) published similar studies made with four varieties of barley, grown both in mixtures and pure stands for
comparison from 1933 until 1948. The results summarized in Figure 1.1 show the distinct competitive superiority of Atlas in a mixture, and the competitive inferiority of Vaughn. The remarkable fact is, however, that in pure stands both in this and other trials, Vaughn is superior in yield to Atlas. The poorer performance of Vaughn in mixtures cannot be ascribed to differences in heading date, height, or disease reaction. Significantly, however, in spite of Vaughn's widely publicized superiority in pure stand trials, it has not been a popular variety, but the reasons for this lack of appeal were not, apparently, researched. As pure conjecture, it may have been that the weed-free conditions in extension trials were not repeated in commercial fields, and the consequent interspecific competition adversely affected the yield of Vaughn. It remains that the type of variety that seemed to do best in commercial conditions was the variety that overwhelmingly predominated in the mixtures.

Working on this result, Suneson (1956) reasoned that bulk-population breeding methods could be used to isolate potentially agronomically desirable varieties from a fairly heterogeneous original population. Suneson (1949, 1955) grew four Composite Cross varieties in separate bulk-population breeding programs. While all four commenced with yields well below that of Atlas (76 to 88%), the gains made by the four varieties, and the eventual, if erratic, superiority demonstrated by one of them was taken by Suneson (1956) to implicate bulk-population breeding as an inexpensive but effective
Figure 1.1. Graph of the proportion of each of four varieties of barley grown in a bulk population from 1933 to 1948 (Suneson, 1949)
way to develop breeding material from which good lines could be de-

Before leaving this discussion we refer to a paper by Donald
(1968) where the concept of breeding for a crop "ideotype" was dis-

cussed. Donald (1968) was mainly concerned with breeding wheat, but
his remarks may be applicable to cereal breeding generally. Donald
argues that a plant within a community will express its potential for
yield most fully if it suffers minimum interference from its
neighbors, i.e., its neighbors should be weak competitors. Thus the
crop ideotype that Donald envisages will yield well, not because the
individual plants compete strongly against each other, but the plants,
densely grown, but not being strongly suppressed by neighbors, exert
a strong pressure on the total resources. He points out that such a
crop must be kept weed-free.

Genotypes that are relatively poor competitors may express two
characteristic negative relationships; firstly, that between the
yield per unit area of the genotype at high density and at low
density, and secondly, that between the genotype yield in a mixture, and
its yield in a pure stand. The studies by Suneson (1949), already
discussed, were cited by Donald (1968) as a demonstration of this. Other
work (Jennings and de Jesus, 1968; Wiebe, Stevens and Petr, 1963) are
also held in support of Donald's thesis.

Returning, finally, to Allard's commentary (1960), we are reminded
that bulk-population breeding refers to cereal-type crops, where the yield is the reproductive seed, rather than some other organ of growth (roots, tubers, leaves, fruit flesh, etc.).

1.2. Selection Under Competition

The second aspect develops from the problems faced by Wiebe et al. (1963) who found that, owing to competitive effects, selection of high yielding types from a relatively heterogeneous genetic background often resulted in a negative response when the selected genotypes were assessed against their own genetic background. In fact, Wiebe et al. (1963) arrived at the somewhat paradoxical conclusion that "where high yield is the criterion for, say in the F6; and the selection is intended for use in pure stands, the instructions from the present study are that one should save the poorest plants from the F6 rather than the good ones". This paradox in the work just described was seized upon by Donald (1968) as evidence in support of his concept of breeding for crop varieties possessing low competitive ability. The same paradox was used by Griffing (1967) in the development of his selection in reference to biological groups. Griffing's work is discussed in more detail in the following work.

1.3. Outline of the Study

The study commences with a review of work published on inter- and intraspecific competition (Sakai, 1955). Later formulations (Schutz, Brim and Usanis, 1968; Schutz and Usanis, 1969) enabled
computer simulations to be made which showed the long term implications of competition (Allard and Adams, 1969).

A different approach used by Nei (1971) enabled some estimate of the number of generations necessary for a gene conferring greater competitiveness to replace another, or, alternatively, the excess fertility the more competitive genotype must possess in order for gene substitution to occur.

Chapter 3 extends the pair-wise concept of Nei (1971) to a three-way competition model, i.e., individuals compete in groups of three, and it is shown that the triplet group may possess traits (existence of equilibrium conditions) not evidenced by the model incorporating only pair-wise competition. Two models are developed and discussed, one in which the population fitness varies, and the other, closer to Nei's (1971) concept, with the mean fitness of the population fixed at unity.

A simplification of the model allows both a greater understanding of the dynamics of the model, and its extension to higher numbered groups.

A different approach to the problem of dealing with the complex interactions associated with groups of four or more individuals is introduced in Chapter 4 where the model is restricted to genotypes competing within indefinitely long rows. This model reflects field practice with many crops, particularly crops being assessed in a plant breeding program. With this model, ideas due to
Griffing (summarized in Griffing, 1977) are extended from his concept of finite fixed sized groups to indefinitely long rows. In generalizing the model in this fashion consequences of usual plant breeding practices (selection within full-sib rows, or half-sib rows, etc.) may be assessed. Furthermore, conditions for equilibrium, and the productive rate at equilibrium, may also be estimated, allowing an alternative, analytical approach to the study of bulk population breeding.
2. LITERATURE REVIEW

2.1. Model Due to Sakai (1955)

In the majority of the genetical studies, the assumption is made, usually implicitly, that genotypes do not interact (Mather, 1969). An attempt to move away from this assumption was made by Sakai (1955) in his study of competitive interaction between and within varieties of rice. Sakai arranged plants in groups of seven; a central plant and a hexagonal arrangement of surrounding competing plants. In Figure 2.1 the plant being measured is marked o, the similar competing plants similarly marked, and the dissimilar competing plants marked ●.

![Figure 2.1. Arrangements of competing plants in groups of seven. Dissimilarly marked plants are of dissimilar varieties (Sakai, 1955)](image)

Sakai was able to express the production of the central plants as a linear function of the number of dissimilar competing plants. Plants of poor competitive ability declined in production, while those with good competitive ability increased as the number of dissimilar competing plants rose.

Sakai (1955) distinguishes between competitive ability and propagation rate. Propagation rate is defined to be the reproductive
rate in a pure sward or stand, while competitive ability is a function of the reduction or increase in reproductive rate according to the number of surrounding competitive plants. In developing the model, Sakai does not consider the number of noncompetitive (similar) surrounding plants.

Let genotype A have a reproductive rate in a pure stand of $x$ relative to B, which is given 1 as its reproductive rate. If A has a poorer propagation rate then $x < 1$; if, in competition with B, A incurs an additive increment of $p$ in propagation rate, while B suffers a similar decrement, A is a stronger competitor than B.

In a usual experimental design, the arrangement of plants is regular, and each plant is surrounded by a constant number, $k$, of other plants. If the proportions of genotype A and B in the population are $a$ and $b$ respectively, then each plant is surrounded, on the average, by $ka$ plants of genotype A and $kb$ plants of genotype B. From this we get the following recursive expression for the proportion of genotypes A and B in the $(n+1)$th generation.

$$a_{n+1} = \frac{a_n (x + kb_n p)}{w_n}$$

$$b_{n+1} = \frac{b_n (1 - ka_n p)}{w_n}$$

where

$$w_n = a_n (x + kb_n p) + b_n (1 - ka_n p)$$

$$= \frac{a_n x}{n} + \frac{b_n}{n}$$
The term $p_k$ is referred to as the competitive ability, and it is negative for poor competitors. So, setting $p_k = p'$, Sakai lists the propagation rate and competitive ability of two rice varieties.

More generally, Sakai describes the interaction of several competing genotypes by the model: (in Sakai's notation)

Genotypes

\[ A \ B \ C \ \ldots \ J \ \ldots \ M \]

Frequencies

\[ a \ b \ c \ \ldots \ j \ \ldots \ m \]

Reproductive rate

\[ x_a x_b x_c \ldots x_j \ldots x_m \quad (x_m = 1 \text{ by definition}) \]

Competitive ability

\[ p_a p_b p_c \ldots p_j \ldots p_m \quad (\sum_{n=a}^{m} p_n = 0) \]

In this instance, as with others which will be discussed later, only pairwise competition is considered. The frequency of genotype $J$ in the $(n+1)$th generation is shown to be

\[
j_{n+1} = j_n \left\{ x_j + ka_j (p_j - p_a) + kb_j (p_j - p_b) + \ldots \right. \\
+ k_j (p_j - p_j) + \ldots + km_j (p_j - p_m) \right\} \frac{1}{w_n}
\]

This is calculated as the frequency of $J$ multiplied by a sum which is its relative reproductive rate plus $k$ times the frequency $J$ competes against $A$ times the difference in competitive ability of $J$ over $A$ plus similar terms for all other genotypes, the entire expression being scaled by $w_n$ to ensure that frequencies sum to unity.

\[
w_n = a_n \left\{ x_j + ka_j (p_j - p_a) + \ldots + km_j (p_j - p_m) \right\} + b_n \left\{ x_j + \ldots \right\} + \ldots \\
+ m_n \left\{ 1 + ka_j (p_j - p_a) + \ldots + km_j (p_j - p_m) \right\}
\]
When \( kp = l-x \) for all genotypes, competitive superiority and propagation rate balance so that equilibrium, if it exists, is established. However, it is not clear from Sakai's paper how to analytically determine this point. In the case of two genotypes, however, an equilibrium exists when \( x + kb_p = a_n x + b_n \). Since \( a_n + b_n = 1 \), we obtain

\[
\begin{align*}
x + kb_p &= (l-b)x + b \\
b(kp + x-1) &= 0 \\
k_p &= 1-x \\
p &= \frac{1-x}{k}
\end{align*}
\]

This expression bears out, in essence, that the model may be unsatisfactory: the equilibrium, if it exists, is neutral if the value of the competitive ability is \( p = (1-x)/k \).

Sakai (1955) examines the components of variance associated with hybrid populations undergoing competition. Genes at a single locus affect two characters pleiotropically (see Figure 2.2).

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Character Expression</th>
<th>Competitive Ability</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>d</td>
<td>c</td>
</tr>
<tr>
<td>Aa</td>
<td>0</td>
<td>h</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>g</td>
</tr>
</tbody>
</table>

Figure 2.2. Genotypic values for character expression and the pleiotropically associated trait, competitive ability (Sakai, 1955)
The value of a character expression in a row of genotypes can be calculated from consideration of Figure 2.3.

\[
\begin{array}{cccccccc}
    & h & -d & h & d & h & d & -d & -d \\
g+c & -2(c+g) & +2g & 2(c-g) & -2(c-g) & 3c-g & -4c & 4c & -2c \\
\end{array}
\]

Figure 2.3. Total genotypic values for competitively interacting genotypes in a row (Sakai, 1955)

To explain this, we consider the third plant, Aa, surrounded by aa and AA. The character expression has a value h, but in competition with the weaker competitor, aa, it gains in competitive ability (g+c). However, in competition with the stronger competitor AA, it loses in competitive ability -(c-g), the difference between the competitive ability value for AA and that for Aa. The sum of these values, h+2g, is that given in the figure.

In the case of regular plantings, each plant may be surrounded by n other plants. We consider first the case where n=1, i.e., the plants compete in pairs only. The situation is best explained in terms of a table from which the F2 variance may be calculated (Table 2.1).

From the table we may determine the following parameters.

Mean = 0.5h

Variance = \((1/16)d^2 + (1/8)(h+c+g)^2 + \ldots + (1/16)d^2 - (1/4)h^2\)

= \(0.5d^2 + 0.25h^2 + c^2 + 0.5g^2 + dc + 0.5hg\)
Table 2.1. Frequency and genotypic value of interacting F2 genotypes (Sakai, 1955)

<table>
<thead>
<tr>
<th>Genotype under Competition</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>0.25</td>
</tr>
<tr>
<td>Aa</td>
<td>0.5</td>
</tr>
<tr>
<td>AA</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Competitor Frequency

<table>
<thead>
<tr>
<th>Competitor</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>0.25</td>
</tr>
<tr>
<td>Aa</td>
<td>0.5</td>
</tr>
<tr>
<td>AA</td>
<td>0.25</td>
</tr>
</tbody>
</table>

This shows that the effect of competition is to increase the variance, provided that \( h \) and \( g \) are of the same sign or that \( |0.5hg| \) is less than \( c^2 + 0.5g^2 \).

Summed over independent loci, the variance may be expressed as

\[ V_{F2} = 0.5D + 0.25H + C + 0.5G + M + 0.5N + E \]

Note that \( M \) or \( N \) may be negative should there be a negative association between gene expression and competitive effect.

Sakai (1955) attempts to allow for differing \( c \) and \( g \) effects for the \( n \) surrounding competing plants, but the algebra quickly becomes intractable and expressions for variances are given in the form above. In a similar manner, variances for F3, F3 progeny means, F4, and so on, can be expressed in terms of \( D, H, C, G, M, N \) and \( E \).
Covariances between parents and offspring are necessary, though, to
determine M and N separately from C and G respectively. From these
estimates of parameters, Sakai calculates estimates of heritability,
which differ markedly from estimates calculated when competition
effects are not included in the model.

2.2. Model Due to Schutz, Brim and
Usanis (1968)

A further approach to the problem of competing genotypes was
published by Schutz, Brim and Usanis (1968). In their model a
number of genotypes comprise the population. Table 2.2 lists the
genotypic frequencies and reproductive rates (Schutz et al. (1968)
preferred the term "reproductive value").

Table 2.2. Genotypes, genotypic frequencies and reproductive rates of
the interacting components of a competing population
(Schutz, Brim and Usanis, 1968)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Frequency</th>
<th>Reproductive rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>X_1</td>
<td>f_1</td>
<td>( w_1 = Y_1 + C_1 )</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>X_j</td>
<td>f_j</td>
<td>( w_j = Y_j + C_j )</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>X_n</td>
<td>f_n</td>
<td>( w_n = 1 + C_n )</td>
</tr>
</tbody>
</table>


The reproductive rates are scaled so that $Y_n = 1$. Here

$$\bar{w} = \sum_{i=1}^{n} f_i(Y_i + C_i),$$

where the $Y_i$'s are the reproductive rates in a pure stand of $X_i$'s relative to that for $X_n$. The net effect of inter-genotypic competition between $X_i$ and the $n$-1 other competing genotypes is denoted by $C_i$, i.e.,

$$C_i = \sum_{j \neq i}^n f_j b_{i/j}$$

where

$$b_{i/j} = b_{i/j} - Y_i$$

and $b_{i/j}$ = output or reproductive rate of $X_i$ under conditions of maximum competition with $X_j$. This underlies the pairwise nature of the competition model.

If we denote the frequency of genotype $X_j$ in the next generation by $f_j'$, then

$$f_j' = \frac{f_j(Y_j + C_j)}{\bar{w}}$$

and hence

$$\Delta f_j = f_j'((Y_j + C_j) - \sum_{i} f_i(Y_i + C_i)/\bar{w}$$

Equilibrium conditions specify that $\Delta f_i = 0$  $\forall i$, and an analytical solution for this set of equations was given by Cockerham and Burrows (1971), which will be discussed more fully later.

Schutz, Brim and Usanis (1968) discuss four identifiable types of competitive interactions:
(1) complementary \( b_{(j/i)} + b_{(i/j)} = 0 \)

(2) overcompensatory \( b_{(j/i)} + b_{(i/j)} > 0 \)

(3) undercompensatory \( b_{(j/i)} + b_{(i/j)} < 0 \)

(4) neutral \( b_{(j/i)} = b_{(i/j)} = 0 \).

They claim that a good competitor is ". . . a genotype whose reproductive value when competing with all other genotypes in the population is equal to, and in some cases greater than, its reproductive value in pure stand".

In the model, they claim that this means that a good competitor has \( b_{(j/i)} \) not less than zero, and greater than zero for at least one competing situation. By symmetry, Schutz et al. (1968) concluded, poor competitors have relatively better pure stand production than in competition. They point out that in a mixture a good competitor is more vigorous when at low frequency, since at high frequency its environment is becoming more like its pure stand. They argue that a poor competitor may, therefore, be retained in a stable mixture against a good competitor provided that its pure stand reproductive rate is good enough. They further allow that bulk population breeding, if it leads to a stable system, has a good chance of retaining poor competitors which will ". . . invariably have outstanding pure stand reproductive capacity". They rightly conclude that this result is contrary to the common belief that bulk breeding always leads to the selection of good competitors which must also have superior pure stand reproductive capacities, (if the
method is to be successful).

This argument is specious. A good competitor only needs to have a reproductive rate, \( w_j \), in competition with all other genotypes that is greater than the average reproductive rate of all other genotypes in competition, \( \bar{w} \). Pure stand performance alone does not enter into the argument.

The reproductive rate of \( X_j, w_j \), can be rewritten:

\[
    w_j = Y_j + C_j = Y_j + \sum_{i \neq j} f_i b^{(j/i)} \\
    = Y_j + \sum_{i \neq j} f_i (b'_{(j/i)} - Y_j) \\
    = \sum_{i
    \text{where } Y_j = b'_{(j/j)}, \text{ and } b'_{(j/i)} \text{ is the reproductive rate of } X_j \text{ in full competition with } X_i, \text{ and } b'_{(j/j)} \text{ is the reproductive rate of } X_j \text{ in a pure stand. Clearly, we can imagine values of } b'_{(i/j)} \text{ such that } w_j > b'_{(j/j)} \text{ and vice versa. The condition for a good competitor, } w_j > \bar{w} \text{, may be met by having } b'_{(j/i)} \text{ greater than } b'_{(k/i)}, \text{ for all } i: \text{ a sufficient but not a necessary condition. We can see that the pure stand reproductive rate, } b'_{(j/j)}, \text{ need not necessarily be greater or less than } b'_{(j/i)}', \text{ the reproductive rate of the genotype in competition.}

It is difficult to understand why Schutz et al. used the form (2.2.1) rather than the more obvious form (2.2.2) which would not lead to the false conclusions above.

We can see what (2.2.2) implies in terms of the types of competitive
interactions:

(1) complementary, \( b_{(j/i)} + b_{(i/j)} = 0 \)

\[ \Rightarrow b'_{(j/i)} + b'_{(i/j)} = b'_{(j/j)} + b'_{(i/i)} \]

i.e., the sum of the productions under competition is equal to the sum of the productions in pure stand.

(2) overcompensatory, \( b_{(j/i)} + b_{(i/j)} > 0 \)

\[ \Rightarrow b'_{(j/i)} + b'_{(i/j)} > b'_{(j/j)} + b'_{(i/i)} \]

which implies that the sum of the productions under competition is greater than that in pure stands. A similar argument can be made for undercompensatory interaction (3), and

(4) neutral, \( b_{(j/i)} = 0 \)

\[ \Rightarrow b'_{(j/i)} = b'_{(i/j)} \]

i.e., production is unaffected by competition.

A few implications concerning the model was discussed by Schutz and Usanis (1969) but, probably owing to the somewhat convoluted algebra of the model, no clear conclusions were drawn.

2.3. Model Due to Allard and Adams (1969)

Allard and Adams (1969) published a study of computer simulations of a model based heavily on that of Schutz, Brim and Usanis (1968). Their terminology differs (\( K_{j/i} \) replaces \( b_{(j/i)} \), for example) but the basic algebra is the same. Allard and Adam denote by \( C_i \) the net effect
of intergenotypic competition between \( X_i \) and the \( n-1 \) other competing genotypes. In the case of four genotypes,

\[
C_j = \sum_{i \neq j} K_{j/i} + \sum_{i \neq j, k} K_{j/i,k} f_i f_k + \sum_{i \neq j, k, l} K_{j/i,k,l} f_i f_k f_l + K_{j/i,k,l,m} f_i f_k f_l f_m. \tag{2.3.1}
\]

The terms \( K_{j/i} \), etc., are defined as follows:

\[
K_{j/i} = K'_{j/i} - X_j
\]

\[
K_{j/i,k} = K'_{j/i,k} - 0.5(K'_{j/i} + K'_{j/k})
\]

\[
K_{j/i,k,l} = K'_{j/i,k,l} - (1/3)(K'_{j/i,k} + K'_{j/i,l} + K'_{j/k,l})
\]

\[
K_{j/i,k,l,m} = K'_{j/i,k,l,m} - 0.25(K'_{j/i,k,l} + K'_{j/i,k,m} + K'_{j/i,l,m} + K'_{j/k,l,m})
\]

\( K' \) is the yield of \( X_j \) surrounded by 1, 2, 3 or 4 genotypes. Allard and Adams (1969) claim that in experiments where estimates of second-, fourth- and eighth-order interactions were available, they were too small to be significant. Allard and Adams cited the results published by Sakai (1957) as being consistent with their claim, but it is not clear, though, that Sakai was making the same measurement.

Even so, a point ignored by Allard and Adams (1969) is that in expression (2.3.1) the maximum coefficient for \( K_{j/i,k} \), being \( f_i f_k \), is less than 0.25, that for \( K_{j/i,k,l} \) less than 0.037, and that for
\( K_{j/i,k,l,m} \) less than 0.004. For example, the coefficient for
\( K_{j/i,k,l,m} \) is \( f_i f_k f_l f_m \), which is equal to \( f_i f_k (1-f_i-f_k-f_l) = H \), say.

\[
\frac{\partial H}{\partial f_i} = f_k f_l (1-2f_i-f_k-f_l).
\]

Partial differentiation with respect to \( f_k \) and \( f_l \) yield two further expressions. We assume that \( f_i, f_k, f_l, f_m \neq 0 \), and equate the differentiations to zero, leading to the equation:

\[
\begin{pmatrix}
2 & 1 & 1 \\
1 & 2 & 1 \\
1 & 1 & 2
\end{pmatrix}
\begin{pmatrix}
f_i \\
f_k \\
f_l
\end{pmatrix}
= 
\begin{pmatrix}
1 \\
1 \\
1
\end{pmatrix}
\]

which gives the solution, \( f_i = f_k = f_l = f_m = 0.25 \). Hence, the coefficient for \( K_{j/i,k,l,m} \leq (0.25)^4 \). A similar argument exists for the other coefficients.

Considering only the \( K_{j/i} \) terms, Allard and Adams (1969), following Schutz, Brim and Usanis (1968), make a similar classification of four types of competitive interaction. Allard and Adams attempt to give a description of each classification which is not satisfactory. In the undercompensatory case they claim \( K_{i/j} + K_{j/i} < 0 \) implies "... competition to mutual disadvantage". Similarly, the overcompensatory case implies "... cooperation to mutual advantage". It can be seen, though, that \( K_{j/i} + K_{i/j} < 0 \) may occur when one is greatly enhanced, but the other placed at an even greater disadvantage, and vice versa for the latter case, so that mutual
disadvantage or mutual advantage may not necessarily describe correctly
the situation, even though, in the former instance the reproductive
capacity of the mixed population is depressed, and in the latter
instance elevated. Aside from noting that Schutz, Brim and Usanis
(1968) avoid this ambiguity, we shall leave the discussion.

2.4. Model Due to Cockerham and Burrows (1971)

The model of Schutz et al. (1968) was given a more rigorous
mathematical treatment by Cockerham and Burrows (1971) who showed not
only the existence of equilibria by analytical methods, but also the
conditions necessary for these equilibria to be stable.

If the fitness of genotype \( X_i \) in competition with \( X_j \) is \( w_{ij} \),
and we consider that the pairwise interactions occur in frequencies
proportional to the genotype frequencies, \( p_i \), then the frequency of
the \( i \)th genotype following one reproductive cycle is

\[
p'_i = \frac{p_i w_i}{w_{..}}, \quad i = 1, \ldots, r,
\]

where

\[
w_i = \sum_j p_j w_{ij}, \quad \text{and} \quad w_{..} = \sum_i p_i w_{i.}.
\]

At equilibrium, therefore,

\[
\hat{p}_i w_{..} = \hat{p}_i w_i, \quad \forall i,
\]

\[
w_{..} = w_i.
\]
Let \{w_{ij}\} be the \(r \times r\) matrix \(W\), then

\[
\tilde{\hat{\beta}}' \tilde{W} \tilde{\beta} = \tilde{W} \tilde{\beta}.
\]

For this equation to have a solution \(\text{rank}(W, I) = \text{rank}(W)\). A unique solution exists if \(\text{rank}(W) = r\). The solution is given by

\[
\hat{\beta} = \frac{W^{-1} 1}{1'W^{-1} 1}.
\]

If \(\text{rank}(W) < r\), then there may be many solutions, which are given by

\[
\hat{\beta} = \frac{W^{-1} 1}{1'W^{-1} 1},
\]

where \(W^{-}\) is any conditional inverse of \(W\); \(W W^{-} W = W\). If any solution gives \(\hat{\beta}_i = 0\) for some \(i\), the system may be examined with the corresponding component deleted.

2.4.1. Equilibrium stability

If \(\hat{\beta}\) is a solution to the Equation (2.4.1), the equilibrium may be tested for local stability by consideration of a Taylor expansion about the equilibrium point. Let

\[
p_i^{(u+1)} = \frac{p_i^{(u)} w_i}{w_{..}}.
\]

be represented by \(p_{u+1} = F(p_u)\). Then

\[
p_{u+1} = \hat{\beta} + \delta_{u+1} = F(p_u) = F(\hat{\beta} + \delta_u),
\]
where $\delta_u$ is the vector of deviations of $p_u$ from the equilibrium value, $\hat{p}_u$, i.e., $\delta_u = p_u - \hat{p}_u$.

A Taylor expansion about $\hat{p}$ gives

$$\hat{p} + \delta u \approx F(p) + \left. \frac{\partial F(p)}{\partial p} \right|_{\hat{p}} \delta p + O(\delta p'')$$

where $O(\delta p'')$ represents terms small enough to be neglected. Let

$$m_{ij} = \left. \frac{\partial F(p)}{\partial p_j} \right|_{\hat{p}}$$

and the $r \times r$ matrix $M = \{m_{ij}\}$, then, approximately,

$$\delta u = M \delta p^{-1} = \delta p$$

Because the sum of the frequencies must be unity, i.e., $\sum p = 1$,
then $\sum \delta p = 0$. The matrix $M$ can therefore be reduced to a

$(r-1) \times (r-1)$ matrix $N = \{n_{ij}\}$, where

$$n_{ij} = m_{ij} - m_i, i,j, = 1,...,r-1$$

If $\delta^*$ represents only the first $r-1$ elements of $\delta$, we may write, approximately,

$$\delta^* = N \delta^* = N \delta^*$$

It then can be shown that

$$n_{ii} = 1 + \hat{p}_i (w_{ii} - w_{ir} - \hat{w}_{ir} + \hat{w}_{i}) \hat{w}_{ir}^{-1}$$

$$n_{ij} = \hat{p}_i (w_{ij} - w_{ir} - \hat{w}_{ir} + \hat{w}_{r}) \hat{w}_{ir}^{-1}$$

where $\hat{p}_i$ is the equilibrium frequency,
\[ \hat{w}_j = \sum_i \hat{p}_i w_{ij}, \quad \text{and} \quad \hat{w}_j = \sum_j \hat{p}_j \hat{w}_j. \]

In general, Cockerham and Burrows (1971) claim, there exists an \((r-1) \times (r-1)\) matrix \(C\), such that

\[ \Sigma C = \Lambda, \]

where \(\Lambda\), a diagonal matrix, contains the eigenvalues of \(N\) along its diagonal. Let

\[ g_u = C \delta^u, \]

\[ \delta^u = C^{-1} g_u, \]

\[ g_u = \Lambda^u g_0. \quad (2.4.2) \]

Now, \(\lim_{u \to \infty} \delta^u = 0 \leftrightarrow \lim_{u \to \infty} g_u = 0\), which occurs if and only if

\[ \lim_{u \to \infty} \lambda_i^u = 0, \quad i = 1, \ldots, r-1. \]

If \(|\lambda_i| < 1, \forall i\), \(\lim_{u \to \infty} \lambda_i^u = 0\), so this is a necessary and sufficient condition for stability. If \(\lambda_i\) is complex the system may oscillate, but the equilibrium is stable for \(||\lambda_i|| < 1\), where \(||\cdot||\) denotes the spectral radius. In the case of 3 or 4 components, \(N\) may have two complex conjugate eigenvalues, \(\lambda_1 = \alpha + i\beta, \lambda_2 = \alpha - i\beta\), where \(i^2 = -1\).

Expressing these roots as \(\lambda_1 = \rho (\cos \theta + i \sin \theta), \lambda_2 = \rho (\cos \theta - i \sin \theta)\), and solving for \(\theta\) gives an estimate of the periodicity of the oscillation, \(k^* = 2\pi/\theta\). Estimates were found to agree
closely with the figures given by Schutz et al. (1968).

Cockerham and Burrows (1971) failed to notice that $N$ may not necessarily be diagonalizable, although the statement concerning the relationship between the local stability of the equilibrium point and the absolute value of the eigenvalues is still valid. A review of the case where $CNC^{-1} = J$, where $J$ is a Jordan canonical form, has been given by Lewis (1978) and is reproduced in the Appendix of this dissertation.

2.4.2. Maximizing the mean

The values for $p$ that give the maximum value for the population mean, $p' \tilde{W} p$, were derived by Cockerham and Burrows (1971) from considerations of the symmetrical form $U = (\tilde{W} + \tilde{W'})/2$.

$$\tilde{p}' U \tilde{p} = \tilde{p}' \tilde{W} \tilde{p} = \tilde{w}.$$  
Now, using the restriction that $\tilde{p}' \tilde{l} = 1$, we may employ Lagrangian multipliers, so that

$$Q(\tilde{p}) = \tilde{p}' U \tilde{p} - 2k(\tilde{p}' \tilde{l} - 1),$$

$$\frac{\partial Q(\tilde{p})}{\partial \tilde{p}} = 2\tilde{U}p - 2kl.$$  
Equate this to zero to determine the turning points of the population mean. This gives us

$$U \tilde{p} = kl,$$
\[ \tilde{p} = \frac{u^{-1} \tilde{1}}{1' u^{-1} \tilde{1}}, \text{ provided } u^{-1} \text{ exists.} \]

To find if this is a maximum, we form the symmetric \((r-1) \times (r-1)\) matrix \(\mathbf{v} = \{v_{ij}\}\), where

\[ v_{ij} = u_{ij} - u_{ik} u_{ki} u_{kk} = v_{ji}, \text{ for one value of } k \neq i. \]

Cockerham and Burrows (1971) claim that \(\tilde{w}_{..}\) is a maximum provided \(\mathbf{v}\) is a negative definite matrix.

2.5. Model Due to Mather (1969)

The first competition model described by Mather (1969) is similar to that described by Nei (1971) (discussed later), in that for any competing pair, an advantage in fitness incurred by one member is equal to the disadvantage incurred by the other. The population mean fitness remains at 1, and there is no equilibrium for genotype frequencies for populations of either haploids or autogenous diploids.

Differing from Nei, however, Mather attempts to develop a model incorporating a more complex competitive system than the simple pairwise case that Nei erroneously claims is applicable to interacting groups of size greater than 2. Mather (1969) assumes that a plant is surrounded by a number of competing plants, each one in a separate zone. In Figure 2.4 the central plant is surrounded by four plants, each one in a separate zone. The competitive effect of the plant in zone 1 on the central plant is \(k_1\), of that in zone 2 on the
Figure 2.4. Competitive interaction (model due to Mather (1969))
(The effect of the genotype in zone 4 on the central plant is $k_4$. If genotype $P$ is type $A$ and that in zone 4 $B$ then the fitness of $P$ is increased by $k_4$. If, however, the genotypes are $B$ and $A$ respectively, then the fitness of $P$ is reduced by $k_4$. The individuals in zones 1 and 3 interact with one another in their competitive effects on $P$, but $k_3$ is defined to allow for their effect)
central plant is $k_2$, and so on. Interactions between plants in different zones, such as zone 1 and zone 3 in Figure 2.4, is allowed for by the definition of $k_3$. While the zones are depicted as equal-area concentric circular bands, this, in practice, need not be the case.

Let the frequencies of genotypes $A$ and $a$ be $u$ and $v$ respectively. Then if the central plant is $A$, we may calculate its fitness to be

$$1 + vk_1 + vk_2 + \ldots = 1 + v \sum_i k_i = 1 + vk,$$

while if $a$ is the central plant its fitness is

$$1 - uk_1 - uk_2 - \ldots = 1 - u \sum_i k_i = 1 - uk.$$

The model is thus reduced to the simple pairwise case. Mather (1969) does, however, raise the question concerning the additivity of the $k_i$'s, but does not resolve it.

In considering the case of intermating two-allele diploids under competition, Mather (1969) develops a model similar to that due to Nei (1971). Slight differences in the formulation of relative fitnesses do not change the basic results. For genotypes $AA$ and $Aa$ in competition, the relative fitnesses are $1+k_a$ and $1-k_a$ respectively, for $Aa$ and $aa$, $1+k_b$ and $1-k_b$, and for $AA$ and $aa$, $1+k_c$ and $1-k_c$, respectively. Thus $(1+s_1)/2$ in Table 2.8 may be replaced by $1+k_a$, and so on. Mather (1969) finds that with this model, the change in allele frequency is given by

$$\Delta u = uv(u^2 k_a + v^2 k_b + uv k_c),$$
a result that compares with Nei's (1971) result:

\[ \Delta p = pq(p^2s_1 + pq^2s_2 + q^2s_3), \]

and hence, Mather (1969) draws similar conclusions.

Leaving the pairwise competition formulations, Mather (1969) describes a frequency-dependent selection model, which may be more clearly described by Table 2.3.

Table 2.3. Frequency-dependent selection model (Mather, 1969)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Frequency</th>
<th>Fitness</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>((1-v)^2)</td>
<td>(1-v^2k)</td>
</tr>
<tr>
<td>Aa</td>
<td>(2v(1-v))</td>
<td>(1-v^2k)</td>
</tr>
<tr>
<td>aa</td>
<td>(v^2)</td>
<td>(1-s+(1-v^2)k)</td>
</tr>
</tbody>
</table>

Mather (1969) concludes that an equilibrium (which is neutral) occurs if \(s=k\) (which results in equal fitness for all genotypes). Significantly, Mather's model given in Table 2.3 bears a close resemblance to Wright's (1955) first model of frequency dependent selection reproduced in outline in Table 2.5.

2.6. Model Due to Schutz and Usanis (1969)

A treatment of the diploid case, with mixed selfing and random mating, was published by Schutz and Usanis (1969). Basically, their model was similar to the one published by Schutz et al. (1968) earlier.
In both papers (Schutz et al., 1968; Schutz and Usanis, 1969) the term "reproductive rate" is used. A term more consistent with population genetics nomenclature is "viability" which is proportional to the probability that a zygote survives to adulthood.

Schutz and Usanis (1969) give the relative viability of the genotype AA as (in their notation),

\[ r_{AA} = H_{AA} + f_{AA} b_{AA/AA} + f_{AA} b_{AA/aa}, \]

where

- \( H_{AA} \) is the relative viability of AA in a pure stand,
- \( f_{AA} \) is the frequency of genotype AA,
- \( b_{AA/AA} = b'_{AA/AA} - H_{AA} \)

where

- \( b'_{AA/AA} \) is the relative viability of AA in complete competition with AA.

