A Theoretical Analysis of Plant Host-Pathogen Interactions in a Gene-for-Gene System

Ronald John Czochor

Biomathematics Series No. 8
Inst. of Statistics Mimeo Series No. 1603

ABSTRACT

CZOCHOR, RONALD JOHN. A Theoretical Analysis of Plant Host-Pathogen Interactions in a Gene-for-Gene System. (Under the direction of H. R. VAN DER VAART.)

The explanatory value of the concepts of selection against unnecessary virulence in the pathogen and selection against unnecessary resistance in the host was investigated using a mathematical model first proposed by Leonard (1977). The model was changed from a simultaneous pair of difference equations to a sequential pair of difference equations in order to better reflect the way the frequencies change in nature. A linear analysis was conducted on both pairs and the simultaneous model was shown to have a locally unstable internal equilibrium point while the analysis of the sequential model produced inconclusive results. Numerical simulation of the sequential model produced inward spirals which led credence to the suspicion that stable limit cycles or a stable equilibrium point existed. The analysis of the higher order terms of the Taylor expansion required the development of a technique for discrete time difference equations that was analagous to that developed by Poincaré (1885) for continuous time differential equations. For parameter values in which stability could be expected biologically, the internal equilibrium point was found to be locally stable.

Also discussed was the potential application of a multiple niche model to the study of gene frequency change in the pathogen on a multiline crop. The model was analyzed and the internal equilibrium point was determined to be stable in some regions of the parameter space. The fact that a multiline crop could produce such a polymorphism in the pathogen

said little about the effectiveness of this strategy in preventing an epidemic, since the pathogen population could be increasing at a high rate while still maintaining a polymorphism.

A model was developed to study the increase of the pathogen population on either the multiline or single variety crop. The change in the number of lesions of each pathogen race was modeled with the amount of uninfected leaf area considered as a limiting factor. The relative site of a lesion in each host-pathogen combination was a crucial factor in describing host resistance. The ease with which parameters could be measured was emphasized and a preliminary comparison was made between rotation of two host varieties and simultaneous planting of the two varieties. In this simplistic comparison, the simultaneous planting (multiline crop) appeared to be less effective, except when the reproductive rate or initial size of the pathogen population was very low.

TABLE OF CONTENTS

1. INTRODUCTION

One of the methods used in the control of plant disease is the deployment of resistant cultivars. With modern agricultural practices, this usually means that large areas are planted with one variety of one species of crop. In many instances, this genetic uniformity presents an opportunity for a genetically malleable pathogen population to overcome the resistance and create disease of epidemic proportions. However, in some instances, this so-called breakdown of resistance does not occur. It would then be advantageous to plant crops such that the resistance does not break down or at least such that the time until breakdown is prolonged.

Before we look further into this issue, we must realize that there are two types of disease resistance. According to Van der Plank (1963), vertical resistance (VR) occurs "when a variety is resistant to some races of the pathogen" and horizontal resistance (HR) occurs "when the resistance is evenly spread against all races of the pathogen". Unfortunately, these two terms have caused quite a stir in the plant pathology literature, so that now they have many different meanings. Van der Plank (1963) actually started the problem when he stated in his discussion of the effect of resistance on an epidemic that VR would have the effect of reducing the initial amount of disease and that HR would normally have the effect of reducing the rate of disease increase. This, however, caused some to consider these effects as the actual definitions of the terms. Then in his next book Van der Plank (1968) decided that since races of the pathogen could differ in aggressiveness, the presence of HR could not be easily determined unless

he interpreted the definitions as follows: "vertical resistance implies a differential interaction between varieties (of host) and races (of pathogen)" and "in horizontal resistance there is no differential interaction". This enabled VR to be determined through analysis of variance or through a ranking procedure to test for significant differences between rankings of pathogen aggressiveness compared to pathogen rankings on a known horizontally resistant variety. However, VR no longer implied differential interaction between varieties and races; in his new book, Van der Plank (1975) stated that VR was actually equivalent to this interaction.

Again in 1975, he discussed vertical resistance and horizontal resistance, but this time he used a theory put forth by Flor (1955). Working with the host plant flax, Linum usitatissimum, and its rust causing pathogen, Melampsora lini; Flor was able to show a one-to-one matching in which a specific allele in the flax confers resistance to a specific allele in the rust. This he called the gene-for-gene theory which has since been shown to be valid for many other plant disease systems (Flor, 1971). Van der Plank said that VR would be resistance for which a gene-for-gene hypothesis is conceivable and that HR would be resistance for which it would not be possible.

Problems with the analysis of variance test for VR and with inconsistencies with the original definitions were brought out by Robinson (1976), Parlevliet and Zadoks (1977), and Nelson (1978). These eventually caused Van der Plank (1978) to further dilute the dichotomy to VR with specificity, VR without specificity, and HR. This was quite different from the original definition in which VR was such resistance as is specific for some races only.

Thus there is more than one type of resistance and there is disagreement on the definitions of the different types of resistance. It will then be necessary to limit our discussion to one of the hypothesized types of resistance, and so, we shall confine our study in the following chapters to vertical resistance with specificity in a theoretical gene-for-gene system. Of ultimate interest is the genetic relationship between the different varieties of host plant and the different races of one species of pathogen and hence all external influences in the form of inputs to and outputs from this subsystem will not be considered.

In a gene-for-gene system, Van der Plank theorized that there must be two types of selection going on. He called them directional selection, by which he meant selection favoring the resistance gene in the host or selection favoring the virulence gene in the pathogen; and stabilizing selection, by which he meant selection against unnecessary resistance or virulence. Since the term stablizing selection has been used differently in the genetic literature, we shall use the term selection against unnecessary resistance or against unnecessary virulence. Such selective forces had to be in effect since gene-for-gene systems would never have been discovered if selection inevitably favored resistance in the host population and virulence in the pathogen population. A gene for resistance cannot be recognized except in comparison with its allele for susceptibility, and genes for resistance will not be recognized if the corresponding gene for avirulence is missing from the pathogen population. Therefore, Van der Plank argued that vertical resistance would not have been found unless genes for avirulence were sometimes selectively favored over genes for virulence.

There have been numerous arguments against this concept of selection against unnecessary resistance. These have been summarized by Crill (1977) and Nelson (1972, 1973), but neither Nelson nor Crill dispute the logic behind the concept. In Chapter 3, we shall look at an explanatory model representing this concept and analyze it to determine whether the observed polymorphic populations of host and pathogen can be explained by this type of selection. A problem arising in the analysis of the model necessitates the development of a new analytic technique which is described in Chapter 4.

The implementation of vertical resistance in an agricultural system of just one variety will provide almost total resistance to disease for a while until a virulent race of the pathogen develops. Then a calamitous breakdown of the resistance occurs due to directional selection in favor of the virulent pathogen race. Hence we have the boom and bust cycle so prevalent with use of vertically resistance varieties in a monoculture.

This presents an especially difficult problem since the creation of suitable, resistance varieties is a time consuming process. A method for circumventing this cycle has been proposed by Jensen (1952) and Borlaug (1953) in the form of multiline cropping practices, in which many varieties of the same crop species would be randomly planted together. It is thought that this would promote a polymorphic pathogen population with subsequent low levels of disease comparable to what is observed in natural co-evolutionary systems. A possible model for this type of agricultural practice along with an analysis of its stability will be presented in Chapter 5.

The ultimate question to be addressed through mathematical modeling of this type of system is whether multiline cropping practices or development of the varieties separately in a sequence of pure line cultivars would be more effective in preserving the longevity, in terms of usefulness, of the resistant varieties. A model to address this question will be developed in Chapter 6.

There have been a number of models proposed to describe the coevolution of the host and its pathogen in a gene-for-gene system as well as models for describing multiline systems. In the next chapter, which parallels Leonard and Czochor (1980) some of these models will be presented. This review will present a historical background for the studies undertaken in this dissertation.

2. LITERATURE REVIEW

The models that have been proposed to describe the genetic interactions between populations of plants and their pathogens can be classified in one of two categories. In the first section of this chapter, a review of the models of gene-for-gene interactions, their assumptions, and conclusions will be conducted. Later, models that have been developed to describe multiline cropping systems will be discussed.

Mode (1958) was one of the first to use a mathematical model to analyze genetic interactions between populations of plants and their pathogens in gene-for-gene relationships. He based his model on a system with two genes for resistance that are alleles at a single locus and two corresponding genes for virulence that occur at independent loci where they are distinguished from alleles for avirulence. He assumed that the fitness of a host in a particular host-pathogen combination varies inversely with the fitness of the pathogen, and that in genetically mixed populations, the fitness of a particular pathogen genotype is determined by weighting its fitness in each combination with a host genotype by the frequency of that host genotype. The model also assumed continuous change in genotype frequencies instead of the discrete change necessitated by the assumption of distinct generations. His analysis of this model showed that a host-pathogen system could reach a stable equilibrium provided that certain conditions were met. These conditions were presented in mathematical terms, so it was not easily apparent what the biological basis for the stable equilibrium might be. In fact, the example that Mode used to

illustrate a stable equilibrium had no fitness values greater than zero, which is biologically absurd. Also Mode's model could not account for the persistence of genes for susceptibility.

In his second model, Mode (1960) studied the change in the frequencies of the various cultivars and races rather than gene frequencies as in his previous model. This model was just a special case of the model Mode (1961) presented the following year. In these models, as in his previous model, he assumed continuous increase of the various populations in order to simplify the analysis. He did, however, recognize that in fact, the generations of the host at least were discrete. Since the assumption of continuous increase usually allows greater chance for stability than increase in discrete generations (see May, 1973), this assumption could very well lead to a conclusion of stability for an equilibrium point that actually is not stable. Other assumptions were also the same as in his first model.

In his general model, Mode (1961) considered four classes of stationary states: (i) when population number in both the host and pathogen populations is constant; (ii) when population number in the host population is constant, but that in the pathogen population is variable; (iii) when population number in the host population is variable, but in the pathogen population is constant; and (iv) when population number in both the host and the pathogen is variable. If selection coefficients in this model are allowed to vary, a stable equilibrium can occur in all four of the above classes. However, if selection coefficients are constant, a stable equilibrium can occur in classes (ii), (iii), and (iv), but not in class (i). Class (i) is just Mode's (1960) second model. Although Mode gives the conditions under

which stability is expected, it is not apparent what they are in terms of biological constraints. He also suggested that in agriculture, a host population consisting of a mixture of cultivars would be superior to a population consisting of a single cultivar only if the mixture tended to slow down the selection of new pathogenic races. This type of result will be addressed in Chapter 6.

In analyzing his models, Mode found that the frequencies of pathogen races at equilibrium are determined by the fitness of host cultivars and that the frequencies of host cultivars at equilibrium are determined by the fitness of pathogen races. These features were also found by Jayakar (1970) and Leonard (1977) in subsequent host-pathogen models.

Jayakar (1970) developed a model principally with the interaction of bacteria and bacteriaphage in mind, but which is applicable to most host-pathogen interactions. He labeled the probability that a given host individual is infected if it is susceptible, as x_-. He then started with a simple model, assuming that once a host is infected by a pathogen, it dies and the pathogen reproduces so that an average of n offspring are produced. He also assumed that if the host cannot be infected because of resistance, then the host individual reproduces and the pathogen dies. The model is for a single locus for resistance in the host and a single locus for virulence in the pathogen. The A genotype ofthe host is susceptible to both the B and the b genotype of the pathogen. The a genotype of the host is resistant to the B (avirulent) genotype and susceptible to the b (virulent) genotype of the pathogen. The frequencies of A and a genotypes in the

host population, respectively, are p_1 and $1-p_1$; and the frequencies of B and b in the pathogen are p_2 and $1-p_2$, respectively.

Jayakar showed that this model will not produce a polymorphism, since it is selecting only for the virulent pathogen genotype which will eventually make the two host types effectively selectively neutral.

Next, Jayakar added the concept put forth by Van der Plank in the study of plant epidemics of some other type of selection that could balance with this directional selection in favor of the virulent pathogen. In fact, he generalized his model to be able to incorporate this type of selection by letting f be the inherent fitness of a relative to A and g the inherent fitness of b relative to B. This new model produced a non-trivial equilibrium point at $p_1 = g$ and $p_2 = (1-x)(1-f)/xf$, if the f and g were such that $0 < p_1 < 1$ and $0 < p_2 < 1$. Jayakar found this internal equilibrium point to be locally unstable, but his computer simulation showed that the trajectories actually cycled around in a closed ellipse that was probably some type of limit cycle. He also pointed out that the addition of mutation did not materially change this behavior. This model is very much the same as a model that Leonard (1977) later developed specifically to describe Van der Plank's concept of "stabilizing selection" in plant disease. We shall see more of this in Chapter 3.

Once again in the plant pathogen literature, Person, et al. (1976) and Groth and Person (1977) considered a model for the effects of selection on pathogen populations in gene-for-gene systems. The fitness values assigned to the avirulent and virulent pathogen on the susceptible host were 1 and $1-s_a$, respectively, and the fitness values assigned to the avirulent and virulent pathogen on the resistant host were $1-s_a$ and

1, respectively. For the frequence of susceptible hosts equal to m and the frequency of resistant host equal to n, they found that the fitness of the virulent and avirulent genotypes would be equal when $n\ s_A = m\ s_a$. Since s_A is likely to be near one, the point of equal fitness occurs when $s_a = n/m$.

Leonard (1969) used essentially the same type of model, but the fitness values assigned to the avirulent and virulent pathogen on the susceptible host were 1 and 1-s, respectively, and on the resistant host the fitnesses were 0 and 1-s , respectively. The point of equal fintess for his model occurred when s = n.

The difference between the models of Person, et al. (1976, 1977) and Leonard (1969) was mainly in the question of whether the virulent pathogen genotype reproduces better on a resistant host than on a susceptible host (Person's model) or equally well on both (Leonard's model). In fact, there is a third choice that Nelson (1978, 1979) suggested, that is that the virulent race would reproduce better on the susceptible host than on the resistant host. One question that may be answered through the analysis of a more general model is which of the above ideas can better produce a stable polymorphism. Such a more general model will be discussed in the next chapter.

The use of multiline varieties was suggested by Jensen (1952) and Borlaug (1953) as a method of preventing the rapid shifts in virulence in pathogen populations which have lead to epidemics in pure line cultivars. Barlaug advocated the "clean crop approach" with its goal of keeping the multiline free of disease by immediately replacing any diseased variety. Browning and Frey (1969) advocated the "dirty crop approach" which is based on the assumptions that the multi-line varieties

can stabilize the race structure of pathogen populations and that this can be done with sufficient resistance in the host mixture to prevent significant damage by the stabilized pathogen population. The validity of these assumptions will be studied in Chapter 5 and through a model developed in Chapter 6.

Kiyosawa (1972) compared theoretical calculations for the annual increase of disease in cultivars grown either as a mixture or separately in sequence with each cultivar being replaced when disease on it reached a critical level. The usefulness of this approach depended upon knowledge of the annual rates of disease increase in a multiline λ_m , and a pure line cultivar, λ . He showed that the mixture will have greater longevity than the sequence of pure line cultivars if $n \lambda_m < \lambda$, where n is the number of cultivars. The estimation of these rates from daily rates is possible, but when disease increase is limited by available host tissue, the problem becomes most complex. The simulation model developed in Chapter 6 will address this problem.

Crill (1977) stated that in gene-for-gene systems, rotations of pure line cultivars would be superior to multiline varieties. However, he apparently based his conclusions on the dubious assumption that selection against unncesssary genes for virulence in the pathogen occurred only in the pure line system and not the multiline system.

Kiyosawa and Yabuki (1978) developed a model for gene frequency in the pathogen population on a mixed host and concluded that the equilibrium among the pathogen races could not occur without the operation of selection against unnecessary virulence genes. An analogous model with similar results will be discussed in Chapter 5.

Other models for multipline cropping systems proposed by Groth (1976), Marshall and Prior (1978, 1979), Marshall and Weir (1982), and Barret and Wolfe (1978) have all addressed the question of the multiline and the number of varieties needed in the multiline to obtain a certain level of effectiveness.