As before, we can express this, as Cockerham and Burrows (1972) did, more simply as

\[ r_{AA} = f_{AA} b'_{AA/AA} + f_{AA} b'_{AA/AA} + f_{AA} b'_{AA/aa}. \]

Similar expressions may be written out for \( r_{Aa} \) and \( r_{aa} \). The frequency of the genotype AA, \( f_{AA}^{(t+1)} \), among adults in the next generation will depend on the amount of selfing and outcrossing that occurs in the population, as well as the net relative viabilities of the respective genotypes.
\begin{align*}
\hat{f}_{AA}(t+1) &= \frac{r_{AA}p_{AA}}{R''}, \\
\hat{f}_{Aa}(t+1) &= \frac{r_{Aa}p_{Aa}}{R''}, \\
\hat{f}_{aa}(t+1) &= \frac{r_{aa}p_{aa}}{R''},
\end{align*}

where

\begin{align*}
p_{AA}(t) &= s(f_{AA} + \frac{1}{4}f_{Aa})(f_{AA} + \frac{1}{2}f_{Aa})^2, \\
p_{Aa}(t) &= \frac{1}{2}s(f_{Aa} + \frac{1}{4}f_{AA})(f_{Aa} + \frac{1}{2}f_{AA})^2, \\
p_{aa}(t) &= s(f_{aa} + \frac{1}{4}f_{Aa})(f_{aa} + \frac{1}{2}f_{Aa})^2.
\end{align*}

The parameter $s$ is the proportion of the population that selfs, and $c$ is the proportion of the population that is involved in random mating. The scaling factor, $R'' = \sum r_i(t) p_i(t)$, and the population mean relative viability for the $(t+1)$-th generation is

\[ R' = \sum r_i(t+1) \hat{f}_i(t+1). \]

A number of computer simulations showed the existence of stable equilibrium frequencies for the heterozygote, even for populations largely self-fertilizing. Furthermore, heterozygotes intermediate for competition effect as well as pure stand relative viability, were maintained in the population. This implies that "heterosis" per se need not be invoked in order to explain the presence of polymorphisms in natural
populations. Particularly important in maintaining mixtures in a population possessing both random mating and selfing, was the "over-compensatory" competitive effect, where \( b_{(i/j)} + b_{(j/i)} > b_{(i/i)} + b_{(j/j)} \), i.e., the two genotypes in a mixture outperformed their pure stand equivalents.

2.7. Model Due to Cockerham et al. (1972)

Developing the model of Schutz and Usanis (1969), Cockerham, Burrows, Young and Prout (1972) considered the case of pairwise interaction between random mating, one locus, two allele diploids. The main thrust of the paper was to show that although the fitness value of a genotype interacting with another remained constant, the effect with changing gene frequencies is reflected in a change in the mean fitness value of a genotype averaged over the whole population. This results in the appearance of frequency-dependent selection. This differs from the frequency-dependent selection models where the fitness value of a genotype is formally written as a function of gene- or genotype frequencies, e.g. Wright (1955). An interesting result of this (not stressed by Schutz and Usanis (1969)), is that purely additive models can produce results that mimic heterotic models of noninteracting systems. This is explained more fully as follows. Table 2.4 sets out the fitness values of the pairwise interacting genotypes.

The mean fitness value \( \bar{W}_i = p^2W_{i1} + 2pqW_{i2} + q^2W_{i0} \), where
Table 2.4. Fitness values for competing diploids (Cockerham et al., 1972)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Competing with</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AA</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>$W_{22}$</td>
<td>$\bar{W}_2$</td>
</tr>
<tr>
<td>Aa</td>
<td>$W_{12}$</td>
<td>$\bar{W}_1$</td>
</tr>
<tr>
<td>aa</td>
<td>$W_{02}$</td>
<td>$\bar{W}_0$</td>
</tr>
</tbody>
</table>

$p (=l-q)$ is the frequency of allele A. By use of relationships between fitness values, i.e., expressing $W_{11} = (1-k)W_{12} + kW_{10}'$ and $W_{1j} = (1-h)W_{2j} + hW_{0j}$, Cockerham et al. (1972) were able to calculate and list conditions for the existence and stability of equilibria in terms of $k$ and $h$.

Some special cases were discussed. If $W_{22} = 1-t$, $W_{00} = 1-s$, $0 < k = h < 1$, and $W_{20} = W_{02} = 1$, only one equilibrium exists. If also, however, $h=0$, the model degenerates to the second frequency-dependent selection model discussed by Wright (1955) which is outlined in Table 2.5, where $t$ and $s$ are replaced by $-s_1$ and $-s_2$ respectively.

Table 2.5. Frequency-dependent selection model II (Wright, 1955)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Frequency</th>
<th>Selective value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-</td>
<td>$1-q^2$</td>
<td>$1+s_1(1-q^2)$</td>
</tr>
<tr>
<td>aa</td>
<td>$q^2$</td>
<td>$1+s_2q^2$</td>
</tr>
</tbody>
</table>
In Wright's (1955) model II the equilibrium point, \( q = \frac{1}{\sqrt{s_1(s_1+s_2)}} \), is stable if both \( s_1 \) and \( s_2 \) are negative, and unstable if both are positive. This agrees with intuitive reasoning: in the former case the more rare genotype has a selective advantage, in the latter case the more common genotype has a selective advantage. The mean population fitness, \( \bar{W} = 1 + s_1(1-q^2)^2 + s_2 q^4 \), is at its maximum at the equilibrium point if the point is stable, and at its minimum at the point if it is unstable.

Of more interest, as Cockerham et al. (1972) point out, is the additive model where \( h = \frac{1}{2} \). Here the population mean \( \bar{W} = 1-tp^2-sq^2 \), and the equilibrium value, \( \hat{p} = s/(s+t) \), (stable for \( t, s > 0 \)) are exactly those of the classical heterotic model (see Table 2.1 of Falconer, 1960).

This, as was mentioned, implies that it may not be necessary to invoke the heterotic model necessarily as an explanation of stable polymorphism.

A second special case occurred if \( W_{20} = 1+s, W_{02} = 1+t, h=k, \) and \( W_{22} = W_{00} = 1 \). If, then \( h=0 \), this reduces further to the first frequency-dependent model discussed by Wright (1955), outlined in Table 2.6, where \( s \) and \( t \) are replaced by \( s_1 \) and \( s_2 \) respectively.

In Wright's (1955) model I the equilibrium point, \( q = \sqrt{s_2(s_1+s_2)} \), is stable if \( s_1 \) and \( s_2 \) are both positive, and unstable if they are both negative. In this case, though, the mean population fitness, \( \bar{W} = 1+(s_1+s_2)q^2(1-q^2) \), is at its maximum at \( q = \sqrt{0.5} \), which will not
Table 2.6. Frequency-dependent selection model I (Wright, 1955)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Frequency</th>
<th>Selective value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-</td>
<td>1-q^2</td>
<td>1+s_1q^2</td>
</tr>
<tr>
<td>aa</td>
<td>q^2</td>
<td>1+s^2(1-q^2)</td>
</tr>
</tbody>
</table>

coincide with the equilibrium point unless s_1=s_2. In this model we may have the case that the population mean fitness decreases as the population undergoes one cycle of selection.

To return to the second special case of Cockerham et al. (1972) being discussed (W_{20} = 1+s, W_{02} = 1+t, W_{22} = W_{00} = 1, h = k), if we put h = 1/2, we have the equilibrium solution, \( \hat{\theta} = s/(s+t) \), the same equilibrium as for the first special case discussed. Here, though, the population mean at equilibrium differs, being less than the maximum population mean \( \bar{W}_{\text{equilibrium}} = 1 + st/(s+t) \); \( \bar{W}_{\text{maximum}} = 1 + (s+t)/4 \) unless t = s.

The observation that \( \bar{W}_{\text{maximum}} \) does not necessarily coincide with \( \bar{W}_{\text{equilibrium}} \) brings up an important aspect of models exhibiting variable fitnesses. Generally, if a genotype, A_iA_j, has a constant fitness, \( W_{ij} \), then if the population is not at equilibrium the population mean, \( \bar{W}(=\Sigma_1^n \hat{\pi}_j W_{ij}) \), will increase from one generation to the next (Kingman, 1961). At a stable equilibrium, therefore, \( \bar{W} \) is, at least, a local maximum. In sharp contrast to this result, we have shown in the examples above (model I of Wright (1955), and the second special
case of the model of Cockerham et al. (1972), that where the fitnesses are expressed as functions of gene frequency, \( \bar{W}_{\text{maximum}} \) may not necessarily coincide with \( \bar{W}_{\text{equilibrium}} \).

As a final comment we note that if we express the fitness, or reproductive value, of the heterozygote as

\[
\bar{W}_1 = p^2 \bar{W}_{12} + 2p(1-p)[(1-k)\bar{W}_{12} + kW_{10}] + (1-p)^2 \bar{W}_{10}
\]

\[
= p^2 \alpha \beta + p(1-\alpha) \beta + \delta,
\]

(2.7.1)

where \( \alpha = 2k-1, \beta = \bar{W}_{12} - \bar{W}_{10}, \) and \( \delta = \bar{W}_{10}, \) then this quadratic may be used to describe the data of Harding et al. (1966), which shows a negative relationship between the frequency of heterozygous plants in a stand and their reproductive values. Allard and Adams (1969) used the expression

\[
W_2 = 1.06 + 4.3 \exp\{-23.52f_2^{(n)}\}
\]

to relate the observed reproductive value of the heterozygote, \( W_2 \), to its frequency, \( f_2^{(n)} \), in a population. The homozygotes were given an arbitrary value of 1. This expression is somewhat arbitrary, and something based more on biological considerations, however roughly, would be preferable.

We have calculated an alternative, simpler expression, which fits the data reasonably well \( r^2 = 0.74, \)

\[
W_2 = 0.858 + 0.06486/f_2^{(n)}.
\]
This expression may face two objections; one, that it is still an arbitrary function, cannot be denied, while the second objection, that for very low frequencies the reproductive values become impossibly high, can be countered by making the not unreasonable assumption that the number of homozygous plants that can simultaneously affect one heterozygous plant must be small, say 50. A large population, therefore, is divided up into areas of 50 individuals each. Where no heterozygote appears, obviously the calculation is not applicable. Elsewhere the minimum frequency is 0.02 (about the minimum value tabulated by Harding et al. (1966)), and the reproductive value of about 4 for that frequency is a quite reasonable maximum.

A least-squares fit to the data for the quadratic (2.7.1) however, does not give satisfactory values, viz,

\[ w_{10} = 2.83, \quad w_{11} = -1.64, \quad \text{and} \quad w_{12} = -5.08. \]

It will be noticed, however, that the value for the heterozygote lies within the range of the two homozygotes - no heterozygote superiority is necessary to explain the persistence of the heterozygote in the population.

2.8. Model Due to Nei (1971)

From a different viewpoint, Nei (1971, 1975) looks at the change in gene frequency in a population occurring in two stages; first, where the population grows without interaction between individuals, i.e., where the density is low, and the intrinsic reproductive rates of
respective genotypes determine the change in gene frequency, and second, where the population becomes "saturated", individuals interact and competitive ability of genotypes determine the change in gene frequency. Insofar as selection in one direction, due intrinsically to the genotype, may become countered by selection in the reverse direction due to interacting genotypes, a nontrivial stable equilibrium may exist. In certain instances equilibrium in population numbers, (implying density equilibrium) may also exist for a density less than that for full competition.

We shall start by considering Nei's (1971) formulation.

2.8.1. Haploids competing in pair groups

The first development considered a haploid model, a single locus, and two alleles.

\[ n_1 \text{ A adults; } n_2 \text{ a adults; } N = n_1 + n_2, \text{ and} \]

\[ p = n_1/N, \quad q = n_2/N \]

In the noncompetitive case the increase in the number of each genotype will be determined by the intrinsic growth rate or reproductive rate. Thus

\[ n_1' = n_1 r_1 = n_1 k_1 v_1 \]

where \( k_1 \) is the fertility of A individuals and \( v_1 \) their viability. A similar expression may be made for the a genotypes.

\[ n_2' = n_2 r_2 = n_2 k_2 v_2 \]
Nei (1971) considers competition occurring in pairs. So in a fully competing situation, with a random choice of pairs of competitors, competition occurs between

\[ \begin{align*} 
A \text{ and } A & \text{ with frequency } p^2, \\
A \text{ and } a & \text{ with frequency } 2pq, \text{ and} \\
a \text{ and } a & \text{ with frequency } q^2. 
\end{align*} \]

Each competing pair group has one survivor. In each of the homogeneous cases the outcome is the same regardless of which member of the pair is the survivor. In the heterogeneous case, however, Nei assumes that the survivor of the competing pair group is A with probability \((1+s)/2\), and is a with probability \((1-s)/2\). This is set out formally in Table 2.7.

<table>
<thead>
<tr>
<th>Competition between</th>
<th>Frequency</th>
<th>Probability of survival of survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>A and A</td>
<td>(p^2)</td>
<td>(A) (1) (a) ((1+s)/2) ((1-s)/2)</td>
</tr>
<tr>
<td>A and a</td>
<td>(2pq)</td>
<td>(A) ((1+s)/2) (a) ((1-s)/2)</td>
</tr>
<tr>
<td>a and a</td>
<td>(q^2)</td>
<td>(A) (1) (a) (1)</td>
</tr>
</tbody>
</table>

Under pure competition we get the result (assuming \(k_1 = k_2\))

\[ n_1' = n_1(1+sq), \text{ and} \]

\[ n_2' = n_2(1-sp) \]
If we put the two situations together we have that a proportion, $c$, of the population is experiencing competitive growth, while $1-c$ is experiencing noncompetitive growth. So

$$n_1' = n\_1[(1-c)r_1 + c(l+sq)], \text{ and}$$

$$n_2' = n\_2[(1-c)r_2 + c(l-sp)].$$

In this expression, $c$ is a function of $n\_1$ and $n\_2$. For example, $c = N/K$, where $K$ is the population size when only competitive interactions occur. This is also referred to as the "carrying capacity" of the system supporting the population.

Nei (1971, 1975) defines as the Wrightian fitness parameters

$$w_1 = n\_1'/n\_1, \quad w_2 = n\_2'/n\_2,$$

for the respective alleles, $A$ and $a$. These may be expressed

$$w_1 = (1-c)r_1 + c(l+sq)$$

$$w_2 = (1-c)r_2 + c(l-sp)$$

Of these, Nei (1975) notes that "... for a given value of $c$ the relative fitness of a genotype is higher when its frequency is low" (sic). This is true of $w_1$ but not $w_2$. This aside, we note that

$$\Delta n_1 = n\_1[a_1 - c(a_1 - sq)]$$

$$\Delta n_2 = n\_2[a_2 - c(a_2 + sp)]$$

$$\Delta N = a(n - c)$$

where $a_1 = r_1 - 1$, $a_2 = r_2 - 1$, and $a = pa_1 + qa_2$. 
The restriction \(0 < a < 1\) that Nei (1975) claims is necessary to avoid divergence of population size does not seem necessary on the one hand, while on the other hand it precludes from the model all species that may more than double their population size in one generation.

Returning to the expressions (2.8.1) above, we note that

\[
\Delta p = \frac{pq[(1-c)(a_1-a_2)+cs]}{1+(1-c)a} \quad (2.8.2)
\]

with \(c = N/K\). We may show the derivation thus:

\[
\Delta p = \Delta \frac{n_1}{N} = \frac{n_1 + \Delta n_1}{N+\Delta N} - \frac{n_1}{N}
\]

\[
= \frac{n_1 N + \Delta n_1 N - n_1 N - n_1 \Delta N}{N(N+\Delta N)}
\]

\[
= \frac{\Delta n_1 - n_1 \bar{a}(1-c)}{N+\Delta N}
\]

\[
= \frac{n_1 [a_1 - c(a_1 - sq) - \bar{a}(1-c) - \bar{a}(1-c)]}{N[1+\bar{a}(1-c)]}
\]

\[
= \frac{n_1}{N} \cdot \frac{[(a_1 q-a_2 q)(1-c)+scq]}{1+\bar{a}(1-c)}
\]

\[
= \frac{pq[(1-c)(a_1-a_2)+cs]}{1+\bar{a}(1-c)}
\]

In an unsaturated population, i.e., where some of the population is not under competition, \(p\) will not necessarily increase if the sign of \((a_1-a_2)\) and \(s\) differ. At saturation, \(c=1\), and \(\Delta p=pqs\), which is the classical result. If \(a_1-a_2 = s\) then we get the same result as Kimura and Crow (1969), but this does not seem to be a realistic condition.
In a saturated population $N$ is constant, $p = n_1/N$, $q = n_2/N$. We now relax the condition of equal fertilities and assume that the fertilities are $k_1$, $k_2$ for the genotypes $A$ and $a$ respectively. 

$k = pk_1 + qk_2$. Therefore the proportions of competing alleles comes out to be

$p_{k_1}/k$ $A$ alleles, and $q_{k_2}/k$ $a$ alleles.

Therefore, under the assumption of product fertility for interacting pairs,

$p' = (pk_1/k)^2 + [pqk_{1}k_{2}(1+s)]/k^2$

$= pk_1(p_{k_1}+q_{k_2}+qk_{2}s)/k^2$

$\Rightarrow \Delta p = pk_1 \frac{(k+q_{k_2}s)}{k^2} - \frac{pk^2}{k^2}$

$\Rightarrow k^2 \Delta p = pk_1\bar{k} + pqk_{1}k_{2}s - \bar{p}k^2 = \bar{p}(k_1-k) + pqk_{1}k_{2}s$

$= pk(pk_1+q_{k_1}-pk_1-q_{k_2}) + pqk_{1}k_{2}s$

$= pkq(k_1-k_{2}) + pqk_{1}k_{2}s$

$= pqk[(k_{1}k_{2}) + k_{1}k_{2}s/\bar{k}]$

$\Rightarrow \Delta p = pq[(k_{1}k_{2}) + k_{1}k_{2}s/\bar{k}]/\bar{k}$

An equilibrium exists if $(k_{1}k_{2}) + k_{1}k_{2}s/\bar{k} = 0$. For this $k_{1} < k_{2}$, a necessary but not sufficient condition. We may show that this equilibrium is locally unstable. Assume the approximation:
\[
\frac{dp}{dt} = \Delta p = p(1-p) \left[ \frac{(k_1-k_2)}{k} + \frac{k_1 k_2 s}{k^2} \right]
\]

where \( \bar{k} = pk_1 + (1-p)k_2 = k_2 + p(k_1 - k_2) \).

Formally, we write, using \( \bar{p} \) to denote the equilibrium value of \( p \),

\[
p' = \bar{p} + \delta'p = F(p) = F(p+\delta p) \approx F(\bar{p}) + \delta pF'(p)_{\bar{p}},
\]

where \( F'(p)_{\bar{p}} \) denotes the first differential of \( F(p) \) w.r.t. \( p \), evaluated at \( \bar{p} \). Approximately, therefore, the ratio \( \delta'p/\delta p \approx F'(p)_{\bar{p}} \)
so if the absolute value of \( F'(p)_{\bar{p}} \) is less than unity, the equilibrium point is locally stable.

\[
F(p) = p + \Delta p = p + p(1-p) \left[ \frac{k_1 - k_2}{k} + \frac{k_1 k_2 s}{k^2} \right]
\]

\[
F'(p) = 1+(1-p) \left[ \frac{k_1 - k_2}{k} + \frac{k_1 k_2 s}{k^2} \right] - p \left[ \frac{k_1 - k_2}{k} + \frac{k_1 k_2 s}{k^2} \right]
+ p(1-p) \left[ -\frac{(k_1 - k_2)(k_1 - k_2)}{k^2} + \frac{2k_1 k_2 s(k_1 - k_2)}{k^3} \right].
\]

At equilibrium \( \Delta p = 0 \Rightarrow \frac{k_1 - k_2}{k} + \frac{k_1 k_2 s}{k^2} = 0 \)

\[\Rightarrow (k_1 - k_2) = -k_1 k_2 s/\bar{k}.\]

Therefore, \( F'(p)_{\bar{p}} = 1+0-0(1-p) \left[ \left[ \frac{k_1 - k_2}{k} \right] \left[ \frac{k_1 - k_2}{k} + \frac{k_1 k_2 s}{k^2} \right] \right] \]

\[= 1-p(1-p) \left[ \frac{k_1 - k_2}{k} \right] \left[ \frac{k_1 k_2 s}{k^2} \right].\]

But \( k_1 < k_2 \), so that \( |F'(p)_{\bar{p}}| > 1 \), and the equilibrium point is unstable.
2.8.2. Haploids competing at the genotypic level

Nei (1971) then considers competition at the genotypic level. This type of competition may occur when, to quote Nei's example, one genotype, because it grows or germinates faster than the other, preempts a resource, leaving the other genotype with a lower probability of survival. In fact, the model may be set up without recourse to competitive effects at all, except to impose the condition of constant population size.

What Nei (1971) refers to as the absolute fitness of A is \((l+s)/2\), and that for a is \((l-s)/2\). Hence,

\[
p' = \frac{p(l+s)/2}{p(l+s)/2 + q(l-s)/2} = \frac{p(l+s)}{l+s(p-q)}
\]

\[
pw_1
\]

and

\[
q' = \frac{q(l-s)}{l+s(p-q)} = qw_2
\]

We can rewrite \(w_1\) as \(1+2sq/[l+s(p-q)]\), and \(w_2\) as \(1-2sp/[l+s(p-q)]\).

Now in the unsaturated case

\[
n'_1 = n_1[(1-c)r_1 + c(1+2sq/[l+s(p-q)])],
\]

\[
n'_2 = n_2[(1-c)r_2 + c(1-2sp/[l+s(p-q)])],
\]

which leads to

\[
\Delta p = \frac{pq[(1-c)(a_1-a_2) + 2cs/[l+s(p-q)])}{1 + (1-c)a}
\] (2.8.3)
When \( c=1 \), \( \Delta p = \frac{2pq}{1+s(p-q)} = 2pq \) if \( s \) is small. However, under the situation where competition acts between pairs of individuals, (i.e., a "competing pairs" situation is implied), \( \Delta p = pq \) (cf. (2.8.2)). Nei (1971) concludes, "... competitive selection between groups of genotypes is twice as effective as that between individuals if \( s \) is the same".

### 2.8.3 Diploids competing in pair groups

Here we have \( PAA + QAa + Raa \) for the genotype array. \( N = n_1 + n_2 + n_3 \), \( P = n_1/N \), \( Q = n_2/N \), \( R = n_3/N \), \( p = P+Q/2 \), and \( q = Q/2+R \).

Under noncompetitive growth, with constant fertility, we have,

\[
\begin{align*}
n_1' &= p^2Nv_1k, \\
n_2' &= 2pqNv_2k, \text{ and} \\
n_3' &= q^2Nv_3k.
\end{align*}
\]

For the case of competition at the individual level we have the following table (Table 2.8).

As before, from each competing pair the total number of offspring is constant, and the fitness of the entire population remains constant. The frequency of \( AA \)'s is given by \( P \),

\[
\begin{align*}
P &= p^4 + 2p^3q(1+s_1) + p^2q^2(1+s_2) \\
&= p^2(1+2pq + q^2s_2) \\
&= p^2w_1
\end{align*}
\]
Table 2.8. Frequencies of competition occurring between the same and different genotypes and probabilities of survival of the three genotypes in the diploid model

<table>
<thead>
<tr>
<th>Competition between</th>
<th>Frequency of competition</th>
<th>Conditional probability of survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA and AA</td>
<td>$p^4$</td>
<td>1</td>
</tr>
<tr>
<td>AA and Aa</td>
<td>$4p^3q$</td>
<td>$(1+s_1)/2$</td>
</tr>
<tr>
<td>AA and aa</td>
<td>$2p^2q^2$</td>
<td>$(1+s_2)/2$</td>
</tr>
<tr>
<td>Aa and Aa</td>
<td>$4p^2q^2$</td>
<td>1</td>
</tr>
<tr>
<td>Aa and aa</td>
<td>$4pq^3$</td>
<td>$(1+s_3)/2$</td>
</tr>
<tr>
<td>aa and aa</td>
<td>$q^4$</td>
<td>1</td>
</tr>
</tbody>
</table>

The frequency of Aa's is

$$Q = 2p^3q(l-s_1) + 4p^2q^2 + 2pq^3(1+s_3)$$

$$= 2pq(l-p^2s_1 + q^2s_3)$$

$$= 2pqw_2$$

And that for aa is

$$R = p^2q^2 + 2pq^3 + q^4 - p^2q^2s_2 - 2pq^3s_3$$

$$= q^2(1-p^2s_2 - 2pq_3)$$

$$= q^2w_3$$

To examine the change in the number of genes, we first consider
\[ n_1' = p^2 N v_1 k (1-c) + c N p^2 (1+2pqs_1+q^2 s_2) \]

\[ n_2' = 2pq N v_2 k (1-c) + c N 2pq (1-p^2 s_1+q^2 s_3) \]

This leads to

\[ n_A' = 2n_1' + n_2' \]

Thus

\[ \Delta n_A = n_A' - n_A = 2Np[(l-c)k v_1 p + (l-c)k v_2 q + cp + \]

\[ 2p^2 qcs_1 + q^2 pcs_2 + cq - cqp^2 s_1 + \]

\[ cq^2 s_3 - p - q] \]

\[ = n_A[k(p v_1 + q v_2) - 1 - c(k(p v_1 + q v_2) - 1 - \]

\[ q(p^2 s_1 + pqs_2 + q^2 s_3))] \]

\[ = n_A[a_A - c(a_A - q s)] \]

Similarly

\[ \Delta n_a = n_a[a_a - c(a_a + p s)] \]

where

\[ a_A = k(p v_1 + q v_2) - 1 \]

\[ a_a = k(p v_2 + q v_3) - 1 \]

\[ a = p a_A + q a_a \]

\[ s = p^2 s_1 + pqs_2 + q^2 s_3. \]

Now, \( \Delta N = 0.5(\Delta n_A + \Delta n_a) \)

\[ = 0.5(n_A a_A - c n_A a_A + n_A c q s + n a a - n a a - c p s) \]

\[ = n a (1-c) \quad (2.8.4) \]
In the saturated case

\[ p' = p^4 + 2p^3q + 2p^2q^2 + 2p^3q^1 + p^2q^2s^2 + p^3q - p^3q^1 + 2p^2q^2 + pq^3 + pq^3s^3 \]  

(2.8.5)

from which we get

\[ \Delta p = pq(p^2s_1 + pq^2s_2 + q^2s_3) \]  

(2.8.6)

Under genic selection, we may assume that \( s_1 = 0.5s_2 = s_3 = s \), so

\[ \Delta p = pqs. \]

The equilibrium position is given by \( \Delta p = 0 \). This occurs when

\[ p = 1, \text{ or } p = 0, \text{ or } p^2s_1 + pq^2s_2 + q^2s_3 = 0. \]  

(2.8.7)

The solution to this quadratic is

\[ p = \frac{-2s_3 - s_2 + \sqrt{s_2^2 - 4s_1s_3}}{2s(s_1 - s_2 + s_3)} \]  

(2.8.8)

The expression implies that \( s_2^2 > 4s_1s_3 \) for an equilibrium to exist, but this in itself may not necessarily imply that the solution for (2.8.8) will lie in the interval \((0,1)\). If \( p, q \in (0,1) \) then there is no solution to (2.8.7) for \( s_1, s_2, s_3 > 0 \). Thus one or more of \( s_1, s_2, s_3 \), must be negative for a solution. An approach towards understanding the dynamics of this model is given in Chapter 3.

2.8.4. Diploids competing at the genotypic level

In this case the fitnesses of AA, Aa, and as are \( w_1 = s_1/s \), \( w_2 = s_2/s \), and \( w_3 = s_3/s \) respectively, where \( s = p^2s_1 + 2pq^2s_2 + q^2s_3 \).

\[ w = p^2w_1 + 2pqw_2 + q^2w_3 = 1. \]
In a saturated population

\[ p' = p(p_1^* + q_2^*)/s \], which leads to

\[ \Delta p = p[p_1^* + q_2^* - p^2 s_1 - 2pq s_2 - q^2 s_3]/s \]

\[ = pq[p(w_1 - w_2) + q(w_2 - w_3)]. \]

Genotypic competition gives us the following

\[ n_1' = (1-c)Np^2 p_1 + cNp s_1/s \]

\[ n_2' = (1-c)N2pq p_2 + cN2pq s_2/s \]

\[ n_A' = 2n_1' + n_2' \]

so

\[ \Delta n_A = n_A [k(p v_1 + q v_2) - 1 - c(k(p v_1 + q v_2) - 1 - ps_1/s]

\[ - q s_2/s + s/s)] \]

\[ = n_A [a_A - c(a_A - q s)] \]

where

\[ \bar{s} = [p(s_1 - s_2) + q(s_2 - s_3)]/s = [p(w_1 - w_2) + q(w_2 - w_3)]. \]

The change in the number of A alleles is at the same rate as for individual competition.

2.8.5. Fertility excess for gene substitution

Nei (1971) now looks at the problem of determining the fertility excess and the genetic variance necessary for gene substitution. He uses only a haploid model which he claims is applicable to diploids in the absence of dominance, i.e., there is genic selection.
Most natural populations are near saturation so that

\[ p' = p(1+sq), \text{ and} \]

\[ \frac{dp}{dt} = \Delta p = pqs. \]

For gene substitution to proceed at this rate, the fitness of genotype A must be at least \( 1+sq \), i.e., \( k \) (fertility) \( \geq 1 + sq \).

For \( 1 < k < 1 + sq \), \( \Delta p < pqs \), and if \( k < 1 \) then the population size will decrease.

"Cost of natural selection" is equivalent to "substitutional load" or the total amount of reduction in mean fitness during gene substitution compared with the fitness of the advantageous genotype, which is assumed to be constant. In this formulation, mean fitness remains constant and the fitness of the advantageous genotype varies. This constancy of fitness was commented on earlier. The model has the same number of offspring surviving from each competing pair, an arguably unrealistic assumption. That aside, however, we have

\[ \Delta p = \frac{dp}{dt} \Rightarrow pqs dt = dp \]

\[ \Rightarrow E = \int_{0}^{\infty} (w_1 - \bar{w}) dt \]

\[ = \int_{0}^{\infty} (w_1 - pw_1 - qw_2) dt \]

\[ = \int_{0}^{\infty} [(1-p)w_1 - qw_2] dt \]

\[ = \int_{0}^{\infty} q(w_1 - w_2) dt \]

\[ = \int_{0}^{\infty} q(1+sq-l+sp) dt \]
This is "accumulated fertility excess necessary for a gene substitution" or "fertility excess required". We can compare this with Crow and Kimura's (1970, p. 252) "cost" which is defined by \(
\int_{0}^{\infty} \frac{(w_1 - \bar{w})}{\bar{w}} dt,
\)
and is described as the cost in excess in survival and fertility that the favored genotype must have to carry out the gene substitution at a specified rate while the entire population remains roughly constant.

Flake and Grant (1974) prefer the term "cost" to Nei's (1971) phrase in that they consider the integral as the integral of the relative loss of the gene that is replaced during substitution. Here the integral is scaled for a constant sized population. At this point we are in a mild dilemma: Flake and Grant (1974) refer to the integral, \(\int (m_A - \bar{m}) dt\), as the cost of substitution where \(m_A\) and \(\bar{m}\) are Malthusian parameters and \(\bar{m} = m_A p + m_A q\). The relationship between Malthusian parameters defined by \(\frac{dN_A}{dt} = m_A N_A\), and Wrightian fitness defined by \(\Delta N_A = N_A (1 - w_A)\) may be approximated by \(m_A \approx 1 - w_A\). We see, therefore, that the above integral may be written, as Nei (1971) writes it, as \(\int (w_A - \bar{w}) dt\), where \(\bar{w} = w_A p + w_A q\).

This, however, is for an increasing population. Where the population is held constant the integral has to be scaled thus: \(\int \frac{(w_A - \bar{w})}{\bar{w}} dt\).

Insofar as Nei's (1971) formulation has \(\bar{w}=1\) for all values of \(p\), this problem did not arise.

We now consider the case of noncompetitive growth. When \(c=0\),
\[ w_1 = (1-c)r_1 + c(1+sq) = r_1, \text{ and } w_2 = r_2. \] Then \[ E = \int_0^\infty (w_1 - w) dt = \int_0^\infty (w_1 - pw_1 - qw_2) dt = \int_0^\infty [(1-p)w_1 - qw_2] dt = \int_0^\infty q(r_1 - r_2) dt. \]

Previously we showed that \[ \Delta p = pq(a_1 - a_2)/(1+a) \text{ (cf. (2.8.2))}, \] i.e., \[ \Delta p = pq(r_1 - r_2)/(p+q+pa_1 + qa_2). \] The denominator may be re-written as \[ p+q+p(r_1 - 1) + q(r_2 - 1) = pr_1 + qr_2. \] So

\[
\frac{dp}{dt} = \frac{pq(r_1 - r_2)}{pr_1 + qr_2}
\]

so \[ q(r_1 - r_2) dt = [(pr_1 + qr_2) dp]/p = (r_1 - r_2) + (1/p)r_2 dp. \] Therefore,

\[ E = (r_1 - r_2)(1-p_0) + r_2(-\ln p_0), \]
and making the assumption that \(1-p_0 \approx 1\), we get

\[ E \approx (r_1 - r_2) - r_2\ln p_0. \]

The average fitness during gene substitution under complete competition \((c = 1)\) of the genotype \(A, \bar{f}\), is

\[
\frac{\int_0^t w_1 dt}{t} = \frac{\int_0^t (1+sq) dt}{t} = \frac{\int_0^t sq dt}{t} = \frac{\int_0^{p_1} (1/p) dp}{1/s \int_0^{p_1} (\frac{1}{p} + \frac{1}{1-p}) dp}
\]

\[
\frac{1}{s} \int_0^{p_1} (\frac{1}{p} + \frac{1}{1-p}) dp = \frac{1}{s} \left[ \ln p - \ln (1-p) \right]_{p_0}^{p_1} = \frac{2}{s} \left( \ln \frac{1-p_0}{p_0} \right) \approx -\frac{2}{s} \ln p_0,
\]

by use of the assumption that \(p_0 \approx 1-p_1\) and \(1-p_0 \approx 1\), i.e., that the gene frequency goes from a very small value to one very near to unity such that \(p_0 = 1-p_1\).
\[
\int_{p_0}^{p_1} \left( \frac{1}{p} \right) dp = \ln \left( \frac{p_1}{p_0} \right) = \ln \frac{1-p_0}{p_0}.
\]

So \( \bar{f} = 1 + 0.5s \).

The number of independent gene changes that can occur simultaneously, \( n \), is given by

\[ k \geq \left( \frac{1+s}{2} \right)^n \equiv \exp(ns/2) \text{ if } s \text{ is small.} \]

i.e., the fertility must be at least great enough for the gene changes to occur at the rate given by the above expression.

The expression, \( (1+s/2) \), is the average reproductive rate during gene substitution, and \( k \) is the average reproductive capacity. So \( n \leq 2(1/s) \ln k n \).

The number of generations is obtained from \( \int_{t_0}^{t_1} dt = (2/s) \ln (1/p_0) \),

(with the assumption that \( 1-p_0 \equiv 1 \)), and therefore the number of gene substitutions per generation is

\[ \frac{(2/s) \ln k n}{(2/s) \ln p_0} = \frac{\ln k}{\ln p_0} = v. \]

From Table 2.7 we calculate the variance of the fitness as

\[ p(1+sq)^2 + q(1-sp)^2 - [p(1+sq) + q(1-sp)]^2 \]

\[ = s^2 pq \]

The accumulated genetic variance is

\[ \int_{t_0}^{t_1} [(1+sq-1)^2 p + (1-sp-1)^2 q] dt = s \int_{p_0}^{p_1} dp = s(p_1 - p_0) \equiv s. \]
Average variance is

$$\bar{V} = \frac{\int_{t_0}^{t_1} s^2 pq \, dt}{t_1 - t_0} = \frac{s(p_1 - p_0)}{-(2/s)\ln p_0} = \frac{s^2}{-2\ln p_0}$$

The total genetic variance $V = n\bar{V}$ leads to the value, $v$, for the maximum number of independently assorted gene substitutions per generation, viz,

$$v = \frac{\ln k}{(-\ln p_0)} = 2\bar{V} \frac{\ln k}{s^2} = 2\bar{V} \frac{\ln k}{s^2} = \frac{V}{s}.$$ 

Thus the number of gene substitutions per generation may be expressed as a function of variance and selection intensity. If the gene substitutions were not independent then a genetic covariance term would appear in the expression as well as the genetic variance.

Nei (1971) fails to realize that his formulation rests heavily on the fact that the population competes only between pairs. His demonstration that the formulation is valid for more than two individuals in a competing group bases its argument on decomposing multiple member groups into competing pairs.

2.9. Truncation Selection

In order to understand the consequences of truncation selection, we shall review the treatment of this subject given by Griffing (1960) who based his work on that of Kimura (1958).

Both of these authors exploited a concept originally developed
by Fisher (1918) that the genetic variance of a character is due to the cumulative effect of a large number of independently segregating loci, each having a small effect. A further observation noted by Haldane (1932) is that the change in the probability that an individual will be selected, $\Delta w$, is proportionately related to the change in the measured character, $\Delta y$, due to a gene substitution, if both $\Delta y$ and $\Delta w$ are small. A treatment of the case where genes have large effects has been given by Latter (1965).

We assume that there is random mating among surviving adults, so that if allele $A_i$ has a frequency $p_i$ then $A_i A_i$ has a frequency $p_i^2$, and $A_i A_j$ has a frequency $2p_i p_j$ (because $A_i A_j = A_j A_i$). Also we have the genotypic value of $A_i A_j = d_{ij}$, which is scaled so that the original population mean

$$\mu_0 = \sum_{i,j} p_i p_j d_{ij} = 0.$$  

Now, $d_{ij}$ may be partitioned so that

$$d_{ij} = \alpha_i + \alpha_j + \delta_{ij},$$

where $\alpha_i$ is the additive effect of allele $A_i$, 

$$\alpha_i = \sum_j p_j d_{ij},$$

and $\delta_{ij}$ is the dominance deviation associated with genotype $A_i A_j$,

$$\delta_{ij} = d_{ij} - \alpha_i - \alpha_j.$$  

The total genotypic variance may be partitioned as follows, with the
usual properties that \( \sum_{i} p_{i} \alpha_{i} = \sum_{i} p_{i} \delta_{ij} = \sum_{j} p_{j} \delta_{ij} = 0. \)

\[
\sigma_{G}^{2} = \sum_{i} \sum_{j} p_{i} p_{j} d_{ij}^{2} = \sum_{i} \sum_{j} p_{i} p_{j} (\alpha_{i} + \alpha_{j} + \delta_{ij})^{2} \\
= \sum_{i} \sum_{j} p_{i} p_{j} \alpha_{j}^{2} + \sum_{i} \sum_{j} p_{i} p_{j} \alpha_{i}^{2} + \sum_{i} \sum_{j} p_{i} p_{j} \delta_{ij}^{2} \\
+ 2 \sum_{i} \sum_{j} p_{i} p_{j} \alpha_{i} \alpha_{j} + 2 \sum_{i} \sum_{j} p_{i} p_{j} \alpha_{i} \delta_{ij} + 2 \sum_{i} \sum_{j} p_{i} p_{j} \delta_{ij} \delta_{ij} \\
= 2 \sum_{i} p_{i} \alpha_{i}^{2} + \sum_{i} \sum_{j} p_{j} \delta_{ij}^{2} \\
+ 2 \sum_{i} p_{i} \sum_{j} p_{j} \alpha_{i} \alpha_{j} + 2 \sum_{i} p_{i} \sum_{j} p_{j} \alpha_{i} \delta_{ij} + 2 \sum_{i} p_{i} \sum_{j} p_{j} \delta_{ij} \delta_{ij} \\
= \sigma_{A}^{2} + \sigma_{D}^{2}
\]

We assume that the character under selection is controlled by many loci, each of which has a small effect. The phenotypic variability is distributed \( \mathcal{N}(0, \sigma^{2}) \).

We can consider, therefore, the population subdivided into groups, each corresponding to one genotype. The individuals in that subdivision associated with genotype \( A_{i} A_{j} \) are distributed \( \mathcal{N}(0, \sigma_{ij}^{2}) \).

To complete the development two further assumptions are invoked.

(i) \( d_{ij} \) is considerably less than \( \sigma \), so that \( (d_{ij}/\sigma)^{2} \) and \( (d_{ij}/\sigma^{2})^{2} \) are negligible.

(ii) Genotypic variance due to each locus is small relative to the phenotypic variance so that \( \sigma_{ij}^{2} \approx \sigma^{2} \).

If we impose truncation selection, i.e., we select all individuals whose phenotypic variance is above a certain point (\( x_{0} \) in the diagram),
then we find that in terms of the subdivisions the situation is better depicted as

The selection value, \( w_{ij} \), for the genotype \( A_iA_j \) is defined to be proportional to the probability that an individual of the genotype \( A_iA_j \) survives selection.

\[
P(x > x_0) = \frac{1}{\sqrt{2\pi\sigma}} \int_{x_0}^{\infty} \exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\} dx
\]

\[
= \frac{1}{\sigma_i j(2\pi)^\frac{1}{2}} \int_{x_0}^{\infty} \exp\left\{-\frac{1}{2}\frac{(x-d_{ij}+d_{ij})^2}{2\sigma_i^2}\right\} dx
\]

\[
= \frac{1}{\sigma(2\pi)^\frac{1}{2}} \int_{x_0}^{\infty} \exp\left\{-\frac{1}{2}\frac{(x-2xd_{ij}+d_{ij})^2}{2\sigma^2}\right\} dx
\]
where v is the proportion of the original population which is selected and $i$ is the selection differential.

$$i = \frac{1}{\sigma(2\pi)^{\frac{1}{2}}} \int_{-\infty}^{\infty} x \exp[-\frac{x^2}{2\sigma^2}] \frac{1}{v} dx$$

The selection differential, $i$, is defined as the average amount by which the individuals chosen as parents for the next generation exceed the population from which they were selected. A property of the normal curve is that the mean of all individuals beyond a truncation...
point (e.g., \( x_0 \) in the above diagram) is given by \( z/v \), where \( z \) is the height of the ordinate at that point, and \( v \) is the proportion selected. If \( z \) is the ordinate of a standard normal curve, then \( \bar{\theta} = z/v \) is the selection differential in standard deviations. To convert it to the selection differential expressed in terms of the units of measurement, it must be multiplied by \( \sigma \), the phenotypic standard deviation. Hence, \( \bar{\theta} = \bar{\theta} \sigma \). Thus the probability that \( A_iA_j \) survives selection is proportional to

\[
1 + \left( \frac{\bar{\theta}}{\sigma^2} \right) \bar{\theta} = 1 + \left( \frac{\bar{\theta}}{\sigma^2} \right) \bar{\theta}.
\]

Hence the selection value for \( A_iA_j \) is

\[
w_{ij} = 1 + \frac{1}{\sigma^2} d_{ij}.
\]

We shall designate the allelic frequencies after the \( k \)-th cycle of selection as \( p_i^{(k)} \), so initially the frequency of \( A_iA_j \) is \( p_i^{(0)} p_j^{(0)} \). The frequency of \( A_iA_j \) following selection is

\[
p_i^{(0)} p_j^{(0)} w_{ij} = p_i^{(0)} p_j^{(0)} + \frac{1}{\sigma^2} p_i^{(0)} p_j^{(0)} d_{ij}.
\]

The total frequency of all selected genotypes is, of course, one, so that

\[
\sum_{i} \sum_{j} [p_i^{(0)} p_j^{(0)} + \frac{1}{\sigma^2} p_i^{(0)} p_j^{(0)} d_{ij}] = 1
\]

The genotypic mean of selected parents is
where $h_B^2$ is heritability in the broad sense. The frequency of allele $A_i$ in the selected population is

$$p_i^{(1)} = \frac{1}{\sigma^2} \sum_{j} p_i^{(0)} p_j^{(0)} d_{ij}$$

$$= p_i^{(0)} \left( 1 + \frac{1}{\sigma^2} \alpha_i \right).$$

Therefore the gene array produced by the selected parents is

$$\sum_{i} p_i^{(1)} A_i = \sum_{i} p_i^{(0)} A_i \left( 1 + \frac{1}{\sigma^2} \alpha_i \right).$$

The progeny array is

$$\left[ \sum_{i} p_i^{(1)} A_i \right]^2 = \sum_{i} \sum_{j} p_i^{(0)} p_j^{(0)} A_i A_j + \frac{1}{\sigma^2} \sum_{i} \sum_{j} p_i^{(0)} p_j^{(0)} A_i A_j (\alpha_i + \alpha_j)$$

$$+ \frac{1}{\sigma^2} \sum_{i} \sum_{j} p_i^{(0)} p_j^{(0)} A_i A_j \alpha_i \alpha_j,$$

so the mean is

$$\mu_1 = 0 + \frac{1}{\sigma^2} \sum_{k} \sum_{j} p_i^{(0)} p_j^{(0)} d_{ij} (\alpha_i + \alpha_j) + \frac{1}{\sigma^2} \sum_{i} \sum_{j} p_i^{(0)} p_j^{(0)} \frac{\alpha_i}{\sigma} \frac{\alpha_j}{\sigma} d_{ij},$$

and if $\sigma$ is sufficiently small the last term is negligible, so that

$$\mu_1 = \frac{1}{\sigma^2} 2 \sum_{i} p_i \alpha_i^2 = \frac{1}{\sigma^2} \sigma_A^2 = h_N^2$$

where $h_N^2 = \sigma_A^2 / \sigma^2$, is heritability in the narrow sense.
2.10. Selection with Reference to Biological Groups

Most of the work discussed have been concerned with how genotypes react against competition from neighboring genotypes, particularly with respect to the change in gene frequencies from one generation to the next. Sakai (1955), in work already discussed in Section 2.1, developed expressions for calculating genetic variances, from which estimates of heritability can be made. The major study, however, in selection under competitive conditions was developed by Griffing (1967, 1968a, 1968b, 1969, 1976a, 1976b) which was summarized by Griffing (1977).

Griffing was primarily motivated by the report by Wiebe, Petr and Stevens (1963) showing that a positive selection for yield in barley resulted in a negative response. They concluded their study with:

Significant reversals in relative yield were found to exist in comparisons between the same genotype, WW or vv, when grown in pure stand and in an advanced generation, thus, indicating that the poorest plants should be saved from an advanced hybrid population rather than the good ones when yield is the criterion for selection. If this phenomenon has a degree of universality, then it may explain why breeding for increased yield has progressed so slowly.

Griffing (1967) commenced by considering groups of size two. The genotypic array of groups can be represented as

\[ \sum \sum \sum \sum \sum_{i,j,x,y} p_{i,j} p_{x,y} (A_i A_j, A x y). \]

We take \( A_i A_j \) to be the genotype being measured and \( A x y \) to be the genotype in competition with \( A_i A_j \). The genotypic value of \( A_i A_j \) in the
group is denoted \( ij^d_{xy} \), and coded so that

\[
\sum \sum \sum p_i p_j p_x p_y (ij^d_{xy}) = 0
\]

We may describe the genotypic value as

\[
ij^d_{xy} = d_{i} + d_{j} + d_{ij} + a_{x} + a_{y} + \delta_{xy} \\
+ da(\alpha\alpha)_{ix} + da(\alpha\alpha)_{iy} + da(\alpha\alpha)_{jx} + da(\alpha\alpha)_{jy} \\
+ da(\alpha\delta)_{ixy} + da(\alpha\delta)_{jxy} + da(\delta\alpha)_{ijx} + da(\delta\alpha)_{ijy} \\
+ da(\delta\delta)_{ijxy}.
\]

The direct additive effect of allele \( A_i = d_i = \cdot d.. = \)

\[
\sum p_j p_x p_y (i^d_{jxy}).
\]

With similar definitions for the dot summation notation, we may write the other effects;

direct dominance effect of \( A_i A_j = d_{ij} = i^d.. - d_i - d_j \),

associative additive effect of allele \( A_x = a_x = ..d.x. \),

associative dominance effect of \( A_x A_y = \delta_{xy} = ..d_{xy} - a_x - a_y \)

additive x additive interaction between alleles \( A_i \) and \( A_x \)

\[
= da(\alpha\alpha)_{ix} = i^d.. - a_i - a_x.
\]

additive x dominance interaction between allele \( A_i \) and \( A_x A_y \)

\[
= da(\alpha\delta)_{ixy} = i^d.. - a_i - a_x - a_y - \delta_{xy}.
\]
dominance x dominance interaction between $A_iA_j$ and $A_xA_y$ 

$$= \sigma_{ijxy}^2 = i^d_{xy} - \sigma_{i} - \sigma_{j} - \sigma_{ij} - \sigma_{i}^x - \sigma_{i}^y - \sigma_{xy}^d$$

$$- \sigma_{ij}^x - \ldots - \sigma_{ijxy}^d.$$ 

In his model, however, Griffing (1967) considered only the additive effects, in order to keep the system relatively simple. This simplicity was also incorporated into the covariance between the two genotypes in a group which was confined to the additive genetic covariance between direct and associate effects. Thus the only genetic variances and covariances incorporated into the model were:

$$\sigma_G^2 = \sum p_ip_jp_xp_y(i^d_{xy})^2$$

$$\sigma_A^2 = 2\sum p_i\sigma_i$$

$$\sigma_D^2 = \sum p_ip_j\sigma_{ij}$$

$$\sigma_A^2 = 2\sum p_x\sigma_x$$

$$\sigma_D^2 = \sum p_y\sigma_y$$

$$\sigma_{ij}^2 = 2\sum p_i\sigma_d$$

Other terms may be readily identified; $\sigma_{AA}$, $\sigma_{DA}$, $\ldots$. Griffing (1967) makes no use of them apart from noting that the total genotypic variance may be partitioned thus:

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_A^2 + \sigma_D^2 + \sigma_{AA}^2 + \sigma_{AD}^2 + \ldots$$
2.10.1. Consequences of individual selection

Following Griffing (1960) (see Section 2.9) we can see that the selection value for $A_i A_j$ when summed over all groups is

$$w_{ij} = 1 + (\frac{i}{\sigma^2})_{\text{ind.}}(\alpha_i d_{ij}),$$

where the subscript "ind." implies that selection is practiced on the individual. The expected gametic array produced by $A_i A_j$ is $\frac{1}{2}(A_i + A_j)$. Therefore, the expected gametic array from all selected individuals is

$$\frac{1}{2} \sum_{i,j} p_i p_j w_{ij} (A_i + A_j) = \sum_i p'_i (A_i)$$

where

$$p'_i = p_i [1 + (i/\sigma^2)_{\text{ind.}}(\alpha_i)].$$

The group population mean in the generation is

$$\mu_1 = \sum p_i p'_i p'_i [\frac{1}{2} p i d_{ij}],$$

$$= \sum p_i p'_i p'_i (i d_{xy})$$

$$= \sum p_i p_j p'_i (1 + (i/\sigma^2)_{\text{ind.}}[(\alpha_i + \alpha_j) + (\alpha_i + \alpha_j)])(i d_{xy})$$

$$= (\frac{i}{\sigma^2})_{\text{ind.}}(\sigma^2 A + d A).$$

(2.10.1)

Now, we can see that if the covariance $d A$ is negative and large enough in magnitude i.e., if it is less than $-\frac{1}{2} \sigma^2 A$, then the response to selection will be negative. Even if this is not the case, a negative value of $d A$ will reduce the response from selection.
Griffing (1967) points out that this anomaly may be ameliorated by group selection.

2.10.2. Consequences of group selection

In group selection, the entire group is accepted or rejected on the basis of the average group mean. In this case the selection value for the group \((A_i A_j, A_x A_y)\) is

\[
W_{ij, xy} = 1 + \left(\frac{i}{\sigma^2}\right)_{gr.} \frac{1}{2} \left(\frac{d_{ij}}{x y} + d_{ij}\right),
\]

where the subscript "gr." implies the selection is based on the groups.

The expected gametic array from this group is

\[
\frac{1}{4} (A_i + A_j + A_x + A_y),
\]

and the total expected gametic array after selection is

\[
\frac{1}{4} \sum_{i,j} p_i p_j x_{ij} (w_{ij, xy} (A_i + A_j + A_x + A_y) = \Sigma p'_i (A_i)
\]

where

\[
p'_i = p_i \left[1 + \left(\frac{i}{\sigma^2}\right)_{gr.} \frac{1}{2} (d_{i1} + a_{i1})\right].
\]

The group mean of the population in the next generation is

\[
\mu_1 = \Sigma p'_i p'_j p'_k x_{ij} \left[\frac{1}{2} d_{ij} + a_{ij}\right] = \frac{1}{2} \left(\frac{i}{\sigma^2}\right)_{gr.} (d_{iA}^2 + 2 d_{aA} + a_{iA}^2).
\]

The observation that this last expression can be re-expressed as

\[
\mu_1 = \left(\frac{i}{\sigma^2}\right)_{gr.} [\Sigma p_i (d_{i1} + a_{i1})^2]
\]

shows that the effect of group selection is always to have a positive response.
A similar approach can be made considering groups of size $n$. In this case, Griffing (1967) showed that individual selection in large groups would be negative even if the covariance is only slightly negative. The expression (2.10.1) becomes, for groups of size $n$,

$$
\mu_1 = \frac{1}{n^{2}} \sigma^2 \text{ind.} \left[ \sigma^2_A + (n-1) \sigma_A^A \right],
$$

which shows that for $\sigma^2_A < \frac{\sigma^2_A}{n-1}$ the response to selection will be negative. On the other hand, selection based on group means will always result in a positive response. The group mean of the population following one cycle of selection can be seen to be

$$
\mu_1 = \frac{1}{n} (1/n) \left( \frac{1}{\sigma^2} \right) \text{gr.} \left( \sigma^2_A + (n-1)^2 \sigma^2_A \right).
$$

2.10.3. Selection in full-sib groups

If we cross $A_i A_j \times A_k A_l$ then the full-sib array produced is

$$
\frac{1}{4} (A_i A_k + A_i A_l + A_j A_k + A_j A_l),
$$

hence the array of full-sib groups of size two is

$$
\frac{1}{4} (A_i A_k + A_i A_l + A_j A_k + A_j A_l) \times \frac{1}{4} (A_i A_k + A_i A_l + A_j A_k + A_j A_l)
$$

$$
$$
If the array of groups is designated by \( M(ijkl)^2 \) then the entire conceptual population of full-sub groups of size two may be given as

\[
\sum_i \sum_j \sum_k \sum_l p_i p_j p_k p_l [M(ijkl)^2]^2.
\]

Griffing (1976a) in developing this model considered only the additive genetic variances and covariances, so we may write the genotypic value of \( A_i A_k \) in the full-sib group \( (A_i, A_k, A_j, A_l) \) as

\[
\text{ik}^d_{jl} = \delta_i + \delta_k + \delta_j + \delta_l,
\]

where the terms are those previously defined.

The selection value of \( A_i A_k \) in groups \( (A_i, A_k, A_j, A_l) \) is

\[
\text{ik}^w_{jl} = 1 + (i/\sigma^2) (\text{ik}^d_{jl}),
\]

and the selection value of \( A_i A_k \) averaged over full-sib groups generated by the cross \( A_i A_j x A_k A_l \) is \( \frac{1}{4} (\text{ik}^w_{ik} + \text{ik}^w_{il} + \text{ik}^w_{jk} + \text{ik}^w_{jl}) \).

Hence, the selection value of \( A_i A_k \) averaged over all crosses capable of producing \( A_i A_k \) progeny is

\[
\text{ik} = \sum_j p_j p_k \left\{ \frac{1}{4} (\text{ik}^w_{ik} + \text{ik}^w_{il} + \text{ik}^w_{jk} + \text{ik}^w_{jl}) \right\},
\]

\[
= 1 + (i/\sigma^2) \left[ \alpha_i + \alpha_k + \frac{1}{2} (\alpha_i + \alpha_k) \right].
\]

The selected gametic array then becomes

\[
\sum_i \sum_k p_i p_k (\text{ik}) \left( \frac{1}{2} (A_i + A_k) \right) = \sum_i p_i (A_i),
\]

where
\[ p'_i = p_i \left\{ 1 + \left( \frac{1}{\sigma^2} \right) \left[ d_i^2 a_i + \frac{1}{2a_i} \right] \right\} . \]

The change in gene frequency can now be shown to be

\[ \Delta p_i = \left( \frac{1}{\sigma^2} \right) \left( p_i \right) \left( d_i^2 a_i + \frac{1}{2a_i} \right). \]

If, having selected an individual on the basis of its performance in a full-sib group, as we have done here, we evaluated it with random groups the change in population mean can be calculated as

\[ \Delta \mu = \sum \sum \sum \sum \sum \left( p_i' p_j' p_k' p_l' \right) \left( d_i d_j d_k d_l \right) \]

\[ = \left( \frac{1}{\sigma^2} \right) \left[ \left( \frac{\sigma^2}{d} \right) + \left( \frac{3}{2a} \right) \right]. \]

Alternatively, we could select on the basis of the group performance, rather than individuals within the group. We could take either the entire group, or an individual within the group, for parental material for the next generation.

In the former case the full-sib group \((A_i A_i, A_j A_j)\) has a selective value \(w_{ik,jl}\). The gametic array produced by this group is

\[ \frac{1}{4} (A_i A_i + A_j A_k), \]

and the expected gametic array from the entire selected population is

\[ \sum (p_i p_j) (p_k p_l) (1/16) \left[ \frac{1}{4} w_{ik,jl} (A_i A_i + A_j A_k) + \frac{1}{4} w_{kl} (A_i A_i + A_j A_k) \right. \]

\[ + \frac{1}{4} \sum w_{ik,jl} (A_i A_i + A_j A_k) + \left. \frac{1}{4} w_{kl} (A_i A_i + A_j A_k) \right] \]

If we sum this over \((j,k,l)\) then the selected gametic array can be
written as (with the usual definition of the dot notation, e.g.,

\[ \sum_{k} \sum_{l} \varphi_{kl} w_{ikl} = w_{ij..} \]

\[ \sum_{i} (1/16) (2w_{i..i.} + w_{i..i.} + w_{...i.}) (A_{i}) \]

= \[ \sum_{i} (1/16) 3(\alpha_{i} + \alpha_{i})(A_{i}) \].

However, all four alleles contribute, so the total selected gametic array can be written as

\[ \sum_{i} \varphi_{i} (A_{i}) \]

where,

\[ \varphi_{i} = p_{i} [1 + (i/\sigma^{2})^{gr} (3/4) (\alpha_{i} + \alpha_{i})] \].

To turn to the second case, that is where the parental material for the next generation is the individual within a selected full-sib group, we consider the full sib groups produced by the mating \( A_{i} A_{j} \times A_{k} A_{l} \), and containing an individual, \( A_{i} A_{k} \). Then the frequency of \( A_{i} A_{k} \) following selection is

\[ (1/8) (2w_{ik},ik^{+w}ik,il^{+w}ik,jk^{+w}ik,jl^{+w}ik,ik^{+w}jk,ik^{+w}jl,ik) \].

From this we can calculate the overall selection value associated with \( A_{i} A_{k} \) by summing over all crosses capable of producing \( A_{i} A_{k} \).

\[ w_{ik} = (1/8) (2w_{ik},ik^{+w}ik,i^{+w}ik,,k^{+w}ik,,i^{+w}ik,,ik^{+w}k,ik^{+w}..,ik) \]

= \[ 1 + (i/\sigma^{2})^{gr} [(3/4)(\alpha_{i} + \alpha_{i})] + [(3/4)(\alpha_{i} + \alpha_{i})] \]
The total selected gametic array is

$$\sum_{i} \sum_{k} p_i p_k (w_{ik}) \left( \frac{1}{2} (a_i + A_i) \right) = \sum_{i} p'_i (A_i),$$

where,

$$p'_i = p_i \left[ 1 + \left( \frac{i}{\sigma^2} \right) \right] \left( 3/4 \right) \left( \alpha_i + \sigma_i^2 \right),$$

which is the same result as before with the entire group forming the parental material, hence the change in the population mean is the same, viz,

$$\Delta \mu = \left( \frac{i}{\sigma^2} \right) \left( 3/4 \right) \left( \sigma_i^2 + 2 \alpha_i \sigma_i + \sigma_i^2 \right),$$

which may be written in the form,

$$\left( \frac{i}{\sigma^2} \right) \left( 3/4 \right) \left( 2 \right) \left( \alpha_i + \sigma_i^2 \right)^2,$$

which shows that the change in population mean is nonnegative.

Griffing (1976a) extends this development to consider groups of size $n$. Following an argument similar to that outlined above, he shows that if the entire full-sib group is selected, or if a random individual from a selected full-sib group is taken for parental material, then the change in the population mean is approximately

$$\Delta \mu = \left( \frac{i}{\sigma^2} \right) \left( \frac{n+2}{2n} \right) \left( 2 \right) \left( \sigma_i^2 + 2 \sigma_i \sigma_i^2 \right) \left( \alpha_i + (n-1) \sigma_i^2 \right)^2,$$

$$= \left( \frac{i}{\sigma^2} \right) \left( \frac{n+2}{2n} \right) \left[ \sigma_i^2 + 2(n-1) \sigma_i \sigma_i^2 \right] \left[ \alpha_i + (n-1)^2 \sigma_i^2 \right],$$

the former expression implies that $\Delta \mu \geq 0$. 

2.10.4. **Clonal groups**

The use of clonal groups for breeding purposes has been, until recently, confined to plants and lower organisms. A clonal group of size n has the genotypic array,

\[(A_i A_j, A_i A_j, \ldots, A_i A_j).\]

Griffing (1976b) shows, by methods identical to those already presented, that for either individual or group selection,

\[\Delta \mu \approx (i/\sigma^2)_{\text{ind.}} \left[ d \sigma_A^2 + 2(n-1) da_a \sigma_a + (n-1)^2 da_a \sigma_a^2 \right].\]

Similarly, Griffing (1976b) also shows that the same approach applied to clonal groups of size n of homozygous genotypes leads to the result,

\[\Delta \mu \approx 2(i/\sigma^2) \left[ d \sigma_A^2 + 2(n-1) da_a \sigma_a + (n-1)^2 da_a \sigma_a^2 \right].\]

Computer simulations presented by Griffing (1976b) show that, with this genetic model, group selection has a marked effect on the efficacy of selection. Negative responses (noted in individual selection practiced on randomly mating and inbred groups) became positive, and selection responses that were positive under individual selection (those for full-sib and clonal groups) became higher under the group selection regime.
3. COMPETITION IN GROUPS GREATER THAN SIZE TWO

3.1. Generalization of the Pairwise Model

In order to extend the pairwise model of Nei (1971) to higher-order groups we first simplify the pairwise model discussed in the previous chapter (Table 2.8, Chapter 2), by placing the model in the setting of the model similar to that due to Sakai (1955) discussed earlier in Chapter 2 (Figure 2.2). In this case the competitive values may be illustrated as in Figure 3.1.

\[ \begin{array}{ccc}
\text{Genotype} & AA & Aa & aa \\
\text{Competitive value} & s & hs & -s \\
\end{array} \]

Figure 3.1. Model adapted from Sakai (1955) showing the competitive values of diploid genotypes

To explain we say that AA in competition with Aa gains \( \frac{1}{2}s(1-h) \), while Aa loses \( \frac{1}{2}s(1-h) \) in their respective probabilities of survival. Likewise, AA in competition with aa gains \( \frac{1}{2}(2s) \), and aa loses \( \frac{1}{2}(2s) \). Aa in competition with aa gains \( \frac{1}{2}s(1+h) \) and aa loses \( \frac{1}{2}s(1+h) \). Making the following substitutions,

\[ S_1 = s(1-h), \]
\[ S_2 = 2s \]
\[ S_3 = s(1+h), \]

we can rewrite Table 2.8 as Table 3.1.
Table 3.1. Frequencies of competition occurring between the same and different genotypes and probabilities of survival of the three genotypes in the diploid model

<table>
<thead>
<tr>
<th>Competition between</th>
<th>Frequency of occurrence</th>
<th>Conditional probability of survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA, AA</td>
<td>$p^4$</td>
<td>AA, AA, AA</td>
</tr>
<tr>
<td>AA, Aa</td>
<td>$4p^3q$</td>
<td>AA, Aa, Aa</td>
</tr>
<tr>
<td>AA, aa</td>
<td>$2p^2q^2$</td>
<td>AA, a, Aa, aa</td>
</tr>
<tr>
<td>Aa, Aa</td>
<td>$4p^2q^2$</td>
<td>AA, a, Aa, aa</td>
</tr>
<tr>
<td>Aa, aa</td>
<td>$4pq^3$</td>
<td>AA, a, Aa, aa</td>
</tr>
<tr>
<td>aa, aa</td>
<td>$q^4$</td>
<td>AA, a, Aa, aa</td>
</tr>
</tbody>
</table>

Hence,

$$p' = p^4 + 2p^3q + p^2q^2 + 2p^2q^2 + pq^3 + p^3q + 2p^3qs(1-h)$$

$$+ 2p^2q^2s - p^3qs(1-h) + pq^3s(1+h)$$

$$= p(p+q)^3 + pq(2p + q^2(1+h))$$

$$= p[l + qs(l+(q-p)h)]$$

(3.1.1)

Thus,

$$\Delta p = pq(1+(q-p)h).$$

(3.1.2)
Hence the equilibrium value for $p$ is

$$\overline{p} = \frac{1+h}{2h},$$

(3.1.3)

which implies that only a heterotic situation ($h>1$) will lead to a nontrivial equilibrium value.

We note that the equilibrium value is independent of $s$.

The stability of this system close to the equilibrium may be investigated formally:

$$p' = F(p) = p\{1+qs[1-h(p-q)]\} \quad (3.1.4)$$

$$p = F(\overline{p})$$

$$p' = \overline{p} + \delta'p = F(p) = F(\overline{p}+\delta p) = F(p) + \delta \frac{\partial F}{\partial p} \bigg|_{\overline{p}},$$

where $\frac{\partial F}{\partial p} \bigg|_{\overline{p}}$ is the first derivative of $F(p)$ w.r.t. $p$, evaluated at $\overline{p}$.

So

$$\delta'p = \delta p \frac{\partial F}{\partial p} \bigg|_{\overline{p}}.$$

$$\frac{\partial F}{\partial p} = \{1 + qs[1-h(p-q)]\} + p\{-s[1-h(p-q)]-2qsh\}$$

$$\frac{\partial F}{\partial p} \bigg|_{\overline{p}} = 1 + 0 - 2\overline{p}(1-\overline{p})sh$$

$$= 1 - \frac{2(h+1)(h-1)sh}{(2h)^2} = 1 - \frac{s(h^2-1)}{2h}.$$

So, for the equilibrium to be stable,

- if $h>1$ then $0<s<4h/(h^2-1)$,
- if $h<-1$ then $4h/(h^2-1)<s<0$.

If, in the former case, $h>1$ but $s>4h/(h^2-1)$, then the system will have a
stable limit cycle about the equilibrium point. A similar result occurs if, in the second case, \( h < -1 \) but \( s < 4h/(h^2 - 1) \), except that for some values of \( p_0 \) (the starting value) that are well above the equilibrium value \( p' < 0 \) and the system will go to fixation in one step. We may summarize this analysis in the following table (Table 3.2).

Table 3.2. Value boundaries for the parameters \( s, h \), and the equivalent \( S_1, S_2, S_3 \), for which the model described by expression (3.1.1) is stable or unstable.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>s</td>
</tr>
<tr>
<td>&gt;0</td>
<td>&gt;1</td>
</tr>
<tr>
<td>&gt;0</td>
<td>&lt;1</td>
</tr>
<tr>
<td>&lt;0</td>
<td>&gt;1</td>
</tr>
<tr>
<td>&lt;0</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>

However, since probabilities of survival must lie in the interval \([0,1]\), we can see that in Table 3.1, if \( h > 1 \) then \( s < 1/(1+h) \). This implies that

\[
s < \frac{1}{1+h} \cdot 4 < \frac{1}{1+h} \cdot \frac{4}{1-1/h} = \frac{4h}{h^2 - 1},
\]

so the oscillatory condition never arises.

Figure 3.2 depicts the change in gene frequency (expression (3.1.)) for different parameters.
Figure 3.2a. Graph of gene frequency for a population of individuals in competition (Model 3.1.1). This figure shows the effect of different values for \( h \) (\(+++\): \( s = 0.1, h = 1.9 \); \( ---\): \( s = 0.1, h = 0.5 \)). In the former case the change in gene frequency is initially rapid but lessens as the equilibrium, \( p = 0.76 \), is approached. In the latter case the change in gene frequency will continue to approach fixation.
Figure 3.2b. Graph of gene frequency for a population of individuals in competition (Model 3.1.1). This figure shows the effect of different values for $s$ (+++: $s = 0.1$, $h = 1.1$; ---: $s = 0.3$, $h = 1.1$). In both cases the equilibrium value is $p = 0.95$. 
The graphs bear out what is implied by Expressions (3.1.1) and (3.1.2). While $p$ is small (circa 0.2, say) $A_p$ is fairly large, but decreases as $p$ increases and the equilibrium point is approached. $A_p$ is also higher for $h = 1.9$ than for $h = 0.5$ (cf. 3.1.1). Expression (3.1.3) shows, however, that a larger $h$ implies an equilibrium value less close to $p = 1.0$, again born out by the curves.

The second graph shows what is implied by (3.1.1). A larger $s$ implies a more rapid initial $A_p$, even though the equilibrium point is the same.