3. A MODEL FOR COEVOLUTION IN GENE-FOR-GENE SYSTEMS.

In this chapter, we shall discuss a model developed by Leonard (1977) that extended his earlier model (Leonard, 1969) to include the interactions affecting selection pressures in the host population as well as in the pathogen population. This model attempts to incorporate Van der Plank's concept of selection against unnecessary resistance and against unncecessary virulence into a gene frequency model of a gene-for-gene host-pathogen system. It was hoped that incorporation of these types of selection into such a model would be enough to ensure genetic polymorphisms in both the host and pathogen populations. The observed coexistence of virulent and avirulent races of the pathogen as well as of resistant and susceptible varieities of the host, would then be explainable by Van der Plank's hypothesized selective forces.

Although Leonard developed the model, his analysis of it was limited to computer results for specific parametric values. In this chapter and the next, we shall attempt a more general and more mathematical analysis and by doing so in a fair amount of detail, we shall hope to gain the ability to determine whether Van der Plank's hypothesis can explain the observed polymorphisms. However, first we must describe the model.

The model consists of two difference equations, one describing the change in gene frequency at a locus with two possible alleles in a diploid host and the other describing the change in gene frequency at a locus with two possible alleles in a diploid pathogen. Difference equations will be used instead of differential equations because the host is assumed to have distinct generations separated by a part of the

instead of differential equations because the host is assumed to have distinct generations separated by a part of the year which is not part of the growing season (for most plants, this is winter). Also, this part of the year requires that the pathogen have an overwintering phase and hence, with reproduction once a year, the pathogen must also have disctinct generations. In many disease systems, the pathogen is known to have more than one generation per growing season. Since some of these generations may be overlapping, a differential equation for the gene frequency change may be called for; however, if we assume no overlapping generations, difference equations may also be used for the case of multiple pathogen generations per growing season.

The one locus, two allele case is, of course, the simplest one, but its analysis is as complex as we would need for this hypothesized selection regime. The model is a general explanatory model and as such it makes quite a few simplifying assumptions. It assumes random mating in the host, and it assumes that the populations are large enough so that random drift will not be a factor. Also, it is purely a selection model and as such it assumes that both alleles in both the host and pathogen are already present in the respective populations. Thus mutation is not considered, but it could easily be worked into the model, at least in any computer simulations. Also, since the model is for gene frequencies and drift is not considered, the influence of coevolution on population size is overlooked.

The development of Leonard's (1977) model can best be described by Table 1 and Table 2, which are exactly as they appear in his paper. For a haploid pathogen, with either an avirulent allele (v) or a

Table 1. Relative Fitness of Pathogen Genotypes on Different Hosts.

Pathogen genotype and reaction resistant host (RR or Rr)	Relative fitnesses of pathogen genotypes on: rr(susceptible)	R_(resistant)
v(avirulent)	1	1-t
V(virulent)	1-k	1-k+a
	Relative fitnesses of pathogen genotypes on mixed host population.	
v	$W_v = q^2 + (1-q^2)(1-t)$	
V	$W_V = q^q(1-k) + (1-q^2)(1-k+a)$	

p = frequency of R

q = 1 - p

Table 2. Relative Fitness of Resistant and Susceptible Host Genotype in the Presnece of Different Pathogen Genotypes.

Host genotype and reaction to avirulent pathogen (v)	Relative fitness of host genotypes in the presence of avirulent (v) or virulent (V).	
	v	V
R_(resistant)	1-c-s(1-t)	1-c-s(1-k+a)
rr(susceptible)	1-s	1-s(1-k)
	Relative fitnesses of host genotypes in the presence of a mixed pathogen.	
R_	$W_{R\_} = m[1-c-s(1-t)] + n[1-c-s(1-k+a)]$	
rr	$W_{rr} = m(1-s) + n[1-s(1-k)]$	

n = frequency of V

m = 1-n

virulent allele (V), we can express the relative fitnesses of these genotypes on the different host genotypes as in Table 1. By relative fitness, we mean the proportionate contribution of offspring to the next generation. Thus if the relative fitness of an avirulent pathogen on a susceptible (rr) host is arbitrarily put equal to 1, then the relative fitness of an avirulent pathogen on a resistant host is 1-t, where t is the effectiveness of resistance. Notice that we are considering the resistant allele R to be completely dominant so that both genotype RR and genotype Rr are resistant. Leonard estimated t to be very close to 1 (between 0.98 and 1.00), since the avirulent pathogen has relatively little or zero reproductive success if it is on a resistant host.

The fitness of a virulent pathogen on the susceptible host will be 1-k, where k is the cost of virulence. This is where Van der Plank's concept of selection against unnecessary virulence is incorporated. If $k > 0$, then the avirulent pathogen will be more fit than the virulent pathogen on the susceptible host where virulence is unncessary. Leonard estimated that k would be in the range of 0.1 to 0.4. Notice that the virulent pathogen on the resistant host still has this cost, k, but also involves another parameter a so that the relative fitness of this virulent genotype on a resistant host has a value of $1-k+a$. Depending on the value of the parameter a the model for the pathogen fitness could be the same as a number of the models or ideas that we have discussed in the previous chapter. For $a = k$, we would have a pathogen fitness array comparable to Person's (1976); for $a = 0$, it would be comparable to Leonard's (1969) model; for $a < 0$, it would be comparable to Nelson's (1978, 1979) ideas; and

for $a > 0$, we would have a fitness array comparable to that suggested by Denward's (1967) data for late blight. Since analysis of the stability of these models was not undertaken, a general analysis of Leonard's model might help to discern the relative merits of the above authors' positions.

The relative fitnesses of the pathogen genotype on a mixed host population were then calculated, where q is the frequency of allele r and hence the frequency of the susceptible host is q^2 and the frequency of the resistant host is $1-q^2$. The average over hosts of the relative fitness of the pathogen in each combination will give us the relative fitness on a mixed host population. Hence, the frequency of the virulent gene, V, in generation $\tau + 1$, $n(\tau+1)$, will be the frequency of the virulent gene, V, in generation τ times its fitness on the mixed host population divided by the average fitness (Crow & Kimura, 1970, p.179). Thus,

$$n(\tau+1) = n(\tau)W_V/\{n(\tau)W_V + [1-n(\tau)]W_v\}$$

where W_V and W_v are the fitnesses on a mixed host population of the virulent and avirulent pathogen genotypes, respectively.

Since the W_V and W_v are not functions of τ, this is an autonomous difference equation and can be written as

$$n' = nW_V/\{nW_V + [1-n]W_v\}$$

So we now have a recurrent equation for the frequency of the virulent pathogen gene in the next generation.

Next we look at the relative fitness of the resistant and susceptible host genotype in the presence of the different pathogen genotypes as displayed in Table 2. The fitness for the resistant host

in the presence of the avirulent pathogen is given by 1-c-s(1-t). In this instance, the parameter, c , is the cost of resistance and like k for the pathogen, it incorporates Van der Plank's concept of selection against unnecessary resistant into the model. Also, the relative fitness of the resistant host in the presence of the avirulent pathogen is lessened by s(1-t), the disease severity rating. The parameter, s, concerns the suitability of the environment for disease development and when multiplied by the relative fitness of the pathogen in that combination of host variety and pathogen race, expresses the fitness loss in the host due to disease.

Also, in Table 2, we have the relative fitness of the resistant host in the presence of a virulent pathogen to be 1-c-s(1-k+a) where once again the fitness is reduced by the cost of resistance and by the disease severity rating. For the susceptible host, the loss in fitness is just due to disease severity; thus we have 1-s for its fitness in the presence of an avirulent pathogen and 1-s(1-k) for its fitness in the presence of a virulent pathogen.

The relative fitnesses of the host genotype in the presence of a mixed pathogen population can then be calculated by averaging the different fitnesses over the different pathogen genotypes. Then the relative fitness of a resistant host genotype in the presence of a mixed pathogen population will be designated W_{RR} or W_{Rr} and the relative fitness of a susceptible host genotype in the presence of a mixed pathogen will be designated W_{rr} . Now, the frequency of the dominant resistant allele in generation $\tau + 1$, $p(\tau + 1)$, will be the frequency of the homozygous resistant genotype times its fitness, plus

one-half of the frequency of the heterozygous resistant genotype times its fitness divided by the average fitness (Crow and Kimura, 1970, P. 179).

$$p(\tau+1) = \frac{p^2(\tau)\ W_{RR} + p(\tau)\ [1-p(\tau)]W_{Rr}}{p^2(\tau)\ W_{RR} + 2p(\tau)[1-p(\tau)]W_{Rr} + [1-p(\tau)]^2 W_{rr}}$$

Once again the fitness functions are not functions of time so that we have an autonomous difference equation. The model then consists of the following pair of recurrence equations for the frequency of the virulent allele in the pathogen, n , and the frequency of the resistant allele in the host, p:

$$n' = \frac{nW_V}{nW_V + (1-n)W_v}$$

$$= \frac{n[1-(2p-p^2)t\,] + n\,[(2p-2^2)(a+t)-k]}{[1-(2p-p^2)t\,] + n[(2p-p^2)(a+t) -k]}$$

$$p' = \frac{p^2 W_{RR} + p(1-p)W_{Rr}}{p^2 W_{RR} + 2p(1-p)W_{Rr} + (1-p)^2 W_{rr}}$$

$$= \frac{p(1-s+nks) + p[ts-c-ns(a+t)]}{(1-s+nks) + (2p-p^2)[\ s-c-ns(a+t)]}$$

(3.1)

The algebraic derivation of these expressions is displayed in Appendix 8.1.

Now that we have this system of recurrence equations for gene frequency in the host and pathogen populations, we are interested in

whether there is a stable internal equilibrium point or whether one of the alleles in either or both the host and pathogen will become fixed at some trivial equilibrium point. In order for this selection regime to yield genotypically polymorphic populations of the host and pathogen, it would be necessary to have a stable internal equilibrium point or (at least) a stable limit cycle about such an internal point. This would prohibit the fixation of either allele in both the host and the pathogen populations.

Thus we must determine whether the system (3.1) has an internal equilibrium point and if it is stable. The point (n^*, p^*) will be an equilibrium point of the system (3.1) if $n^{*\prime} = n^*$ and $p^{*\prime} = p^*$. We call (n^*,p^*) an internal equilibrium point if $0 < n^* < 1$ and $0 < p^* < 1$, and a trivial equilibrium point if both n^* and p^* are equal to either 0 or 1. If either n^* or p^* (both not both) were equal to either 0 or 1, the point (n^*,p^*) would not qualify as an equilibrium point of (3.1). This, along with the existence of the trivial equilibrium points (0,0), (1,0), (0,1), and (1,0) is obvious from the equations of system (3.1). From the first equation of the system (3.1), it can be seen that if p^* is such that $2p^* - p^{*2} = k/(a+t)$, then $n' = n$ and from equation two of system (3.1), it can be seen that if $n^* = (ts-c)/s(a+t)$, then $p' = p$. Thus (n^*,p^*), where $n^* = (ts-c)/s(a+t)$ and $2p^* - p^{*2} = k/(a+t)$ will be an internal equilibrium point of (3.1) if $0 < n^* < 1$ and $0 < p^* < 1$.

Leonard suggested the following for the ranges of the various parameter values:

$0.1 < k < 0.4$ $\quad 0.0 < a < 0.8$ $\quad 0.01 < c < 0.05$

$0.95 < t < 1.0$ $\quad 0.1 < s < 0.8$

It is obvious then that the point (n^*, p^*) where $n^* = (ts-c)/s(a+t)$ and $p^* = 1-\sqrt{(a+t-k)/(a+t)}$ is an internal equilibrium point. The other solution for $p^* = 1 + \sqrt{(a+t-k)/(a+t)}$ would give an unacceptable value for p^*, namely a value greater than one. Hence there is only one internal equilibrium point and all that needs to be determined is if it is stable.

Now that we have described Leonard's model, we can start our analysis by first writing the model in the following more general form, where we include the possibility of multiple pathogen generations per year. It can easily be verified by mathematical induction that the following equation for j generations of pathogen per growing season reduces to Leonard's model, equation (3.1), when $j = 1$ (see Appendix 8.2).

$$n' = n(A+B)^j/\{A^j+n[(A+B)^j - A^j]\} = f(n,p)$$

$$p' = p(C+D)/\{C+(2p-p^2)D\} = g(n,p) \qquad (3.2)$$

where

$A(p) = 1-(2p-p^2)t$,

$B(p) = (2p-p^2)(a+t)-k$,

$C(n) = 1-s+nks$,

$D(n) = ts-c-ns(a+t)$.

In order to determine the local stability of any of the various singular points of this nonlinear system, we must first expand the system (3.2) via a Taylor series expansion about the singular point of interest

and then we must analyze the linear part of the system. Thus we want to calculate the Jacobian matrix for this system (3.2) which is just the following coefficient matrix for the first order terms of the Taylor expansion about (n^*,p^*)

$$\begin{bmatrix} \frac{\partial f(n,p)}{\partial n} & \frac{\partial f(n,p)}{\partial p} \\ \frac{\partial g(n,p)}{\partial n} & \frac{\partial g(n,p)}{\partial p} \end{bmatrix}_{(n^*,p^*)}$$

From the definitions of A, B, C, and D, we have the following:

$$\frac{\partial A}{\partial n} = 0 \qquad \frac{\partial A}{\partial p} = -(2-2p)$$

$$\frac{\partial B}{\partial n} = 0 \qquad \frac{\partial B}{\partial p} = (2-2p)(a+t)$$

$$\frac{\partial C}{\partial n} = sk \qquad \frac{\partial C}{\partial p} = 0$$

$$\frac{\partial D}{\partial n} = -s(a+t) \qquad \frac{\partial D}{\partial p} = 0$$

Hence from (3.2) and the above and after some simplification, we get the following:

$$\frac{\partial f}{\partial n} = \frac{(A+B)^j A^j}{\{A^j + n[(A+B)^j - A^j]\}^2}$$

$$\frac{\partial f}{\partial p} = \frac{j(2-2p)n(1-n)(A+B)^{j-1} A^{j-1}\{aA + t(A+B)\}}{\{A^j + n[(A+B)^j - A^j]\}^2}$$

$$\frac{\partial g}{\partial n} = \frac{-s\,p(1-p)^2\,[Dk + C(a+t)]}{[C + (2p-p^2)D]^2}$$

$$\frac{\partial q}{\partial p} = \frac{(C + p^2D)(C+D)}{[C + (2p-p^2)D]^2} \qquad (3.3)$$

The eigenvalues of this Jacobian matrix evaluated at the singular point in question will now be calculated. If the absolute value of all eigenvalues is less than one, the singular point will be locally asymptotically stable. If the absolute value of all eigenvalues equals one, the singular point of the linear system is a center; but the singular point of the nonlinear system will be either a center or a focus. If the absolute value of all the eigenvalues is not either less than or equal to one, then the singular point will be said to be locally unstable. Note that whenever the study of a linearized system is used to determine the stability or instability of a singular point, this is just local behavior for the nonlinear system.