3.2. Triplet Diploid Model with Constant Fitness

3.2.1. Model 1

We shall develop the model keeping the original concept of Nei (1971) that there is a constant number, viz, one, of survivors from each competing group. In the immediate development there is exactly one survivor from a competing triplet. In this case it is assumed that, other things being equal, each individual of a triplet has an equal chance (1/3) of survival.

The table of probabilities of survival are listed in Table 3.3. Adults mate at random, so that the genotypic array of offspring is $p^2A A + 2pqAa + q^2aa$, where $p (= 1 - q)$ is the frequency of allele $A$ in the parent population.

The frequencies of the surviving genotypes $AA$, $Aa$, and $aa$ are $P$, $Q$, and $R$ respectively, where
\[ P = p^6 + (4+6S_1)p^5q + (2+3S_2)p^4q^2 + (4+12S_3)p^4q^2 + (4+12S_4)p^3q^3 + (1+3S_5)p^2q^4 \]

\[ = p^2[1 + 3\{2p^3qS_1 + p^2q^2(S_2 + 4S_3) + 4pq^3S_4 + q^4S_5\}] \]

\[ Q = (2-6S_1)p^5q + (8-12S_3)p^4q^2 + 12(1+S_6)p^3q^3 \]

\[ + (8+12S_7)p^2q^4 + (2+6S_8)pq^5 \]

\[ = 2pq[1 - 3(p^4S_1 + 2p^3qS_2 - 2p^2q^2S_6 - 2pq^3S_7 - q^4S_8)] \]

\[ R = (1-3S_2)p^4q^2 + (4-12(S_4+S_6))p^3q^3 \]

\[ + (2-3S_5 + 4-12S_7)p^2q^4 + (4-6S_8)pq^5 + q^6 \]

\[ = q^2[1 - 3(p^4S_2 + 4p^3q(S_4+S_6) + p^2q^2(S_5+4S_7) + 2pq^3S_8)] \]

Thus, the allelic frequency, \( p' \), for the next generation is given by

\[ p' = P + 0.5Q \]

\[ = p^2 + pq + (6p^5q - 3p^5q)S_1 + 3p^4q^2S_2 + (12p^4q^2 - 6p^4q^2)S_3 \]

\[ + 12p^3q^3S_4 + 3p^2q^4S_5 + 6p^2q^3S_6 + 6p^2q^4S_7 + 3pq^5S_8 \]

\[ = p + 3pq[p^4S_1 + p^3q(S_2 + 2S_3) + 2p^2q^2(2S_4 + S_6) \]

\[ + pq^3(S_5 + 2S_7) + q^4S_8] \]  \hspace{1cm} (3.2.1)
Similarly

\[ q' = 0.5Q + R \]

\[ = q - 3pq\{p^4S_1 + p^3q(S_2 + 2S_3) + 2p^2q^2(2S_4 + S_6) \}
+ pq^3(S_5 + 2S_7) + q^4S_8 \} . \]

From this we have that the change in allelic frequency

\[ \Delta p = 3pq\{p^4S_1 + p^3q(S_2 + 2S_3) + 2p^2q^2(2S_4 + S_6) + pq^3(S_5 + 2S_7) \]
+ q^4S_8 \} . \quad (3.2.2) \]

Equilibrium values exist for \( p=0 \), or \( p=1 \), and for \( p(= 1 - q) \) equal to the roots of the quartic

\[ p^4S_1 + p^3q(S_2 + 2S_3) + 2p^2q^2(2S_4 + S_6) + pq^3(S_5 + 2S_7) + q^4S_8 = 0 . \quad (3.2.3) \]

Expressed in terms of \( p \) only, this is

\[ a_4p^4 + a_3p^3 + a_2p^2 + a_1p + a_0 = 0 \]

where

\[ a_4 = S_1 - S_2 - 2S_3 - 4S_4 + 2S_6 - S_5 - 2S_7 + S_8 \]
\[ a_3 = S_2 + 2S_3 - 8S_4 - 4S_6 + 3S_5 + 6S_7 - 4S_8 \]
\[ a_2 = 4S_4 + 2S_6 - 3S_5 - 6S_7 - 6S_8 \]
\[ a_1 = S_5 + 2S_7 - 4S_8 \]
\[ a_0 = S_8 \]
Table 3.3. Frequencies of competition occurring between the same and different genotypes in groups of three, and probabilities of survival of the three genotypes in the diploid model

<table>
<thead>
<tr>
<th>Members of competing group</th>
<th>Frequency of competition</th>
<th>Conditional probability of survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA, AA, AA</td>
<td>$p^6$</td>
<td>$1$</td>
</tr>
<tr>
<td>AA, AA, Aa</td>
<td>$6p^5q$</td>
<td>$2/3 + S_1$</td>
</tr>
<tr>
<td>AA, AA, aa</td>
<td>$3p^4q^2$</td>
<td>$2/3 + S_2$</td>
</tr>
<tr>
<td>AA, Aa, Aa</td>
<td>$12p^4q^2$</td>
<td>$1/3 + S_3$</td>
</tr>
<tr>
<td>AA, Aa, aa</td>
<td>$12p^3q^3$</td>
<td>$1/3 + S_4$</td>
</tr>
<tr>
<td>AA, aa, aa</td>
<td>$3p^2q^4$</td>
<td>$1/3 + S_5$</td>
</tr>
<tr>
<td>Aa, Aa, Aa</td>
<td>$8p^3q^3$</td>
<td>$1$</td>
</tr>
<tr>
<td>Aa, Aa, aa</td>
<td>$12p^2q^4$</td>
<td>$2/3 + S_7$</td>
</tr>
<tr>
<td>Aa, aa, aa</td>
<td>$6pq^5$</td>
<td>$1/3 + S_8$</td>
</tr>
<tr>
<td>aa, aa, aa</td>
<td>$q^6$</td>
<td>$1$</td>
</tr>
</tbody>
</table>
In order to get some understanding of the nature of this model we shall employ the following simplifications.

1. Heterozygote superiority; \( -S_1 = -S_3 = S_6 = S_7 = S_8 = S_{(1)} > 0 \)

2. AA genotype more successful than aa when both are competing against each other; \( S_2 = S_4 = S_5 = S_{(2)} > 0 \).

We make the further assumption that \( S_{(1)} = 0.5S_{(2)} = S \). This is similar to the assumption made by Nei (1971).

The change in allelic frequency (expression (3.2.2) may now be written

\[
\Delta p = 3Sp(1-p)(6p^4-12p^3+4p^2+1) \tag{3.2.4}
\]

An equilibrium exists for \( p = 0 \), \( p = 1 \), and for \( p \) equal to the roots of the quartic,

\[
6p^4 - 12p^3 + 4p^2 + 1 = 0,
\]

i.e., \( p = 0.770564316 \), and

\[ p = 1.510148508. \]

Two complex conjugate roots also exist;

\[ p = -0.13735 + i0.35265. \]

To test the stability of the equilibrium points we may use a numerical method that shows that the nontrivial equilibrium, \( p = 0.77056 \ldots \), is stable for a value of \( S = 0.1 \). More formally we may use the following analytical approach.

Express (3.2.1) as
\[ p' = F(p) \]
\[ = p - 3S(6p^6 - 18p^5 + 16p^4 - 4p^3 + p^2 - p) \]
\[ \frac{\partial F}{\partial p} = 1 - 3S(36p^5 - 90p^4 + 64p^3 - 12p^2 + 2p - 1) \]

So

\[ \frac{\partial F}{\partial p} \bigg|_{p=0} = 1 + 3S \]

which will be greater than 1 for all values of \( s \) greater than zero. The equilibrium point is unstable. Similarly,

\[ \frac{\partial F}{\partial p} \bigg|_{p=1} = 1 + 3S \]

and the same conclusion holds. For these two equilibrium points, the system is unstable.

\[ \frac{\partial F}{\partial p} \bigg|_{p=0.77} = 1 - 3S(0.747892671) \]

So a stable equilibrium exists at \( p = 0.77056 \ldots \) for \( 0 < s < 0.89139 \ldots \).

3.2.2. **Triplet diploid model with constant fitness model 2**

The previous formulation contains eight parameters which make it difficult to gather some understanding of the way the model behaves, short of incorporating some rather special conditions. Therefore, we rearrange the model to incorporate assumptions that will yet allow a reasonable understanding of the system. This arrangement parallels the
simplification of the pairwise competitive model, discussed at the be-
ginning of this chapter. We have the following:

(1) In any triplet, two identical genotypes are mutually
antagonistic, more so if they are AA, less so if Aa, and least if aa.

(2) If, in any triplet, two members are dissimilar genotypes, then
AA competes most strongly against aa, less strongly against Aa, and
Aa competes more strongly against aa than against AA.

We may diagram the ten different triplets as follows;

Figure 3.3. Diagrams of all possible competing triplets showing
competitive interactions.

Table 3.4 tabulates the competitive effects between genotypes. In the
second diagram, for example, the probability that an AA genotype will
survive is \(2[1/3-gc+s(1-h)]\), and the probability that an Aa genotype
will survive is \([l/3+2(gc-s(1-h))]\).

We may list the pairwise competitive effects as follows (Table
3.4).
Table 3.4. Pairwise competitive effects on survival probabilities of members of competing triplet groups

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Genotype in competition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AA</td>
</tr>
<tr>
<td>AA</td>
<td>-gc</td>
</tr>
<tr>
<td>Aa</td>
<td>-s(1-h)</td>
</tr>
<tr>
<td>aa</td>
<td>-2s</td>
</tr>
</tbody>
</table>

We argue that if from the homogeneous AA group the probability that exactly one AA will survive is 1, i.e., the probability for a particular AA to survive is 1/3, then the competitive effect suffered by any one individual must be exactly countered by the mutual antagonism between the other two individuals. This leads to the following concept: In the absence of competition the probability that a particular AA will survive is 1/3. Competitive antagonism against the two neighboring AA individuals reduces this probability by 2(-gc) = -2gc. The mutual antagonism between the two neighboring individuals is gc, so this must be multiplied by an arbitrary constant, 2, and added to the probability of survival of the AA genotype.

To exemplify further, in the second group, AA gains in competition against Aa by s(1-h). (Compare this to the competing pairs model discussed earlier in this chapter.) AA, however, suffers in competition with the other AA by -gc. Aa in competition with AA loses in competition by s(1-h), but gains the gc suffered by the AA genotype in its competition.
against the other AA. Thus the probability that an AA survives the competing triplet is \(2[1/3-gc+s(1-h)]\), while the probability that an Aa will be the survivor is \(\{1/3+2(gc-s(1-h))\}\). The previous model, which used eight parameters, may now be expressed in terms of five, to wit, s, h, c, g and k. The table of survival probabilities is set out in Table 3.5.

Table 3.5. Survival probabilities of competing genotypes in groups of size three (constant fitness model)

<table>
<thead>
<tr>
<th>Competing group of occurrence</th>
<th>Frequency</th>
<th>AA</th>
<th>Aa</th>
<th>aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA, AA, AA</td>
<td>(p^6)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA, AA, Aa</td>
<td>(6p^5q)</td>
<td>(2[1/3-gc+s(1-h)])</td>
<td>(1/3+2gc-2s(1-h))</td>
<td></td>
</tr>
<tr>
<td>AA, AA, aa</td>
<td>(3p^4q^2)</td>
<td>(2[1/3-gc+2s])</td>
<td>(1/3+2gc-4s)</td>
<td></td>
</tr>
<tr>
<td>AA, Aa, Aa</td>
<td>(12p^3q^3)</td>
<td>(1/3+2kc+2s(1-h))</td>
<td>(2[1/3-kc-s(1-h)])</td>
<td></td>
</tr>
<tr>
<td>AA, Aa, aa</td>
<td>(12p^3q^3)</td>
<td>(1/3+2s+s(1-h))</td>
<td>(1/3+s(1+h)-s(1-h))</td>
<td>(1/3-2s-s(1+h))</td>
</tr>
<tr>
<td>AA, aa, aa</td>
<td>(3p^2q^4)</td>
<td>(1/3+2c+4s)</td>
<td>(2[1/3-c-2s])</td>
<td></td>
</tr>
<tr>
<td>Aa, Aa, Aa</td>
<td>(8p^3q^3)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aa, Aa, aa</td>
<td>(12p^2q^4)</td>
<td>(2[1/3-kc+s(1+h)])</td>
<td>(1/3+2kc-2s(1+h))</td>
<td></td>
</tr>
<tr>
<td>Aa, aa, aa</td>
<td>(6pq^5)</td>
<td>(1/3+2c+2s(1+h))</td>
<td>(2[1/3-c-s(1+h)])</td>
<td></td>
</tr>
<tr>
<td>aa, aa, aa</td>
<td>(q^6)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
From this table we may calculate the change in gene frequency for allele A.

\[ p' = p^6 + 6p^5q\{2[1/3-gc+s(1-h)] + 1/2[1/3+2gc-2s(1-h)]\} \]

\[ + 3p^4q^2\{2[1/3-gc+2s]\} \]

\[ + 12p^4q^2\{[1/3+2kc+2s(1-h)] + [1/3-kc-s(1-h)]\} \]

\[ + 12p^3q^3\{[1/3+s(3-h)] + 1/2[1/3+2sh]\} \]

\[ + 3p^2q^4\{1/3+2c+4s\} \]

\[ + 8p^3q^3(1/2) \]

\[ + 12p^2q^4\{1/3-kc+s(1+h)\} \]

\[ + 6pq^5\{1/2[1/3+2c+2s(1+h)]\} \]

\[ = p\{p^5+6p^4q(5/6-gc+s-sh) + 3p^3q^2\{2/3-2gc+4s\} \]

\[ + 12p^3q^2\{2/3+kc+s-sh\} + 12p^2q^3\{1/2+3s\} \]

\[ + 3pq^4\{1/3+2c+4s\} + 4p^2q^3 + 12pq^4\{1/3-kc+s(1+h)\} \]

\[ + 6q^5\{1/6+c+s+sh\} \]

\[ = p\{p^5+5p^4q+10p^3q^3+10p^2q^3+5pq^5+q^5 \]

\[ + c(6pq^4+6q^5) - gc(6p^4q+6p^3q^2) \]

\[ + s(6p^4q+12p^3q^3+12p^2q^3+36p^2q^3+12pq^4+12p^2q^4+6q^5) \]

\[ + sh(-6p^4q-12p^3q^2+12pq^4+6q^5) \]

\[ + ck(12p^3q^2-12pq^4) \]

\[ = p[1+6q(c(q^3-gp^3) + s + sh(q-p) + 2ckpq(p-q))]. \quad (3.2.5) \]
Hence,
\[ \Delta p = 6pq[c(g - gp^3) + s + sh(q-p) + 2ckpq(p-q)] \]  \hspace{1cm} (3.2.6)

An equilibrium value for p is given by the solution to the cubic,
\[ c(q^3 - gp^3) + s + sh(q-p) + 2ckpq(p-q) = 0, \]  \hspace{1cm} (3.2.7)

which in terms of p may be written as
\[ c(l+g+4k)p^3 - 3c(l+2k)p^2 + (3c+2sh+2ck)p - (c+s+sh) = 0. \]  \hspace{1cm} (3.2.8)

The following figures (Figures 3.4a, 3.4b) graph the change in gene frequency from one generation to the next over 70 generations. While the expression (3.2.5) may be used with a wide range of competition parameters, it can be seen that some values will, when substituted into Table 3.5, produce probabilities of survival either greater than unity, or negative. So, while the set of values, s = 0.5, h = 1.0, c = 0.6, k = 1.9, and g = 2.9, will produce a stable limit cycle using (3.2.5), these values in Table 3.5 give a negative probability of survival for genotype aa in the group AA, aa, aa. The computer simulations use only admissible values, those that do not produce probabilities outside the interval [0,1].

The first figure (Figure 3.4a) shows that for a recessive heterozygote (h = -1), as k increases from 1.1 to 1.9, the equilibrium value of p falls and the initial rate of change in p also falls. If k = 1, the mutual antagonism between the heterozygotes (Aa) equals that between the recessive homozygotes (aa), whereas, if k = g (and for most of our simulation studies, g = 2), the mutual antagonism between the heterozygotes equals that of the dominant homozygotes (AA). If k
Figure 3.4a. Graph of gene frequency from one generation to the next for the constant fitness model for groups of size 3 (Expression 3.2.5). In this figure $s = 0.01$, $c = .1$, $g = 2$, for all curves (++++: $h = -1$, $k = 1.1$; x-x-x: $h = -1$, $k = 1.9$; ---: $h = .2$, $k = 1.1$)
Figure 3.4b. Graph of gene frequency from one generation to the next for the constant fitness model for groups of size 3 (Expression 3.2.5). In this figure $s = 0.01$, $h = 0.9$, $k = 1.1$, $g = 2$, for both graphs (+++: $c = 0.1$; ---: $c = 0.3$).
is held constant \((k = 1.1)\) and \(h\) increases from \(-1\) to \(0.2\), the equilibrium value rises only slightly, as does the rate of change in \(p\). These equilibria were found, numerically, to be stable.

The second figure (Figure 3.4b) shows how the change in the mutual antagonism between like genotypes \((c)\) affects the system. For \(c = .1\), the equilibrium point is higher \((\bar{p} = 0.48)\) but the rate of change in \(p\) is initially less than the case where \(c = 0.3\) \((\bar{p} = 0.43)\). Both equilibria are shown, numerically, to be stable.

3.3. Constant Fitness Model for Groups of Size \(n\)

The basic setting of the constant fitness model for competing triplets can be exploited in the development of models of competition between individuals in groups of size four. In this case there are fifteen distinct groups, and the interactions between individuals in a group may be depicted by a series of tetrahedrons, as shown in Figure 3.5.

The gains and losses to the probability of survival for an individual in pairwise competition are the same as those set out in Table 3.4. Because the assumption of constant fitness ensures that the probability of there being exactly one survivor from each quadruplet is unity, we can argue that any disadvantage suffered by a member of a group must necessarily be accounted to other members of the group. So, in the first group represented in the figure four identical genotypes compete; each genotype suffers from mutual interference from the other three,
Figure 3.5. Three of the fifteen distinct groups of competing genotypes in quadruplets. The interference due to similar genotypes is $-2c$ between AA genotypes, and $-kc$ and $-c$ between Aa and aa genotypes respectively.
this reducing its probability of survival by $-3gc$. The total mutual antagonism between the three neighboring individuals $= 3(-gc) = -3gc$. In this case the arbitrary constant is one, and the genotype in question gains from the mutual antagonism between neighbors $3gc$. The net probability is unchanged, and each individual has a probability of survival $\frac{1}{4}$.

Calculations like those for groups of size three can be made for groups of sizes four and five. These reveal a pattern that suggests a more unified approach may be developed for groups of size $n$.

We shall consider that in a homogeneous group of size $n$ the effects of neighbors on an individual must be balanced by the mutual antagonism between those neighbors. The direct effect of AA in a group of $n$ AA's is

$$(n-1)(-gc).$$

The total number of pairwise interactions occurring between the $n-1$ neighbors is $\frac{1}{2}(n-1)(n-2)$. Each interaction is $gc$, so the total interaction occurring between the neighbors, $\frac{1}{2}(n-1)(n-2)gc$, must be multiplied by a constant, $K$, to balance the direct interaction on the individual being assessed. Hence,

$$gc(n-1) = K \frac{1}{2}(n-1)(n-2)gc$$

$$K = \frac{2}{n-2}.$$  

In a population undergoing random mating we shall assume that the gametic array is $pA + qa$, and the genotypic array is $p^2AA + 2pqAa +$
\( q^2 a = PAA + 2QAA + Raa. \)

The probability that a group of \( n \) genotypes contains \( rAA, tAa, \) and \( (n-r-t)aa \) genotypes is given by the multinomial probability function;

\[
P( AA^{(1)}, AA^{(2)}, \ldots, AA^{(r)}, Aa^{(1)}, \ldots, Aa^{(t)}, aa^{(1)}, \ldots, aa^{(n-t-r)}) = \frac{n!}{r!t!(n-r-t)!} \cdot p^r (2q)^t \cdot r^{n-t-r}.
\]

In such a group the probability that an \( AA \) survives is

\[
P( AA | rAA, tAa, (n-r-t)aa) = r\left\{ \frac{1}{n} + ts(1-h) + 2s(n-r-t) - (r-1)g\right\} \\
+ \frac{1}{n} \left\{ \frac{1}{2} (r-1)(r-2)g + \frac{1}{2} t(t-1)k + \frac{1}{2} (n-r-t)(n-r-t-1)c \right\}
\]

\[
= r\left\{ \frac{1}{n} + ts(1-h) + 2s(n-r-t) - (r-1)g + \frac{c}{n-2}\right\}
\]

Likewise

\[
P( Aa | rAA, tAa, (n-r-t)aa) = t\left\{ \frac{1}{n} - rs(1-h) + (n-r-t)s(1+h) - (t-1)k\right\},
\]

and

\[
P( aa | rAA, tAa, (n-r-t)aa) = (n-r-t)\left\{ \frac{1}{4} - 2rs - ts(1+h) - (n-r-t-1)c\right\} \\
+ \frac{c}{n-2}\left\{ r(r-1)g + t(t-1)k \right\} \\
+ (n-r-t-1)(n-r-t-2)\right\}.
\]

Therefore the frequency of \( AA \)'s that survive in the population is
\[ p' = \sum_{r=0}^{n} \sum_{t=0}^{n-r} \frac{n!}{r!t!(n-r-t)!} \left[ (2Q)^r t^r R^{n-r-t} \right]_{r=1} \left[ \sum_{s=1}^{t-s} (1-h) + 2s(n-r-t) \right] \]

\[ - (r-1)gc + \frac{c}{n-2}[(r-1)(r-2)g + t(t-1)k + (n-r-t)(n-r-t-1)] \}

\[ = \sum_{r=1}^{n} \sum_{t=0}^{n-r} \frac{(n-1)!}{(r-1)!t!(n-r-t)!} p^{r-1}(2Q)^{t} R^{n-r-t} \]

\[ + P(2Q)n(n-1)s(1-h) \sum_{r=1}^{n-1} \sum_{t=1}^{n-r-t} \frac{(n-2)!}{(r-1)!t!(n-r-t)!} p^{r-1}(2Q)^{t} R^{n-r-t} \]

\[ + PR(n-1)(2s) \sum_{r=1}^{n-1} \sum_{t=0}^{n-r-t} \frac{(n-2)!}{(r-1)!t!(n-r-t-1)!} p^{r-1}(2Q)^{t} R^{n-r-t-1} \]

\[ - P^2 n(n-1)gc \sum_{r=2}^{n} \sum_{t=0}^{n-r} \frac{(n-2)!}{(r-2)!t!(n-r-t)!} p^{r-2}(2Q)^{t} R^{n-r-t} \]

\[ + \frac{c}{n-2}(P^3 n(n-1)(n-2)g \sum_{r=3}^{n} \sum_{t=0}^{n-r-t} \frac{(n-3)!}{(r-3)!t!(n-r-t)!} p^{r-3}(2Q)^{t} R^{n-r-t} \]

\[ + P(2Q)^2 kn(n-1)(n-2) \sum_{r=1}^{n-2} \sum_{t=2}^{n-r-t} \frac{(n-3)!}{(r-1)!t!(n-r-t)!} p^{r-1}(2Q)^{t-2} R^{n-r-t} \]

\[ + PR^2 n(n-1)(n-2) \sum_{r=1}^{n-2} \sum_{t=0}^{n-r-t-2} \frac{(n-3)!}{(r-1)!t!(n-r-t-2)!} p^{r-1}(2Q)^{t} R^{n-r-t-2} \]

\[ = P + n(n-1)\{P(2Q)s(1-h) + PR2s - P^2 gc + P^3 gc + P(2Q)^2 kc + PR^2 c\}. \]

We may explain the limits of summation by example. In the third summation in Expression (3.3.1) we see that

\[ 1 \leq r, \quad 0 \leq t, \quad 1 \leq n-r-t, \]

which imply that

\[ n-r-t \leq n-1, \quad r \leq n-t-1, \quad t \leq n-r-1, \]
and therefore,

\[1 \leq r \leq n-t-1 \leq n-1, \quad 0 \leq t \leq n-r-1.\]

Likewise, the penultimate summation of the Expression (3.3.1) has that

\[1 \leq r, \quad 2 \leq t, \quad 0 \leq n-r-t,\]

so that,

\[2 \leq t \leq n-r, \quad 1 \leq r \leq n-2.\]

The frequency of the Aa's that survive in the population is given by

\[
2Q' = \sum_{r=0}^{n} \sum_{t=0}^{n-r} \frac{n!}{r!(n-r-t)!} p^{r} (2Q) t^{-1} n^{-r-t} \left\{ \begin{array}{c}
rs(1-h) + (n-r-t)s(1+h) \\
-(t-1)kc + \frac{c}{n-2} [r(r-1)g+(t-1)(t-2)k+(n-r-t)(n-r-t-1)]
\end{array} \right\}
\]

\[
= \sum_{r=0}^{n-1} \sum_{t=1}^{n-r} \frac{(n-1)!}{r!(n-r-t-1)!} \frac{n!}{r!(n-r-t)!} p^{r} (2Q) t^{-1} n^{-r-t}
\]

\[
- p(2Q)n(n-1)s(1-h) \sum_{r=1}^{n-1} \frac{(n-2)!}{r!(n-r-t-1)!} p^{r-1} (2Q) t^{-1} n^{-r-t}
\]

\[
+ (2Q)n(n-1)s(1+h) \sum_{r=0}^{n-2} \frac{(n-2)!}{r!(n-r-t-1)!} p^{r} (2Q) t^{-1} n^{-r-t-1}
\]

\[
- (2Q)^{2} n(n-1)kc \sum_{r=0}^{n-2} \frac{(n-2)!}{r!(n-r-t-1)!} p^{r} (2Q) t^{-1} n^{-r-t}
\]

\[
+ \frac{c}{n-2} \left( p^{2} (2Q)n(n-1) \sum_{r=0}^{n-1} \frac{(n-3)!}{r!(n-r-t-2)!} p^{r-2} (2Q) t^{-2} n^{-r-t} \right.
\]

\[
\left. + (2Q)^{3} n(n-1) \sum_{r=0}^{n-3} \frac{(n-3)!}{r!(n-r-t-3)!} p^{r} (2Q) t^{-3} n^{-r-t} \right)
\]

\[
+ (2Q)^{2} Rn(n-1) \sum_{r=0}^{n-3} \frac{(n-3)!}{r!(n-r-t-2)!} p^{2} (2Q) t^{-1} R^{-1} n^{-r-t-2}
\]

\[
= 2Q + n(n-1) \left\{ -p(2Q)s(1-h) + (2Q)Rs(1+h) - (2Q)^{2} kc + p^{2} (2Q) cg \\
+ (2Q)^{3} ck + (2Q)R^{2} c \right\}.
\]
Now, the allele frequency of A in the next generation is

\[ p' = p' + Q' \]

\[ = P + Q + n(n-1)\{s(l-h)PQ + PR2s + RQs(l+h) + gc(P^3-P^2+P^2Q) \]

\[ + kc[P(2Q)^2 - \frac{1}{2}(2Q)^2 + \frac{1}{2}(2Q)^3] + c(PR^2+QR^2) \} . \]

Because, under Hardy-Weinberg assumptions, \( P = p^2 \), \( Q = pq \), \( R = q^2 \),

\[ p' = p^2 + pq + n(n-1)\{s(l-h)p^3q + 2sp^2q^2 + s(l+h)pq^3 + gcP^4(p^2-l+pq) \]

\[ + kc4p^2q^2(p-\frac{1}{2}+pq) + cq^4(p^2+pq) \}

\[ = p + n(n-1)\{s(l-h)p^3q + 2sp^2q^2 + s(l+h)pq^3 + gcP^4(p-l) \]

\[ + kc4p^2q^2(\frac{1}{2}) + cq^4p \}

\[ = p + n(n-1)pq\{s(p^2+2pq+q^2) + sh(q^2-p^2) - gcP^3 + kc2pq + cq^3 \}

\[ = p + n(n-1)pq\{s + (q-p)sh + 2pq(p-q)kc + c(q^3-gp^3) \} . \quad (3.3.3) \]

So we may conclude that the change in gene frequency for a population composed of competing groups of size \( n \) is

\[ \Delta p = n(n-1)pq\{s + (q-p)sh + 2pq(p-q)kc + c(q^3-gp^3) \} , \quad (3.3.4) \]

so the equilibrium frequency is independent of the group size, and is given by the solution to the cubic given in (3.2.7) or (3.2.8).

The analysis of the stability of the equilibrium point is less tractable than for other models examined.

\[ p' = \bar{p} + \delta' = F(p) = F(\bar{p} + \delta p) \approx F(\bar{p}) + \delta pF'(p) \mid \bar{p} . \]
\[ F(p) = p + n(n-1)pq(s-(gp^3-q^3)c-(p-q)sh+2pq(p-q)ck) \]
\[ = p + n(n-1)p(1-p)[-c(1+g+4k)p^3+3c(1+2k)p^2-(3c+2sh+2ck)p \]
\[ + s+c+sh] \]
\[ F'(p)|_{p} = 1 - n(n-1)pq[3c(1+g+4k)p^2-6c(1+2k)p+(3c+2sh+2ck)] \]
\[ = 1 - n(n-1)q[3c(1+g+4k)p^2-9c(1+2k)p^2+3(3c+2sh+2ck)p \]
\[ - 3(s+c+sh) + 3c(1+2k)p^2-2(3c+2sh+2ck)p+3(s+c+sh)] \]
\[ = 1 - n(n-1)q[3c(1+2k)p^2 - 2(3c+2sh+2ck)p + 3(s+c+sh)] \]
\[ = 1 + n(n-1)[3c(1+2k)p^2 - (9c+4sh+10ck)p^2 \]
\[ + (9c+3s+7sh+4ck)p - 3(s+c+sh)] . \]

For local stability the cubic (a) must be negative, and (b) must be not less than \(-\frac{2}{n(n-1)}\).

Apart from noting that the value for \(g\) does not affect the stability of the system, we are not able to make much further comment on the relationship between the parameters \(c, k, s, h,\) and the group size \(n,\) and the system stability.

The graph bears out what one would expect from considerations of Expression (3.3.3). The size of the group, \(n,\) has no effect on the equilibrium point, \((p = 0.47)\) but the initial rate of change of gene frequency, \(\Delta p,\) is noticeably increased for the larger \(n.\)
Figure 3.6. Graph of the constant fitness model (3.3.3) showing the effects of different values of n (in the graph s = .01, h = .2, c = .1, k = 1.5, g = 2; +++: n = 3; ---: n = 5)
3.4. Haploid Genotypes in Competing Groups with Constant Fitness

Having developed the model to deal with groups of two, three, four or five interacting diploid genotypes, we are able to deal summarily with the cases for interacting haploids. As for the diploid cases, we assume that the interaction between the like A genotypes is \(-gc\), while that between like a genotypes is \(-c\). When A competes against a the former gains \(s\), and the latter loses a like amount to their respective probabilities of survival. This is shown more clearly in Table 3.6. For triplet groups, Table 3.7 may be drawn up.

Table 3.6. Pairwise competition effects between haploids

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Genotype in competition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A (-gc) (s)</td>
</tr>
<tr>
<td></td>
<td>a (-c)</td>
</tr>
</tbody>
</table>

Table 3.7. Frequencies of competition occurring between the same and different genotypes, and probabilities of survival of the two haploid genotypes competing in triplets

<table>
<thead>
<tr>
<th>Members of competing group</th>
<th>Frequency</th>
<th>Probability of survival</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(p^3)</td>
<td>(1)</td>
</tr>
<tr>
<td>A, A, A,</td>
<td>(3p^2q)</td>
<td>(2[1/3-gc+s]) ([1/3-2s+2gc])</td>
</tr>
<tr>
<td>A, A, a</td>
<td>(3pq^2)</td>
<td>([1/3+2s+2c]) (2[1/3-s-c])</td>
</tr>
<tr>
<td>A, a, a</td>
<td>(q^3)</td>
<td>(1)</td>
</tr>
<tr>
<td>a, a, a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
From this we may calculate the gene frequency for the next generation to be

\[
p' = p^3 + 3pq^2(\frac{1+2s+2c}{3}) + 3p^2q[\frac{1}{3} - gc + s] = p^3 + 2p^2q + pq^2 - gc(6p^2q) + s(6pq^2 + 6p^2q) + 6pq^2c = p(p+q)^2 + 6pq[s+c(q-gp)] = p[1+6q[s+c(q-gp)]]
\]  

(3.4.1)

We may calculate the equilibrium value of \( p \),

\[
\frac{s+c}{c(1+g)} = \frac{1+s/c}{1+g} \tag{3.4.2}
\]

We may draw a few conclusions from this result. There will be no equilibrium value for \( p \) in the interval \((0,1)\) unless \( s/c \) is less than \( g \). If the equilibrium does exist, then its value will increase with increasing \( s \), which is intuitively obvious, decrease with increasing \( c \), and decrease with increasing \( g \), which again is intuitively correct.

Consideration of the quadruplet case reveals a pattern not unlike that for the diploid case, and a similar unified approach is also possible. The pairwise interactions are given in Table 3.6.

As before, we have an arbitrary constant, depending on group size, scaled so that each individual in a homogeneous group has an equal probability of survival, \( 1/n \). In the homogeneous group of \( n \) \( A \) genotypes, the competitive interaction effects between individuals is \(-gc\). For each individual,
\( P(\text{individual survives}) = \{1/n-(n-1)gc+K'(n-1)(n-2)gc\}. \)

For the probability to be \(1/n\),

\[(n-1)gc = K_2'(n-1)(n-2)gc,\]

\[K = \frac{2}{n-2},\]

which is the same coefficient as for the diploid case.

We take the genotypic array, \(pA + qa\), and calculate the probabilities of having \(rA\) and \((n-r)a\) genotypes in a group of size \(n\).

\[P(rA, (n-r)a) = \binom{n}{r} p^r q^{n-r}.\]

For this group, the probability that \(A\) survives is given by

\[P(A|rA, (n-r)a) = r\left[\frac{1}{n}-(r-1)gc + (n-r)s + \frac{c}{n-2}\right],\]

while that for \(a\) is

\[P(a|rA, (n-r)a) = (n-r)\left[\frac{1}{n}rs-c(n-r-1) + \frac{c}{n-2}(r-1)(r-2)g\right] + (n-r-1)(n-r-2)].\]

Therefore the frequency of \(A\) in the next generation is

\[p' = \frac{n}{\sum_{r=0}^{n} \binom{n}{r} p^r q^{n-r}} \binom{n}{r} p^r q^{n-r}\left[\frac{1}{n}-(r-1)gc + (n-r)s + \frac{c}{n-2}\right].\]
\[ \Delta p = n(n-1)pq(s+c(q-gp)), \tag{3.4.4} \]

Therefore the change in gene frequency per generation is

and the equilibrium value, which is independent of group size, is

\[ \bar{p} = \frac{1+s/c}{1+g}, \] provided that the value lies in the interval \((0,1)\).

We may also formally investigate the stability of the equilibrium as follows

\[ p' = \bar{p} + \delta p = F(p) = F(\bar{p} + \delta p) = F(\bar{p}) + \delta p F'(p). \]

\[ F(p) = p\{1+n(n-1)(1-p)s+c(1-p(1+g))\}. \]

\[ F'(p) = \{1+n(n-1)(1-p)[s+c(1-p-gp)]\} \]

\[ + p\{-n(n-1)[s+c(1-p-gp)]\} \]

\[ + p\{n(n-1)(1-p)(-c)(1+g)\}. \]

At equilibrium, \( s+c(1-\bar{p}(1+g)) = 0 \), so that

\[ F'(p) \bigg|_{\bar{p}} = 1 - \bar{p}(1-\bar{p})cn(n-1)(1+g) \]

\[ = 1 - c \frac{(1+\frac{s}{c})(1-\frac{s}{c})}{(1+g)} n(n-1), \]

which must be less than unity in absolute value for stability of the equilibrium. The fraction is positive for nontrivial \( \bar{p} \), therefore

\[ 1 - \frac{(1+\frac{s}{c})(gc-s)}{(1+g)} n(n-1) < 1, \]

which implies that if
\[ n(n-1) < \frac{2(1+g)}{(c+s)(cg-s)}, \]

then the equilibrium will be stable.

The following graph (Figure 3.7) shows the change in gene frequency for the haploid model, (3.4.3), for various values of \( n, s, c, \) and \( g. \) While a wide range of parameters may be used with Expression (3.4.3), only those that give probabilities in the interval \([0,1]\) in Table 3.7 are used.

Two comparisons may be made in the first graph (Figure 3.7a). The first is that between the system for \( s = 0.01 \) and \( s = 0.1. \) In the former case both the initial rate of change of gene frequency and the equilibrium value is less \((\bar{p} = 0.37)\) than the latter \((\bar{p} = 0.67). \) In both instances \( n = 3, c = .1, \) and \( g = 2. \) The other comparison shows that increasing \( n, \) while initially increasing the rate of change of gene frequency, does not alter the equilibrium value \((\bar{p} = 0.37). \) The marked increase in \( \Delta p \) is sufficient to induce a slight oscillation (heavily damped) about the equilibrium value.

If \( c \) is increased from \( 0.1 \) to \( 0.2, \) the initial rate of change of gene frequency is elevated, but the eventual equilibrium value is lowered. Figure 3.7b depicts this quite clearly. The equilibrium value is shown to fall from \( \bar{p} = 0.37 \) to \( \bar{p} = 0.35. \) If \( g \) is lowered from \( 2 \) to \( 1.5 \) the difference between the mutual antagonism between \( A \) genotypes and that between \( a \) genotypes is lowered. The effect is seen as a moderate increase in the equilibrium gene frequency \((\bar{p} = 0.42, \) rather than \( 0.35).\)

Numerical studies show that all equilibria displayed are stable.
Figure 3.7a. Graph of gene frequency from one generation to the next for the haploid constant fitness model (3.4.3) showing the effect of different values of $n$ and $s$ (in this figure $c = 0.1$, $g = 2$, for all curves. ---: $n = 3$, $s = 0.01$; +++: $n = 3$, $s = 0.1$; x-x-x: $n = 5$, $s = 0.01$)
Figure 3.7b. Graph of gene frequency from one generation to the next for the haploid constant fitness model (3.4.3) showing the effect of different values of c and g (in this figure, n = 3, s = 0.01 for all curves. ---: c = 0.1, g = 2; +++: c = .2, g = 2; x-x-x: c = .2, g = 1.5)
While it is possible to use values of s, c and g in Expression (3.4.3) that show stable limit cycles, care must be taken to ensure that these values do not give negative probabilities of survival for a genotype. Within this rather tight restriction we are able to show a mildly interesting result. If, in a natural situation, owing to changes in the environment, either temporally or spatially, the group size fluctuates, then the system will change from one with a stable equilibrium, to one with a stable limit cycle, with a consequent change in the trend of change in gene frequency. The following graph shows the effect of n changing from 3 to 47, for values of s, c and g equal to 0.03, 0.07 and 1.8 respectively.

For the first ten generations n=3. The graph shows the gene frequency approaching the stable equilibrium, \( \bar{p} = 0.51 \). The value for \( F'(p) \bigg|_p \) is 0.71. For generations 10 to 19 n = 42. \( F'(p) \bigg|_p \) is -1.06, and a stable limit cycle develops. For generations 20 to 29 n is again set equal to 3, and then equal to 42 for generations 30 to 59, when it is switched back to 3 again. The consequent fluctuations in the stability of the system can be readily followed.

It has not been possible to find values of competition parameters that give rise to stable limit cycles with the other models already discussed, or those to be discussed in Section 3.5.
Figure 3.7c. Graph of gene frequency from one generation to the next for the haploid constant fitness model (3.4.3) showing the effect of different values of n (in this figure s = 0.03, c = 0.07, g = 1.8. n = 3 for generations 0 to 9, 20 to 29, and 60 to 70. Elsewhere, n = 42)
3.5. Variable Fitness Competition Models

If we relax the condition that each competing group has exactly one survivor with probability 1, we encounter frequency dependent selection models. Generally, such models have ignored specifically the concept of competition occurring between genotypes in a group, and have, rather, developed along the lines suggested by Wright (1955). The studies of Harding et al. (1966) have been discussed in Chapter 2. Briefly to recount, they found the reproductive capacity of heterozygotes, in a selfing population, to increase four-fold as the frequency of the genotypes fell from 0.40 to 0.02. Although an exponential function was used to describe the relationship between genotypic frequency and reproductive capacity, we find that a reciprocal model, of the form $y = a+bx$, fits the data of Harding et al. (1966) reasonably well ($r^2 = 0.76$). We shall use this in the following model, see Table 3.8. $P$, $2Q$, and $R$ refer to the genotypic frequencies of genotypes AA, Aa, and aa, respectively, before selection.

Table 3.8. Frequency dependent reproductive capacity model for autogenous (self-fertilizing) diploid

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Frequency</th>
<th>Reproductive capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>$P$</td>
<td>1</td>
</tr>
<tr>
<td>Aa</td>
<td>$2Q$</td>
<td>$1 + r + \frac{s}{2Q}$</td>
</tr>
<tr>
<td>aa</td>
<td>$R$</td>
<td>1</td>
</tr>
</tbody>
</table>
From this we may calculate the change in genotype frequency among young individuals, prior to selection, to be

\[ P' = \frac{\{P + \frac{1}{2}Q(1 + r + \frac{s}{2Q})\}}{\bar{w}} \]

\[ 2Q' = \frac{Q(1 + r + \frac{s}{2Q})}{\bar{w}} \]  \hspace{1cm} (3.5.1)

\[ R' = \frac{\{R + \frac{1}{2}Q(1 + r + \frac{s}{2Q})\}}{\bar{w}} \]

\[ \bar{w} = 1 + 2Q(r + \frac{s}{2Q}) \]

\[ = 1 + (2Qr + s). \]

Equilibrium occurs when \( Q' = Q \), i.e.,

\[ 2Q = 2Q' = \frac{Q(1 + r + \frac{s}{2Q})}{(1 + 2Qr + s)}, \]

\[ 2Q + (2Q)^2 r + 2Qs = Q + Qr + \frac{1}{2}s \]

and hence,

\[ 2Q = \frac{-s\frac{1}{2}(1-r) + \sqrt{(\frac{1}{2}r)^2 + s - \frac{1}{2}r + \frac{1}{4}}}{2r} \]  \hspace{1cm} (3.5.2)

As mentioned in Chapter 2, the relationship

reproductive value = 0.8468 + 0.0652/f,

where \( f \) is the frequency of the heterozygotes, fitted the data of Harding et al. (1966) fairly well. This leads to the following estimates for values of \( r \) and \( s \).

\[ r = -0.1532 \]

\[ s = 0.0652 \]

The equilibrium value for the heterozygote frequency is calculated to be
$2q = 0.051$,

and numerical studies show this equilibrium to be stable. The following table (Table 3.9) shows the frequency obtained from the model compared with values taken from the data of Harding et al. as depicted by Allard and Adams (1969).

Table 3.9. Observed heterozygote frequencies (Harding et al. from Allard and Adams (1966), and calculated heterozygote frequencies according to the model described by (3.5.1), for successive generations

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.40</td>
<td>.23</td>
<td>.19</td>
<td>.16</td>
<td>.14</td>
<td>.10</td>
<td>.11</td>
<td>.16</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td>.25</td>
<td>.13</td>
<td>.09</td>
<td>.07</td>
<td>.06</td>
<td>.05</td>
<td>.05</td>
<td>.05</td>
</tr>
</tbody>
</table>

3.6. Variable Fitness Models with Competition Occurring in Groups

In our study, however, we are mainly concerned with identifying competing groups. Although we still arrive at a model that exhibits the characteristics of a frequency dependent model, the approach differs.

3.6.1. Twin groups

In this model, selection, owing to competition, operates on the groups after zygote formation. We shall outline the model for single locus two allele haploids competing in groups of two (Table 3.10).

Here the heterogeneous group has a total fitness of unity, while the other homogeneous groups have fitnesses less than unity (Table 3.10).
Table 3.10. Frequencies of groups and probabilities of survival for genotypes (haploid model)

<table>
<thead>
<tr>
<th>Competing group</th>
<th>Frequency</th>
<th>Conditional Probability of Survival A</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, A</td>
<td>$p^2$</td>
<td>1-gf</td>
<td></td>
</tr>
<tr>
<td>A, a</td>
<td>2pq</td>
<td>$\frac{1}{2} + s$</td>
<td>$\frac{1}{2} - s$</td>
</tr>
<tr>
<td>a, a</td>
<td>$q^2$</td>
<td></td>
<td>1-f</td>
</tr>
</tbody>
</table>

We may take the model to imply that when A competes against A in a group, there is a probability of $f$ that neither will survive, and similarly for a against a, when the probability of no survivor is $h$.

Alternatively, we can consider these tabulated values as relative viabilities, and so, because it allows $f$ and $h$ to have positive values, this shall be the approach we will continue to use.

We can see that the population mean fitness is

$$
\bar{w} = 1 - p^2 g f - q^2 f.
$$

The gene frequency of A after one cycle of selection is

$$
p' = \frac{[p^2 (1-gf) + 2pq(\frac{1}{2}+s)]/\bar{w}}{p (1-pgf+2qs)/\bar{w}} = \frac{p (1-pgf+2qs)/\bar{w}}{w} \tag{3.6.1}
$$

Thus

$$
\Delta p = \frac{p}{w} (2qs-pgf+p^2 gf+q^2 f) = \frac{pq}{w} (2s-pgf+qf) \tag{3.6.2}
$$

so that the equilibrium value for $p$ is that for which

$$
2s - pgf + (1-p) = 0
$$
i.e.,

\[ p = (f+2s)/(f+gf). \]  \hspace{1cm} \text{(3.6.3)}

Clearly, \( 2s < gf \) for \( \bar{p} \) to be a nontrivial solution. To examine the stability of this equilibrium we may formally write

\[ p' = \bar{p} + \delta' p = F(p) = F(\bar{p} + \delta p) = F(\bar{p}) + \delta p F'(p) \]

\[ \delta' p = \delta p F'(p) \]

\[ F(p) = p(1-pgf+2qs)/w \]

\[ = p(1-pgf+2qs)/(1-p^2gf-q^2f) \]

\[ F'(p) = (1-pgf+2qs)/w + p\frac{-3}{w}\frac{\partial}{\partial p}(1-pgf+2qs) - (1-pgf+2qs)\frac{\partial w}{\partial p}/w^2 \]

At equilibrium, \( (1-pgf+2qs) = \bar{w}, \) so

\[ F'(p) \bigg|_\bar{p} = 1 + \frac{\bar{p}}{w}[-(gf+2s)+2p(gf+f)-f] \]

\[ = 1 + \frac{(f+2s)(2s-gf)}{w(f+gf)} \]

Now for \( \left| F'(p) \right|_\bar{p} < 1, \) \( 2s-gf < 0 \) is a necessary condition, i.e., \( 2s < gf. \)

At equilibrium \( \bar{w} = 1+2s-p(gf+2s), \) therefore

\[ F'(p) \bigg|_\bar{p} = 1 - \frac{(f+2s)(gf-2s)}{(f+gf)(1+2s-[gf+2s][f+2s]/[f+gf])} \]

\[ = 1 - \frac{gf^2+2s(gf-f)-4s^2}{f+gf-gf^2-4s^2}. \]
This implies that
\[
0 < \frac{gf^2 + 2s(gf-f) - 4s^2}{f + gf - gf - 4s^2} < 2
\]
for the equilibrium to be stable, i.e.,
\[
4s^2 + 2s(gf-f) + 3gf^2 - 2(f+gf) < 0.
\]
Now, since \(2s < gf\), then
\[
4s^2 + 2s(gf-f) + 3gf^2 - 2(f+gf) < g^2f^2 + f(gf-f) + 3gf^2 - 2(f+gf)
\]
\[
= 2[g^2f^2 + gf^2 - (f+gf)]
\]
\[
= 2[(gf-1)(f+gf)]
\]
\[
< 0 \text{ since } f < 1, \text{ and } g > 0.
\]
So any nontrivial solution is stable.

The formulae for \(\Delta p\), (3.6.2), and \(\bar{p}\) (3.6.3), allow us to make a few generalizations. First, since \(\overline{w}\) changes with the square of the frequency, \(\overline{w}\) will remain fairly constant until \(p\) or \(q\) become quite close to 1, unless \(f\) and \(h\) have rather high values. Secondly, if \(f\) and \(h = 0\), we have Nei's (1971) result that \(\Delta p = 2pq\). Like Nei's result, generally the rate in change in \(p\) is linearly related to \(s\).

3.6.2. **Extension to larger groups**

The question now centers on how to extend this model to higher numbered groups. Essentially, the model must be not unrealistic on the one hand, but also capable of generalization to larger sized groups on the other.
3.6.3. Haploids competing in triplets

In a triplet group we may consider that a series of pairwise competitive interactions occur. Thus, for a triplet consisting of A, A, a, A compete with a frequency 1/3, while A and a compete with a frequency 2/3. The result can be determined from consideration of Table 3.11.

Table 3.11. Frequencies and relative viabilities of pairwise competition between members of triplet A, A, a

<table>
<thead>
<tr>
<th>Competition between</th>
<th>Frequency of occurrence</th>
<th>Conditional relative viabilities</th>
<th>A</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, A</td>
<td>1/3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A, a</td>
<td>2/3</td>
<td>1/2+s</td>
<td>1/2-s</td>
<td></td>
</tr>
</tbody>
</table>

Total probability of survival 2/3(1+s) 1/3(1-2s)

The other heterogeneous group, A, a, a, is handled in a similar manner. We now impose the condition that the relative viability is affected adversely by the mutual antagonism between like genotypes. We assume that where A competes against A, the conditional relative viability is scaled by (1-gf), as in the pairwise case, and where three A genotypes compete, the relative viability is scaled by (1-gf)^2. Using a similar argument for the a genotype groups, we are able to derive Table 3.12.
Table 3.12. Frequencies of groups and relative viabilities for haploids competing in triplets

<table>
<thead>
<tr>
<th>Competing group</th>
<th>Frequency</th>
<th>Conditional relative viabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, A, A</td>
<td>$p^3$</td>
<td>$(1-gf)^2$</td>
</tr>
<tr>
<td>A, A, a</td>
<td>$3p^2q$</td>
<td>$\frac{2}{3}(1+s)(1-gf)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\frac{1}{3}(1-2s)$</td>
</tr>
<tr>
<td>A, a, a</td>
<td>$3pq^2$</td>
<td>$\frac{1}{3}(1+2s)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\frac{2}{3}(1-s)(1-f)$</td>
</tr>
<tr>
<td>a, a, a</td>
<td>$q^3$</td>
<td>$(1-f)^2$</td>
</tr>
</tbody>
</table>

In this model we calculate the mean fitness to be

$$w = 1 - 2p^2gf + p^3g^2f^2 - 2q^2f + q^3f^2 + 2pqsf(q-pg)$$

$$= p(l-pgf)^2 + q(l-qf)^2 + 2pqsf(q-pg).$$

The change in gene frequency after one cycle of selection may be calculated from

$$\bar{wp'} = p^3(1-gf)^2 + 2p^2q(1-gf) + pq^2 + 2pq^2s(l-gf) + 2pq^2s$$

$$= p(l-pgf)(1-pgf+2qs).$$

This leads to the expression

$$\Delta p = \frac{p}{\bar{w}}[(l-pgf)(l-pgf+2qs) - p(l-pgf)^2 - q(l-qf)^2 - 2pqsf(q-pg)]$$

$$= pq[(l-pgf)(l-pgf+2s) - (l-qf)^2 - 2psf(pg-q)]/\bar{w}.$$ 

To compare this to the model for the competing pairs case, we write it as

$$\Delta p = pq(2s-2pgf+2qf+...)/(1-2p^2gf-2q^2f+...),$$

where the "..." implies terms that are products or squares involving
s, h and f, and for this comparison may be ignored if they are relatively small. For the pairwise case,

\[ \Delta p = pq(2s-pgf+qf)/(1-p^2gf-q^2f) \]

which implies that if s is somewhat larger than f and gf, increasing the number of individuals competing in a group will increase the rate of gene change per generation. We may also note from this approximation that the population mean fitness will decrease as the number of competing members in a group increases. This, broadly, is the same as that found with the constant fitness competition models. The equilibrium value, however, is not the same for the triplet case as for the pairwise case. Taking the expression for \( \Delta p \) and equating it to zero we arrive at the following solution for a quadratic.

\[
\frac{-f(g+1)(1+s)-f^2 - \sqrt{f(g+1)[f(g+1)(1+s^2)-4s^2-2gf^2]+g^2f^4}}{f(g+1)(2s+gf-f)}
\]

While it is possible to produce an expression for the conditions for stability, its complexity precludes establishing straightforward expressions for stability conditions.

3.6.4. Haploids competing in groups of n

In the general case, for groups of size n, we let the frequency of A genotype be \( p \), and that for a be \( q (=1-p) \). The probability that a group will contain \( r \) A and \( (n-r) \) a genotypes is

\[ P[rA, (n-r)a] = \binom{n}{r} p^r q^{n-r}. \]
If there are \( r \) \( \text{A} \) genotypes then the number of ways to form pairs between the \( \text{A} \)'s is 

\[
\frac{1}{2}r(r-1).
\]

The number of ways to form pairs between the \((n-r)\) \( \text{a} \) genotypes is 

\[
\frac{1}{2}(n-r)(n-r-1).
\]

The number of ways to form pairs between the \( r \) \( \text{A} \) and the \((n-r)\) \( \text{a} \) genotypes is 

\[
r(n-r).
\]

The total number of ways of forming pairs between the \( n \) members of the group is 

\[
\frac{1}{2}n(n-1).
\]

Hence, we may construct the following table (Table 3.13) showing the relative viabilities of genotypes competing in the various pairs.

Table 3.13. Frequency of competing pair groups and relative viabilities of members

<table>
<thead>
<tr>
<th>Competition</th>
<th>Frequency of occurrence</th>
<th>Relative viabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{A, A} )</td>
<td>( \frac{r(r-1)}{n(n-1)} )</td>
<td>( \frac{1}{2} + s )</td>
</tr>
<tr>
<td>( \text{A, a} )</td>
<td>( \frac{2r(n-r)}{n(n-1)} )</td>
<td>( \frac{1}{2} + s )</td>
</tr>
<tr>
<td>( \text{a, a} )</td>
<td>( \frac{(n-r)(n-r-1)}{n(n-1)} )</td>
<td>( \frac{1}{2} + s )</td>
</tr>
</tbody>
</table>

1

1

1
Where there are $x$ A genotypes competing in a group we assume that the relative viability of the A genotype is reduced by $(1-gf)^{x-1}$ owing to the mutual antagonism between like genotypes.

Similarly, the mutual antagonism between the $y$ a genotypes reduces their relative viability by $(1-f)^{y-1}$. So we can write that

$$P(A \text{ survives} \mid r A'a, (n-r) a's) = \left[ \frac{r(r-1)}{n(n-1)} + \frac{2r(n-r)}{n(n-1)} \left( \frac{1}{2} + s \right) \right] (1-gf)^{r-1}$$

$$= \left[ \frac{r(r-1) + 2r(n-r)}{n(n-1)} s \right] (1-gf)^{r-1}$$

$$P(a \text{ survives} \mid rA's, (n-r) a's) = \left[ \frac{(n-r)(n-r-1)}{n(n-1)} + \frac{2r(n-r)}{n(n-1)} \left( \frac{1}{2} - s \right) \right] (1-f)^{n-r-1}$$

$$= \left[ \frac{(n-r)(n-r-1+r)}{n(n-1)} - \frac{2r(n-r)}{n(n-1)} s \right] (1-f)^{n-r-1}$$

Therefore, the frequency of A in the next generation is

$$p' = \left\{ \sum_{r=0}^{n} \frac{n!}{r!(n-r)!} p^r q^{n-r} \left[ \frac{r}{n} + \frac{2r(n-r)}{n(n-1)} s \right] (1-gf)^{r-1} \right\} /$$

$$\left\{ \sum_{r=0}^{n} \frac{n!}{r!(n-r)!} p^r q^{n-r} \left[ \frac{r}{n} + \frac{2r(n-r)}{n(n-1)} s \right] (1-gf)^{r-1} \right\}$$

$$+ \left[ \frac{(n-r)}{n} - \frac{2r(n-r)}{n(n-1)} s \right] (1-f)^{n-r-1}$$

If we neglect terms in $s$, $h$, and $gf$ involving higher powers than the first we have
\[ p' = \frac{n}{\sum_{r=0}^{n} \binom{n}{r} p^r q^{n-r} \left( \frac{r}{n} + \frac{2r(n-r)}{n(n-1)} s - \frac{r(r-1)}{n} \right) \phi f} \]

\[ = \frac{n}{\sum_{r=0}^{n} \binom{n}{r} p^r q^{n-r} \left( \frac{r}{n} + \frac{2r(n-r)}{n(n-1)} s - \frac{2r(n-r)}{n(n-1)} s \right) \phi f} - \frac{r(r-1)}{n} \phi f - \frac{(n-r)(n-r-1)}{n} \phi f} \].

Now

\[ \sum_{r=0}^{n} \binom{n}{r} p^r q^{n-r} \frac{r}{n} = p \sum_{r=1}^{n-1} \binom{n-1}{r} p^{r-1} q^{n-r} = p, \]

\[ \sum_{r=0}^{n} \binom{n}{r} p^r q^{n-r} \frac{2r(n-r)}{n(n-1)} s = 2spq \sum_{r=1}^{n-2} \binom{n-2}{r} p^{r-1} q^{n-r-1} = 2spq, \]

\[ \sum_{r=0}^{n} \binom{n}{r} p^r q^{n-r} \frac{r(r-1)}{n} \phi f = (n-1)p^2 \phi f \sum_{r=2}^{n-2} \binom{n-2}{r} p^{r-2} q^{n-r} = (n-1)p^2 \phi f, \]

\[ \sum_{r=0}^{n} \binom{n}{r} p^r q^{n-r} \frac{(n-r)(n-r-1)}{n} \phi f = (n-1)q^2 \phi f \sum_{r=0}^{n-2} \binom{n-2}{r} p^r q^{n-r-2} = (n-1)q^2 \phi f. \]

Therefore,

\[ p' = \frac{p+2spq-(n-1)p^2 \phi f}{1-(n-1)p^2 \phi f-(n-1)q^2 \phi f} \]

\[ = p \left\{ \frac{1+2sq-(n-1)pq \phi f}{1-(n-1)p^2 \phi f-(n-1)q^2 \phi f} \right\}. \quad (3.6.4) \]

Thus the change in gene frequency for one generation is approximately given by

\[ \Delta p = \frac{p+2spq-(n-1)p^2 \phi f-p+(n-1)p^3 \phi f+(n-1)pq^2 \phi f}{1-(n-1)p^2 \phi f-(n-1)q^2 \phi f} \]

\[ = \frac{2spq-(n-1)p^2 (1-p) \phi f+(n-1)pq^2 \phi f}{1-(n-1)p^2 \phi f-(n-1)q^2 \phi f} = \frac{pq(2s-(n-1)pq+(n-1)q) \phi f}{1-(n-1)p^2 \phi f-(n-1)q^2 \phi f}. \quad (3.6.5) \]
The equilibrium value for \( p \) is approximated by

\[
2s-(n-1)p_gf + (n-1)q_f = 0,
\]

\[
\hat{p} = \frac{(n-1)f+2s}{f(n-1)(1+g)} = \frac{1}{1+g} + \frac{2s}{f(n-1)(1+g)}.
\]

We have shown that where the number of surviving individuals from a competing group differs from unity according to the competitive effects of the neighboring individuals, we may get results quite different from those either implicated by Nei (1971), or obtained from the constant fitness model. If \( f=0 \), then we have exactly one survivor for each competing pair, and we get the same result as Nei (1971).

Significantly, the equilibrium value for gene frequency varies as the group size increases - a feature not possessed by the other models discussed.

We may show that the fitness, \( \bar{w} \), at equilibrium is not necessarily the maximum fitness. If we write \( \bar{w} \) as

\[
\bar{w} = 1-m(p^2 g_f+q^2 f),
\]

where \( m+1 \) is the number of individuals competing in a group, then

\[
\frac{dw}{dp} = m[-2pf(1+g) + 2f],
\]

and if we equate this to zero, we find the maximum or minimum \( \bar{w} \) at

\[
\bar{p} = 1/(1+g).
\]

The second derivative is negative at \( \bar{p} \) so we may take the value of \( \bar{w} \)
at \( p \) to be a maximum. We may conclude that, within the limits imposed by the approximations we are making, the maximum \( w \) will not necessarily occur at the equilibrium point for small competing group sizes. However, if the group size becomes large, the equilibrium point will approach that point giving the maximum \( w \). The following graphs (Figure 3.8) show the change in gene frequency for different values of \( s, f, g \) and \( n \).

We can see from this graph what may be inferred from Expression (3.6.4) and (3.6.5). The change in \( n \) from 3 to 5 changes the equilibrium point only slightly, but while \( p \) is small (less than 0.4), we can see that the change in gene frequency is more rapid if \( n \) is larger. If the mutual antagonism between similar genotypes is small \( (f = 0.05, gf = 0.01) \) then with \( s \) large \( (s = 0.1) \) the gene frequency will go to fixation.

3.6.5. Diploid model

Using an approach similar to that used for the fixed fitness models, we shall consider first only the case of pairwise interactions. The Table 3.14 outlines the relative viabilities of genotypes involved in various pairwise interactions. We assume that the frequency of the alleles \( A \) and \( a \) are \( p \) and \( q \) \((= 1-p)\) respectively. The frequencies of the genotypes \( AA, Aa \) and \( aa \) are \( P, 2Q \) and \( R \) respectively. Under Hardy-Weinberg assumptions, \( P = p^2, 2Q = 2pq \) and \( R = q^2 \).

We further assume that, owing to mutual antagonism between like genotypes, the relative viabilities of homogeneous groups are reduced.
Figure 3.8. Graph of gene frequency from one generation to the next for the variable fitness haploid model (Expression 3.6.4) (---: n = 3, gf = .1, f = .15, s = .01; +++: n = 5, gf = .1, f = .15, s = .01; x-x-x: n = 3, gf = .01, f = .05, s = .1)
Table 3.14. Frequency of pairwise competition and relative viabilities of genotypes

<table>
<thead>
<tr>
<th>Competition between</th>
<th>Frequency of occurrence</th>
<th>Relative viabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA, AA</td>
<td>$p^4$</td>
<td>1</td>
</tr>
<tr>
<td>AA, Aa</td>
<td>$4p^3q$</td>
<td>$\frac{1}{2} + s(1-h)$</td>
</tr>
<tr>
<td>AA, aa</td>
<td>$2p^2q^2$</td>
<td>$\frac{1}{2} + 2s$</td>
</tr>
<tr>
<td>Aa, Aa</td>
<td>$4p^2q^2$</td>
<td>1</td>
</tr>
<tr>
<td>Aa, aa</td>
<td>$4pq^3$</td>
<td>$\frac{1}{2} + s(1+h)$</td>
</tr>
<tr>
<td>aa, aa</td>
<td>$q^4$</td>
<td>1</td>
</tr>
</tbody>
</table>

Thus, the relative viability of the group, AA, AA, becomes $(1-gf)$, that for the group, Aa, Aa, becomes $(1-kf)$, and that for the group aa, aa, becomes $(1-f)$.

The probability that genotype AA survives is proportional to its net relative viability, i.e.,

$$P(\text{AA survives}) \propto (1-gf)p^4 + \left[\frac{1}{2} + s(1-h)\right]4p^3q + \left[\frac{1}{2} + 2s\right]2p^2q^2$$

$$= p^2(1-gfp^2 + 4qs - 4pqsh).$$

Similarly,

$$P(\text{Aa survives}) \propto \left[\frac{1}{2} - s(1-h)\right]4p^3q + \left(1-kf\right)4p^2q^2 + \left[\frac{1}{2} + s(1+h)\right]4pq^3$$

$$= 2pq\left[1 - 2s(p-q) + 2sh(p^2 + q^2) - 2pqkf\right].$$

$$P(\text{aa survives}) \propto 2p^2q^2\left(\frac{1}{2} - 2s\right) + 4pq^3\left[\frac{1}{2} - s(1+h)\right] + q^4(1-f)$$

$$= q^2(1 - 4ps - 4pqsh - q^2f).$$
Therefore, the total fitness is proportional to
\[
1 - p^4 g f + 4p^2 q s - 4p^3 q s h - 4pq s (p-q) + 4pq s h (p^2 + q^2) - 4p^2 q^2 k f
\]

\[
- 4pq^2 s - 4pq^3 s h - q^4 f
\]

\[
= 1 - (p^4 g + 4p^2 q^2 k + q^4) f
\]

= denom., say.

The frequency of the allele, A, in the next generation is, therefore,

\[
p' = P(AA among adults) + \frac{1}{2} P(Aa among adults), i.e.,
\]

\[
p' = \frac{[p^2 g f p^4 + 4p^2 q s - 4p^3 q s h + pq - 2pq s (p-q) + 2pq s h (p^2 + q^2) - 2p^2 q^2 k f]}{1 - (p^4 g + 4p^2 q^2 k + q^4) f}
\]

\[\text{(3.6.6)}\]

We may calculate the change in gene frequency for each generation to be

\[
\Delta p = \frac{p[1-pf(gp^2 + 2q^2 k) + 2qs - 2qsh (p-q)]}{1 - (p^4 g + 4p^2 q^2 k + q^4) f}
\]

\[\text{(3.6.7)}\]

The equilibrium point is determined from the equation, \(\Delta p = 0\), i.e.,

\[
2s[1-h(p-q)] - p^3 g f + 2pgkf(p-q) + q^3 f = 0,
\]

\[
2s - 2sh(2p-1) - p^3 g f + 2p(1-p)kf(2p-1) + (1-p)^3 f = 0,
\]
\[ p^3 f(g+4k+1) - 3p^2 f(2k+1) + p(4sh+2kf+3f) - 2s(1+h) - f = 0. \]

Under genic selection, we may take \( h=0 \), and \( k = \frac{1}{2}(g+1) \), thus the change in gene frequency may be calculated as

\[
pq\left[-p^3 f(g+2g+2+1) + 3p^2 f(g+1+1) - p(2s+fg+f+3f) + 2s(3/2) + f\right] \\
\frac{1 - f(p^4 g+2p^2 q^2 +2p^2 q^2 g+q^4)}
\]

\[
= pq\frac{[3s-2ps-f(p^3 3g+3) - p^3 3g+2p^3 g+4]-1]}{1-f[p^2 g(p^2 +2q^2) + q^2 (q^2 +2p^2)]} \\
= pq\frac{[s(1+2g)-f(gp(3p^2 -3p+1)-(1-p)(3p^2 -3p+1)]}{1-f[p^2 g(p^2 +2q^2) + q^2 (q^2 +2p^2)]} \\
= \frac{pq[s(1+2g)-f(1-3pg)(gp-q)]}{1-f[p^2 g(p^2 +2q^2) + q^2 (q^2 +2p^2)]},
\]

which may be compared to that expression for the change in gene frequency for the haploid model (Expression 3.6.5). It will be noticed that the haploid model does not necessarily give the same result as the diploid model under the assumption of genic selection.

### 3.6.6. Diploid model for groups of size \( n \)

In a group of \( n \) members the probability that there are \( r \) AA, \( t \) Aa, and \( (n-r-t) \) aa genotypes is given by

\[
P[rAA, tAa, (n-r-t)aa] = \frac{n!}{r!t!(n-r-t)!} p^r (2q)^t R^{n-r-t},
\]

where \( P, Q \) and \( R \) are the frequencies of the genotypes AA, Aa and aa respectively. In this group the total number of pairwise arrangements is \( \frac{1}{2} n(n-1) \). Of these, the number between respective genotypes are as follows:
AA and AA; \( \frac{1}{2}r(r-1) \),

AA and Aa; \( rt \),

AA and aa; \( r(n-r-t) \),

Aa and Aa; \( \frac{1}{2}t(t-1) \),

Aa and aa; \( t(n-r-t) \), and

aa and aa; \( \frac{1}{2}(n-r-t)(n-r-t-1) \).

We assume that, among the homogeneous competing pairs, mutual
antagonism reduces their respective relative viabilities: between AA
genotypes by \((1-gf)^{r-1}\), between Aa genotypes by \((1-kf)^{t-1}\), and between
aa genotypes by \((1-f)^{n-r-t-1}\).