Thus, the local stability of the internal equilibrium point can be investigated by calculating the eigenvalues of the Jacobian matrix evaluated at $n^* = (ts-c)/s(a+t)$ and p^* such that $2p^* - p^{*2} = k/(a+t)$. At this singular point (n^*,p^*), $A = 1 - \{kt/(a+t)\}$, $B = 0$, $C = 1-s + (ts-c)k/(a+t)$, and $D = 0$. Hence the Jacobian matrix evaluated at (n^*,p^*) is:

$$\underline{d} \equiv \begin{bmatrix} A^{2j}/A^{2j} & j2n^*(1-p^*)(1-n^*)(a+t)A^{2j-1}/A^{2j} \\ p^*s[k-(a+t)]C/C^2 & C^2/C^2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & jK_1 \\ L_1 & 1 \end{bmatrix}$$

To calculate the eigenvalues of the Jacobian matrix, we solve the following equation for λ :

$$\begin{vmatrix} 1-\lambda & jK_1 \\ L_1 & 1-\lambda \end{vmatrix} = 0 \Rightarrow (1-\lambda)^2 = jK_1L_1 \Rightarrow \lambda = 1 \pm \sqrt{jK_1L_1}$$

Recall that

$$K_1L_1 = \frac{2n^*(1-p^*)(1-n^*)(a+t)}{1-\frac{kt}{a+t}} \cdot \frac{p^*s[k-(a+t)]}{1-s+\frac{(ts-c)k}{a+t}}$$

$$= \frac{-2p^*(1-p^*)(1-n^*)n^*s(a+t)^2\ (1-\frac{k}{a+t})}{(1-\frac{kt}{a+t})\left[1-s+\frac{(ts-c)k}{a+t}\right]}$$

Since all the parameters are greater than or equal to zero and less than or equal to one, $K_1L_1 < 0$. Also,

$$K_1L_1 \leq \frac{2sp^*n^*(1-p^*)(1-n^*)(a+t)^2}{1-s+\frac{(ts-c)k}{a+t}} .$$

We also have from the fact that $n^*(1-n^*) \leq 0.25$ and $p^*(1-p^*) \leq 0.25$ that

$$K_1L_1 \leq \frac{2(0.25)(0.25)s\ (a+t)^2}{1-s+\frac{(ts-c)k}{a+t}}$$

To find the upperbound for $|K_1L_1|$, we pick the parameter values from the ranges given by Leonard (1977) that give the highest value for the above expression. Thus if we pick a = 0.8 and t = 1.0, we have the following:

$$|K_1L_1| \leq \frac{(.125)(1.8)^2 s}{1-s+\frac{(s-c)k}{(1.8)}}$$

Once again picking the parameter values from the ranges given by Leonard (1977) that will produce the largest upper bound we choose s = 0.8, c = 0.05, and k = 0.1.

Then:

$$|K_1L_1| \leq \frac{(.405)(0.8)}{1-0.8+\frac{(0.8-0.05)(0.1)}{1.8}} = 1.3405$$

and thus

$$-1.3405 < K_1L_1 < 0.0 \ .$$

Thus the eigenvalues of the Jacobian matrix are complex conjugates such that $|\lambda| > 1$ and the internal equilibrium point is locally unstable. The trivial equilibrium points are also locally unstable, except for some values of the parameters for which the point (1,1) could be stable, but in this case, we would not biologically except a polymorphism. The calculation of the eigenvalues of the Jacobians at the trivial equilibrium points and further discussion of the stability of these points can be seen in Appendix 8.3.

Hence, Leonard's model, where j = 1, leads to a locally unstable internal equilibrium. Sedocole (1978) used the same type of analysis as above, but studied only the specific case of j = 1 at the internal equilibrium point and concluded that Leonard's model did not account for the relatively stable polymorphisms that are observed in nature. In rebuttal, Leonard and Czochor (1978) pointed out, as is discussed above, that the trivial equilibrium points were also unstable locally. The only way a trivial point can be reached is along one of the incoming separatrices of the point and if our system were made up of continuous-time differential equations, we would be able to say that there must be at least one limit cycle or limit point. This would be possible because the uniqueness of a solution precludes the intersecting of any two trajectories, thereby prohibiting the boundaries from being reached and allowing application of the Poincaré-Bendixson Theorem to arrive at the result.

However, in discrete difference equations, the so-called solution curve or more properly, solution trajectory, is made up of a sequence of discrete points and although the solutions are still unique, it is not prohibited for one solution's point to lie on the line segment connecting two points of another solution. Hence, it is, a priori, possible for the boundary line to be reached by one trajectory even though another trajectory lies on that boundary line. The Poincaré-Bendixson Theorem then cannot be applied to difference equations by appealing to this noncrossing of solution curves. Unfortunately, this invalidates the application to difference equations of a large body of analytic procedures from differential equations. Thus, we cannot conclude from our analysis whether a limit cycle in the interior of the unit square exists in this

case, but we may find that computer simulation strongly suggests that something like a limit cycle may exist.

Leonard and Czochor (1978) also stated that the intent behind the development of Leonard's (1977) model was not accurately described by Leonard or Sedcole (1978). Leonard actually had a slightly different model in mind. He expected that the frequencies of the resistant and susceptible host plants would remain constant during the growing season and that during this time the frequencies of the virulent and avirulent pathogen genes would change according to the forces of selection. A change in the frequencies of resistant and susceptible plants would occur, depending on the relative numbers and viabilities of the seeds produced by each genotype in the previous season. Since seed production by each of the host genotypes depends upon the amount of disease suffered by each and since the amount of disease suffered by the host is highly dependent on the composition of the pathogen population at the end of the preceding growing season, we shall make the simplifying assumption that the frequency of a certain host allele is a function of the frequency of that host allele in the preceding growing season and the frequency of the virulent pathogen allele at the end of the previous growing season. Thus under this assumption, change in the genetic composition of host and pathogen actually occurs in a series of alternate steps. First, the pathogen adjusts to the host population as it exists during that growing season, and then, the new host population in the next growing season represents an adjustment to the disease damage caused by the pathogen population of the previous growing season. It is assumed that, although

the pathogen may undergo more than one generation in a growing season, the host's gene frequency will be a function of the final pathogen frequencies in the preceding growing season.

In (3.1), n' is the frequency of the virulent allele in the pathogen population at the end of the growing season and therefore, disregarding differential survival, at the start of the next growing season, and p' is the frequency of the resistant allele in the host population in the next growing season. Then we are able to write the alternate steps models with j generations of pathogen per growing season as in (3.2) as follows:

$$n' = f(n,p) = n(A+B)^j/\{A^j + n[(A+B)^j - A^j]\}$$

$$p' = g(n',p) = p[C(n') + D(n')]/[C(n') + (2p-p^2)D(n')] \qquad (3.5)$$

where as before

$$A = [1-(2p-p^2)t] ,$$

$$B = [(2p-p^2)(a+t) - k] ,$$

$$C = [(1-s+n'ks] ,$$

and

$$D = [ts - c - n's(a+t)] .$$

Also as before, we have

$$\partial A/\partial p = -(2-2p)t ,$$

$$\partial A/\partial n = 0$$

$$\partial B/\partial p = (2-2p)(a+t) ,$$

and

$$\partial B/\partial n = 0 .$$

But now we have the following:

$$\partial C/\partial p = ks\cdot[\partial f(n,p)/\partial p],\ \partial C/\partial n = ks\cdot[\partial f(n,p)/\partial n]$$

and

$$\partial D/\partial p = -s(a+t)\cdot[\partial f(n,p)/\partial p]\ ,\ \ \partial D/\partial n = -s(a+t)\cdot[\partial f(n,p)/\partial n]$$

or in other words,

$$\frac{\partial D}{\partial p} = \frac{-\partial C}{\partial p}\,[s(a+t)/ks] \quad \text{and} \quad \frac{\partial D}{\partial n} = -\frac{\partial C}{\partial n}\,[s(a+t)/ks]\ .$$

If we now do the calculations of the partial deviatives that make up the Jacobian for the new system (3.5), we get the following:

$$\frac{\partial f(n,p)}{\partial n} = \frac{(A+B)^j\,A^j}{\{A^j + n[(A+B)^j - A^j]\}^2}$$

$$\frac{\partial f(n,p)}{\partial p} = \frac{j(2-2p)n(1-n)(A+B)^{j-1}A^{j-1}\{aA + t(A+B)\}}{\{A^j + n[(A+B)^j - A^j]\}^2}$$

$$\frac{\partial g(n',p)}{\partial n} = \frac{-sp(1-p)^2[Dk + C(a+t)]}{\{C + (2p-p^2)D\}^2}\cdot\frac{\partial f}{\partial n}$$

$$\frac{\partial g(n',p)}{\partial p} = \frac{(C+p^2D)\ (C+D) + \frac{\partial f}{\partial p}\{-sp(1-p)^2\ [Dk + C(a+t)]\}}{\{\ C + (2p-p^2)D\}^{\ 2}}$$

Thus the local stability of the internal equilibrium point for system (3.5) can be determined by calculating the eigenvalues of the Jacobian evaluated at (n^*,p^*) which is the same internal equilibrium point as in system (3.1). Hence, for $n^* = (ts-c)/(as+st)$ and $2p^*-p^{*2} = k/(a+t)$, we have $A = [1-(kt)/(a+t)]$, $B = 0$, $C = 1-s+[k(ts-c)]/(a+t)$, and $D = 0$. Thus the Jacobian is as follows (see Appendix 8.4):

$$\begin{bmatrix} \frac{A^{2j}}{A^{2j}} = 1 & \frac{jn^*(2-2p^*)(1-n^*)(A+t)A^{2j-1}}{A^{2j}} = jK_1 \\ \\ \frac{-sp^*(1-p^*)^2C(a+t)}{C^2} = L_1 & 1 + jK_1L_1 \end{bmatrix}$$

The eignevalues of this Jacobian are:

$$\lambda = \{ (2 + jK_1L_1) \pm \sqrt{(2 + jK_1L_1)^2 - 4}\}/2 \quad .$$

If $-4 \leq jK_1L_1 \leq 0$, then the λ's are complex conjugates and $|\lambda| = 1$; if $jK_1L_1 > 0$ or $jK_1L_1 < -4$ at least one of the eigenvalues has absolute value greater than one. However, in Leonard's suggested ranges for the various parameters, we have seen that $-1.3405 < K_1L_1 < 0$ and hence the eigenvalues have absolute value equal to one, at least for $j \leq 2$.

Thus the incorporation of this change in gene frequency occurring in a series of alternate steps produces a type of feedback that has a potential stabilizing effect on the internal equilibrium point of this system. It changes the Jacobian matrix by simply changing the term $\partial g/\partial p$ from 1 to $1 + (\partial f/\partial p)(\partial g/\partial n)$.

As can be seen in Appendix 8.5, the local stabiltity of the trivial singular points is unchanged by this additional type of feedback produced in the alternate step type of model (3.5). The results of the stability analysis for the trivial points of (3.5) are exactly the same as the results for the trivial points of (3.2) that were displayed in Appendix 8.3.

Although the absolute values of the eigenvalues for the internal singular point are equal to one for suggested values of the parameters,

they may be greater than one if j, the number of pathogen generations during the growing season, is large enough to make $jK_1L_1 < -4$. It makes sense biologically then that the internal equilibrium point in this case would be unstable, since the pathogen would have a distinct advantage. However, if the absolute values of the eignvalues are equal to 1, we know that the internal singular point of system (3.5) is either a center or a focus and therefore not necessarily an unstable focus as it was in system (3.2).

Hence, our linear analysis is inconclusive and the effect of the higher order terms in the Taylor expansion should be analyzed. However, a glimpse of the first order partial derivates for this system persuaded us first to resort to numerical studies. The following set of figures (Figure 1 through Figure 3) graphically represent the results of a number of those computer simulations of the system (3.5) for various parametric values. Notice that for most of these simulations there is a marked spiralling in toward the equilibrium point with increasing time. Hence, it appears that, for at least some region of the parameter space, the internal singular point is stable or at least there exists some stable limit cycle about that point.

In any event, it does appear that the forces of selection against unnecessary resistance in the host are enough to account for the observed polymorphic populations of hosts and pathogen; if not at a stable equilibrium point, then at least in some stable cycle.

The above analysis would have been enough to show the explanatory power of Van der Plank's concept, but Fleming (1980) published a model which he claimed to be a continuous analog to Leonard's discrete model.

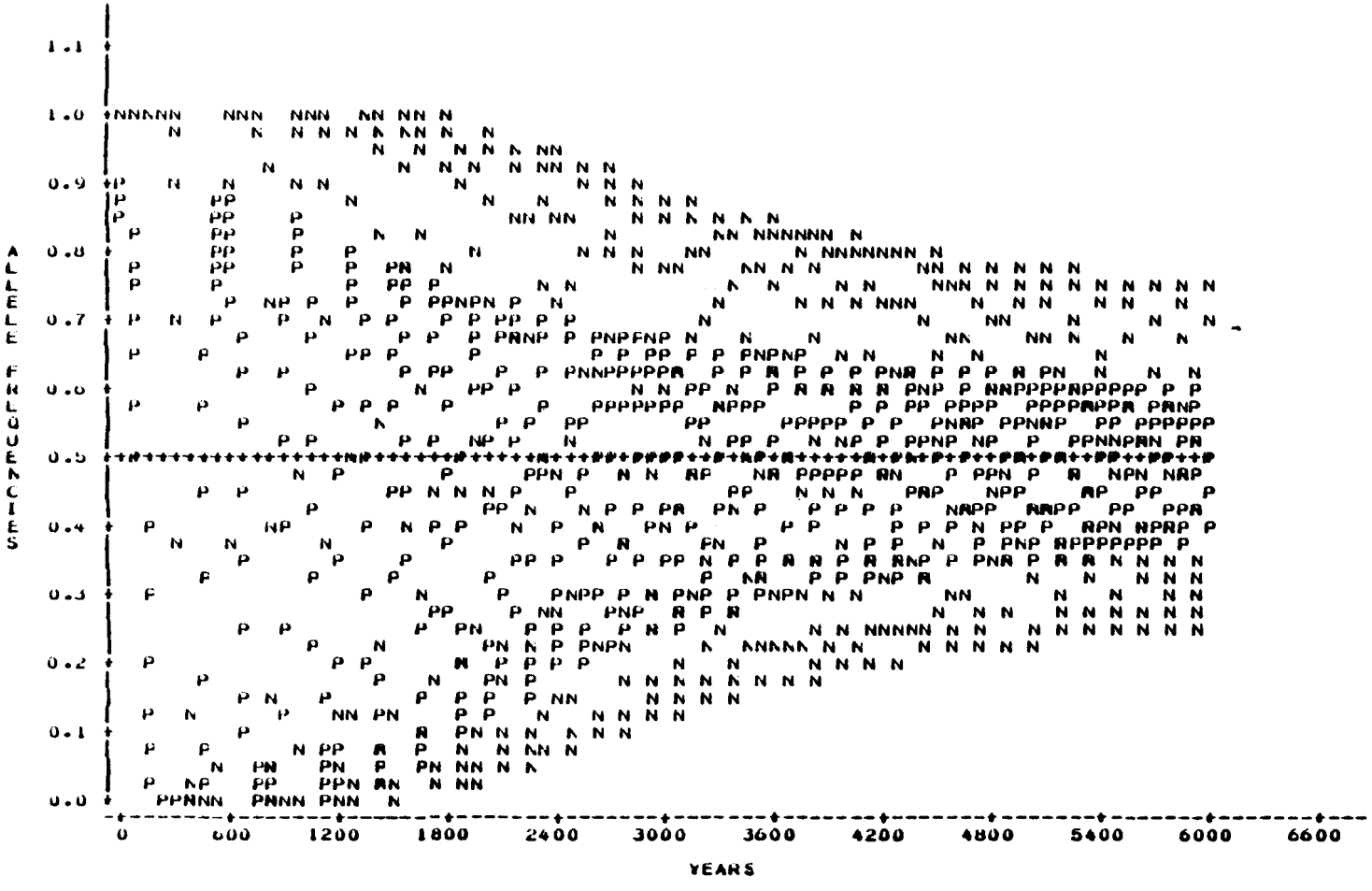

Figure 3.1. The frequency of the virulent allele in the pathogen, N, and the frequency of the resistant allele in the host, P, are plotted with respect to time for the following values of the parameters k = 0.75, t = 1.0, a = 0.0, c = 0.4, and s = 0.8. The "+" signifies the equilibrium value for P and the "-" signifies the equilibrium value for N.