The probability that AA survives is proportional to the net
relative viability, i.e.,

\[
P[AA | rAA, tAa, (n-r-t)aa] \propto \left\{ \frac{r(r-1)}{n(n-1)} + \frac{2rt}{n(n-1)} \right\}(1-gf)^{r-1}
\]

\[+ \frac{2r(n-r-t)(r-1)}{n(n-1)} - \frac{2s(l-h)}{n(n-1)}(1-gf)^{r-1},\]

\[= \left\{ \frac{r(r-1)}{n(n-1)} + \frac{rt}{n(n-1)} + \frac{r(n-r-t)}{n(n-1)} + \frac{2s(l-h)rt}{n(n-1)} + \frac{4sr(n-r-t)}{n(n-1)} \right\}(1-gf)^{r-1}
\]

\[= \left\{ \frac{r}{n} + \frac{2s(l-h)rt}{n(n-1)} + \frac{4sr(n-r-t)}{n(n-1)} \right\}(1-gf)^{r-1}.
\]

Therefore,

\[
P(AA \text{ survives}) \propto \sum_{r=0}^{n} \sum_{t=0}^{n-r-t} \frac{n!}{r!t!(n-r-t)!} \cdot \frac{r}{n} \cdot \frac{2s(l-h)rt}{n(n-1)} + \frac{4sr(n-r-t)}{n(n-1)}(1-gf)^{r-1}
\]

If we assume that products and powers, except the first, of terms involving
\( f \) and \( s \) are negligible, then approximately,
\[
P(\text{AA survives}) = \frac{n}{n-1} \sum_{r=0}^{n} \sum_{t=0}^{n-r-1} \frac{n!}{r!t!(n-r-t)!} p_r^r(2Q)^t R^{n-r-t} \frac{r}{n} + \frac{2s(l-h)rt}{n(n-1)} + \frac{4sr(n-r-t)}{n(n-1)} - \frac{gf^r(r-1)}{n} \]

\[
= \frac{n}{n-1} \sum_{r=1}^{n} \sum_{t=0}^{n-r-1} \frac{n!}{(r-1)!t!(n-r-t)!} p_r^r(2Q)^t R^{n-r-t} \]

\[
+ 2QR[2s(l-h)] \sum_{r=1}^{n-1} \sum_{t=1}^{n-r-1} \frac{n!}{(r-1)!(t-1)!(n-r-t)!} p_r^r(2Q)^t R^{n-r-t} \]

\[
+ PR[4s] \sum_{r=0}^{n-1} \sum_{t=0}^{n-r-1} \frac{n!}{(r-2)!t!(n-r-t)!} p_r^r(2Q)^t R^{n-r-t} \]

\[
- \frac{gf(n-1)p^2}{n} \sum_{r=2}^{n} \sum_{t=0}^{n-r-1} \frac{n!}{(r-2)!t!(n-r-t)!} p_r^r(2Q)^t R^{n-r-t} \]

\[
= P + 4PQs(l-h) + 4PRs - p^2(n-1)gf \]

\[
= P(1+4Qs(l-h)+4Rs-P(n-1)gf). \]

Similarly, we may show that the Aa genotype has a probability of survival given by

\[
P(\text{Aa survives} | \text{rAA, tAa, (n-r-t)aa}) = \left\{ \frac{2rt}{n(n-1)} \right\}^{-1} \left[ 2s(l-h) \right] \]

\[
+ \frac{t(t-1)}{n(n-1)} + \frac{2t(n-r-t)}{n(n-1)} \left[ \frac{1}{2} + s(l+h) \right] (1-kf)^{t-1} \]

\[
= \left\{ \frac{rt+t(t-1)+t(n-r-t)}{n(n-1)} \right\} + \frac{2s(l+h)n(n-r-t)}{n(n-1)} - \frac{2s(l-h)n(n-r-t)}{n(n-1)} \frac{2}{n(n-1)} (1-kf)^{t-1} \]

\[
= \left\{ \frac{t}{n} + \frac{2s(l+h)n(n-r-t)}{n(n-1)} - \frac{2s(l-h)n(n-r-t)}{n(n-1)} \right\} (1-kf)^{t-1}. \]

Therefore, the probability of Aa surviving may be related by
P(Aa survives) = \sum_{r=0}^{n} \sum_{t=0}^{n-r} \frac{n!}{r!t!(n-r-t)!} p^r(2Q)^t R^{n-r-t} \left\{ \frac{t}{n} + \frac{2s(l+h)t(n-r-t)}{n(n-1)} - \frac{2s(l-h)t}{n(n-1)}(1-kf)^{t-1} \right\} \\
\leq 2Q \sum_{r=0}^{n-1} \sum_{t=1}^{n-r} \frac{(n-1)!}{r!(t-1)!(n-r-t)!} p^r(2Q)^t R^{n-r-t} \\
+ 2s(l+h)2QR \sum_{r=0}^{n-2} \sum_{t=1}^{n-r} \frac{(n-2)!}{r!(t-1)!(n-r-t-1)!} p^r(2Q)^t R^{n-r-t-1} \\
- 2s(l-h)2PQ \sum_{r=1}^{n-1} \sum_{t=1}^{n-r} \frac{(n-2)!}{(r-1)!(t-1)!(n-r-t)!} p^{r-1}(2Q)^t R^{n-r-t} \\
- kf(n-1)(2Q)^2 \sum_{r=2}^{n-2} \sum_{t=2}^{n-r} \frac{(n-2)!}{r!(t-2)!(n-r-t)!} p^r(2Q)^t R^{n-r-t} \\
= 2Q[1+2Rs(l+h)-2Ps(l-h)-(n-1)2Qkf] \\

A similar argument exists for the genotype, aa:

P(aa survives|AA, aA, (n-r-t)aa) = \left\{ \frac{2r(n-r-t)}{n(n-1)} \left[ \frac{1}{2} - 2s \right] \\
+ \frac{2t(n-r-t)}{n(n-1)} \left[ \frac{1}{2} - s(l+h) \right] + \frac{(n-r-t)(n-r-t-1)}{n(n-1)}(1-f)^{n-r-t-1} \right\} \\
= \left\{ \frac{n-r-t}{n} - \frac{4sr(n-r-t)}{n(n-1)} - \frac{2s(l+h)t(n-r-t)}{n(n-1)}(1-f)^{n-r-t-1} \right\} \\

Therefore,

P(aa survives) = \sum_{r=0}^{n} \sum_{t=0}^{n-r} \frac{n!}{r!t!(n-r-t)!} p^r(2Q)^t R^{n-r-t} \left\{ \frac{t}{n} - \frac{4sr(n-r-t)}{n(n-1)} - \frac{2s(l+h)t(n-r-t)}{n(n-1)}(1-f)^{n-r-t-1} \right\} \\
\leq R \sum_{r=0}^{n-1} \sum_{t=0}^{n-r} \frac{(n-1)!}{r!(n-r-t-1)!} p^r(2Q)^t R^{n-r-t-1} \\
= R \sum_{r=0}^{n-1} \sum_{t=1}^{n-r} \frac{(n-1)!}{r!(n-r-t)!} p^r(2Q)^t R^{n-r-t} \\
= R \sum_{r=0}^{n-1} \sum_{t=1}^{n-r} \frac{(n-1)!}{r!(n-r-t)!} p^r(2Q)^t R^{n-r-t}
The total fitness is proportional to

\[ P + 4s(1-h)PQ + 4sPR - P^2(n-1)gf + 2Q + 4s(1+h)QR - 4s(1-h)PQ \]

\[ - 4Q^2(n-1)kf + R - 4sPR - 4s(1+h)QR - R^2(n-1)f \]

\[ = 1 - f(n-1)(P^2g + 4Q^2k + R^2). \]

Therefore, the frequency of allele A in the gametic array from these adults is

\[ p(A) = \frac{P + 4s(1-h)PQ + 4sPR - P^2(n-1)gf + 2Q + 2s(1+h)QR - 2s(1-h)PQ - 2Q^2(n-1)kf}{1 - f(n-1)(P^2g + 4Q^2k + R^2)}. \]

Under Hardy-Weinberg assumptions \( P = p^2, Q = pq, R = q^2 \), so the change in gene frequency for one generation can be given by the approximation,
\[
p' = \frac{p+2s(p^3 q+2p^2 q^2+pq^3)-2sh(p^3 q-pq^3)-f(n-1)(p^4 q+2p^2 q^2)}{1-f(n-1)(p^4 g+4p^2 q^2 k+q^4)}
\]

\[
= \frac{p+2spq-2shpq(p-q)-f(n-1)p^2(p^2 g+2q^2 k)}{1-f(n-1)(p^4 g+4p^2 q^2 k+q^4)}
\]

\[
= \frac{p+2spq[1-h(p-q)]-f(n-1)p^2(p^2 g+2q^2 k)}{1-f(n-1)(p^4 g+4p^2 q^2 k+q^4)}
\]  
\hspace{1cm} (3.6.8)

We may compare this to the case for \( n = 2 \) (Expression 3.6.6).

The change in gene frequency can be calculated to be

\[
\Delta p = \frac{p+2spq[1-h(p-q)]-f(n-1)p^2(p^2 g+2q^2 k)-p}{1-f(n-1)(p^4 g+4p^2 q^2 k+q^4)}
\]

\[
= \left\{ \frac{2spq[1-h(p-q)]-f(n-1)p(p^3 q+2q^2 pk-p^4 g-4p^2 q^2 k+q^4)}{(denominator)} \right\}
\]

\[
= \left\{ \frac{2spq[1-h(p-q)]-f(n-1)p(p^3 q-2pq^2 (2p-1)k-q^4)}{(denominator)} \right\}
\]

\[
= \left\{ \frac{2spq[1-h(p-q)]-f(n-1)pq(p^3 q-2pq(p-q)k-q^3)}{(denominator)} \right\}
\]

\[
= \frac{pq[2s[1-h(p-q)]-f(n-1)[p^3 q-2pq(p-q)k-q^3]}{1-f(n-1)(p^4 g+4p^2 q^2 k+q^4)}
\]  
\hspace{1cm} (3.6.9)

We can see that generally an approximate value for the equilibrium point is given by a solution to the cubic

\[
p^3 f(n-1)(g+4k+1) - 3p^2 f(n-1)(2k+1) + p[4sh+f(n-1)(2k+3)] \]

\[
- 2s(l+h) - f(n-1) = 0. \]  
\hspace{1cm} (3.6.10)

The following graphs show the change in gene frequency for various values of the parameters, \( n, s, f, g, h, \) and \( k. \)
Figure 3.9a. Graph of gene frequency for the variable fitness diploid model (3.6.8) (in this figure $s = 0.01$, $h = 0.5$, $g = 2$, $k = 1.5$ for all curves, \(--\): $n = 3$, $f = 0.1$; \(x-x-x\): $n = 3$, $f = 0.15$; \(+++: n = 11$, $f = 0.1$)
Figure 3.9b. Graph of gene frequency for the variable fitness diploid model (3.6.8) (in this figure $n = 3$, $s = 0.01$, $g = 2$ for all curves, $---$: $h = 0.75$, $k = 1.25$, $f = 0.1$; $+++$: $h = 0.5$, $k = 2$, $f = 0.05$)
Figure 3.9b. Graph of gene frequency for the variable fitness diploid model (3.6.8) (in this figure n = 3, s = 0.01, g = 2 for all curves, ---: h = 0.75, k = 1.25, f = 0.1; +++: h = 0.5, k = 2, f = 0.05)
The first figure shows that increasing \( n \) increases the rate of change in gene frequency. The shape of the curves may be taken to imply that the equilibrium point, approximated by the cubic, (3.6.10), is stable. This is shown to be so by running the simulation with starting values for \( p = 0.9 \). These curves are not shown here.

The estimated equilibrium values are:

- for the curve marked "—", \( \hat{p} = 0.47 \);
- for the curve marked "x-x-x", \( \hat{p} = 0.42 \);
- for the curve marked "+++", \( \hat{p} = 0.40 \).

It can be seen that the effect of increasing \( f \) is to accelerate the movement of the system towards the equilibrium position, even though in this case the equilibrium frequency for \( f = 0.15 \) is closer to the starting value than that for \( f = 0.01 \). The effect of increasing \( n \) from 3 to 5 is to cause the system to approach very rapidly the equilibrium. Whereas, for \( n = 3 \) the graph still shows that the frequency is changing after 70 generations, where \( n = 5 \), there is essentially no further change after ten generations.

The second figure shows two curves. The approximation for the equilibrium value for the first curve ("—") is \( \hat{p} = 0.67 \), and that for the second curve ("+++") is \( \hat{p} = 0.81 \). This latter graph shows that even with quite small values for \( f \) (\( f = 0.05 \)) a stable, nontrivial, equilibrium may exist.
4. MODELS OF ROW COMPETITION

4.1. Haploid Individuals Competing in a Row

We have shown in the previous sections that if we relax the condition that individuals compete only in a pairwise fashion, we may obtain results different from what may be inferred from the pairwise results. The difficulty of incorporating in a tractable model all the interactions in even a simple square planting is considerable, and to avoid that difficulty, we shall consider a row-wise system of competitive interactions. This model is not biologically unacceptable since many commercial crops are planted in this fashion. To be fair, inter-row as well as intra-row competition does occur, but the former may be considered rather general, while the latter may be accommodated into a reasonably tractable formulation.

We imagine a number of plants linearly arranged as in Figure 4.1.

\[ X_1 - X_j - X_i - X_k - X_m \]

Figure 4.1. Arrangement of plants in a row

We shall consider first the case of autogamous plants undergoing competition from the nearest neighbor only. In a pure stand, we may state that the reproductive capacity of \( X_i \) is \( b_i + 2d_{ii} \), so the relationship with the previous formulation, after Schutz, Brim and Usanis (1968), is
If \( X_i \) is in competition with \( X_j \) on one side and \( X_k \) on the other, as in Figure 4.1 above, then the reproductive capacity of \( X_i \) is

\[ b_i + d_{ij} + d_{ik}. \]

The parameter, \( d_{ij} \), is the decrement (or increment) in the reproductive capacity of \( X_i \) owing to the competition from \( X_j \). For an extended planting the expected reproductive capacity of \( X_i \) is given by

\[ b_i + \sum_j p_j d_{ij} + \sum_k p_k d_{ik} = b_i + \sum_j p_j d_{ij}. \]

where \( p_i \) is the frequency of the genotype \( X_i \).

The average reproductive capacity of the whole population is

\[ \sum_i p_i (b_i + \sum_j p_j d_{ij}) = \sum_i p_i b_i + \sum_i \sum_j p_i p_j d_{ij}. \]

If we express \( \{d_{ij}\} = D, \{p_i\} = p, \{b_i\} = b \), then the average reproductive capacity is

\[ p'b + 2p'Dp. \]

The frequency of \( X_i \) in the next generation is

\[ p_i' = \frac{p_i (b_i + \sum_j p_j d_{ij})}{p'b + 2p'Dp}, \]

so

\[ \Delta p_i = \frac{p_i (b_i + \sum_j p_j d_{ij} - p'b - 2p'Dp)}{p'b + 2p'Dp}. \]

which leads to a solution for the equilibrium value,
\[ \begin{align*}
\sum_{j} 2 \mathbf{p}_j \mathbf{d}_{ij} &= \mathbf{p}' \mathbf{b} + 2 \mathbf{p}' \mathbf{Dp}, \quad \forall i, \\
\text{or} \\
\mathbf{b} + 2 \mathbf{Dp} &= (\mathbf{p}' \mathbf{b}) \mathbf{l} + (2 \mathbf{p}' \mathbf{Dp}) \mathbf{l},
\end{align*} \tag{4.1.2} \]

\( \mathbf{p}' \mathbf{b} \) and \( \mathbf{p}' \mathbf{Dp} \) are scalars, so the expression may be rewritten

\[ \begin{align*}
\mathbf{b} + 2 \mathbf{Dp} &= \mathbf{l} (\mathbf{p}' \mathbf{b}) + \mathbf{1} (2 \mathbf{p}' \mathbf{Dp}) \\
\Rightarrow (\mathbf{I}-\mathbf{l}\mathbf{p}' \mathbf{b}) &= -2 (\mathbf{I}-\mathbf{l}\mathbf{p}' \mathbf{Dp}).
\end{align*} \]

Because the matrix, \( (\mathbf{I}-\mathbf{l}\mathbf{p}' \mathbf{b}) \), possesses a particular pattern, this allows a number of conclusions to be drawn regarding the nature of an inverse. In particular, we refer to Theorem 8.3.3 on page 170 of Graybill (1969): Let the \( k \times k \) matrix \( \mathbf{C} \) be given by \( \mathbf{C} = \mathbf{D} + \alpha \mathbf{a} \mathbf{b}' \), where \( \mathbf{D} \) is a nonsingular diagonal matrix, \( \mathbf{a} \) and \( \mathbf{b} \) are each \( k \times 1 \) vectors, and \( \alpha \) is a scalar such that

\[ \alpha \neq - \left( \sum_{i=1}^{k} \frac{\mathbf{a}_i \mathbf{b}_i}{\mathbf{d}_{ii}} \right)^{-1}. \]

The inverse of \( \mathbf{C} \) is

\[ \mathbf{C}^{-1} = \mathbf{D}^{-1} + \gamma \mathbf{a} \mathbf{b}' \]

where

\[ \gamma = -\alpha (1 + \alpha \sum \frac{\mathbf{a}_i \mathbf{b}_i}{\mathbf{d}_{ii}})^{-1}; \quad \mathbf{a}_i = \mathbf{a}_i / \mathbf{d}_{ii}; \quad \mathbf{b}_i = \mathbf{b}_i / \mathbf{d}_{ii}; \quad \text{and} \quad \mathbf{d}_{ii} \text{ is the i-th diagonal element of} \mathbf{D}. \]

In our example, however, \( -\mathbf{l} = - (\mathbf{I}-\mathbf{l}\mathbf{p}' \mathbf{b}) \), and no inverse exists, although a conditional inverse of rank \( k-1 \) may be easily derived, provided \( \mathbf{p}' \mathbf{b} \) were known. We see further that if \( (\mathbf{I}-\mathbf{l}\mathbf{p}' \mathbf{b}) \) were invertible we would have \( \mathbf{p}' \mathbf{b} = - (1/2) \mathbf{D}^{-1} \mathbf{b}, \) providing \( \mathbf{D}^{-1} \) existed. It would then
immediately follow that the denominator of the Expression (4.1.1) would be zero, leading to an undefinable result, since the numerator, 
\[ b_i + 2\Sigma p_j d_{ij}, \]
would also be zero at equilibrium.

We can circumnavigate this impasse by re-expressing (4.1.2). We note that, obviously, \( b_i = \Sigma p_j b_i \), hence,

\[ b_i + 2\Sigma p_j d_{ij} = \Sigma p_j (b_i + 2d_{ij}), \]

so if we let \( \{b_i + 2d_{ij}\} = B \), then the Equation (4.1.2) may be written as

\[ Bp = (p'p)^{-1} \]

which may lead to a more tractable result. Following Cockerham and Burrows (1971), we see that (4.1.3) leads to a solution for \( p \),

\[ \hat{p} = kB^{-1} \]

where \( k \) is a scalar, which may be readily determined thus,

\[ 1 = p'\hat{p} = p'B^{-1}k \]

\[ k = \frac{1}{p'B^{-1}p} \]

so

\[ \hat{p} = \frac{B^{-1}}{p'B^{-1}p} \]

To return to (4.1.2) we note that a particular case occurs when 
\( \tilde{D} = -\tilde{D}' \), i.e. when \( d_{ij} = -d_{ji} \). In this case \( p'Dp = 0 \) for all values of \( p \) (Theorem 12.12.19, page 336, Graybill, 1969). This occurs when
the decrement suffered by $X_i$ in competition with $X_j$ is equal to the increment $X_j$ incurs in competition with $X_i$.

We can see that this formulation of the model differs little from that of Schutz, Brim and Usanis (1968) in the rewritten form discussed earlier. To a certain extent it is less useful for estimating the parameters, $b$ and $D$. To accomplish this we need to observe the reproductive capacity of each genotype in a number of stands differing in frequency for the various genotypes. It will be seen, though, that for the autogenous progenies of an F1 hybrid, the stands based on $F2$, $F3$, $F4$, ...etc, while containing differing frequencies of each of the three genotypes, do not lead to estimates for $b$ and $D$. We shall illustrate this by an example.

Wiebe, Petr and Stevens (1963) published data obtained from growing a barley hybrid (in which the homozygous genotypes and the heterozygous genotype for a particular locus were easily identifiable) in stands of $F2$, $F3$, $F4$, "F00", and pure stands, as well as an equiproporionate stand.

We take the expected reproductive capacity of $X_i$ to be

$$b_i + 2\sum_{j} d_{ij}$$

and we take $X_1$, $X_2$ and $X_3$ to be the genotypes $VV$, $Vv$ and $vv$ respectively. The reproductive capacity of $VV$ in the different populations may be given as follows.
\[
\begin{align*}
F_2; \quad & r_{1,F2} = b_1 + \frac{1}{2}d_{11} + d_{12} + \frac{1}{2}d_{13}, \\
F_3; \quad & r_{1,F3} = b_1 + (3/4)d_{11} + \frac{1}{2}d_{12} + (3/4)d_{13}, \\
F_4; \quad & r_{1,F4} = b_1 + (7/8)d_{11} + \frac{1}{4}d_{12} + (7/8)d_{13},
\end{align*}
\]

the equiproportionate population,

\[
F_\infty; \quad r_{1,F_\infty} = b_1 + (2/3)d_{11} + (2/3)d_{12} + (2/3)d_{13},
\]

and the pure stand,

\[
P_1; \quad r_{1,P1} = b_1 + 2d_{11},
\]

and the population formed from equal proportions of the homozygotes,

\[
F_{oo}; \quad r_{1,F_{oo}} = b_1 + d_{11} + d_{13}.
\]

We may express this in the form of a linear model, \( y = Xb + e \), thus,

\[
\begin{bmatrix}
r_{1,F2} \\
r_{1,F3} \\
r_{1,F4} \\
r_{1,F_\infty} \\
r_{1,P1} \\
r_{1,F_{oo}}
\end{bmatrix} = 
\begin{bmatrix}
1 & 1/2 & 1 & 1/2 \\
1 & 3/4 & 1/2 & 3/4 \\
1 & 7/8 & 1/4 & 7/8 \\
1 & 2/3 & 2/3 & 2/3 \\
1 & 2 & 0 & 0 \\
1 & 1 & 0 & 1
\end{bmatrix} \begin{bmatrix}
b_1 \\
d_{11} \\
d_{12} \\
d_{13} \\
e_{1,F2} \\
e_{1,F3} \\
e_{1,F4} \\
e_{1,F_\infty} \\
e_{1,P1} \\
e_{1,F_{oo}}
\end{bmatrix}.
\]

There need be no defined distribution on the e's.

The matrix, however, is of rank 3, which we may show by reducing the matrix to row echelon form.
No unique solution exists, and we need data from another mixed stand, where the proportions of the genotypes are linearly independent to the ones already listed, to be able to identify all four parameters.

The development of the model in this direction, though, is considered necessary to bring in a further aspect of plants competing in rows.

In Figure 4.1 outlined at the beginning of this section, we can see that the nearest-but-one plants to $X_i$ are $X_{i1}$ and $X_{i2}$. Should the presence of $X_{i1}$ induce a decrement in $X_j$, then the latter should have an ameliorated effect on $X_{i1}$, i.e., if $X_{i1}$ induces a decrement on $X_j$, it
indirectly induces an increment on $X_i$ through its effect on $X_j$.

Published reports on the interaction between plants in a row are scarce, and conflicting. Donald (1963) quotes Hozumi, Koyama and Kira (1955) whose work with corn showed that the influence persisted as far as the fourth plant away. Figure 4.2 (taken from Donald, 1963) graphs the correlation coefficients of shoot weight between a plant and the plants adjacent, or two, three, four or five plants away. A noticeable positive correlation exists between the plant measured and the second and fourth plants, while a clear negative correlation exists between the measured plant and the first and third.

In contrast, Lichter (1972), studying sugarbeet, showed that while a highly significant negative correlation coefficient of $-0.97$ existed for the root weight of the central plant and that of its nearest neighboring plant, the correlation between the central and the second neighboring plant was 0.26 (nonsignificant). Regression coefficients may be calculated from Lichter's (1972) data. A change of one unit increase root weight in the nearest plant brings about a decrement of $0.33 \pm 0.08$ in the central plant, but a change of one unit in the second plant away brings about a change of only $0.02 \pm 0.09$ in the central plant. Effectively, a sugarbeet plant affects its nearest neighbors only, the plants further away remain unaffected. This is in distinct contrast to the results of Hozumi, Koyama and Kira (1955).

In Figure 4.1 we depicted genotypes arranged in a linear fashion:

\[-X_1 - X_j - X_i - X_k - X_m - \ldots\]
Figure 4.2. Correlation for shoot weight between a central plant and consecutive neighbors in a row (Donald, 1963; after Hozumi et al., 1955)
The individual being measured is, as before, $x_i$. In the absence of second-neighbor effects its reproductive capacity is

$$b_i + d_{ij} + d_{ik}.$$  

Now, however, we assume that $x_i$ affects $x_j$, and $x_m$ affects $x_k$ sufficiently so that the effect of $x_j$ and $x_k$ on $x_i$ is ameliorated or enhanced. To approach this with a simple additive model we assume that $x_j$ will be reduced or enhanced by $d_{jl}$ owing to the influence of $x_i$. A portion, $-\rho$, of this change is transferred to $x_i$. Thus, the reproductive capacity of $x_i$ becomes

$$b_i + d_{ij} + d_{ik} - \rho d_{jl} - \rho d_{km}.$$

The expected reproductive capacity for $x_i$ competing against all possible neighbors may be expressed as

$$b_i + 2 \sum_{j} p_j d_{ij} - 2 \rho \sum_{j,k} p_j p_k d_{jk}.$$

It can be seen that if we assume that the effect of consecutive neighbors follows an auto-correlation-like pattern, then the reproductive capacity of $x_i$ becomes

$$b_i + 2 \sum_{j} p_j d_{ij} - 2(\rho - \rho^2 + \rho^3 - \ldots) \sum_{j,k} p_j p_k d_{jk}$$

$$= b_i + 2 \sum_{j} p_j d_{ij} - \frac{2\rho}{1+\rho} \sum_{j,k} p_j p_k d_{jk},$$

provided $|\rho| < 1$, which is a reasonable assumption.

We wish to express this in the form $w_i$, so that $w_i$ is the $i$-th row of $\widetilde{W}$, where $\widetilde{W}$ is a matrix and $\tilde{p}$ is the vector of frequencies of the
genotypes. From this consideration we can determine the elements of the matrix, $W$ to be

$$w_{ij} = b_i + 2d_{ij} - \frac{2\rho}{1+\rho} \sum_k p_k d_{kj}$$

$$w_{jj} = b_j + 2d_{jj} - \frac{2\rho}{1+\rho} \sum_k p_k d_{kj}$$

Following an idea due to Cockerham and Burrows (1971) we let

$$t_{ij} = w_{ij} - w_{jj}$$

i.e.,

$$t_{ij} = b_i - b_j + 2(d_{ij} - d_{jj}).$$

Thus the matrix, $T$, ($T = \{t_{ij}\}$) contains elements that do not depend on $p$.

Returning to the $\{w_{ij}\}$, we find that

$$w_{i.} = i\text{-th row of } W_p$$

$$= b_i + 2[i\text{-th row of } D_p] - 2\frac{\rho}{1+\rho} p'D_p$$

and

$$w_{..} = \sum_i w_{i.} = p'W_p$$

$$= p'b + 2p'D_p - 2\frac{\rho}{1+\rho} p'D_p$$

$$= p'b + 2(1 - \frac{\rho}{1+\rho}) p'D_p$$

$$= p'b + 2\frac{1}{1+\rho} p'D_p.$$
population.

Let \( w \) be the vector of the \( w_{jj} \) values. Then

\[
\begin{align*}
\sum_j w'P + \sum_j \bar{w}'T \bar{p} &= \sum_j p_j b_j + 2\sum_j p_j d_{jj} \sum_j - \frac{\rho}{1+\rho} \sum_j p_j \sum_k p_k d_{kj} \\
&+ 2\sum_j p_j p_j d_{ij} - 2\sum_j p_j d_{jj} \\
&= \sum_j p_j b_j + \frac{2}{1+\rho} \sum_j p_j \sum_k d_{kj} = \bar{w}.
\end{align*}
\]

We may express \( p'\bar{T}p \) as \( t.. \), so \( w.. = \bar{w}+t.. \), where \( \bar{w} = w'p \). Now,

\[
\begin{align*}
\sum_j w'P + \sum_j p_j t_{ij} &= \sum_j p_j b_j + 2\sum_j p_j d_{ij} \sum_j - \frac{\rho}{1+\rho} \sum_j p_j \sum_k p_k d_{kj} \\
&+ b_i - \sum_j p_j b_j + 2\sum_j p_j d_{ij} - 2\sum_j p_j d_{jj} \\
&= b_i + 2\sum_j p_j d_{ij} - \frac{\rho}{1+\rho} \sum_k p_j \sum_k d_{kj} \\
&= w_i.
\end{align*}
\]

Now \( w_i - w.. = t_i - t.. \), so we now have the expression describing the change in gene frequency for one generation.

\[
p_i' = p_i w_i / w.. = p_i (\bar{w}+t_i..) / (\bar{w}+t..) \tag{4.1.4}
\]

and

\[
\Delta p_i = p_i (t_i.. - t..) / (\bar{w}+t..) \tag{4.1.5}
\]

We now wish to examine the existence of equilibrium values using the \( w_{ij} \)'s and \( t_{ij} \)'s. Because the matrix \( \bar{T} \) does not depend on \( \bar{p} \), the latter may be more useful in the following discussion.

At equilibrium, \( \Delta p_i = 0 \), for all \( i \), so
\[ \hat{\omega}_i = \hat{\omega}_{..} \text{, and } \hat{\xi}_i = \hat{\xi}_{..} \text{, for all } i. \]

Now, to determine the equilibrium point in terms of the \( w_{ij} \)'s, we have

\[ b + 2Dp - 2\frac{1}{1+\rho} \sim \sim lP' \sim \sim dp = lP'b + 1 \sim \sim 2\frac{1}{1+\rho} \sim \sim p'Dp \]

\[ \Rightarrow (lP'-I)b = 2(lP'-I)Dp \]

because

\[ \frac{1}{1+\rho} + \frac{\rho}{1+\rho} = 1, \]

\[ \Rightarrow \hat{\beta} = k(1/2) \sim \sim \sim \sim \sim \sim (lP'-I) \sim \sim \sim \sim \sim \sim (lP'-I)b \]

for some \( k \) such that \( l'p = 1. \)

\[ l'p = k(1/2) \sim \sim \sim \sim \sim \sim \sim (lP'-I) \sim \sim \sim \sim \sim \sim (lP'-I)b \]

therefore

\[ k = \frac{2}{l'D^{-1}(lP'-I)(lP'-I)b}, \text{ and therefore} \]

\[ \hat{\beta} = \frac{D^{-1}(lP'-I)(lP'-I)b}{l'D^{-1}(lP'-I)(lP'-I)b} \]

One conclusion can be reached: the effect of the nonnearest neighbors does not influence the equilibrium point, if it exists, provided that the effect follows the auto-correlation-like system outlined earlier. The expression, as it stands, remains intractable. However, if we use the fact that \( \Delta p_i = 0, \forall i, \Rightarrow \hat{\xi}_i = \hat{\xi}_{..}, \forall i, \) then this may be expressed as

\[ \sim \sim \sim Tp = p'Tp1 \]
i.e.,
\[ T_p = k l, \quad k \text{ a scalar}, \]
\[ \Rightarrow \tilde{p} = k T^{-1} \]

But
\[ l = 1' \tilde{p} \Rightarrow 1'T^{-1} \tilde{p} = 1/k \]
\[ \Rightarrow \tilde{p} = \frac{T^{-1} \tilde{p}}{l'T^{-1} \tilde{p}} \]

Tallis (1966) gives a proof showing that if
\[ \tilde{p}' = \frac{W \tilde{p}}{\tilde{p}'W} \]
then at equilibrium, where \( \tilde{p}' = \tilde{p} \), then
\[ \tilde{p} = \frac{W^{-1} \tilde{p}}{l'W^{-1} \tilde{p}} \]

His proof follows that due to Wright (1969) which suffers from the criticism that if there are \( n \) alleles at a locus the frequency of the \( n \)-th allele is not independent of that of the \( n-1 \) other alleles. We offer an alternative approach.

Let \( D = \left( \begin{array}{c} p_1 \\ p_2 \\ \vdots \\ p_n \end{array} \right) \) with off-diagonal elements zero.

Then the expression
\[ p_i' = \frac{p_i W_i}{\sum} \quad , \quad i = 1, \ldots, n, \] may be written as
\[ p' = \frac{D W p}{P' W p} \] where

\[ w_i. = \sum_j p_j w_{ij}, \text{ and} \]

\[ W = \{ w_{ij} \}. \]

Equilibrium conditions exist where \( \Delta p = 0 \). So

\[ \Delta p = \left( \frac{D W}{P' W p} - 1 \right) p \]

\[ = \left( \frac{D W - p' W p}{P' W p} \right) p . \]

Thus

\[ \Delta p = 0 \implies W D l = k, \text{ a scalar, then} \]

\[ D W p = (p' W p)p \quad (4.1.6) \]

\[ D l = p, \text{ and } D' = D, \text{ so (4.1.6) becomes} \]

\[ D W D l = (l' D W D l) D l . \]

Let \( (l' D W D l) = k, \text{ a scalar, then} \)

\[ D W D l = k D l . \]

Clearly, for \( p_i > 0, \forall i, D^{-1} \) exists, so

\[ W D l = k l, \text{ or} \]

\[ W p = k l, \text{ and} \]
\[ \hat{p} = kW^{-1} \]

with
\[ k = \hat{p}'W\hat{p} \]
\[ = k1'(W^{-1})'W^{-1}1k, \]
\[ k = [1'(W^{-1})'1]^{-1}, \]

but because it is a scalar,
\[ k = (1'(W^{-1})'1)^{-1}. \]

Therefore (4.1.7) has a solution
\[ \hat{p} = \frac{W^{-1}}{1'(W^{-1})'1} \]

if \( W \) is nonsingular.

This completes the development.

The solution where \( W \) is singular is suggested by Cockerham and Burrows (1971) to be
\[ \hat{p} = \frac{1'(W^{-1})'1}{W^{-1}} \]

We discuss this result in Appendix B.

The graphs, Figure 4.3, show the influence of the parameter \( \rho \) on the rate of change in gene frequency. As the influence of neighbors along the row increases, i.e., \( \rho \) increases, the rate of change in gene frequency increases. The values of \( d_{ij} \) are given in the matrix,
\[ D = \begin{bmatrix} .5 & .5 & .5 & .5 \\ .52 & .52 & .52 & .52 \\ .53 & .53 & .53 & .53 \\ .54 & .54 & .54 & .54 \end{bmatrix} \]

and those for \( b_i \) an
For $p = 0$, the net reproductive value of the genotypes is the same as those given by Allard and Adams (1969).

4.2. Diploid Individuals Competing in a Row

In this development, we consider the case of an intermating population planted in rows. The genotypic array of the population at zygote formation is $\sum \sum p_x p_y A_x A_y$. Each genotype, $A_u A_v$, has the following viability:

$$w_{uv} = b_{uv} + 2 \sum \sum p_x p_y d_{uv,xy} - 2 \frac{p}{1+p} \sum \sum p_x p_y p_z d_{xy,rs}$$

Analogous to the case of autogenous individuals, the relative viability of $A_u A_v$ depends on a constant for that genotype, $b_{uv}$, affected (additively) by the immediate neighbors, $A_x A_y$ and $A_i A_j$, by $d_{uv,xy}$ and $d_{uv,ij}$ respectively, and affected by the more remote neighbors by $2 \frac{p}{1+p} \sum \sum p_x p_y p_z d_{xy,rs}$.

During a cycle of selection, we assume that each surviving adult produces the same number of gametes, so that
Figure 4.3a. Graph of gene frequency for four haploid genotypes competing in a row. This figure, with Figure 4.3b, shows the effect of $p$ on rate of change of gene frequency. The competitive effects are those given by Allard and Adams (1969, p. 354).