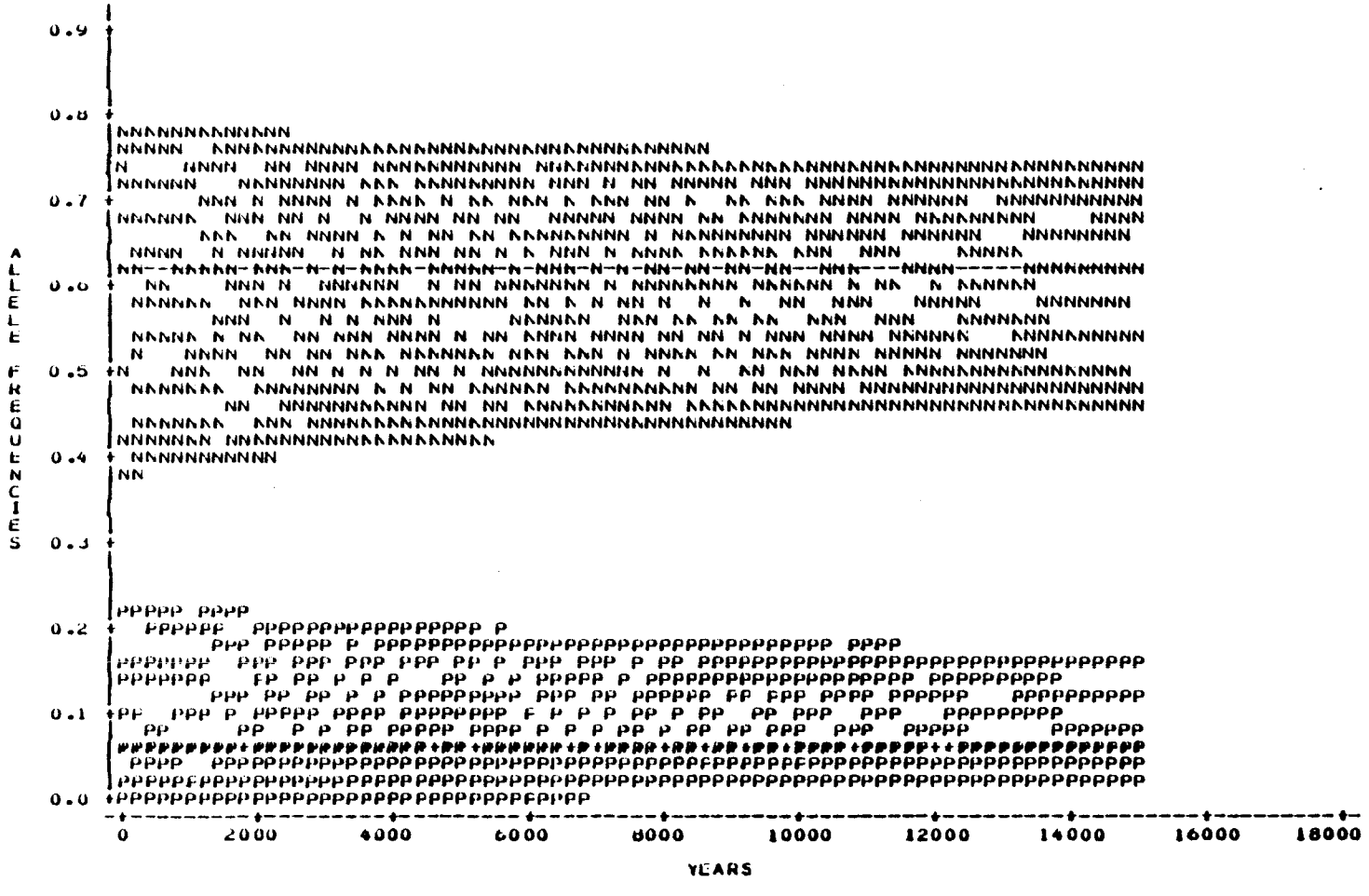

Figure 3.2. The frequency of the virulent allele in the pathogen, N, and the frequency of of the resistant allele in the host, P, are plotted with respect to time for the following values of the parameters k = 0.2, t = 1.0, a = 0.6, c = 0.01, and s = 0.08. The "+" signifies the equilibrium m value for P and the "-" signifies the equilibrium value for N.

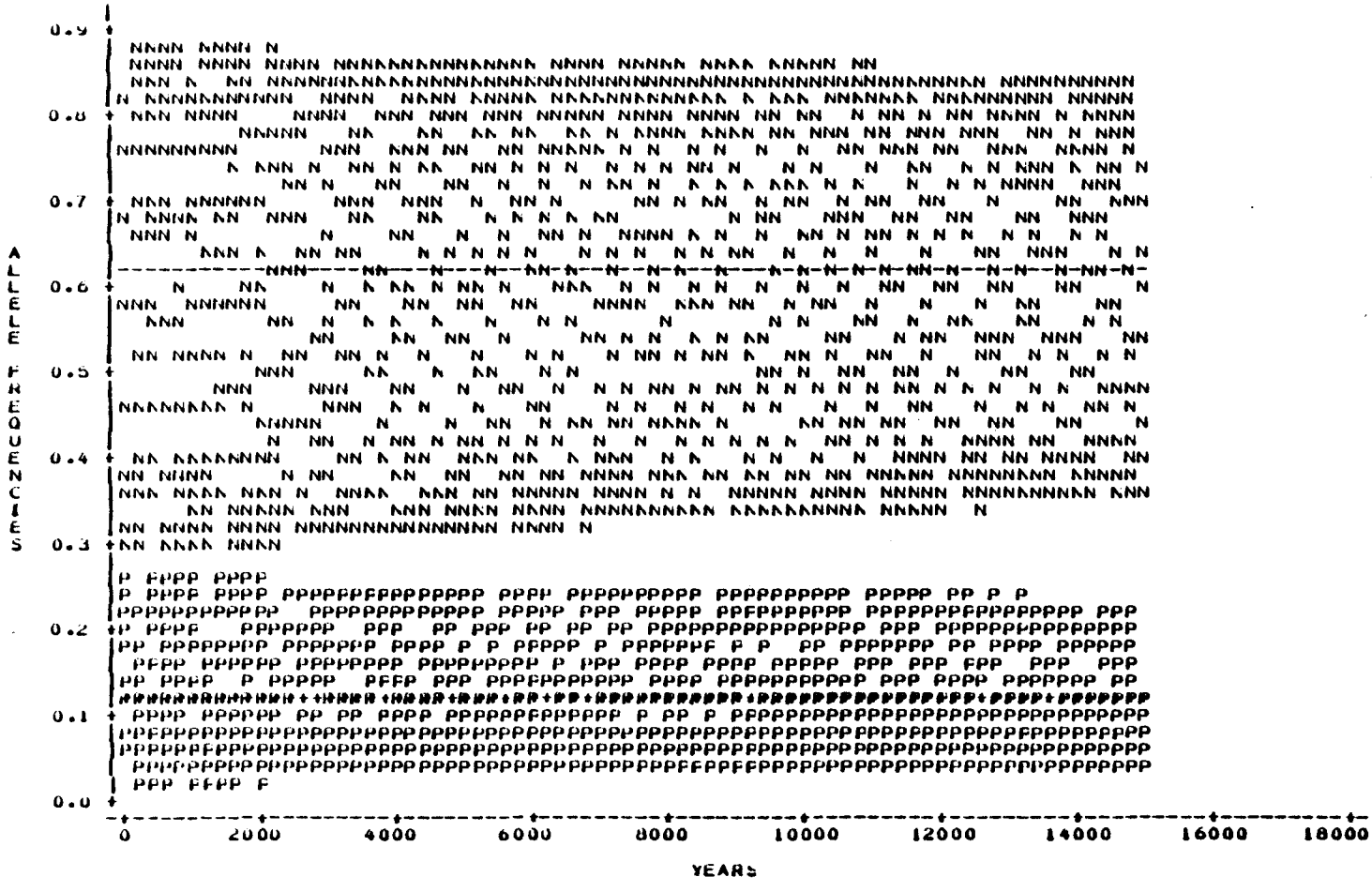

Figure 3.3. The frequency of the virulent allele in the pathogen, N, and the frequency of the resistant allele in the host, P, are plotted with respect to time for the following values of the parameters k = 0.3, t = 1.0, a = 0.4, c = 0.06, and s = 0.5. The "+" signifies the equilibrium value for P and the "-" signifies the equilibrium value for N.

Fleming was able to construct a Lyapunov function for his continuous model and to show that the internal singular point for his model was globally a center. It is not necessarily true that the solution of a differential equation, developed in the same manner as a difference equation, will be qualitatively the same as the solution of that difference equation (see van der Vaart, 1973). In the case of Fleming's model versus Leonard's model, this is obvious from the computer simulation of Leonard's model that demonstrates a spiralling in effect.

However, it is more compelling and mathematically more attractive to obtain an analytical proof of this fact and that will be attempted in the next chapter. We would like to show that our internal singular point for system (3.5) is not a center, but a focus. Although there are many techniques for doing this in differential equations, the only method that appears to be applicable to difference equations as well, is that developed by Poincaré (1885). All the others appeal to the fact that trajectories in continuous-time cannot cross and thereby appeal to the Poincaré-Bendixson Theorem in some form.

In the next chapter, we shall describe Poincaré's method, adjust it for difference equations, and apply it to determine whether the internal singular point in Leonard's model is a center or a focus.

4. THE CENTER-FOCUS PROBLEM

It is wellknown that for a system of two nonlinear differential equations, if a singular point is a center for the linearized system, then it is a center or a focus for the original system (Coddington and Levinson, 1955, p. 382). A method for discerning a focus from a center by using higher order terms of the Taylor series expansion for the nonlinear equations was described by Poincare (1885). The analog of this method for a system of two nonlinear difference equations will be discussed here. Although the methods are analogous, some difficulties will arise in difference equations that do not arise in continuous systems.

Suppose, in the first place, that the system of two nonlinear difference equations has been transformed so that the singular point is at the origin (0,0). Then the system (S) can be written as the following:

$$\begin{aligned} x' = x(t+1) &= g(x(t),y(t)) = g(x,y) = g_1(x,y) + g_2(x,y) + \ldots \\ y' = y(t+1) &= h(x(t),y(t)) = h(x,y) = h_1(x,y) + h_2(x,y) + \ldots \end{aligned} \qquad (S)$$

where g_i and h_i are homogeneous polynomials of degree i in x and y. If the singular point in the linearized system is a center (i.e., the eigenvalues of the Jacobian matrix of g and h, and in fact, of g_1 and h_1, evaluated at the singular point have modulus equal to one), then the singular point of the nonlinear system (S) is either a center of a focus. To discover whether the singular point is a center or a focus, we will try to construct a continuous function

$$f(x,y) = f_2(x,y) + f_3(x,y) + f_4(x,y) + \ldots$$

whose contour lines, $f(x,y) = C$, constitute a family of closed curves such that either

(1) $f(x,y)$ is constant along the solution curves of (S)

or

(2) $f(x,y)$ changes monotonically along the solution curves of (S). Actually, $f(x,y)$ is just a Lyapunov function where the f_i are homogeneous polynomials in x and y of degree i.

In order to judge whether the system is a center or a focus, we must investigate the function, f' , which is defined as follows:

$$f'(x,y) = f(x',y') - f(x,y).$$

As in Lyapunov's direct method, if $f' = 0$ for all values of x and y, then f is constant along the solution curves of (S), and we have a center since f has been constructed to constitute a family of closed curves. On the other hand, if $f' > 0$ (<0) for all values of x and y except the singular points, then f is monotone increasing (decreasing) along the solution curves of (S), and we have an unstable (stable) focus. The big difficulty in using Lyapunov's direct method lies in finding a suitable function f; however, in the following method for the center-focus problem, an algorithm for a stepwise construction of a suitable f is given.

Recall that f' was defined as follows:

$$f'(x,y) = f[g(x,y), h(x,y)] - f(x,y) =$$

$$= \{f_2[g(x,y), h(x,y)] - f_2(x,y)\}$$

$$+ \{f_3[g(x,y), h(x,y)] - f_3(x,y)\} + \ldots$$

For notational convenience, we shall make the following definition. Define $f_{ij}(x,y)$ to be the homogeneous polynomial of degree j in x and y that results from $f_i[g(x,y), h(x,y)]$ such that:

$$f_i[g(x,y), h(x,y)] = \sum_{j=i}^{\infty} f_{ij}(x,y) .$$

Here is where the method for difference equations gets more complex than in Poincare's differential equations. It becomes an extremely difficult accounting problem to keep track of the various terms, yet we are able to solve this problem in the following way:

$$f'(x,y) = \{[\sum_{j=2}^{\infty} f_{2j}(x,y)] - f_2(x,y)\} + \{[\sum_{j=3}^{\infty} f_{3j}(x,y)] - f_3(x,y)\} + \ldots$$

And then we rearrange to obtain:

$$f'(x,y) = \{f_{22}(x,y) - f_2(x,y)\} + \{f_{23}(x,y) + f_{33}(x,y) - f_3(x,y)\}$$

$$+ \{f_{24}(x,y) + f_{34}(x,y) + f_{44}(x,y) - f_4(x,y)\} + \ldots$$

If we then define f'_i as follows:

$$f'_i(x,y) = \{[\sum_{\ell=2}^{i} f_{\ell i}(x,y)] - f_i(x,y)\} \qquad (4.1)$$

we can break f' up into:

$$f' = f'_2 + f'_3 + f'_4 + \ldots$$

where f'_i contains all terms of degree i in x and y in f' .

Thus for a system with a specific $g(x,y)$ and $h(x,y)$ that has a center for its linearized system, we shall try to construct an f such that $f' = 0$ for all values of x and y. Note that $f' = 0$ will not be true for all values of x and y, unless $f'_2 = 0$ for all values of x and y. Otherwise, in a small enough neighborhood of the origin the higher order terms would not be able to compensate for f'_2 not being equal to zero. First then, to start our construction of f, we must choose f_2 so that $f'_2 = 0$. By the same argument as above, once we have f_2 , we need to choose f_3 so that $f'_3 = 0$. If we can continue to choose f_i so that $f'_i = 0$ ad infinitum, we could construct f so that $f' = 0$ and thus prove the existence of a center. This is in general quite impractical and so the main use of this method is in disproving a center. To this end, we hope to arrive at a step in which f_i cannot be chosen so that $f'_i = 0$ and in this case, we shall choose f_i so that f'_i is definite for all values of x and y.

The only other requirement that f must fulfill is that $f(x,y) = f_2(x,y) + f_3(x,y) + f_4(x,y) + \ldots$ be a continuous function whose contour lines $f(x,y) = C$, produce closed curves. Poincaré (1885, p. 178) showed that if $f_2(x,y) = k_2$, with k_2 a constant, represents a closed curve, then within a small enough neighborhood of the origin, $f(x,y) = k$, with k an appropriate constant is also a closed curve. Since in our case, the linearized system gives a center, we know that the contour lines, $f_2(x,y)$ equals k_2, are closed curves, therefore, within a small neighborhood of the origin, $f(x,y) = k$ produces a closed curve.

Hence, we will have constructed a function $f^*(x,y) \stackrel{\text{def}}{=} f_2(x,y) + f_3(x,y) + \ldots + f_i(x,y)$, whose contour line, $f(x,y) = k$, within a small

neighborhood of the origin, is a closed curve and such that $f^{*\prime} = f'_2 + f'_3 + \ldots + f'_i$ is either negative definite or positive definite.

The terms up to degree i that are involved in the function f' will be all present in $f^*(x',y') - f^*(x,y)$. Here x' and y' are expressed by the following truncated version of (S) which we shall call (S*).

$$\begin{aligned} x' &= g^*(x,y) = g_1(x,y) + g_2(x,y) + \ldots g_j(x,y) \\ y' &= h^*(x,y) = h_1(x,y) + h_2(x,y) + \ldots h_j(x,y) \end{aligned} \qquad (S^*)$$

where $j < i$. In fact, in order to include all possible terms of degree i that come from $f^*(x',y')$, we need j such that $j = i - 1$.

This is so since for $f_2(x',y') = a_2(x')^2 + b_2(x')(y') + c_2(y')^2$ the terms of degree i that are produced are $a_2(2h_1h_{i-1}) + b_2(h_1g_{i-1} + g_1h_{i-1}) + c_2{}^*2g_1g_{i-1})$. Thus if we work with a truncated version of (S), it must be (S*).

Now that we've constructed f* so that it represents a closed curve and so that f*' is definite, the questions that remain are how does f*(x,y) behave along the solution curves of (S*) and is it comparable to the behavior of f along the solution curves of (S)? These questions can be answered by the following theorem, the proof of which resembles Theorem 9.14 in LaSalle (1978).

Theorem: If $f^*(x,y) = f_2(x,y) + f_3(x,y) + \ldots + f_i(x,y) = k$ is a closed curve and if $f^{*\prime} = f'_2 + f'_3 + \ldots + f'_i < 0$, then $f(x,y) = f_2(x,y) + f_3(x,y) + \ldots$ is a Lyapunov function for (S) and the origin is an asymptotically stable point of (S).

Proof: For (S*), we have the following:

$f'_{S*}(x,y) \stackrel{def}{=} f*[g*(x,y), h*(x,y)] - f*(x,y) = \{f_2[g*(x,y), h*(x,y)] - f_2(x,y)\} + \{f_3[g*(x,y), h*(x,y)] - f_3(x,y)\} + \ldots + \{f_i[g*(x,y), h*(x,y)] - f_i(x,y)\} = f'_2 + f'_3 + \ldots + f'_i$ + terms of degree larger than i up to terms of degree ij = f*' + (higher order terms).

Since the coefficients of these terms of degree larger than i are all finite, we can choose a small neighborhood of the origin such than for (x,y)in that neighborhood, $f'_{S*}(x,y) < 0$ when $f*' < 0$. Thus f*(x,y) is a Lyapunov function of (S*) in this region of the origin and hence the origin is an asymptotically stable point of (S*). Now, write (S) as follows:

$$x' = g(x,y) = g*(x,y) + g_{j+}$$

$$y' = h(x,y) = h*(x,y) + h_{j+}$$

where h_{j+} and g_{j+} are polynomials in x and y of degree higher than j. Then:

$f'_S(x,y) = \{f[g(x,y), h(x,y)] - f(x,y)\} = \{f_2[g(x,y), h(x,y)] - f_2(x,y)\} + \{f_3[g(x,y), h(x,y)] - f_3(x,y)\} + \ldots + \{f_i[g(x,y), h(x,y)] - f_i(x,y)\} + \{f_{i+1}[g(x,y), h(x,y)] - f_{i+1}(x,y)\} + \ldots$

And so, $f'_S(x,y) = f'_{S*}(x,y)$ + terms in x and y of degree higher than i. This is so since the addition of g_{j+} and h_{j+} to g*(x,y) and h*(x,y) in (S*) adds on a polynomial with lowest degree terms as follows:

$$a_2[2g_1(x,y)g_{j+1}(x,y)] + b_2[g_1(x,y)h_{j+1}(x,y) + g_1(x,y)h_{j+1}(x,y)] + c_2[2h_1(x,y)h_{j+1}(x,y)].$$

The above polynomial of degree i+1 is arrived at in the following manner:

$$f_2[g(x,y), h(x,y)] = a_2(g* + g_{j+})^2 + b_2(g* + g_{j+})(h* + h_{j+}) + c_2(h* + h_{j+})^2$$

$$= a_2[(g^*)^2 + 2(g^*g_{j+}) + (g_{j+})^2]$$

$$+ b_2[(g^*h^*) + (g^*h_{j+}) + (h^*g_{j+})$$

$$+ (g_{j+}h_{j+})] + c_2[(h^*)^2 + 2(h^*h_{j+}) + (h_{j+})^2]$$

+ terms of degree $j + 2$ and higher. Here $j+2 = i+1$.

Once again, we can choose a neighborhood of the origin such that for x and y in the neighborhood, if $f'_{S^*}(x,y) < 0$, then $f'_S(x,y) < 0$. And since we've shown that in a small enough neighborhood of the origin $f'_{S^*}(x,y) < 0$ if $f^{*\prime} < 0$, we need only choose the smaller neighborhood to get $f'_S(x,y) < 0$ and hence the origin is asymptotically stable for (S).

An analogous proof can be used to shown that the origin is an unstable point of (S) if $f^{*\prime} > 0$.

So if we construct $f^*(x,y) = f_2(x,y) + f_3(x,y) + \ldots + f_i(x,y)$ such that $f^{*\prime} = f'_2 + f'_3 + f'_4 + \ldots + f'_i$ is either positive or negative definite, we can show that the origin is either an unstable or asymptotically stable singular point for the system (S). And the region of attraction will at least be the smallest neighborhood of the origin which satisifes the following: (1) $f^*(x,y) = k$ is a closed curve, given that $f_2(x,y) = k_2$ is a closed curve, (2) $f'_{S^*}(x,y)$ has the same sign as $f^{*\prime}$ for values of x and y in that neighborhood, and (3) $f'_S(x,y)$ has the same sign as $f'_{S^*}(x,y)$ in that neighborhood. Notice also that this neighborhood will be larger than that used in the linear analysis, since the neighborhood in the linear analysis is small enough to disregard the higher order terms and here, we are considering a neighborhood in which these higher order terms have an effect.

One problem that arises in the difference equation case, does not arise in Poincaré's study of differential equations. That is, in calculating f'_i for the difference equations, we automatically generate higher order terms, whereas for differential equations in which we have:

$$f'_i = \sum_{j=0}^{i-2} \{[\partial f_{i-j}(x,y)/\partial x] g_{j+1}(x,y) +$$

$$+ [\partial f_{i-j}(x,y)/\partial y] h_{j+1}(x,y)\}$$

the higher order terms are not generated. This makes the study of difference equations very cumbersome and, as has been seen, it complicates the notation. It also requires an additional neighborhood argument (i.e., see point (2) in the preceding paragraph).

Because of the trouble with keeping track of the higher order terms, it is instructive to see the application of this method in all its detail; however, in the interest of readability, most of this detail shall be relegated to Appendix 8.6. The method will be applied to Leonard's model to determine if its internal equilibrium point is a center or a focus. Recall that Fleming (1980) developed a so-called continuous analog to Leonard's model, showing the singular point to be a center, and computer simulation of the alternate steps version of Leonard's model that was presented in the previous chapter, appeared to show a stable focus or at least a stable limit cycle. Thus we would hope to prove by this method that the internal singular point of the alternate steps version of Leonard's model is not a center, unlike that of Fleming's model.

First, we shall write the equations of the alternate steps model in the following expanded form via a Taylor series expansion where x and y are deviations from the singular point at the origin. Using $x' = x(t+1)$, $y' + y(t+1)$, $x = x(t)$, and $y = y(t)$; we have the following system (S):

$$x' = g(x,y) = x + K_1y + K_2xy + K_3y^2 + K_4x^2y + K_5xy^2 + K_6y^3$$

$$y' = h(x,y) = L_1x + (1 + L_1K_1)y + (L_1K_2 + L_2)xy +$$
$$+ (L_1K_3 + L_2K_1)\, y^2 + (L_3K_1 + L_1K_5 + L_4 + L_2K_2)xy^2$$
$$+ [L_2K_3 + L_1K_6 + (L_1K_1^2/C_1) + L_4K_1]y^3$$

where the K_i and L_i as well as C_1 are all functions of the original parameters. The actual expressions are given in Appendix 8.7.

To implement the method, we must first find $f_2(x,y) = a_2x^2 + b_2xy + c_2y^2$ such that equation (4.1) with $i = 2$ yields:

$$f'_2 = a_2(x + K_1y)^2 + b_2(x + K_1y)\,[(1 + L_1K_1)y + L_1x] + c_2[L_1x$$
$$+ (1 + L_1K_1)y]^2 - [a_2x^2 + b_2xy + c_2y^2] = 0 .$$

It is seen in Appendix 8.6 that this will be zero for all values of x and y if we choose $a_2 = L_1$, $b_2 = L_1K_1$, and $c_2 = -K_1$. Thus $f_2(x,y)$ equals $L_1x^2 + L_1K_1xy - K_1y^2$. This will be a closed curve if its discriminant, $(L_1K_1)^2 + 4L_1K_1 < 0$. This is true since for biologically realistic parameter values as suggested by Leonard $L_1 < 0$ and $K_1 > 0$ and $|L_1K_1| < 4$.

So we have found $f_2(x,y)$ such that $f'_2 = 0$ for all values of x and y and $f_2(x,y)$ = "constant" is a closed curve. Now we can evaluate $f_2(x',y')$ to get our terms of higher order (i.e., $f_{23}(x,y)$, $f_{24}(x,y)$, ...).

Next we must find $f_3(x,y) = a_3x^3 + b_3x^2y + c_3xy^2 + d_3y^3$ such that equation (4.1) with i = 3 yields:

$$f'_3 = a_3[(x + K_1y)^3 - x^3] + b_3\{(x + K_1y)^2 [L_1x + (1 + L_1K_1)y] - x^2y\} + c_3\{(x + K_1y) [L_1x + (1 + L_1K_1)y]^2 - xy^2\} + d_3\{[L_1x + (1 + L_1K_1)y]^3 - y^3\} + f_{23}(x,y) = 0 .$$

Appendix 8.6 shows that this will be true if we choose the following:

$$a_3 = (1/3K_1)\{-2L_1c_3 - 3L_1^2d_3 - G_1\}$$

$$b_3 = -\{L_1c_3 + L_1^2d_3\}$$

$$c_3 = -\{1/[(1/3)L_1K_1^2 + K_1]\} \{(2/3)K_1^2G_1 - K_1G_2 + G_3\}$$

$$d_3 = -\{1/[L_1^2K_1 + 3L_1]\} \{G_2 - K_1G_1\}$$

where

$$G_1 = 2L_1K_2 - L_1L_2K_1$$

$$G_2 = 2L_1K_3 + L_1K_1K_2 - 2L_2K_1 + L_1^2 K_1^2K_2$$

$$G_3 = L_1K_1K_3 - L_1L_2K_1^2 - 2L_2K_1^2 .$$

Since $|K_1L_1| < 3$ as seen in the previous chapter, we do not have to worry about zero in the denominator, and so we can choose a_3, b_3, c_3, and d_3 that satisfy $f'_3 = 0$.

Once $f_3(x,y)$ is found in this manner, we calculate the fourth degree terms from $f_3[g(x,y), h(x,y)]$ to get $f_{34}(x,y)$ and try to find $f_4(x,y) = a_4x^4 + b_4x^3y + c_4x^2y^2 + d_4xy^3 + e_4y^4$ such that

$$f'_4 = f_{44}(x,y) - f_4(x,y) + f_{24}(x,y) + f_{34}(x,y) = 0$$

where $f_{44}(x,y)$ is just $f_4\{[x + K_1y], [L_1x + (1 + L_gK_1)y]\}$.

Since we have already found $f_2(x,y)$ and $f_3(x,y)$, we can write the following, where the F_i are known constants depending on the original parameters. In other words,

$$f_{24}(x,y) + f_{34}(x,y) = F_1x^3y + F_2x^2y^2 + F_3xy^3 + F_4y^4 \quad .$$

In order to find $f_4(x,y)$, such that $f'_4 = 0$, the coefficients for each term in x and y must equal zero. In other words, we must choose a_4, b_4, c_4, d_4, and e_4 so that the following is true.

$$A \begin{bmatrix} a_4 \\ b_4 \\ c_4 \\ d_4 \\ e_4 \end{bmatrix} + \begin{bmatrix} 0 \\ F_1 \\ F_2 \\ F_3 \\ F_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad (4.2)$$

where,

$$A = \begin{bmatrix} 0 & L_1 & L_1^2 & L_1^3 & L_1^4 \\ 4K_1 & 0 & 2L_1 & 3L_1^2 & 4L_1^3 \\ 6K_1^2 & 3K_1 & 6L_1K_1 & (3L_1 + 9L_1^2K_1) & (6L_1^2 + 12L_1^3K_1) \\ 4K_1^3 & 3K_1^2 & (6L_1K_1^2 + 2K_1) & (6L_1K_1 + 9L_1^2K_1^2) & (4L_1 + 12L_1^2K_1 + 12L_1^3K_1^2) \\ K_1^4 & K_1^3 & (K_1^2 + 2L_1K_1^3) & (K_1 + 3L_1K_1^2 + 3L_1^2K_1^3) & (4L_1K_1 + 6L_1^2K_1^2 + 9L_1^3K^3) \end{bmatrix}$$

If we pre-multiply A by the following matrix:

$$M = \begin{bmatrix} \frac{1}{L_1} & 0 & 0 & 0 & 0 \\ \frac{-2}{L_1} & 1 & 0 & 0 & 0 \\ \frac{-4K_1}{L_1C} & \frac{-2K_1}{C} & \frac{2}{3(K_1L_1 + 2)} & \frac{2L_1}{C} & \frac{-4L_1}{CK_1} \\ \frac{2K_1^2}{L_1c} & \frac{K_1^2}{C} & 0 & \frac{-2}{C} & \frac{4}{K_1C} \\ 0 & -K_1^2 & K_1 & -1 & 1/K_1 \end{bmatrix}$$

where

$$C = (4 - 2K_1^2L_1^2),$$

we get the following row equivalent augmented matrix to (4.2):

$$MA = \begin{bmatrix} 0 & 1 & L_1 & L_1^2 & L_1^3 & 0 \\ 4K_1 & -2 & 0 & L_1^2 & 2L_1^3 & -F_1 \\ 0 & 0 & 0 & 0 & 0 & -H_2 \\ 0 & 0 & 0 & [-2L_1(1-.25K_1L_1)] & [-4L_1^2(1-.25K_1^2L_1^2)] & -H_4 \\ -K_1^3 & K_1^2 & -K_1 & 1 & 0 & -H_5 \end{bmatrix}$$

Here $H_2 = (K_1L_1 - 2)(F_2/3) + K_1F_1 - L_1F_3 + [2L_1/K_1]F_4$ and H_4 and H_5 are comparable functions of F_1, F_2, F_3, and F_4.

All row operations are legitimate since neither L_1 nor K_1 is zero and $|L_1K_1| < 2$. Thus A is singular and (4.2) cannot be satisfied; so we must attempt to make f'_4 definite. To do this, we first rearrange A so that row 3 is now row 5 and the other rows remain in order. Hence, we have the block matrix A* with A_{1245} consisting of rows 1, 2, 4, and 5 of matrix A, on top and A_3, or row 3, on the bottom, and we have

$$\bar{f}^T = (f_1, f_2, f_3, f_4, f_5) = (0, -F_1, -F_3, -F_4, -F_2).$$

Since A is singular, A* is singular, and we cannot satisfy $A^*\bar{\alpha} - \bar{f} = 0$.