(In the model discussed here $b' = (0, 0, 0, 0),$

$$D = \begin{pmatrix} .5 & .5 & .5 & .5 \\ .52 & .52 & .52 & .52 \\ .53 & .53 & .53 & .53 \\ .54 & .54 & .54 & .54 \end{pmatrix},$$

$p = 0$)
Figure 4.3b. Graph of gene frequency for four haploid genotypes competing in a row. This figure shows the effect of $\rho$ on the rate of change of gene frequency, (cf. Figure 4.3a). In this figure the values for $b$ and $D$ are the same as for Figure 4.3a, but here $\rho = 0.7$
\[ p'_u = \frac{p}{w} (\Sigma p_{vuv}w_{uv}). \]

\[ p'_u = p_u (\frac{\Sigma u_{uv}b_{uv} + 2d_{uv} - 2}{\frac{\rho}{1+\rho} d_{uv}}), \]

where

\[ b_{uv} = \Sigma \Sigma p_{v} p_{uv} \]

\[ d_{uv} = \Sigma \Sigma \frac{\rho}{1+\rho} p_{y} d_{uv,xy} \]

and

\[ b_{..} = \Sigma p_{b_{..}}. \]

Hence

\[ \Delta p = \frac{p}{w} (\frac{b_{uv} + 2d_{uv} - 2}{\frac{\rho}{1+\rho} d_{uv}} - \frac{b_{..} - 2}{\frac{\rho}{1+\rho} d_{..}}). \]

As with the model for autogenous species, we see that the parameter, \( \rho \), does not enter into the numerator of the expression, i.e.,
\[ \Delta p_u = p_u \left( -\frac{u_{-u} + 2d_{-u}}{\bar{w}_{-u}} \right), \]

which implies that even though the population mean \( \bar{w}_{-u} \), and the rate of change in gene frequency \( \Delta p_u \), are affected by the parameter \( \rho \), the gene frequency for equilibrium is independent of \( \rho \).

At equilibrium

\[ b_{-u} + 2d_{-u} = b_{-u} + 2d_{-u}, \forall u. \]

At this point considerable complexity enters into the framework of this model. We shall attempt to simplify the situation by the introduction of \( f_{ij} = p_i p_j \), so that the genotypic array of the population may be written as \( \sum \sum f_{ij} A_i A_j \). If \( f \) is the \((n^2 \times 1)\) column vector of \( \{f_{ij}\} \), then \( f \) is made up of successive columns of the \((n \times n)\) matrix \( pp' \), where \( p' = (p_1, p_2, \ldots, p_n) \). We also have the result:

\[ Hf = p, \text{ where} \]

\[
H = \left[ \begin{array}{ccc}
1' & 1' & \\
n & \ddots & \\
1' & \end{array} \right], \quad f = \left[ \begin{array}{c}
p_1 \\
p_2 \\
\vdots \\
p_n \end{array} \right], \quad p = \left[ \begin{array}{c}
p_1 \\
p_2 \\
\vdots \\
p_n \end{array} \right],
\]

\( 1' \) is a row vector \((1 \times n)\) of 1's, and all off-diagonal block elements are zero. If we let
\[
G = \begin{bmatrix}
P & \frac{P}{P} \\
\frac{P}{P} & \ddots \\
\vdots & \ddots & \ddots \\
\frac{P}{P} & \cdots & \cdots & P
\end{bmatrix},
\]
then we note that the relationship between \(H\) and \(G\) is

\[
\begin{align*}
(1) & \quad HG = I_n \\
(2) & \quad HGH = H \\
(3) & \quad GHG = G \\
(4) & \quad HG = (HG)' \\
(5) & \quad GH \neq (GH)'
\end{align*}
\]

This implies that \(G\) is a conditional inverse of \(H\) and \(H\) is a conditional inverse of \(G\). Were it not for (5) above, they would be generalized or Moore-Penrose inverses.

We now have

\[
w_{uv} = b_{uv} + 2 \sum_{ij} f_{ij} d_{uv,ij} - \frac{\rho}{1+\rho} \sum_{xy} \sum_{ij} f_{ij} f_{xy} d_{ij,xy}.
\]

So we now have the task of setting up a \(n^2 \times n^2\) matrix of \(\{w_{uv,ij}\}\), such that \(W = \{w_{uv,ij}\}\).

\[
w_{uv,ij} = b_{uv} + 2d_{uv,ij} - \frac{\rho}{1+\rho} \sum_{xy} f_{xy} d_{xy,ij}
\]

So \(W_f\) is a column vector of \(w_{uv}\)'s, and

\[
G'W_f = (w_1, w_2, \ldots, w_n)' = \bar{W}(1)
\]

\[
f'W_f = \bar{w}.
\]
\[ w_{uv} = \text{[uv-th element of } W_f] \]

Now, if we let

\[ t_{uv,ij} = w_{uv,ij} - w_{ij,ij} \]

\[ = b_{uv} - b_{ij} + 2d_{uv,ij} - 2d_{ij,ij} \]

which is independent of \( p_x \) or \( f_{ij} \).

In a manner similar to that for the haploid case, we can derive the following correspondence. Let

\[ w = \text{a column vector of } w_{uv,uv}, \text{ then } \tilde{w}_f = \sum_{(uv)} f_w w_{uv,uv} \]

and

\[ \tilde{t}_{uv,..} = \sum_{(xy)} f_{xy} t_{uv,xy} \]

Therefore \( \tilde{w}_f + \tilde{t}_{uv,..} = \sum_{(xy)} f_{xy} w_{xy,xy} + \sum_{(xy)} f_{xy} w_{uv,xy} \]

\[ - \sum_{(xy)} f_{xy} w_{xy,xy} \]

\[ = \tilde{w}_{uv} \]

which is the uv-th element of \( \tilde{w}_f + T_f \).

\[ \tilde{w}_{..} = \tilde{w}_f + \tilde{T}_f = \tilde{w}_f + t_{..},.. \]

and

\[ \Delta p_u = p_u \left( \frac{\sum_v \left( t_{uv, -t,..} \right)}{\tilde{w}_f + \tilde{T}_f} \right) \]

\[ = p_u \left( \frac{\left( t_{u, -t,..} \right)}{\tilde{w}_f + t_{..},..} \right) \]

and at equilibrium,
\[ t_{u,..} = t_{..}, \forall u. \]

But, also \[ t_{u,..} = (0' \ p' 0'). T_f, \]

where the vector \( (0' \ p' 0') \) has zeroes in the first \( n(u-1) \) positions, and the last \( n(n-u) \) positions, so \( (0' \ p' 0'). T_f = f'. T_f. \forall u, \)

\[
\begin{bmatrix}
  p' \\
  p' \\
  \vdots \\
  p'
\end{bmatrix}
\]

i.e.,

\[
\begin{bmatrix}
  p' \\
  p' \\
  \vdots \\
  p'
\end{bmatrix} T_f = [f'. T_f]_1
\]

i.e.,

\[
\begin{bmatrix}
  p' \\
  p' \\
  \vdots \\
  p'
\end{bmatrix} T_f = \begin{bmatrix}
  p' \\
  p' \\
  \vdots \\
  p'
\end{bmatrix} T_f = \begin{bmatrix}
  p' \\
  p' \\
  \vdots \\
  p'
\end{bmatrix}
\]

i.e., \( G'TGp = p'G'TGp_l. \)

The nature of \( G'TG \) can be seen to be

\[
\begin{bmatrix}
  t_{1,1}, t_{1,2}, \ldots, t_{1,n} \\
  t_{2,1}, t_{2,2}, \ldots, t_{2,n} \\
  \vdots \\
  t_{n,1}, t_{n,2}, \ldots, t_{n,n}
\end{bmatrix}
\]

i.e., \( \{t_{i,j}\} \),

where \( t_{i,j} = \sum_{x} \sum_{y} p_{x} p_{y} t_{ix,jy} \). Note that \( t_{iu,jv} = t_{ui,jv} = t_{ui,vj} = t_{ui,vj} \), and \( t_{ij,ij} = 0 \ \forall i,j \), so that summation over the first or third subscript is the same as that over the second or fourth respectively.
This result does not appear to lead to a tractable expression for
the equilibrium value. We may note, in passing, that the matrices
$I \otimes p$ and $I \otimes p'$ ($\otimes$ being the Kronecker product) may be replaced by

\[
\begin{bmatrix}
  p_1 & p_2 & \ldots & p_n \\
  \vdots & \vdots & \ddots & \vdots \\
  p_1 & p_2 & \ldots & p_n \\
\end{bmatrix}
\]

(zeros in the blank positions)

and its transpose, respectively, with the same results.

The matrix $T = \{t_{uv,ij}\}$ possesses some measure of structure. All
diagonal elements are zero, and the relationship outlined above,
$t_{ui,jv} = t_{ui,jv}$ etc., is displayed in the matrix, but in spite of this
structure, the expression does not, apparently, lead directly to an
analytical result.

An iterative procedure for a numerical result is possible.

Given $b_{uv} d_{uv,xy}$ and $\rho$, the values for $t_{uv,xy}$ can be calculated.
A procedure may then be to take $p(0)$ as the vector of first choices.

\[
P(1) = \frac{G'TG}{\sim} = \frac{I_n \otimes p(0)}{\sim}; \quad G = I_n \otimes p(0).
\]

Since $G'TG$, being n x n, is considerably smaller than $T$ which is $n^2 \times n^2$,
this arrangement facilitates computation.

We may rearrange the matrix $T$ to be symmetric but still retain the
results we have.
Let \( T^* = \begin{bmatrix} T_{11}^{*} & T_{12}^{*} & \cdots & T_{1n}^{*} \\ T_{21}^{*} & T_{22}^{*} & \cdots & T_{2n}^{*} \\ \vdots & \vdots & \ddots & \vdots \\ T_{n1}^{*} & T_{n2}^{*} & \cdots & T_{nn}^{*} \end{bmatrix} \)

where, for example, \( T_{12}^{*} \) may be set out in full,

\[
\begin{bmatrix}
 t_{12,11} & t_{12,12} & \cdots & t_{12,1n} \\
 t_{12,21} & t_{12,22} & \cdots & t_{12,2n} \\
 \vdots & \vdots & \ddots & \vdots \\
 t_{12,n1} & t_{12,n2} & \cdots & t_{12,nn} 
\end{bmatrix}
\]

Clearly, \( T_{12}^{*} \) is symmetric, and it equals \( T_{21}^{*} \), so the \( T^* \) becomes a symmetric matrix, composed of blocks which themselves are symmetric.

Thus, the Expression (4.2.1) may be written

\[
G' T^* G \sim = (p'G' T^* Gp) \sim .
\]

This may facilitate computation of the expression, although it does not lead to an analytical result.
5. TRUNCATION SELECTION IN COMPETING DIPLOIDS IN ROWS

5.1. Mass Selection

Instead of considering the model from a population genetics point of view, we may see what happens when truncation selection is applied to a population of individuals competing in an indefinitely long row. We follow, in this discussion, the study by Griffing (1960) of truncation selection, as well as his studies of truncation selection involving interacting genotypes (Griffing, 1967; 1968a; 1968b; 1969; 1976a; 1976b; 1977).

In briefly reviewing Griffing’s (1960) paper which is discussed in more detail in Section 2.10, we note the following. An initial population consists of genotypes $A_iA_y$ with allelic frequencies $p_x$ for allele $A_x$. The genotypic value of the sub-population of the genotypes $A_iA_j$ is normally distributed with mean $d_{ij}$ and variance $\sigma^2$. The probability that $A_iA_j$ survives selection is $w_{ij}^* = 1 + (i/\sigma^2)d_{ij}^*$, where $i$ is the selection differential (the standardized selection differential $i = i/\sigma$).

If $d_{ij}^* = \alpha_i^* + \alpha_j^* + \delta_{ij}^*$, then the additive effect of the allele $A_i = \Sigma p_j d_{ij}^* = \alpha_i^*$. (The usual properties, $\Sigma p_i \alpha_i^* = 0$, $\Sigma p_i \delta_{ij}^* = 0$, hold.) The dominance effect associated with $A_iA_j = \delta_{ij}^*$. The total genetic variance, $\Sigma \Sigma p_i p_j d_{ij}^2 = \sigma_G^2$, and this may be partitioned into the additive genetic variance, $\sigma_A^2 = 2\Sigma p_i \alpha_i^2$, and the dominance variance, $\sigma_D^2 = \Sigma \Sigma p_i p_j \delta_{ij}^2$. The mean of the selected individuals from the
initial population is \( \frac{i}{2} \sigma^2 / \sigma^2 \), and the mean of the progeny of this group is \( \frac{i}{2} \sigma^2 / \sigma^2 \).

In Section 4.2 we introduced the term \( w_{uv} \) to denote the viability of the genotype \( A_u A_v \) in an indefinitely long row. In this section \( w_{uv} \) denotes the mean genotypic value of \( A_u A_v \), with selection being based on this genotypic value. (Note that \( w^*_{uv} \), the probability that \( A_u A_v \) survives selection, is not \( w_{uv} \).) We may treat \( w_{uv} \) in the same manner as we did in Section 4.2, so we may equate terms as follows:

\[
d_{ij}^* = w_{ij}.. = b_{ij} + 2 \sum_{x,y} p_x p_y d_{ij,xy} - \frac{2}{1+\rho} d_{..}..
\]

In this case we take \( d_{ij}^* \) as a deviation from the mean, and so

\[
\sum_{i,j} \sum_{x,y} p_x p_y d_{ij}^* = 0 \Rightarrow b_{..} + \frac{2}{1+\rho} d_{..}.. = 0; \text{ the usual definition of the dot notation is assumed. So } w_{..}.., \text{ the mean of the initial population is arbitrarily set to zero. The probability that } A_i A_j \text{ survives selection is}
\]

\[
w_{ij}^* = 1 + \left( \frac{i}{\sigma^2} \right) d_{ij}^* = 1 + \left( \frac{i}{\sigma^2} \right) w_{ij}.. = 1 + \left( \frac{i}{\sigma^2} \right) \left( b_{ij} + 2d_{ij}.. - \frac{2}{1+\rho} d_{..}.. \right)
\]

We may expand the parameter \( w_{ij,xy} \) as
\[ b_{ij} + 2d_{ij,xy} - \frac{2\rho}{1+\rho} d_{1+xy} \]

\[ = d^\alpha_i + d^\alpha_j + d^\beta_{ij} \]

\[ + 2\left[ d^\gamma_i + d^\gamma_j + d^\eta_{ij} + a^\gamma x + a^\gamma y + a^\eta_{xy} \right] \]

\[ + da^{(\gamma\gamma)_{ix}} + da^{(\gamma\gamma)_{iy}} + da^{(\gamma\gamma)_{jx}} + da^{(\gamma\gamma)_{jy}} \]

\[ + da^{(\eta\eta)_{ixy}} + da^{(\eta\eta)_{jxy}} + da^{(\eta\gamma)_{ijx}} + da^{(\eta\gamma)_{ijy}} \]

\[ + da^{(\eta\eta)_{ij,xy}} \]

\[ - \frac{2\rho}{1+\rho} \left[ a^\gamma x + a^\gamma y + a^\eta_{xy} \right] \]

\[ = (d^\alpha_i + 2d^\gamma_i) + (d^\alpha_j + 2d^\gamma_j) + (d^\beta_{ij} + 2d^\eta_{ij}) \]

\[ + \frac{2}{1+\rho}(a^\gamma x + a^\gamma y + a^\eta_{xy}) \]

\[ + 2\left[ da^{(\gamma\gamma)_{ix}} + da^{(\gamma\gamma)_{iy}} + da^{(\gamma\gamma)_{jx}} + da^{(\gamma\gamma)_{jy}} \right] \]

\[ + da^{(\eta\eta)_{ixy}} + da^{(\eta\eta)_{jxy}} + da^{(\eta\gamma)_{ijx}} + da^{(\eta\gamma)_{ijy}} \]

\[ + da^{(\eta\eta)_{ij,xy}}. \]

We use the prefix, subscript \( d \), to indicate the direct effects, the \( \alpha \)'s and \( \gamma \)'s being additive effects, the \( \beta \)'s and \( \eta \)'s being dominance effects. Because \( d^\gamma_i \) and \( d^\eta_{ij} \) are completely confounded with \( d^\alpha_i \) and \( d^\beta_{ij} \) respectively, we may arbitrarily set \( d^\gamma_i \), \( d^\gamma_j \), and \( d^\eta_{ij} \) equal to zero, simplifying the expression even further. From this we may obtain the following.
The direct additive effect of $A_i = d_i^A = w_i$.

The direct dominance effect of $A_iA_j = d_{ij}^D = w_{ij} - d_i^A d_j^A$.

The associative additive effect of $A_x$ as measured on $A_iA_j = a_{x} = 0.5(1+p)w_{xy}$.

The associative dominance effect of $A_xA_y$ as measured on $A_iA_j = a_{xy} = 0.5(1+p)(w_{xy} - w_{y} - w_{x})$.

The additive x additive interaction effect between direct allele $A_i$ and associative allele $A_x$ is $\frac{1}{1+p} a_{x_i} = 0.5(w_{i} - d_i^A)$.

The additive x dominance interaction effect between direct allele $A_i$ and associative genotype $A_yA_y = a_{xy} = 0.5(w_{i} - d_i^A x_{xy})$.

The dominance x dominance interaction effect between direct genotype $A_iA_j$ and associate genotype $A_yA_y = a_{ij} = 0.5(w_{ij} - d_{ij}^D)$.

We may partition the genetic variance in the following manner.

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \frac{(2}{1+p})^2 \sigma_A^2 + \frac{(2}{1+p})^2 \sigma_D^2 + 4 \sigma_{AA} + 4 \sigma_{AD} + 4 \sigma_{DA} + 4 \sigma_{DD}$$

where

$$\sigma_G^2 = \Sigma \Sigma \Sigma p_i p_j p_x p_y w_{ij,xy}^2$$

$$\sigma_A^2 = 2 \Sigma p_i (d_i^A)^2$$
In this study, however, we shall be concerned with additive effects only.

Returning to the selection study, we have that the selection value for $A_i A_j$, summed over all groups is

$$w^*_{ij} = 1 + \frac{i}{\sigma^2} \text{ind.} d_{ij}$$

$$= 1 + \frac{i}{\sigma^2} \text{ind.} w_{ij}, \ldots$$

$$= 1 + \frac{i}{\sigma^2} \text{ind.} (b_{ij} + 2d_{ij}, \ldots -2\frac{\rho}{1+\rho} d, \ldots)$$

where $\frac{i}{\sigma^2} \text{ind.}$ implies selection on an individual basis. The frequency of $A_i A_j$ following selection is

$$p_i p_j w^*_{ij} = p_i p_j + \frac{i}{\sigma^2} \text{ind.} p_i p_j (b_{ij} + 2d_{ij}, \ldots -2\frac{\rho}{1+\rho} d, \ldots)$$

and the genotypic mean of the selected parents may be written as
\[ \mu_s = \sum \sum p_i p_j [1 + (i/\sigma^2)] \text{ind.} (b_{ij} + 2d_{ij}, -2 \frac{\rho}{1+\rho} d_{ij}, \ldots) w_{ij}, \ldots \]

since the selected parents are measured in the original population.

\[ \mu_s = w, \ldots + \frac{i}{\sigma^2} \sum \sum p_i p_j (\alpha_i + \alpha_j + \beta_{ij})^2 \]
\[ = 0 + \frac{i}{\sigma^2} \text{ind.} \left( \sigma_A^2 + \sigma_D^2 \right) \]

which differs from the total genetic variance \( x (i/\sigma^2) \) of the classical truncation selection model (Griffing, 1960).

The array of gametes from the selected parents can be seen to be

\[ 0.5 \sum \sum p_i p_j [1 + (i/\sigma^2)] (b_{ij} + 2d_{ij}, -2 \frac{\rho}{1+\rho} d_{ij}, \ldots) (A_i + A_j) \]
\[ = 0.5 \left( \sum \sum p_i p_j A_i + \sum \sum p_i p_j A_j + \frac{i}{\sigma^2} \left( \sum \sum p_i p_j d_A A_i \right) \right) \]
\[ + \sum \sum p_i p_j A_i + \sum \sum p_i p_j A_i + \sum \sum p_i p_j d_A A_j \]
\[ + \sum \sum p_i p_j d_B A_i + \sum \sum p_i p_j d_B A_j \}
\[ = \sum p_i A_i + \frac{i}{\sigma^2} \left( \sum p_i d_A A_i \right) \]
\[ = \sum p_i [1 + (i/\sigma^2)] A_i \]
\[ = \sum p_i A_i \]
\[ \Rightarrow p'_i = p_i [1 + (i/\sigma^2)] A_i. \]

The mean of the progeny of the selected parents is

\[ \mu'_1 = \sum \sum p_i p'_i p'_j p'_y w_{ij, xy} \]
\[
\pm \sum_i \sum_j \sum_x \sum_y p_i p_j p_x p_y \left[ 1 + \left( i / \sigma^2 \right) \right] \text{ind.} \left( d_i d_j d_x d_y \right) w_{ij, xy}
\]

\[
= 0 + \left( i / \sigma^2 \right) \left( \sum_i p_i' d_i' w_i, \ldots + \sum_j p_j' d_j' w_j, \ldots + \sum_x p_x' d_x' w_x, \ldots, x \right)
\]

\[
+ \sum_y p_y' d_y' w_y, \ldots, y \right)
\]

\[
= \left( i / \sigma^2 \right) \left( 2\sum_i p_i' d_i' w_i + 2\sum_x p_x' d_x' \frac{2}{1+\rho} \gamma_x \right)
\]

We shall define the covariance between additive direct effects and additive associate effects by

\[
d \sigma_A = 2\sum_i p_i' d_i' (\alpha_i) (\gamma_i).
\]

So, we may write

\[
\mu_1 = \left( i / \sigma^2 \right) \left( \sigma_A^2 + \frac{2}{1+\rho} \right) \text{da}_A
\]

The covariance, \( \text{da}_A \), may take a negative sign, and if also it is of sufficient magnitude, then the consequence of selection is in the opposite direction to selection itself. We may note, however, that if \( w_i, \ldots = w \ldots, i \), i.e., if there is symmetry in the mean relative viabilities for individuals as well as associates, then the consequence of selection is in the same direction as selection. We may also note that as the effect of nonimmediate neighbors decreases, i.e., as \( \rho \) becomes small, the covariance contributes more to the next generation mean. In part, we see a direct relationship to the earlier discussion: the equilibrium frequencies were not necessarily those which gave maximum population mean, and coincidence occurred only, in general, if
there were symmetry of selective values. There is also a correspondence with frequency-dependent selection (Wright, 1955).

5.2. Consequences of Group Selection

To overcome the problem of selection resulting in a negative response, we consider the effect of selection on groups of competing plants. We shall consider the case where an individual is mated to a random sample of the original population, and its half-sib offspring sown in a row. In a plant breeding context, this is a common situation: plants which have been open pollinated in a field are selected and the seed from each sown. The maternal parent is known, but the pollen may be considered to be a random sample of pollen from the entire population in which the maternal plant grew. There do remain in practice, however, the nonrandom effects owing to the greater contribution of pollen from neighboring plants than would be expected from random sampling, and the difficulties incurred by differences in flowering time.

Parenthetically, we may note that a litter of some animal species may form a half-sib group, e.g., dogs, while other animal species, particularly some insects, possess marked monogamy.

We shall take $A_A^h g$ as the common parent and mate it at random to the population, and place the progeny in an open row. The progeny have the genotypic array
\[ 0.5A_i \sum p_i A_i + 0.5A_h \sum p_i A_i = 0.5(A_h + A_g) \sum p_i A_i \]

If we take an individual, \( A_i A_i \), from this group, the genotypic value is

\[ w_{hi} = b_{hi} + 2(0.5d_{hi,h} + 0.5d_{hi,g}) - 2\frac{\rho}{1+\rho}(0.25d_{h,h} + 0.25d_{h,g} + 0.25d_{g,h} + 0.25d_{g,g}) \]

which may be expressed in terms of the additive genetic parameters,

\[ w_{hi} = \alpha_h + \alpha_i + [\gamma_h + \gamma_g - 0.5(1+\rho)(\gamma_h + \gamma_h + \gamma_g + \gamma_g)] \]

\[ = \alpha_h + \alpha_i + \frac{1}{1+\rho}(\gamma_h + \gamma_g) \]

The mean of a half-sib row can now be calculated as

\[ \bar{w}_{hg} = \sum_i p_i \frac{1}{2} w_{hi} + \sum_i p_i \frac{1}{2} w_{gi} \]

\[ = \frac{1}{2} \sum_i p_i (w_{hi} + w_{gi}) \]

\[ = \frac{1}{2} \sum_i p_i \left\{ \alpha_h + \alpha_i + \alpha_g + \frac{2}{1+\rho} (\gamma_h + \gamma_g) \right\} \]

\[ = \left( \frac{1}{2} \right) \left\{ \alpha_h + \alpha_i + \frac{2}{1+\rho} (\gamma_h + \gamma_g) \right\} \]

If we select the common parents of half-sib groups on the basis of group performance then the gametic array of the selected parents becomes

\[ \frac{1}{2} \sum p_h p_g d_{hg} (A_h + A_g) = \sum p_h A_h \]

where

\[ d_{hg}^* = 1 + \left( \frac{i}{\sigma^2} \right) \bar{w}_{gh} \]
i.e.,
\[ \frac{1}{2} \sum_{h} p_h p_g \{ 1 + \frac{1}{2} \left( \frac{i}{\sigma^2} \right) \{ d_h + d_g + \frac{2}{1+\rho} a_{h} + a_{g} \} \} (A_h + A_g) \]

\[ = \sum_{h} h^p A_h \]

Expanding this we get

\[ \frac{1}{2} \{ \sum_{h} \sum_{g} p_h p_g A_h + \sum_{h} \sum_{g} p_h p_g A_g + \frac{1}{2} (i/\sigma^2) \{ \sum_{h} \sum_{g} p_h p_g d_h A_h \]

\[ + \sum_{h} \sum_{g} p_h p_g d_g A_g + \sum_{h} \sum_{g} p_h p_g d_h A_h + \sum_{h} \sum_{g} p_h p_g d_g A_g \]

\[ + \frac{2}{1+\rho} \{ \sum_{h} \sum_{g} p_h p_g a_{h} A_h + \sum_{h} \sum_{g} p_h p_g a_{g} A_g + \sum_{h} \sum_{g} p_h p_g a_{h} A_h \]

\[ + \sum_{h} \sum_{g} p_h p_g a_{g} A_g \} \}

\[ = \sum_{h} h^p \{ 1 + \frac{1}{2} (i/\sigma^2) \{ d_h + \frac{2}{1+\rho} a_{h} \} \} A_h \]

i.e.,
\[ p_h^' = p_h \{ 1 + \frac{1}{2} (i/\sigma^2) \{ d_h + \frac{2}{1+\rho} a_{h} \} \}. \]

The mean of the progeny resulting from intermating the selected parents can be written as

\[ \mu_s = \sum_{i} \sum_{j} \sum_{x} \sum_{y} p_{i} p_{j} p_{x} p_{y} w_{i,j,x,y} \]

\[ = \sum_{i} \sum_{j} \sum_{x} \sum_{y} p_{i} p_{j} p_{x} p_{y} \{ 1 + \frac{1}{2} (i/\sigma^2) \{ d_i + d_j + d_x + d_y \}

\[ + \frac{2}{1+\rho} \{ a_i + a_j + a_x + a_y \} \} \} w_{i,j,x,y} \]
This last expression shows that the mean of the progeny of selected parents will always be greater than the original population mean.

5.2.1. Selection of an individual within a half-sib group

We may be somewhat interested in the situation where an individual is randomly selected from a truncation-selected half-sib group. Classical genetic theory shows that the gain from half-sib progeny testing is twice that of half-sib testing (Sprague, 1966).

The genotypic array of a random individual from a half-sib group with a common parent $A_A$ is

$$\sum_{i} p_i A_i \frac{1}{2} (A_A)$$

or

$$\frac{1}{2} \sum_{i} p_i A_i (A_A)$$

The selection value of the group under truncation selection is

$$\{1 + (i/\sigma^2)\frac{\bar{w}}{gh}\}$$

where $\frac{\bar{w}}{gh}$ is defined as before. The genetic array of
selected individuals is

\[ \frac{1}{2} \sum_{i \leq g} \sum_{h \geq h} p_i p_g p_{A_i} (A_g + A_h) \{ 1 + \frac{1}{\sqrt{\sigma^2}} \} \}

\[ = \frac{1}{2} \{ \sum_{i \geq g} \sum_{h \leq h} p_i p_g p_{A_i} (A_g + A_h) \{ 1 + \frac{1}{\sqrt{\sigma^2}} \} \} \}

\[ + \frac{1}{1+\rho} (\gamma_h + \gamma_g) \}

\[ = \frac{1}{2} \sum_{i \geq g} \sum_{h \leq h} p_i p_g \{ 1 + \frac{1}{\sqrt{\sigma^2}} (d_h + 1) \} \}

\[ + \frac{1}{1+\rho} (\gamma_h + \gamma_g) \}

Therefore the gametic array produced by the random individuals from

selected groups is

\[ \frac{1}{4} \sum_{i \leq g} \sum_{h \geq h} p_i p_g \{ 1 + \frac{1}{\sqrt{\sigma^2}} (d_g + \gamma_g) \} \}

\[ \{ A_i + A_g \}

\[ + \frac{1}{4} \sum_{i \leq h} \sum_{h \geq h} p_i p_h \{ 1 + \frac{1}{\sqrt{\sigma^2}} (d_h + \gamma_h) \} \}

\[ \{ A_i + A_h \}

\[ = \frac{1}{4} \sum_{i \leq g} \sum_{h \geq h} p_i p_g \{ 1 + \frac{1}{\sqrt{\sigma^2}} (d_g + \gamma_g) \} \}

\[ \{ A_i \}

\[ + \frac{1}{4} \sum_{i \leq g} \sum_{h \geq h} p_i p_g \{ 1 + \frac{1}{\sqrt{\sigma^2}} (d_g + \gamma_g) \} \}

\[ \{ A_i \}

\[ + \frac{1}{4} \sum_{i \leq h} \sum_{h \geq h} p_i p_h \{ 1 + \frac{1}{\sqrt{\sigma^2}} (d_h + \gamma_h) \} \]

\[ \{ A_i \}

\[ + \frac{1}{4} \sum_{i \leq h} \sum_{h \geq h} p_i p_h \{ 1 + \frac{1}{\sqrt{\sigma^2}} (d_h + \gamma_h) \} \]

\[ \{ A_i \}
The next generation arising from mating random individuals from
selected half-sib groups has a mean

\[ \mu_s = \sum_{i} \sum_{j} \sum_{x} p_i' p_j' x' w_{ij,xy} \]

where

\[ p_i' = p_i \]
\[ p_j' = p_j \left[ 1 + \frac{1}{2} (i/\sigma^2) (\alpha_j + \frac{2}{1+\rho} a_j) \right] \]
\[ p_x' = p_x \]
\[ p_y' = p_y \left[ 1 + \frac{1}{2} (i/\sigma^2) (\alpha_y + \frac{2}{1+\rho} a_y) \right]. \]

Hence the mean of the next generation, as a deviation from the
original population mean, is

\[ \mu_s = \sum_{i} \sum_{j} \sum_{x} p_i' p_j' x' \left( 1 + \frac{1}{2} (i/\sigma^2) (\alpha_j + \frac{2}{1+\rho} a_j) \right) \times \]
\[ \left[ 1 + \frac{1}{2} (i/\sigma^2) (\alpha_y + \frac{2}{1+\rho} a_y) \right] w_{ij,xy} \]
\[ = \sum_{i} \sum_{j} \sum_{x} p_i' p_j' x' p_y' \left( 1 + \frac{1}{2} (i/\sigma^2) (\alpha_j + \frac{2}{1+\rho} a_j) \right) \times \]
\[ \left( 1 + \frac{1}{2} (i/\sigma^2) (\alpha_y + \frac{2}{1+\rho} a_y) \right) w_{ij,xy} \]
\[ \mu_1 = \frac{1}{2} (i/\sigma^2) \left( \frac{1}{2} \sigma_A^2 + \frac{2}{1+\rho} \frac{1}{2} \sigma_A^2 + \frac{2}{1+\rho} \right) \]

which ensures that selection results in a positive response. We may express this mean as

\[ \mu_1 = \frac{1}{2} (i/\sigma^2) \left( \frac{1}{2} \sigma_A^2 + \frac{2}{1+\rho} \sigma_A^2 + \frac{2}{1+\rho} \right) \]

which is half that of the gain due to half-sib progeny test selection.

It should be noted, however, that in a practical plant breeding program the extra gain due to the half-sib progeny test selection is more than offset by the increased generation interval compared to half-sib selection.

5.2.2. Truncation selection in full-sib groups

In this case we may consider the case of a number of pairs of parents selected at random from the original population and mated to produce a series of full-sib groups which are planted in rows.

Selection is based on the performance of a full-sib group, and since the gene array of the mating parents and their respective full-sib groups are the same, it does not matter whether we select as a parent for
the next generation the parent of a successful full-sib group or an individual randomly selected from it (cf. Griffing, 1976a).

The mating $A_iA_j \times A_kA_l$ produces the offspring array

$$\frac{1}{4}(A_iA_k + A_iA_l + A_jA_k + A_jA_l).$$

The mean of this full-sib group is

$$(1/16)(w_{ik,ik} + w_{ik,il} + w_{ik,jk} + w_{ik,jl} + \ldots + w_{jl,jl})$$

Rewriting this in terms of additive genetic parameters, the mean of the full-sib group may be written as

$$\overline{w}_{ijkl} = (1/16)\{8(\alpha_i A_i + \alpha_j A_j + \alpha_k A_k + \alpha_l A_l)$$

$$+ 8 \frac{2}{1+\rho} (\gamma_i A_i + \gamma_j A_j + \gamma_k A_k + \gamma_l A_l)$$

If we select the entire group on the basis of group performance then we have the following gametic array.

$$\frac{1}{4}(A_i + A_j + A_k + A_l).$$

Therefore over all selected full-sib groups, the gametic array will be
The mean of the population arising from intermating the selected full-sib groups may be calculated as

\[ \sum \sum \sum p_i p_j p_k p_l \frac{1}{4}(A_i + A_j + A_k + A_l) \left[ 1 + \bar{w}_{ij} \bar{w}_{kl} \frac{1}{\sigma^2} \right] \]

\[ = \frac{1}{4} \sum \sum \sum p_i p_j p_k p_l (A_i + A_j + A_k + A_l) \]

\[ + \frac{(1/\sigma^2)}{16} \sum i p_i \alpha_i A_i + \frac{(4)(8)}{16} \cdot \frac{2}{1+\rho} \sum i p_i \gamma_i A_i \}

\[ = \sum p_i A_i + \frac{(1/\sigma^2)}{2} \sum i p_i \alpha_i A_i + \frac{2}{1+\rho} \sum p_i \gamma_i A_i \]

\[ = \sum p_i (1 + \frac{1}{2}(1/\sigma^2) (\alpha_i + \frac{2}{1+\rho} \gamma_i) A_i) \]

The penultimate expression demonstrates that the consequence of selection will always be positive. We may note that the response
is the same as that for the half-sib progeny test, i.e., where the common parent of a half-sib group is selected on the basis of its half-sib offspring performance. This is analogous to classical selection theory (Sprague, 1966).