We can see that A_{1245} is a 5x4 matrix of rank 4 by performing elementary row operations as indicated in Appendix 8.8. With A_{1245} of rank 4, we can choose $\bar{\alpha}$ such that $(A^*\bar{\alpha})_i - f_i = 0$ for i = 1, 2, 3, 4 or, in other words, such that

$$\sum_{j=1}^{5} a_{ij}\alpha_j = f_i \quad \text{for } i = 1, 2, 3, 4. \tag{4.3}$$

Thus we can choose α such that the coefficients of y^4, y^3x, x^3y, and x^4 are all zero. The value of the coefficient of x^2y^2 is:

$$(A^*\bar{\alpha})_5 - f_5 = \sum_{j=1}^{5} a_{5j}\alpha_j - f_5 \tag{4.4}$$

We have from MA that the following is true with $g_1 = m_{31}$, $g_2 = m_{32}$, $g_3 = m_{34}$, and $g_4 = m_{35}$, where the m_{3j} for j = 1,2,4,5 are from matrix M.

$$0 = \sum_{i=1}^{4} g_i a_{ij} + \{2/[3(K_1L_1 + 2)]\}a_{5j} ,$$

which implies that

$$a_{5j} = \sum_{i=1}^{4} [-3(K_1L_1 + 2)/2]g_i a_{ij} .$$

Thus, we can calculuate (4.4) as follows:

$$\sum_{j=1}^{5} \{ \sum_{i=1}^{4} [-3(K_1L_1 + 2)/2]g_i a_{ij}\}\alpha_j - f_5 = \sum_{i=1}^{4} \{[-3(K_1L_1 + 2)/2]g_i\}[\sum_{j=1}^{5} a_{ij}\alpha_j] - f_5$$

which from equation(4.3) is equal to the following:

$$-\frac{3}{2} (K_1L_1 + 2) \sum_{i=1}^{4} g_i f_i - f_5 = [-3(K_1L_1 + 2)/2]$$

$$\{ \sum_{i=1}^{4} g_i f_i + [2/3(K_1L_1 + 2)]f_5\} = -\frac{3}{2} (K_1L_1 + 2)(-H_2)$$

Since $-2 < K_1L_1 < 0$ for most parameter values that we would expect to produce a stable equilibrium point (see Chapter 3), we can see that the coefficient of x^2y^2 has the same sign as H_2 . Thus, we have chosen f_4 such that f'_4 is definite with the sign of H_2 . The system's singular point will be a locally stable focus if $H_2 < 0$, and a locally unstable focus if $H_2 > 0$. If $H_2 = 0$, then, we must go on finding $f_i(x,y)$ such that $f'_i = 0$ until we can disprove our center. Unfortunately, the general expression for H_2 in terms of the parameters requires an intractable number of terms and is therefore impossible to investigate by analytical means. Hence, it needs to be numerically calculated for various values of the parameters in order to get an idea ofits sign. The values for differential parametric values are given in Table 4.1.

Table 4.1. Values of H_2 for various combinations of parameter values. The system's singular point will be a locally stable focus if $H_2 < 0$ and a locally unstable focus if $H_2 > 0$.

t	k	c	s	a	H_2
1.0	0.2	0.01	0.1	0.0	-.0051
"	"	"	0.2	"	-.0051
"	"	"	0.4	"	-.0039
"	"	"	0.6	"	+.0004
"	"	"	0.8	"	+.0159
1.0	0.2	0.01	0.2	0.2	-.0280
"	"	"	"	0.4	-.0538
"	"	"	"	0.6	-.0815
"	"	"	"	0.8	-.1104
"	"	"	0.8	0.2	-.0776
"	"	"	"	0.4	-1.1356
"	"	"	"	0.6	-3.5185
"	"	"	"	0.8	-7.4954
1.0	0.05	0.01	0.2	0.0	-.0011
"	0.1	"	"	"	-.0024
"	0.4	"	"	"	-.0119
"	0.6	"	"	"	-.0204
"	0.05	"	0.8	"	+.0179
"	0.1	"	"	"	+.0221
"	0.4	"	"	"	-.0078
"	0.6	"	"	"	-.0290
1.0	0.2	0.05	0.2	"	-.0224
"	"	0.1	"	"	-.0309
"	"	0.05	0.8	"	+.0533
"	"	0.1	"	"	+.0543
"	"	0.2	"	"	-.0198

The calculations for Table 4.1 were done in double precision using sixteen significant figures, so that most significant rounding errors would be negated. In all instances, H_2 is not zero. Hence, the singular point is not a center as in Fleming's (1980) model. In fact, for most parametric values, H_2 is negative and thus the singular point is a stable focus. For a few of the calculated values, H_2 is positive, but this happens only with values of the parameter, s, so high that we would not biologically expect stability, especially when coupled with low cost of unnecessary virulence and values for the parameter, a, such that $1-k+a < 1$. This is highly favorable for the virulent pathogen since it lessens the fitness of the resistant host which could easily cause a tendency toward fixation of the virulent gene and hence instability of the internal equilibrium point. In other words, for this type of parameter combination, we would biologically expect an unstable equilibrium.

Thus for most values of the parameters, other than those in which we would biologically expect instability, Leonard's model has a stable internal equilibrium point and so van der Plank's concept of selection against unnecessary virulence can in fact produce the polymorphic populations of resistant and susceptible hosts and virulent and avirulent pathogen that we observe and that have led to the gene-for-gene theory. That this type of selection is the only explanation for the observed polymorphism is open to question, but, nevertheless, it is a possible explanation, which is all this model was set up to show.

5. A POTENTIAL MULTILINE MODEL FOR GENE FREQUENCY.

Now that we have thoroughly analyzed a model for coevolution in a natural gene-for-gene system, it is necessary to study a model that could potentially be applied to the evolution of the pathogen in a multiline cropping system. This will get us one step closer to determining the relative merits of the multiline cropping strategy as compared to the single variety rotation strategy. It is important, in addressing a question through modeling, that any existing model with possible applicability to this problem be studied. If there is no direct use for the existing models, they may at least be instructive in the construction of a new model to answer the question.

For the question of whether a multiline cropping strategy will yield a stable polymorphism in the pathogen population, the classic model for a subdivided population by Levene (1953) and its analog for haploids by Gliddon and Strobeck (1975) could be of use. Gliddon and Strobeck's model for haploids describes necessary and sufficient conditions for existence of a stable multiple niche polymorphism involving just two alleles at a locus. Although Strobeck (1979) later expanded the model to include multiple alleles, we shall concentrate on the two allele cases, since it is more intuitive and its application to pathogen evolution in a multiline crop is easier to follow. Gliddon and Strobeck's adaptation of Levene's model concerns two alleles, A and a, that occur with frequencies of $1-q$ and q respectively. Thus the evolution of the pathogen with A being the virulent allele and a being the avirulent allele can be studied through this model. According to the model, the genotypes in this haploid population are distributed randomly

among N niches with the relative fitnesses of the genotypes A and a in the i^{th} niche having the values of 1 and w_i respectively. After selection through differential survival, the population undergoes reproduction with the i^{th} niche contributing a constant proportion, c_i, of individuals to the total population ($\Sigma c_i = 1$).

In applying this model to the evolution of a pathogen on a multiline crop, we take the various varieties in our multiline to be our niches and since the varieties of the host are randomly distributed within the multiline, the assumption of random dissemination of the haploid population among the N niches is easily satisfied. The various fitnesses for A and a genotypes for each niche brings into account the resistance or susceptibility of the host. If host variety i is resistant to the pathogen, then w_i will be very low, but if the host is susceptible, then w_i will be close to 1 or maybe larger than 1 if it is assumed the virulent pathogen is less fit on the susceptible host than is the avirulent pathogen. The only difficulty that arises in applying this model to the evolution of a pathogen in a plant disease system with a multiline crop is the assumption that the niche i or variety i contributes a constant proportion of individuals to the total population. It is possible that the proportion contributed by variety i could really be a function of q (i.e., $c_1(q)$), the frequency of the virulent allele and this would complicate the model.

Levene (1953) stated that his model was the worst possible case for the maintenance of multiple-niche polymorphisms since the random dispersal of the pathogen forbids the possibility of the pathogen selectively choosing a favorable niche. On the other hand, Dempster (1955) argued that Levene's model was not the worst possible case since

the assumption of a constant c_i , forces an implied frequency dependent selection to occur within the niches. This is a biological argument, since it is the only way that a constant proportion can be contributed by each variety, but the mathematical representation of Levene's model as well as Gliddon and Strobeck's model does not explicitly state this implied assumption. However, Christiansen (1975) developed a model using $c_i(q)$ and in comparing it to Levene's model found that it required more stringent restrictions on the parameters in order to obtain stability. Thus if we are just interested in the possibility of a stable equilibrium, we can study the model of Gliddon and Strobeck (1975), since if we cannot achieve stability with it, we cannot achieve stability with Christiansen's model.

The model they proposed produces the following recurrence equation for the new gene frequency of the avirulent genotype when grown on a multiplinecrop of N varieties.

$$q' = q \sum_{i=1}^{N} \frac{c_i w_i}{1 + q(w_i - 1)} \tag{5.1}$$

In order to determine the necessary and sufficient conditions for the existence of a stable, multiple-niche polymorphism in haploids, Gliddon and Strobeck carried out an analysis of (5.2) similar to that conducted by Levene (1953). The change in the gene frequency (5.2) is derived from (5.1) in the following manner.

$$q' - q = q \left\{ \sum_{i=1}^{N} \frac{c_i w_i}{1 + q(w_i - 1)} - \sum_{i=1}^{N} \frac{c_i [1 + q(w_i - 1)]}{[1 + q(w_i - 1)]} \right\}$$

$$= q\{(1-q)\sum_{i=1}^{N}\frac{c_i w_i}{[1+q(w_i-1)]} - (1-q)\sum_{i=1}^{N}\frac{c_i}{[1+q(w_i-1)]}\}$$

$$= q(1-q)\sum_{i=1}^{N}\frac{c_i(w_i-1)}{[1+q(w_i-1)]}$$

Thus we have

$$\Delta q = \sum_{i=1}^{N}\Delta q_i c_i = pq\sum_{i=1}^{N}\frac{c_i(w_i-1)}{1+q(w_i-1)} \qquad (5.2)$$

where Δq_i is the change in gene frequency in the i^{th} niche and $p = 1-q$. In order for an internal equilibrium point to exist, there must be a value of q between 0 and 1 such that

$$h(q) = \sum_{i=1}^{N}\frac{c_i(w_i-1)}{1+q(w_i-1)} = 0 .$$

Levene (1953) observed that such a point exists and is stable if $h(0) > 0$ and $h(1) < 0$, since h(q) is a continuous function of q. These conditions are met when

$$\sum_{i=1}^{N} c_i w_i > 1 \qquad (5.3)$$

and

$$\sum_{i=1}^{N} c_i/w_i > 1 \qquad (5.4)$$

Gliddon and Strobeck (1975) correctly identified conditions (5.3) and (5.4) as necessary and sufficient conditions for the existence of an internal equilibrium point, but their proof that the internal equilibrium point is unique and stable was not correct. Although Strobeck (1979) used a different method of proof in his analysis of the multiple model, the method of proof attempted by Gliddon and Strobeck (1975) is far more lucid and, therefore. it is of value to present a corrected version of it.

The internal equilibrium point (p^*,q^*) is defined as:

$$\Delta q^* = p^*q^* \sum_{i=1}^{N} \frac{c_i(w_i-1)}{1 + q^*(w_i-1)} = 0 \qquad (5.5)$$

In their proof, Gliddon and Strobeck attempted to show that Δq must be negative for all $q > q^*$ and positive for all $q < q^*$. They showed that for $q > q^*$ and $(w_i-1) < 0$, then $1 + q(w_i-1) < 1 + q^*(w_i-1)$ and $(w_i-1)/(1 + q(w_i-1)) < (w_i-1)/(1 + q^*(w_i-1))$. Similarly they noted that for $q > q^*$ and $(w_i-1) > 0$, then $1 + q(w_i-1) > 1 + q^*(w_i-1)$ and $(w_i-1)/(1 + q(w_i-1)) < (w_i-1)/(1 + q^*(w_i-1))$. However, they incorrectly stated that from this it must follow that $\Delta q_i < \Delta q^*_i$ when $q > q^*$. That this is not necessarily true can be seen by comparing equations (5.2) and (5.5). Since pq may be greater than p^*q^* , it is possible that $\Delta q_i > \Delta q^*_i$, even when $(w_i-1)/(1 + q(w_i-1)) < (w_i-1)/(1 + q^*(w_i-1))$.

Once Gliddon and Strobeck had shown that for $q > q^*$, $(w_i-1)/(1 + q(w_i-1)) < (w_i-1)/(1 + q^*(w_i- 1))$, they should have argued that since $c_i > 0$,

$$\sum_{i=1}^{N} \{c_i(w_i-1)/[1 + q(w_i-1)]\} < \sum_{i=1}^{N} \{c_i(w_i-1)/[1 + q^*(w_i-1)]\} .$$

Therefore, since $pq > 0$ and $\Delta q^* = 0$, $q < 0$. By parallel argument, when $q < q^*$, $\Delta q > 0$.

To show that there can be only one internal equilibrium point, suppose that there is a second internal equilibrium point, $q^{*'}$, such that $q^{*'} > q^*$ and there are no other equilibrium points between them. By the above argument, if $q = q^* + \varepsilon$, then $\Delta q < 0$, and if $q = q^{*'} - \varepsilon$, then $\Delta q > 0$. Since Δq is a continuous function of q, there must be an equilibrium point between q^* and $q^{*'}$. However, this contradicts the initial condition; therefore, the internal equilibrium point is unique, and it is stable if there are no oscillations about it.

Gliddon and Strobeck (1975) showed that for a small perturbation, ε, from the equilibrium, by writing $q' = q^* + \varepsilon'$ and $q = q^* + \varepsilon$ in (5.1),

$$\varepsilon' = \varepsilon \sum_{i=1}^{N} \frac{c_i w_i}{[1 + q^*(w_i - 1)]^2}$$

and that

$$\sum_{i=1}^{N} \frac{c_i w_i}{[1 + q^*(w_i - 1)]^2} > 0 .$$

Thus, a small perturbation will not result in an oscillation about the equilibrium point. However, this analysis does not apply to large perturbations, and does not rule out the possibility of a limit cycle about the equilibrium point.

To show that a limit cycle does not occur, assume that ε is any perturbation from the equilibrium so that $q = q^* + \varepsilon$ is between 0 and 1. In the following generation,

$$q^* + \varepsilon' = \sum_{i=1}^{N} \frac{(q^* + \varepsilon)c_i w_i}{1 + (q^* + \varepsilon)(w_i - 1)}$$

so that

$$\varepsilon' = \sum_{i=1}^{N} \frac{q^* c_i w_i + \varepsilon c_i w_i}{1 + q^*(w_i-1) + \varepsilon(w_i-1)} - \sum_{i=1}^{N} \frac{q^* c_i w_i}{1 + q^*(w_i-1)} \tag{5.6}$$

and from equation (5.1) and the definition of q* (5.5)

$$\varepsilon' = \sum_{i=1}^{N} \frac{q^* c_i w_i + c_i w_i - q^* c_i w_i \left[1 + \frac{\varepsilon(w_i-1)}{1 + q^*(w_i-1)}\right]}{1 + (q^* + \varepsilon)(w_i - 1)}$$

$$= \varepsilon \sum_{i=1}^{N} \frac{c_i w_i \left[1 - \frac{q^*(w_i-1)}{1 + q^*(w_i-1)}\right]}{[1 + (q^*-\varepsilon)(w_i-1)]} \tag{5.7}$$

For all values of i , $c_i w_i > 0$ and $1 - \frac{q^*(w_i - 1)}{1 + q^*(w_i-1)} > 0$

and $1 + (q^* + \varepsilon)(w_i-1) > 0$, so ε' has the same sign as ε. Thus the internal equilibrium point is stable within the entire interval (0,1) and the conditions (5.3) and (5.4) guarantee the instability of the trivial equilibrium points.

If it is possible then, to describe the frequency change of a haploid pathogen on a multiline crop by the model proposed by Gliddon and Strobeck (1975), and if we can choose our varieties and their respective proportions so that conditions (5.3) and (5.4) are satisfied, then the multiline would produce a stable genetic polymorphism in the pathogen population.

As we noted earlier, the one troublesome assumption in the Levene-type model proposed by Gliddon and Strobeck is the assumption that each niche, i, produces a constant proportion, c_i of the new pathogen generation each year. In general, one would expect the proportion produced by each niche, i, to depend upon the frequency of the virulent allele in the pathogen population. Christiansen (1975) analyzed such a model with $c_i(q)$, as the proportion of the pathogen produced in niche i . His analysis concluded that in both the model using c_i and the model using $c_i(q)$ the qualitative results were the same, that is, that population subdivision enhanced the possibility of accumulated variation and that increased isolation intensified this effect.

Although increased isolation might intensity the accumulation of variation, within the pathogen, we would not want to plant our multiline crop in patches of single varieties to enhance this isolation since this would lessen the effectiveness of the multipline crop in its retardation of the rate of disease increase. This point shall be discussed more thoroughly in Chapter 6.

The model by Gliddon and Strobeck (1975) then, can be used to tell us whether a multiline cropping strategy can produce a stable genetic polymorphism in the pathogen population. The analysis of the model has proven that such a stable polymorphism is possible in a haploid pathogen population.