5.2.3. Clonal selection

The technique of assessing plants by their performance as competing ramets of the same clone was suggested as long ago as 1931. Jenkin (1931) outlined a grass plant assessment technique which consisted of dividing plants into ramets and close planting these ramets in beds that approximated sward conditions. This aspect of assessment under competitive stress was largely ignored or considered unnecessary by plant breeders until about the mid-sixties when studies such as those by Lazenby and Rogers (1964) questioned the previously held contention that assessing forage plants at wide spacings (typically circa 60 cm) identified genotypes that would be superior performers in the closer spacings encountered by plants in swards or crops.

In our model we shall take the mean value of the clonal group of the genotype $A_iA_j$ to be $w_{ij,ij}$, which in genetic parameters may be written as

$$d_i^a + d_j^a + \frac{2}{1+\rho}(a_i^y + a_j^y)$$

For the convenience of a tractable model, we shall again ignore the interaction terms. The probability that $A_iA_j$ survives selection is, therefore, $1 + (i/\sigma^2)w_{ij,ij}$. The gametic array following selection is
\[
\begin{align*}
\frac{1}{2} \sum_{i,j} p_i p_j (A_i + A_j) \{1 + (i \sigma^2) w_{ij}, i_j \} \\
= \frac{1}{2} \sum_{i,j} p_i p_j (A_i + A_j) + \frac{1}{2} (i \sigma^2) \sum_{i,j} p_i p_j (A_i + A_j) w_{ij}, i_j \\
= \sum_{i} p_i A_i + (i \sigma^2) \left\{ \frac{1}{2} \sum_{i,j} p_i p_j (d_i + d_j) \\
+ \frac{2}{1+\rho} (a_i + a_j) A_i \right\} \\
+ \frac{1}{2} \sum_{i,j} p_i p_j (d_i + d_j + \frac{2}{1+\rho} (a_i + a_j) A_i) \\
= \sum_{i} p_i A_i + (i \sigma^2) \left\{ \sum_{i} p_i (d_i + \frac{2}{1+\rho} a_i) A_i \right\} \\
= \sum_{i} p_i (1 + (i \sigma^2) (d_i + \frac{2}{1+\rho} a_i)) A_i .
\end{align*}
\]

The mean of the next population formed by intermating the selected genotypes is, therefore,

\[
\mu_1 = \sum_{i,j} \sum_{x,y} p_i^p_j^p r_{ij,xy} w_{ij,xy} \\
= \sum_{i,j} \sum_{x,y} p_i p_j p_x p_y \{1 + (i \sigma^2) (d_i + d_j + d_x + d_y) \\
+ \frac{2}{1+\rho} (a_i + a_j + a_x + a_y) \} \{d_i + d_j \} \\
+ \frac{2}{1+\rho} (a_x + a_y) \}
\]
\[ f(W, \{\alpha_i + z_p \beta_j\}) \]

\[ = 0 + \left(\frac{1}{\sigma^2}\right) \left\{ \sum_i p_i d_{\alpha_i}^2 + \sum_j p_j d_{\beta_j}^2 + \frac{2}{1+\rho} \sum_x p_x d_{\alpha_x} a^y_x \right. \]

\[ + \frac{2}{1+\rho} \sum_y p_y d_{\beta_y} a^y_a \gamma_y + \frac{2}{1+\rho} \sum_i p_i d_{\alpha_i} a_i \gamma_i + \frac{2}{1+\rho} \sum_j p_j a_j \gamma_j d_{\beta_j} \]

\[ + \left(\frac{2}{1+\rho}\right)^2 \sum_x p_x d_{\alpha_x} a^2_x + \left(\frac{2}{1+\rho}\right)^2 \sum_y p_y a^2_y \right\} \]

\[ = \left(\frac{1}{\sigma^2}\right) \left\{ 2\sum_i p_i d_{\alpha_i}^2 + 4\frac{2}{1+\rho} \sum_i p_i d_{\alpha_i} a_i \gamma_i + 2\left(\frac{2}{1+\rho}\right)^2 \sum_i p_i \gamma_i^2 \right\} \]

\[ = \left(\frac{1}{\sigma^2}\right) \left\{ \sigma^2_A + 2\frac{2}{1+\rho} \sigma_A + \left(\frac{2}{1+\rho}\right)^2 \sigma_A^2 \right\}, \]

which may be seen to be positive, (i.e., the gain from selection will be positive) by the re-expression

\[ 2\left(\frac{1}{\sigma^2}\right) \sum_i p_i \left( d_{\alpha_i}^2 + \frac{2}{1+\rho} a_i \gamma_i \right)^2. \]

It should be remembered that the term in the denominator, \( \sigma^2 \), differs from one selection method to another, and the above result does not imply, per se, that the response is the same as that from full-sib selection. For example, the different phenotypic variances for different modes of selection (full-sib, half-sib, etc.) are given in Sprague and Eberhart (1977) and Empig et al. (1972).
6. COMPETITION IN SQUARE PLANTINGS

6.1. Unrestricted Square Plantings

In this case, the plants are spaced a distance $a$ apart in a square formation. Each plant will have four immediate neighbors and four neighbors $a\sqrt{2}$ away. Each of these neighbors are similarly surrounded by eight plants.

![Figure 6.1. Interactions between plants in a square disposition](image)

Whereas in the row, the effect of a plant is felt only through its lineal neighbors, in a square planting the effect of a plant, say A in the figure, on X, can occur simultaneously through a large number of paths. Insofar as the algebra that was exploited satisfactorily with the lineal case becomes rapidly intractable here, we have to resort to a different artifice in developing a model.

We notice that there are four plants $a$ from the central plant, four $a\sqrt{2}$, four $2a$, eight $a\sqrt{5}$, and so on. The areas around each plant may be divided into concentric zones, see Figure 6.2.

This, to a point, follows Mather (1969) except that Mather had
only one plant per zone - an arrangement applicable to a nonregular planting arrangement. The competitive effects on \( X_i \) of the individuals in zone 1 is \( k_1 \), in zone 2; \( k_2 \), and so on, with the rider that the interactions between plants in a zone, or between plants in different zones, that affect the competitive effect on \( X_i \) are allowed for by the definition of the \( k \)'s.

We assume generally that \( k_1 > k_2 > k_3 \), etc., so the reproductive capacity of \( X_i \) is

\[
\omega_i = b_i + 4\sum p_j k_{1,i} + 4\sum p_j k_{2,i} + 4\sum p_j k_{3,i} + \ldots
\]

\[
= b_i + \sum p_j (4k_{1,i} + 4k_{2,i} + 4k_{3,i} + \ldots)
\]

\[
= b_i + \sum p_j k_{-i,j}.
\]  

(6.1.1)

In this manner we have "defined out" the contribution due to interaction, i.e., as interaction changes between plants not \( X_i \), then
the k's change, rather than some other parameter. The model reduces to a pairwise consideration which has been covered previously.

6.2. Immediate Neighbor Model

Alternatively, we may approach this problem by considering the influence of only the eight immediate neighboring individuals surrounding the central individual.

6.2.1. Haploid model

In the two-allele haploid case, if the frequencies of alleles A and a are \( P_A \) and \( P_a \), respectively, \( (P_A + P_a = 1) \) then the probabilities that, of the eight individuals surrounding any central individual, none are genotype A \( (P_0) \), only one is genotype A \( (P_1) \), ..., all eight are genotype A \( (P_8) \), are given by the binomial probability function;

\[
P_r = \frac{8!}{r!(8-r)!} P_A^r P_a^{8-r}.
\]

It may be quickly seen, however, that between groups possessing the same numbers of genotypes, the arrangements of these genotypes may differ. Thus the competitive effect of one set of, say, four A genotypes and four a genotypes on the central member may differ from another set if there are interactions between neighbors (Figure 6.3).

The genotype A may be surrounded by eight a genotypes, or seven a and one A genotype, ..., or eight A genotypes, i.e., nine groups. Genotype a may be associated with a similar set of nine groups. Each group may be constructed in a number of arrangements. The total number
Figure 6.3. Diagram showing a genotype surrounded by four A genotypes and four a genotypes. The mutual antagonism between like genotypes may result in the total competitive effect on the central genotype in (a) differing from that on the central genotype in (b)

of arrangements can be found, by tedious enumeration, to be 100. Thus one could be led to specify 100 different reproductive capacities, a number that would clearly make it difficult to determine an underlying trend in the model due to changes in the reproductive capacities. We shall, therefore, introduce a simplifying concept, similar to those already employed, so that the influence of intra- and intergenotypic interactions may be more easily understood.

We shall first assume that the influence of a neighbor is the same regardless of whether it is placed at the corner, or along the side, of a square. Effectively, this means that we may consider the neighbors lying on the points of a regular octagon about the central individual. The problem of the ways in which the two genotypes may form arrangements in the octagon reduces to what has become known as "Whitworth's bracelet problem" (David and Barton, 1962, p. 94).

The type of interaction model used, which is explained later,
makes use of this approach because the total competitive effect of surrounding genotypes on the central genotype depends on the number of like genotypes placed contiguously, i.e., in a cluster (David and Barton (1962) use the term "group" in lieu of "cluster"). In other words, a cluster may be defined as all those individuals of one genotype lying between consecutive individuals of another genotype.

In the following discussion, we shall, for simplicity, revert to the classical probabilistic concepts of balls of various colors, rather than individuals of various genotypes.

6.2.2. Derivation of probability of t clusters of white balls

(1) Let there be $r_1$ white balls and $r_2$ red balls; $r_1 + r_2 = r$. These are arranged in a ring. We want to find the probability that there are $t$ clusters of white individuals in such a ring.

(2) Each cluster of white balls is bounded on each side by a red ball. Thus if, to fix ideas, we let the first position be occupied by a white ball, which is the first element of its cluster, then this cluster is followed by a cluster of red balls. The same holds for any cluster of white balls. Thus, if there are $t$ clusters of white balls there must be $t$ clusters of red balls.

(3) Let the positions in the ring be numbered in a clockwise fashion. Since a particular configuration of clusters is not changed if each element is moved one unit clockwise, we may, without loss of generality, assume the $r$-th position to be occupied by a red ball.
The remaining $r-1$ positions then contain $r_1$ white balls and there are \( \binom{r-1}{r_1} \) ways to choose $r_1$ positions among $r-1$ to be occupied by the white balls.

(4) The red ball in the $r$-th position may be followed by either a white ball or a red ball in the first position. Any of the other $r_2-1$ red balls may be followed by a ball of either color. Thus each of the $r_2$ red balls may be the last ball of a cluster, and, if there are $t$ clusters, $t$ of the red balls are the last balls in the cluster. There are \( \binom{r_2}{t} \) ways to choose $t$ balls among $r_2$ balls as those which end the $t$ clusters.

(5) Because the $r$-th ball in the ring is a red ball, the $r_1$-th white ball is necessarily the last ball in the $t$-th cluster of white balls. Among the remaining $r_1-1$ white balls $t-1$ are the last balls of the remaining $t-1$ clusters. There are \( \binom{r_1-1}{t-1} \) ways to choose $t-1$ white balls among $r_1-1$ white balls as those which end $t-1$ clusters.

(6) For each of the \( \binom{r_2}{t} \) ways to choose $t$ red balls to end clusters there are \( \binom{r_1-1}{t-1} \) ways to choose $t-1$ white balls to end the clusters of white balls.

(7) From (3) and (6) we can say that

\[
P(t \text{ clusters of white balls in a ring } | \; r_1 \text{ white balls, } r_2 \text{ red balls}) = \frac{\binom{r_1-1}{t-1} \binom{r_2}{t}}{\binom{r-1}{r_1}} = \frac{r_2!}{t!(r_2-t)!} = \frac{(r-1)!}{r_1!(r_2-1)!}
\]
The first expression is that given in David and Barton (1962), while the last appears to be that used in their paper (Barton and David, 1958). The multinomial term is the denominator, \( \binom{r}{r_2} \), and this is used in the calculation of Table 6.1.

Table 6.1. Ways in which clusters of 1, 2, 3 or 4 A genotypes may surround a central genotype. The conditional p.d.f. is obtained from dividing the number in the third, fourth, fifth or sixth column by the number in the second column.

<table>
<thead>
<tr>
<th>Number of A genotypes</th>
<th>( \binom{r}{r_1} )</th>
<th>Number of clusters of A genotypes, ( t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>28</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>56</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>70</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>56</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>28</td>
<td>8</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

We now turn to the question of the competitive interaction associated with these groups of varying clusters. The direct competitive effects of the neighboring genotypes is given in Table 6.2.
Table 6.2. Direct competitive effects on a central individual due to each individual of the surrounding neighbors

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Neighboring genotype</th>
<th>$A$</th>
<th>$a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-gc</td>
<td>s</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>-s</td>
<td>-c</td>
<td></td>
</tr>
</tbody>
</table>

Indirect effects are also incorporated into this model. The surrounding individuals effect each other as well as the central genotype, and this mutual interaction affects their competitive effect on the central genotype. We shall confine this influence only to the immediate neighbors of the individuals forming the surrounding neighbors.

![Diagram](image)

Figure 6.4. Influence of neighboring individuals directly and indirectly on group central members, $A_c$

In Figure 6.4(a), $A_c$ gains $s$ due to its competitive superiority over $a_1$. However, $a_1$ incurs a loss of $-c$ from $a_2$, and a loss of $-s$ from $A_1$. We assume that a constant proportion, $\rho$, of each of these effects influence the competitive effect between $a_1$ and $A_c$. Thus the net gain by $A_c$ due to direct effects of $a_1$ and the indirect
effects through $a_1$ of $a_2$ and $A_c$ is $s + sρc + sps = s(1+ρ[c+s])$.

In the second diagram (Figure 6.4(b)) $A_c$ suffers $-gc$ in direct competition with $A_1$, but gains through $A_1$ gcρgc due to the antagonism suffered by $A_1$ in competition with $A_2$. $A_1$, however, gains $s$ in its competition with $a_1$, and this makes $A_1$ more competitive against $A_c$. Hence the net gain to $A_c$ from $A_1$ directly, and $a_1$ and $A_2$ indirectly is $-gc + gcρgc - gcsp = -gc[1-ρ(gc-s)]$.

We now consider the net competitive effect of different groups of $A$ and $a$ on a central $A$ genotype.

(1) Surrounding genotypes consist of 8 $A$ genotypes (Figure 6.5).

![Figure 6.5. Arrangement of 8 A genotypes surrounding a central A genotype](image)

The net effect on $A$ is $8[-gc(1-2ρgc)] = -8gc(1-2ρgc)$.

(2) 7 $A$ genotypes and 1 $a$ genotype (Figure 6.6).

![Figure 6.6. Arrangement of 7 A genotypes and 1 a genotype surrounding a central A genotype](image)
The net effect on A is \(5[-gc(l-2pqc)] \) (considering \(A_1\) to \(A_5\))
\[\quad +2[-gc(l-pqc+ps)] \) (considering \(A_6\) and \(A_7\))
\[\quad + [s(1+2ps)] \) (considering a)
\[= -7gc + 12gcqc - 2gcps + s + 2sps. \]

(3) 6 A genotypes and 2 a genotypes (Figure 6.7(a), (b), (c) and (d)).

Figure 6.7. Arrangements of 6 A genotypes and 2 a genotypes surrounding a central A genotype

In the first figure (Figure 6.7(a)) there are two clusters, while in the other three figures there are four clusters in each. The competitive effect in the first figure may be calculated thus: The net effect of competition on A is

\[4[-gc(l-2pqc)] \) (considering \(A_1\) to \(A_4\))
\[\quad + 2[-gc(l-pqc+ps)] \) (considering \(A_5\), \(A_6\))
\[\quad + 2[s(1+ps-pc)] \) (considering a, a)
\[= -6gc + 2s + 10gcqc - 2gcps + 2sps - 2pspc. \]

The net effect of competition on A in arrangements b, c and d are the same, viz,
-6gc + 2s + 8gcpgc - 4gcps + 4sps.

We need only, therefore, determine the conditional probability of the number of clusters occurring in a group of $r_A$ and $(8-r_A)$ a genotypes, which we have in Table 6.1.

Similar considerations lead to Table 6.3.

Table 6.3. Reproductive capacities, $S_{i; j}$, for genotype $i$ surrounded by $j A$ genotypes, $i = A, a; j = 0, 1, \ldots, 8$

<table>
<thead>
<tr>
<th>$S_{A;0}$</th>
<th>$= 1 + 8s + 16spc$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{A;1}$</td>
<td>$= 1 + 7s - gc + 12spc + 2sps - 2gcps$</td>
</tr>
<tr>
<td>$S_{A;2}$</td>
<td>$= 1 + 6s - 2gc + (60/7)spc + (24/7)sps + (4/7)gcpgc - (24/7)gcps$</td>
</tr>
<tr>
<td>$S_{A;3}$</td>
<td>$= 1 + 5s - 3gc + (40/7)spc + (30/7)sps + (12/7)gcpgc - (30/7)gcps$</td>
</tr>
<tr>
<td>$S_{A;4}$</td>
<td>$= 1 + 4s - 4gc + (24/7)spc + (32/7)sps + (24/7)gcpgc - (32/7)gcps$</td>
</tr>
<tr>
<td>$S_{A;5}$</td>
<td>$= 1 + 3s - 5gc + (12/7)spc + (30/7)sps + (40/7)gcpgc - (30/7)gcps$</td>
</tr>
<tr>
<td>$S_{A;6}$</td>
<td>$= 1 + 2s - 6gc + (4/7)spc + (24/7)sps + (60/7)gcpgc - (24/7)gcps$</td>
</tr>
<tr>
<td>$S_{A;7}$</td>
<td>$= 1 + s - 7gc + 2sps + 12gcpgc - 2gcps$</td>
</tr>
<tr>
<td>$S_{A;8}$</td>
<td>$= 1 - 8gc + 16gcpgc$</td>
</tr>
<tr>
<td>$A_{a;0}$</td>
<td>$= 1 - 8c + 16cpc$</td>
</tr>
<tr>
<td>$S_{a;1}$</td>
<td>$= 1 - s - 7c + 2cpc - 2sps + 12cpc$</td>
</tr>
<tr>
<td>$S_{a;2}$</td>
<td>$= 1 - 2s - 6c + (24/7)cpc - (24/7)sps + (60/7)cpc + (4/7)spgc$</td>
</tr>
<tr>
<td>$S_{a;3}$</td>
<td>$= 1 - 3s - 5c + (30/7)cpc - (30/7)sps + (40/7)cpc + (12/7)spgc$</td>
</tr>
<tr>
<td>$S_{a;4}$</td>
<td>$= 1 - 4s - 4c + (32/7)cpc - (32/7)sps + (24/7)cpc + (24/7)spgc$</td>
</tr>
<tr>
<td>$S_{a;5}$</td>
<td>$= 1 - 5s - 3c + (30/7)cpc - (30/7)cps + (12/7)cpc + (40/7)spgc$</td>
</tr>
<tr>
<td>$S_{a;6}$</td>
<td>$= 1 - 6s - 2c + (24/7)cpc - (24/7)sps + (4/7)cpc + (60/7)spgc$</td>
</tr>
<tr>
<td>$S_{a;7}$</td>
<td>$= 1 - 7s - c + 2cpc - 2sps + 12spgc$</td>
</tr>
<tr>
<td>$S_{a;8}$</td>
<td>$= 1 - 8s + 16spgc$</td>
</tr>
</tbody>
</table>
We shall assume, further, that $ps$ is some parameter $p$ times $s$, so $sp^2s$ may be written $s^2p$. Similar expressions exist for the other values. This means that we have now reduced the possible 100 different reproductive capacities to 18, each of which may be expressed in terms of just four parameters, $s$, $g$, $c$, and $p$. Further, these four parameters are reasonably comprehensible in terms of the competitive interactions.

It is now straightforward to express the genotypic frequency after one cycle of reproduction as

$$P'_A = \frac{\sum_{i=0}^{8-1} \frac{8!}{i!(8-i)!} P_A^i \, p^{8-i} (P_A S_A; i)}{\sum_{i=0}^{8-1} \frac{8!}{i!(8-i)!} P_A^i \, a^{8-i} (P_A S_A; i + P_A S_a; i)}.$$  \hspace{1cm} (6.2.1)

While this does not lead to any analytic results, we are able to investigate numerically the consequences of varying the parameters. The following graphs (Figure 6.8) show the gene frequency over a number of generations for populations with different values for the competition parameters, $s$, $c$, $g$ and $p$. Restrictions on the parameters are necessary, as with the other models, to avoid negative values of the relative reproductive capacities, or relative viabilities. Because no analytical results are available, we shall discuss the implications of the graphs more than we have with other models.

The first graph (Figure 6.8a) shows that for the values of $s$, $c$, and $g$ equal to .01, .1 and 2 respectively, the value of $p$ has only a slight influence on either the rate of change of gene frequency or the equilibrium value. Comparisons between Figure 6.8a and Figure 6.8b show that either increasing $s$ (from .01 to .1) or reducing $c$ (from .1 to
.01) increases the equilibrium value quite markedly. However, with both $s$ and $c$ small (.01) the change in gene frequency is considerably reduced.

In Figure 6.8b the value of $p$ had an imperceptible effect on the gene frequency when $s$ and $c$ are small (0.01). This is expected, since the products, $scp$, $s^2p$ and $c^2p$ must be very small. In the same figure the curve for $s$, $c$ both equal to 0.1 shows a damped oscillatory effect.

Oscillations have not been the rule with the numerical studies of the models in this investigation, and here we find that in Table 6.3 some values of $S_{i;j}$ are negative. If we restrict ourselves to values that ensure positive $S_{i;j}$ values, the oscillation vanishes. On the other hand, negative $S_{i;j}$ values do not imply oscillations. In Figure 6.8c, where $c = 0.1$, $g = 2$, some $S_{i;j}$ are negative, but the curve designated "x-x-x" indicates no oscillatory tendency.

The final figure, Figure 6.8c, shows the effect of reducing the difference between the effect of mutual antagonism between like genotypes. The effect between $a$ and $a$ is 0.05, i.e., $c$, and that between $A$ and $A$ is 0.075 ("---") or 0.1 ("+++"), i.e., $gc$. As the difference decreases, both the equilibrium and the initial rate of change of gene frequency is raised. This, in effect, is what we had with the haploid variable fitness model (cf. Figure 3.8).

6.2.3. Diploid model

In turning to develop a diploid version of this two-dimensional model, we are faced with considerable complexity. Basically, the situation may be illustrated by considering a ring formed of beads: $n_1$ black, $n_2$ white and $n_3$ red. We need to know the probabilities of the various
Figure 6.8a. Graph of gene frequency for the two-dimensional model (Expression 6.2.1). (In this figure \( s = 0.01, c = 0.1, g = 2 \) for all curves. +++: \( \rho = 0.1 \); ---: \( \rho = 0.3 \); x-x-x: \( \rho = 0.7 \))
Figure 6.8b. Graph of gene frequency for the two-dimensional model (Expression 6.2.1).
(In this figure ρ = 0.3, g = 2 for both curves. +++: s = 0.1, c = 0.1;
x-x-x: s = 0.01, c = 0.01)
Figure 6.8c. Graph of gene frequency for the two-dimensional model (Expression 6.2.1).
(In this figure $s = 0.01$, $\rho = 0.3$ for all curves. $\ldots$: $c = 0.05$, $g = 1.5$; $+++$: $c = 0.05$, $g = 2$; $x-x-x$: $c = 0.1$, $g = 2$)
unique arrangements of these colors. David and Barton (1962) comment on this problem, and while tables are given for the number of clusters formed of each color (Barton and David, 1958), no solution to this particular problem is given. Nor is it immediately evident (H. A. David, Department of Statistics, Iowa State University, Ames, Iowa, personal communication, 1978).

While this may appear to preclude further comment on this model, we may make the observation that for s, c, and \( \rho \) small, or even only \( \rho \) very small, the indirect effects of neighbors of neighbors becomes negligible and a simple pairwise model (e.g., Expressions (3.6.6), (3.6.7)) would be a satisfactory description of this case, providing the values of s and c are adjusted to accommodate that each genotype in being influenced, in an additive manner, by eight neighbors rather than one.
7. CONCLUSION

7.1. Population Genetics Models of Competition

7.1.1. Constant fitness models

In the models we have developed, we were able to produce results that showed Nei's (1971) treatment of the subject of genotypes in competition to be incomplete. The major thrust of our model building, developing models incorporating relatively few parameters, enables the consequences of competitive effects to be quantified relatively easily. Counter to this simplification, however, remains the argument that this simplification may render a model unrealistic. In particular, the constant fitness models may be considered to be unrealistic because, regardless of the genotypes composing a group, there is exactly one survivor. It could be that an adaptation of Wallace's (1970) concept of "soft" selection may be more realistic, if also more complicated.

Leaving these criticisms aside, we are able to show that if we allow intragenotypic interactions stable polymorphisms may exist - a feature not apparent with Nei's (1971) interpretation. No stable limit cycles resulted if we restricted the competitive effects to reasonable values, i.e., values that did not result in negative probabilities of survival or negative viabilities. Figure 3.7b, however, does show a damped oscillatory response. This may suggest that natural systems may not necessarily possess stable gene frequencies.

An interesting feature of the fixed fitness models is that the equilibrium value is independent of group size. We are then left with the result that the rate of change in gene frequency increases with
group size. Heuristically, this is expected: if we increase the group size, but still only one member of a group survives, we are effectively increasing the selection intensity. That the rate of change in gene frequency increases with group size is a feature of all the group-wise models developed in this study.

7.1.2. **Variable fitness models**

The variable fitness models avoid the criticism levelled at the constant fitness models by allowing the probability of one survivor from a group to be a function of the genotypes of the group. Basically, similar results arise from this model as from the constant fitness model, except that changes in group size causes a change in the equilibrium gene frequency. We have shown, however, that the introduction of intragenotypic interactions into the model of competition results in the existence of stable nontrivial equilibria. The basically similar two-dimensional interaction models possess the same characteristics.

7.1.3. **Quantitative genetics models**

We have discussed Donald's (1968) concept of an ideotype plant. While the physical characteristics of such a plant (or an ideotype animal) may be written down, the plant under competitive effects encounters interactions that may not be physically observable. Hence, in any plant breeding program some screening technique must be developed for assessing the plants under the conditions similar to those the bred variety, or selected genotype, will encounter. Griffing's (1977) approach gives a strong suggestion as how this may be done, but the problem of deciding
what sort of groups, and how many members in each group, remains somewhat insoluble. We have shown in this study, however, that the concept of indefinitely long rows, a typical crop situation, serves well to outline a workable plant breeding screening method. The conclusions regarding selecting individuals from closely related groups, or selecting individuals at random from groups that perform well on the whole, can be equally applied to individuals within rows.

In this context, it is interesting to note that a widely used corn breeding technique is to grow the progeny of F2 plants in rows and to select on the basis of the average row performance. Seed selected (perhaps randomly) from selfed plants in selected rows are used to raise F3 plants, and so on, until a breeder may have a series of rows of plants representing, say, F8 progeny, but each row identifiable with an original F2. This procedure effectively apes the selection procedure outlined in Chapter 5 where selection between rows in half-sibs, full-sibs, and homozygous individuals was shown theoretically to be an effective measure. The well-known quote from Molière's "Le Bourgeois Gentilhomme" seems appropriate, "Faire de la prose sans le savoir".
BIBLIOGRAPHY


Griffing, B. 1976b. Selection in reference to biological groups VI. Use of extreme forms of nonrandom groups to increase selection efficiency. Genetics 82:723-731.


Sakai, K. 1957. Studies on competition in plants. VII. Effect on competition of varying number of competing and noncompeting individuals. J. Genet. 55:227-234.


9. ACKNOWLEDGMENTS

I record with pleasure and deep appreciation the assistance and encouragement given to me by Dr. Ed Pollak. I thank also Dr. Oscar Kempthorne for his invitation to study at Iowa State University and the financial assistance from the National Institute of Health. Further financial assistance came from the Department of Scientific and Industrial Research, New Zealand, for which I am grateful.

To my wife, Marion, and children, Richmond and Peter, I offer not only heartfelt thanks, but also profuse apologies. I would also like to record my appreciation to my parents whose verbal and material encouragement made the completion of this study possible.

Finally, I am indebted to Mrs. Pat Gunnells, who accurately and promptly typed this thesis.
Lewis (1978) presented a succinct treatment of stability criteria for the case where the frequencies of the components, at time $t+1$, of a multispecies community are described by the $N$ recurrence equations

$$p_i(t+1) = F_i[p_1(t), p_2(t), \ldots, p_N(t)], \quad i = 1, \ldots, N.$$ 

In Section 2.4 we were concerned with the frequencies of genotypes in a population. Cockerham and Burrows (1971) showed that the stability of an equilibrium point, defined by

$$\hat{p}_i(t+1) = F_i[\hat{p}_i(t), \ldots, \hat{p}_r(t)] = \hat{p}_i(t), \quad i = 1, \ldots, r,$$

may be investigated by considering the following.

If the recurrence equations are expressed as

$$\hat{p}(t+1) = F[\hat{p}(t)],$$

then let the $r \times r$ matrix $M = \{m_{ij}\}$, where

$$m_{ij} = \left[ \frac{\partial F_i[\hat{p}]}{\partial \hat{p}_j} \right].$$

Because there is a row dependency ($\hat{p} \cdot \hat{p} = 1$) we may form the $(r-1) \times (r-1)$ matrix $N = \{n_{ij}\}$, where

$$n_{ij} = m_{ij} - m_{ir}.$$

Cockerham and Burrows (1971) claimed that there exists a matrix, $C$,
such that $\text{CNC}^{-1} = \Lambda$, where $\Lambda$ is a diagonal matrix containing the eigenvalues of $N$ along the diagonal. If the maximum spectral radius of the eigenvalues is greater than one, the equilibrium is unstable. Section 2.4 details the development and the rationale for this, and therefore it is not repeated here. Lewis (1978) gives the correct, more general development.

There exists a matrix $C$ with the property that $\text{CNC}^{-1}$ is the Jordan canonical form of $N$ (Bellman, 1960, p. 191).

$$
\Lambda = \text{CNC}^{-1} = \begin{bmatrix}
I_{k_1}(\lambda_1) & & \\
& I_{k_2}(\lambda_2) & \\
& & \ddots & \\
& & & I_{k_t}(\lambda_t)
\end{bmatrix},
$$

where $\sum_{j=1}^{t} k_j = r-1$, the $\lambda$'s are the eigenvalues (not necessarily distinct) and

$$
I_k(\lambda) =
\begin{bmatrix}
\lambda & 1 & 0 & \ldots & 0 & 0 \\
0 & \lambda & 1 & \ldots & 0 & 0 \\
0 & 0 & \lambda & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & \lambda & 1 \\
0 & 0 & 0 & \ldots & 0 & \lambda
\end{bmatrix}
$$

a $k \times k$ matrix.
Exactly as we had for Expressions (2.4.2), we may let

\[ g_u = C \hat{\delta}^*, \]

\[ \hat{\delta}^* = C^{-1} g_u \]

\[ g_u = \Lambda^u g_0. \]

Clearly, when \( \Lambda \) is diagonal (which occurs when the eigenvectors are linearly independent), the equilibrium is locally stable or unstable according to whether \( \max |\lambda_i| \) is less than or greater than one, respectively. When \( \max |\lambda_i| = 1 \), the linear approximation does not converge or diverge. The \( F_i \)'s are then determined by the higher order terms.

When \( \Lambda \) is not diagonal, the sufficient conditions are still \( \max |\lambda| < 1 \) for stability and \( \max |\lambda_i| > 1 \) for instability. To see this, consider \( I^n_k(\lambda) \). It is easy to verify using Pascal's Triangle

\[
\binom{n}{i} + \binom{n}{i+1} = \binom{n+1}{i+1}
\]

that

\[
I^n_k(\lambda) = \begin{bmatrix}
\lambda^n & (\binom{n}{1})\lambda^{n-1} & (\binom{n}{2})\lambda^{n-2} & \cdots & (\binom{n}{k-1})\lambda^{n-k+1} \\
0 & \lambda^n & (\binom{n}{1})\lambda^{n-1} & \cdots & (\binom{n}{k-2})\lambda^{n-k+2} \\
0 & 0 & \lambda^n & \cdots & (\binom{n}{k-3})\lambda^{n-k+3} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \lambda^n
\end{bmatrix}
\]
where \( \binom{n}{k} = 0 \) when \( k > n \).

When \( |\lambda| \geq 1 \) there is obviously instability. Thus even when \( \max |\lambda_i| = 1 \) we have proved instability if the Jordan form includes an \( L_k(1) \) with \( k > 1 \).
We are concerned about the nature of the values of $p_i$ when
\[ p_i' = \frac{\sum_j p_{ij} w_{ij}}{\sum_j p_{ij} p_{ij} w_{ij}}, \quad \forall i, \]
is constant, i.e., $p_i' = p_i = \hat{p}_i$. Provided $\hat{p}_i \neq 0$ for all $i$, then we may write this as
\[ \sum_j p_{ij} p_{ij} w_{ij} = \sum_j p_{ij} w_{ij}, \quad \forall i. \]

This leads to the matrix form of the solution
\[ W p = k \hat{p}, \text{ for some scalar, } k. \]

We have discussed the solution where $W$ is invertible. The question remaining is what happens when $W$ is not invertible. One possibility is that there is no admissible solution. If, alternatively, a solution does exist, then it may be any one of a set of possible solutions given by
\[ \hat{p} = k W^{-1} l + (I - W^{-1} W) z \]
for any conditional inverse, $W$, such that $W W W = W$, and arbitrary $z$ (Theorem 7.3.1, page 142, Graybill, 1969).

If we impose the condition that $\sum_l p = 1$, then we can write the solution set as
In the case of a two-component system, we could have a singular \( W \) of the form
\[
\begin{pmatrix}
1 & 1 \\
1 & 1
\end{pmatrix},
\]
which clearly would not alter the gene frequencies during a generation. A neutral equilibrium exists for any values of \( p_1 \) and \( p_2 \) (\( p_1 + p_2 = 1 \)).

A different situation occurs when the matrix is of the form
\[
\begin{pmatrix}
w_{11} & w_{12} \\
w_{21} & w_{22}
\end{pmatrix},
\]
which it must be for a singular \( W \). In this instance, the genotype \( X_2 \) must have either a greater reproductive capacity than \( X_1 \), or a lesser, when both are in competition either with \( X_1 \) or \( X_2 \). So either \( X_2 \) replaces completely \( X_1 \), or \( X_2 \) is completely replaced by \( X_1 \). There is obviously no stable equilibrium for the two-component case. This may not be so for the three-component case.