6. A MULTILINE MODEL FOR PATHOGEN POPULATION GROWTH.

The two models that we have looked at so far have studied the change in gene frequencies of the pathogen and, in Leonard's model, the change in gene frequencies of the host as well. Both models are useful in determining whether it is possible that the genetic makeup of the respective populations will at all maintain a polymorphic population and, if so, for what parameter values this will hold. However, this is not very helpful in determining the value of a particular cropping strategy. Although a particular strategy, for rotating resistant varieties or for a particular multiline crop may produce a stable set of frequenices for the various races of the pathogen, the absolute numbers of the pathogen attacking the crop could still be large enough to destroy the crop. Hence, if we are interested in the effectiveness of a particular strategy, we must be able to integrate the gene frequency models with some type of population growth model.

Roughgarden (1979) has done some work along these lines. In his studies of coevolution, he considered simultaneous equations of population growth and of gene frequency change; however, in order to do any analysis of these systems he had to assume that the gene frequencies were constant and therefore, he studied only the equilibrium population sizes instead of population growth. Barret (1978) did look at the multiline strategy from the point of view of population growth, but then used his model to study frequencies of the races and not to study the actual number of pathogens produced. Also, Barret did not consider the amount of available tissue as a limiting factor. Trenbath (1977) did attempt to consider available tissue as a limiting factor in his study of the growth of various races

of the pathogen. He even included host growth, but, he did not consider dead lesions as taking up available space and hence limiting pathogen population growth.

In order to compare the rotation method with the multiline method, it is necessary to develop a model of population growth within the pathogen races that includes the limiting effect of available plant tissue and to introduce and study a measure for crop damage. Recall that in the rotation method, we plant one variety of the host for as many growing seasons in a row as it will stay effective or minimize crop damage and then we switch to another variety. In the multiline method, we plant all of our available varieties together, thereby lessening the amount of available susceptible tissue for each race of the pathogen and hence lowering the growth rate of the pathogen population. Thus the necessity of incorporating this limiting effect into the model is apparent.

The comparison would then be done, as Kiyosawa (1972) did, by setting some maximum number of pathogen allowable, NMAX, and then by determining the number of growing seasons that the multiline will last until NMAX is reached as compared to the number of growing seasons the rotated varieties would last if a new variety is introduced each time the pathogen population reaches NMAX. This could be considered a measure of the maximum tolerable level of crop damage, but of course, there can be other such measures.

Of course, the evaluation of the value of NMAX is highly dependent upon the situation. A small acreage farmer would necessarily set a much lower value than would a farmer with a large farm, since the small acreage farmer would have fewer plants. Another important consideration is the level of tolerance the host varieties have, some crop varieties

can support a relatively large number of lesions with little loss in yield. Yield and other economic measures would also be important considerations in determining a value for NMAX to use in the comparison of a multiline cropping strategy with a rotation cropping strategy.

In order to make this comparison, we must first develop a new model for population growth of the various pathogen races on the mixed host. This general model for any number of host varieties could then be specialized to study the pathogen development on a single variety, as is used in the rotation strategy.

The first question that needs to be addressed is just what are we going to measure or count in our population growth model; do we want to count spores or lesions? Since lesion number is more easily measured and since lesions are the space occupying aspect of the pathogen, we shall let $N_{ij}(t)$ be the number of lesions of race i on host variety j at time t.

We shall, once again as in the previously discussed models, assume that the pathogen develops in distinct generations and hence, we shall use difference equations to model the growth of the pathogen population. This assumption is not necessarily true, since some pathogen populations are known to have multiple pathogen generations per growing season and possibly overlapping generations. This model then, may be considered a first attempt at comparing the multiline strategy with the rotation strategy; since, although its assumptions may not be correct in all instances, it is fairly intuitive and legitimate for some diseases.

Let the parameter, w_{ij}, be the number of possible new lesions left to the next generation by an individual of pathogen race i on host variety j, assuming that leaf area is not limiting. If we did not

assume that the pathogen is haploid and that like produces like, we would need another subscript for w_{ij} to signify the type of new lesions that are produced. If we then take the number of lesions of race i on host variety j at generation t and multiply by w_{ij}, we will get the total number of additional lesions left to generation t + 1 by pathogen race i on host variety j.

The word "possible" in the above paragraph is important. The number of new lesions of race i that are formed on host variety j in generation t+1 is limited by the amount of available or uninfected leaf area of host variety j. The total number of possible lesions of race i produced from lesions at time t on a sufficiently large field so that leaf area is not limiting is:

$$\sum_j w_{ij} N_{ij}(t).$$

This can be multiplied by the proportion of available leaf area of host variety j to get $N_{ij}(t+1)$. As in the Levene type model described in the previous chapter, we obviously assume random dispersal of new lesions and this is a legitimate assumption since the host varieties are randomly distributed.

The expression for the percentage of available leaf area of host variety j is fairly complex and will therefore necessitate a stepwise construction. First, we have the proportion of the field that is planted in host variety j to be c_j. Next, we must account for the proportion of leaf area of host j that is uninfected at time t. We shall assume that the amount of leaf area the host plant has is a constant and that the amount of uninfected leaf area varies due to infection only. This is admittedly a very simplistic assumption since the host does grow during the season,

but it could be acceptable if we consider a disease that only develops on mature plants, where most of the growth has already occurred, we shall denote the constant amount of leaf area that host variety j has if planted throughout the field, in terms of the possible number of virulent lesions as K_j. This parameter is measured by taking the amount of leaf area for the whole field planted with variety j and dividing it by the size of the average virulent lesion on that host; this produces the number of possible virulent lesions that could be supported on a field of host j. Actually, K_j is the number of possible virulent lesions that could be supported on an entire field of host j and is usually less than the measurement of K_j calculated above. This is so since the number of lesions are not necessarily closely packed.

Notice in describing K_j, we called it the number of possible <u>virulent</u> lesions that could be supported on an entire field of host variety j. It is quite apparent that, in many instances, the size of a lesion made by a pathogen that is virulent on a specific host will be significantly larger than the size of a lesion made by pathogen that is avirulent on that host. Thus the size of each lesion of pathogen race i on host variety j can be described by a measure relative to the size of a virulent race j on host variety j. The measure m_{ij}, for pathogen race i on host variety j describes the relative size of the lesion and is therefore a number between zero and one with $m_{ij} = 1$. Hence, if we have $N_{ij}(t)$ new lesions of avirulent race i on host variety j at generation t, then the number of virulent size lesion spaces that these lesions fill is $m_{ij}N_{ij}(t)$.

Although we have assumed distinct generations so that only lesions formed in the preceding generation live and reproduce, the dead lesions

of all the preceding generations in that growing season still remain on the host and inhabit leaf area which cannot be reinfected. Therefore, in order to express the proportion of uninfected leaf area of host j, we need the number of lesions both living and dead that take up space on host variety j at time t , namely:

$$\left(\sum_{k=1}^{t} N_{ij}(k) \right) .$$

Now, we can write the number of virulent lesion sized spaces that have been filled up on host j by generation t as:

$$\left[\sum_{i} m_{ij} \sum_{k=1}^{t} N_{ij}(k) \right]$$

Hence the proportion of host j that have been infected up until generation t can be expressed as follows:

$$\sum_{i} [(m_{ij}) \sum_{k=1}^{t} N_{ij}(k)]/(c_j K_j)$$

All this allows us to write the following expression for the proportion of originally available lesion spots on the whole field that is still open and made available by variety j at time t:

$$c_j\{1 - \sum_{i} [(m_{ij}) \sum_{k=1}^{t} (k)]/[c_j K_j]\} .$$

If we now multiply the total number of possible lesions of race i produced from lesions at generation t by the percentage of uninfected leaf area of host j that is available to these possible lesions of race i in generation t+1, we get the number of new lesions of race i on host variety j at generation t+1 as follows:

$$N_{ij}(t+1) = [\sum_{j} w_{ij}N_{ij}(t)][c_j - \{\sum_{i} (m_{ij}) [\sum_{k=1}^{t} N_{ij}(k)]\}/K_j] \qquad (6.1)$$

The above difference equation represents the growth of the pathogen population during the growing season, but when $N_{ij}(t)$ is the last generation of pathogen before the crop is harvested, some type of overwintering of the pathogen must occur and if needed in the measure of crop damage the size of the crop would be measured. Thus if T is the number of generations the pathogen goes through in one growing season, then we reset the time index as follows:

$$N_{ij}(0) = PH_{ij}N_{ij}(T) \qquad (6.2)$$

Here, PH_{ij} is the percentage of lesions of race i on host variety j surviving over the winter. This is usually less than ten percent.

The equation (6.1) looks vaguely like the discrete pseudo-logistic model, but does not suffer from the problems of the pseudo-logistic, such as oscillation about the carrying capacity and the possibility of chaos (May, 1974). This is so because of the following extra parts to the model. The parameter, c_jK_j, is not the carrying capacity, but the amount of available space and as such limits the number of new lesions that can form on host variety j. That is, $\sum_{i}(m_{ij})N_{ij}(t+1)$ must be limited so that:

$$\sum_{i} (m_{ij})[\sum_{k=1}^{t+1} N_{ij}(k)] < c_jK_j \qquad (6.3)$$

Hence, we need a third part to the model. That is, if (6.3) is not true, then we must apportion the remaining uninfected leaf area on host variety j proportionately to all pathogen races. The proportion of uninfected leaf area assigned to each pathogen race i is computed to be

$N_{ij}(t+1)/\Sigma_i N_{ij}(t+1)$ where $N_{ij}(t+1)$ is the number of possible new lesions of pathogen race i on host variety j in generation t+1 as given in equation (6.1). Thus the number of virulent sized lesion spaces on host variety j filled with pathogen race i at generation t+1 is:

$$[N_{ij}(t+1)/\Sigma_i N_{ij}(t+1)][c_j K_i - \Sigma_i m_{ij} \Sigma_{k=1}^{t} N_{ij}(k)] \qquad (6.4)$$

Although this appears to be the most equitable method for dividing up the number of available spaces on host variety j among the different races of new lesions, it does allow much empty space to exist after the lesions have developed since the new lesions are not all as big as the spaces (i.e., the avirulent lesions are smaller than the virulent lesions). Nevertheless, since the new lesions all start at the same size, it is legitimate to divide them up this way. Once this is done, there is theoretically no more uninfected leaf area and thus this generation t+1 must necessarily be the last generation for pathogen race i on host variety j in the present growing session. Expression (6.4) is then renamed $N_{ij}(T)$ and applied to equation (6.2) to determine the amount surviving until the next growing season.

Since the model can at times require three different steps and since step one, equation (6.1), is quite complex; the most practical method for study of this model is through the use of computer simulation. This simulation will be done for a single variety planting in a rotation type of strategy and for a multiline cropping strategy.

Although most of the parameters may be fairly easily measured for an individual field and host-pathogen system, we shall do a general study of the two cropping models, and we shall leave any actual measure-

ments of the parametric values for a further study. However, a description of how these parameters could be measured is a necessary attribute of any biological model. This requires the modeler to stop and think about the biological application of his model and keeps him from developing esoteric models of little use to the biologist.

The ease with which the parameters of the model can be measured stems from the fact that we are considering lesions as our reproductive unit and that we are considering available leaf area in terms of lesion size. We have already described the measurement of K_j to be the total leaf area of host variety j divided by the size of the average virulent lesion on host j. The parameter, m_{ij}, is measured by the ratio of the average lesion size of race i on host variety j to the average virulent lesion size on host variety j. The parameter, w_{ij}, is measured by taking a whole field of host variety j and creating one lesion of race i at time t, and at a fixed time period later, called the infectious period or generation length, the number of lesions at that time t+1 are counted. This count measurement of w_{ij} is legitimate only if the infectious period is shorter than the incubation period, since otherwise, we would have overlapping generations, thereby preventing an accurate measurement of w_{ij} in this manner. Also, the overwintering rate, PH_{ij}, for pathogen race i on host variety j, can be measured by once again looking at our single variety field and by counting the number of pathogen race i lesions at the end of the season; then by counting the number of lesions of race i at the start of the next growing season, we can calculate PH_{ij} by dividing this number by the number at the end of the last growing season. Finally, we must know T, the number of generations the pathogen goes through in one growing season.

If we calculate all the parameters as listed above, we need only plug these values into our computer simulation model to determine the growth of the pathogen for the various cropping strategies. Table 6.1 contains a copy of the computer program that we shall use for our comparison.

The computer simulation was run for two pathogen races with arbitrary, but relatively realistic values for the parameters in order to compare the rotation scheme with the multiline scheme. The initial value for the number of lesions of race 1 and race 2 at the start was set, NR(1) = NR(2) = 100. Unequal values for NR(1) and NR(2) would probably make little difference since changes in the cropping strategy could also be used. For these values, the simulation was conducted for various values of T, the number of pathogen generations per growing season, W(I,J), the number of new lesions produced by race I on host J and, KS, the number of virulent sized spaces available in the entire field. The rotation scheme used C(1) = 1.0 until the maximum allowed number of lesions had formed and then switched to C(2) = 1.0, and then once the maximum allowed number of lesions had formed, the varieties were said to be useless although rotating back to variety 1 would have been feasible. The multiline scheme set the proportions of each variety equal, C(1) = C(2) = 0.5 and, once the maximum allowed number of lesions had been reached, the varieties were said to be useless.

The results of these simulations are given in Table 6.2 with the actual output for some of the simulations given in Table 6.3, Table 6.4, and Table 6.5. Notice that the multiline cropping strategy was consistently less favorable than the rotation strategy for all the various simulation runs with initial lesion numbers of 100 for each race.

However, with the reproductive rates and initial size low enough for the pathogen to stay less than NMAX while growing on a single variety, the population may not grow enough to survive overwintering on the multiline. In this instance, the multiline is more effective as is seen in the simulation output of Table 6.5.

Thus the relative merits of the multiline cropping strategy as compared with the rotation cropping strategy depend upon the rate of increase of the pathogen population during the year as compared with the survival rate over the winter. If the mutiline slows the rate of increase down enough so that the overwinter survival rate gradually lessens the number of pathogen to zero, then the multiline will be superior to the rotation scheme. However, if the rate of increase of the pathogen on the multiline is large with respect to the overwinter survival rate, then the rotation scheme will be superior to the multiline scheme and so in this case, the model gives the same results as Kiyosawa's (1972) study.

The model discussed in this chapter is admittedly very sketchy and preliminary, as are the results of the simple computer simulation. There are many aspects of the problem that should be taken into consideration. Among these, migration into and out of the system could have a profound effect on the results. One would expect migration into the field to enhance the relative effectiveness of the multiline cropping strategy. Another aspect of the problem should be incorporated into the model is Van der Plank's selection against unncessary virulence which could be included in the term w_{ij}. Finally, if this were to be used as a decision model to determine whether to use the rotation strategy or the multiline strategy, many other variables would have to be considered. As we have

mentioned in our discussion of the value of NMAX, both crop yield and cost-effectiveness of the two cropping strategies would need to be considered, as would any other economic aspects that would influence the decision. The effect of the environment would also need to be considered and would probably require the incorporation of a stochastic term into model.

Thus there is much work to be done with this model. Since the parameters of the model are fairly easy to measure and since its application as a decision model could be of value, a further study of this model should and will be conducted in a later work.

Table 6.1. Computer program for the simulation model to determine the growth of the pathogen for various cropping strategies.

```
//PHOST1  WATFIV NCS.ES.G7139,CZOCHOR
      REAL M,NR,N,KS,NL,KST,NMAX
      INTEGER T,TMAX
      DIMENSION NR(2),N(2,2,21),C(2),CT(3),SJN(2,3),SKN(2,2,21),SN(3,2),
     1PH(2),W(2,2),KS(2),CST(2),NL(3,2),IN(2,2,21),SM(2,21)
C THE FOLLOWING ARE THE NUMBER OF HOSTS AND PATHOGEN
      NHOST=2
      NPATH=2
      ROT=0.0
      WRITE(3,141)
  141 FORMAT(4X,'N(1,1) IS :', 3X,'N(1,2) IS :', 3X,'N(2,1) IS :',
     14X,'N(2,2) IS :')
C "NGEN" IS USED TO COUNT THE NUMBER OF GROWING SEASONS
      NGEN=0
C "M" IS THE PROPORTIONATE SIZE OF VIRULENT TO AVIRULENT LESION
      M=3.0
C "T" IS THE NUMBER OF PATHOGEN GENERATIONS PER GROWING SEASON
      T=3
C "NR(I)" IS THE NUMBER OF PATHOGEN RACE "I" AT THE START
      READ(1,120) NR(1),NR(2)
C "PH(I)" IS THE SURVIVAL RATE OVER THE WINTER OF PATHOGEN RACE "I"
      READ(1,121) PH(1),PH(2)
C "KS(J)" IS THE NUMBER OF VIRULENT SIZED LESIONS THAT A FIELD OF HOST
C "J" CAN SUPPORT
      READ(1,122) KS(1),KS(2)
C "W(I,J)" IS THE NUMBER OF NEW LESIONS THAT COULD BE PRODUCED BY
C A PATHOGEN RACE "I" LESION ON A HOST VARIETY "J"
      READ(1,130) W(1,1),W(1,2),W(2,1),W(2,2)
C "C(J)" IS THE PROPORTION  OF THE FIELD PLANTED WITH HOST VARIETY "J"
      READ(1,131) C(1),C(2)
  122 FORMAT(2F10.2)
  120 FORMAT(2F10.2)
  121 FORMAT(2F10.8)
  130 FORMAT(4F10.3)
  131 FORMAT(2F10.8)
C BELOW WE COUNT UP THE NUMBER OF LESIONS OF PATHOGEN RACE "I" ON HOST
C VARIETY "J" AT PATHOGEN GENERATION 1 , "N(I,J,1)"
      DO 7 I=1,NPATH
      DO 6 J=1,NHOST
    6 N(I,J,1)=NR(I)*C(J)
    7 CONTINUE
   30 CONTINUE
      DO 31 J=1,NHOST
   31 CT(J)=C(J)
   23 CONTINUE
      LV=0
C BELOW WE SUM UP "N(I,J,K)" IN A NUMBER OF DIFFERENT WAYS
```

Table 6.1 (Continued).

```
C FIRST WE INITIALIZE ALL VALUES
      DO 11 K=1,T
      DO 13 I=1,NPATH
      SJN(I,1)=0.0
      DO 13 J=1,NHOST
      SKN(I,J,1)=0.0
      SN(1,J)=0.0
      NL(1,J)=0.0
C HERE WE SUM UP THE NUMBER OF LESIONS OF RACE "I" ON HOST "J" THAT
C HAVE OCCURRED OVER THE "K" NUMBER OF PATHOGEN GENERATIONS IN THIS
C GROWING SEASON, "SKN(I,J,K+1)". THUS THE NUMBER THAT HAVE OCCURRED
C BY THE END OF THE GROWING SEASON IS "SKN(I,J,T+1)".
      SKN(I,J,K+1)=N(I,J,K)+SKN(I,J,K)
      IF (I.EQ.J) GO TO 14
      SN(I+1,J)=SKN(I,J,K+1)/M + SN(I,J)
      NL(I+1,J)=N(I,J,K)+NL(I,J)
      IF(N(I,J,1).EQ.0) NL(NPATH+1,J)=1.0
      GO TO 16
   14 SN(I+1,J)=SKN(I,J,K+1)+SN(I,J)
      NL(I+1,J)=N(I,J,K)+NL(I,J)
      IF(N(I,J,1).EQ.0.0) NL(NPATH+1,J)=1.0
      SM(J,1)=0.0
C "SJN(I,J+1)" IS THE SUM OVER ALL HOST VARIETIES "J" OF "N(I,J,K)"
   15 SJN(I,J+1)=N(I,J,K)+SJN(I,J)
   13 CONTINUE
      KST=0.0
      SNT=0.0
C BELOW WE SUM UP ALL VIRULENT LESION SIZED SPACES THAT HAVE BEEN
C OCCUPIED BY BOTH VIRULENT AND AVIRULENT LESIONS DURING THE SEASON
      DO 29 J=1,NHOST
      SNT=SNT+SN(NPATH+1,J)
   29 KST=KST+KS(J)*C(J)
      DO 15 I=1,NPATH
      DO 15 J=1,NHOST
      SM(J,K+1)=SN(NPATH+1,J)
C "SM(I,K+1)" IS THE SUM OF ALL VIRULENT SIZE LESIONS ON HOST "J"
      KL=K+1
C "CST(J)" IS THE PROPORTION OF AVAILABLE VIRULENT LESION SIZED
C SPACE ON HOST VARIETY "J"
      CST(J)=(C(J)-SN(NPATH+1,J)/KS(J))
      IF (C(J).EQ.0.0) GO TO 26
      IF (CST(J).LE.0.0) GO TO 20
C BELOW WE CALCULATE THE SIZE OF THE PATHOGEN POPULATION FOR THE
C NEXT GENERATION
      N(I,J,K+1)=W(I,J)*SJN(I,NHOST+1)*CST(J)
      IF (K.GE.T) GO TO 25
      GO TO 17
   26 N(I,J,K+1) = 0.0
      GO TO 17
C IF THE TOTAL NUMBER OF NEW LESIONS IS CALCULATED TO BE MORE THAN
C A CERTAIN VARIETY HAS SPACE FOR, WE DIVIDE THE NUMBER OF AVAILABLE
C LESION SPACES UP PROPORTIONATELY AND THEN GO TO OVERWINTERING
   20 N(I,J,K)=(N(I,J,K)/NL(NPATH+1,J))*(C(J)*KS(J)-SM(J,K))
C BELOW WE INTRODUCE THE EFFECT OF OVERWINTERING SURVIVAL
   25 N(I,J,K+1)=PH(J)*N(I,J,K)
      LV=1
   17 IN(I,J,K)=N(I,J,K)
```

Table 6.1 (Continued).

```
C THIS SETS N EQUAL TO ZERO IF IT EVER GETS LESS THAN ONE LESION
   15 IF(IN(I,J,K).EQ.0) N(I,J,K+1)=0.0
   21 WRITE(3,140)((IN(I,J,K),J=1,2),I=1,2)
C "KST*(.75)" IS THE MAXIMUM NUMBER OF LESIONS WE WILL TOLERATE, NMAX
      NMAX=KST*(.75)
      IF (SNT.GE.NMAX) GO TO 12
      IF (LV.EQ.1) GO TO 12
   11 CONTINUE
   12 NGEN=NGEN+T
      IF (NGEN.GT.30) GO TO 22
C BELOW WE REINITIATE AND RETURN FOR ANOTHER GROWING SEASON
      DO 24 I=1,NPATH
      DO 24 J=1,NHOST
   24 N(I,J,1)=N(I,J,KL)
      IF (SNT.LT.NMAX) GO TO 23
C IF OUR RESISTANT VARIETIES ARE NO LONGER USEFUL WE PRINT THIS
      DO 35 J=1,NHOST
   35 IF(C(J).EQ.0.0) ROT=1.0 + ROT
      IF(ROT.EQ.0.0.OR.ROT.GT.NHOST) GO TO 32
C BELOW WE ROTATE TO ANOTHER VARIETY TO START THE NEXT SEASON
C SINCE THE OLD VARIETY HAS SUSTAINED TOO MUCH DAMAGE
   27 DO 28 J=1,NHOST
      DO 28 I=1,NPATH
      N(I,J,1)=N(I,NHOST+1-J,KL)
   28 C(J)=CT(NHOST+1-J)
      GO TO 30
C IF WE HAVE MORE LESIONS THAN WE ARE WILLING TO TOLERATE,
C THE MULTILINE IS NO LONGER OF USE.
  143 FORMAT(1X,'THE RESISTANT VARIETIES ARE NOW USELESS',3X,'TMAX=',I3)
C "TMAX" IS THE NUMBER OF GROWING SEASONS THAT THE VARIEIES LAST
   32 TMAX=NGEN/T
      WRITE(3,143)TMAX
   22 WRITE(3,142) (KS(J),J=1,2),((W(I,J),J=1,2),I=1,2),(C(J),J=1,2)
  142 FORMAT(1X,'KS(1)=',F12.2,1X,'KS(2)=',F12.2,1X,'W(1,1)=',F8.2,1X,
     1'W(1,2)=',F8.2,1X,'W(2,1)=',F8.2,1X,'W(2,2)=',F8.2,1X,'C(1)=',
     1F8.6,1X,'C(2)=',F8.6)
  140 FORMAT(1X,4I15)
      WRITE(3,144) T,NHOST,NPATH,M
  144 FORMAT(1X,'T=',I3,1X,'NO. OF HOSTS IS:',I3,1X,'NO. OF PATHOGEN
     1 RACES ARE:',I3,1X,'M=',F6.2)
      STOP
      END
$DATA
  100.0     100.0
   .01        .01
100000.00 100000.00
   40.0       2.0       2.0       40.0
   1.0         0.0
/*
//
```

Table 6.2. Results of computer simulation for two pathogen races where the initial number of lesions of race 1 and of race 2 were set to be NR(1) = NR(2) = 100.

T	W(1,1)	W(1,2)	W(2,1)	W(2,2)	KS	C(1)	C(2)	Results
3	40	2	2	40	10,000,000	1.0	0.0	No pathogen after 7 years.
3	40	2	2	40	10,000,000	0.5	0.5	Useless after 5 years.
4	40	2	2	40	10,000,000	1.0	0.0	No pathogen after 7 years.
4	40	2	2	40	10,000,000	0.5	0.5	Unless after 2 years.
3	40	2	2	40	1,000,000	1.0	0.0	No pathogen after 5 years.
3	40	2	2	40	1,000,000	0.5	0.5	Useless after 3 years.
4	40	2	2	40	1,000,000	1.0	0.0	Useless after 4 years.
4	40	2	2	40	1,000,000	0.5	0.5	Useless after 1 year.
4	50	1	2	60	1,000,000	1.0	0.0	No pathogen after 3 years.
4	50	1	2	60	1,000,000	0.5	0.5	Useless after 1 year.
4	40	4	4	40	1,000,000	1.0	0.0	Useless after 4 years.
4	40	4	4	40	1,000,000	0.5	0.5	Useless after 1 year.
3	40	2	2	40	100,000	1.0	0.0	No pathogen after 7 years.
3	40	2	2	40	100,000	0.5	0.5	Useless after 1 year.
3	50	1	2	60	1,000,000	1.0	0.0	No pathogen after 4 years.
3	50	1	2	60	1,000,000	0.5	0.5	Useless after 2 years.
3	10	.1	.1	10	10,000	1.0	0.0	No pathogen after 2 years.
3	10	.1	.1	10	10,000	0.5	0.5	No pathogen after 4 years.
For NR(1) = 50 and NR(2) = 50								
3	10	.1	.1	10	10,000	1.0	0.0	Still pathogen after 10 years.
3	10	.1	.1	10	10,000	0.5	0.5	No pathogen after 3 years.

Table 6.3. Number of lesions from the simulation for KS(1) = KS(2) = 10,000,000, W(1,1) = W(2,2) = 40, W(1,2) = W(2,1) = 2, and T = 4.

Rotation from host variety 1 to host variety 2 after NMAX has been reached.

N(1,1) is:	N(1,2) is :	N(2,1) is :	N(2,2) is :
100	0	100	0
3999	0	199	0
159930	0	399	0
6292139	0	786	0
62921	0	7	0
2501018	0	15	0
7436052	0	2	0
0	74360	0	0
0	148352	0	0
0	294501	0	0
0	578849	0	0
0	5788	0	0
0	11574	0	0
0	23136	0	0
0	46209	0	0
0	462	0	0
0	924	0	0
0	1848	0	0
0	3696	0	0
0	36	0	0
0	73	0	0
0	147	0	0
0	295	0	0
0	2	0	0
0	5	0	0
0	11	0	0
0	23	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0

Multiline with 50% of the field planted in host variety 1, and 50% in host variety 2.

N(1,1) is:	N(1,2) is :	N(2,1) is :	N(2,2) is :
50	50	50	50
1999	99	99	1999
41981	2099	2099	41981

Table 6.3(continued).

873720	43686	43686	873720
8737	436	436	8737
183155	9157	9157	183155
3696186	184809	184809	3696186

The resistant varieties are now useless.

Table 6.4. Number of lesions from the simulation for KS(1) = KS(2) = 10,000,000, W(1,1) = W(2,2) = 40, W(1,2) = W(2,1) and T = 3.

Rotation from host variety 1 to host variety 2 after NMAX has been reached.

N(1,1) is:	N(1,2) is:	N(2,1) is:	N(2,) is:
100	0	100	0
3999	0	199	0
159930	0	399	0
1599	0	3	0
63961	0	7	0
2541704	0	15	0
25417	0	0	0
1014097	0	0	0
8960485	0	0	0
0	89604	0	0
0	178674	0	0
0	354152	0	0
0	3541	0	0
0	7082	0	0
0	14159	0	0
0	141	0	0
0	283	0	0
0	566	0	0
0	5	0	0
0	11	0	0
0	22	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0

Multiline with 50% of the field planted in host variety 1, and 50% in host variety 2.

N(1,1) is:	N(1,2) is :	N(2,1) is :	N(2,2) is:
50	50	50	50
1999	99	99	1999

Table 6.4(continued).

41981	2099	2099	41981
419	20	20	419
8815	440	440	8815
184775	9238	9238	184775
1847	92	92	1847
38788	1939	1939	38788
807824	40391	40391	807824
8078	403	403	8078
169364	8468	8468	169364
3428328	171416	171416	3428328
34283	1714	1714	34283
714929	35746	35746	714929
4036476	201823	201823	4036476

Resistant varieties are now useless.

Table 6.5. Number of lesions from the simulation for KS(1) = KS(2) = 10,000,000, W(1,1) = (2,2) = 10, W(1,2) = W(2,1) = 0.1, and T = 3.

Rotation from host variety 1 to host variety 2 after NMAX has been reached.

N(1,1) is :	N(1,2) is :	N(2,1) is :	N(2,2) is:
50	0	50	0
496	0	4	0
4686	0	0	0
46	0	0	0
466	0	0	0
4424	0	0	0
44	0	0	0
440	0	0	0
4191	0	0	0
41	0	0	0
417	0	0	0
3982	0	0	0
39	0	0	0
396	0	0	0
3793	0	0	0
37	0	0	0
377	0	0	0
3621	0	0	0
36	0	0	0
360	0	0	0
3465	0	0	0
34	0	0	0
345	0	0	0
3322	0	0	0
33	0	0	0
331	0	0	0
3190	0	0	0
31	0	0	0
318	0	0	0
3068	0	0	0
30	0	0	0
305	0	0	0
2956	0	0	0

Multiline with 50% of the field planted in host variety 1, and 50% in host variety 2.

N(1,1) is :	N(1,2) is :	N(2,1) is :	N(2,2) is :
25	25	25	25
248	2	2	248

Table 6.5 (continued).

1183	11	11	1183
11	0	0	11
50	0	0	59
293	0	0	293
2	0	0	2
14	0	0	14
73	0	0	73
0	0	0	0
0	0	0	0

7. CONCLUSION.

In the previous chapters, we have attempted to study a question of paramount importance to modern agriculture. We have used mathematical models to analyze how best to deploy the available cultivars of a crop to prolong their cumulative resistance to plant pathogens. In order to study this question, an explanation of what was happening in the natural coevolution of a gene-for-gene system was needed. Van der Plank (1963) provided such an explanation with his concept of "stabilizing selection" or selection against unnecessary virulence and against unnecessary resistance.

A number of mathematical models had been developed to study the development of the pathogen races in a gene-for-gene system. Leonard (1977) developed a model for the express purpose of studying Van der Plank's concept and analyzing if it was a viable explanation of what was happening in the natural coevolution of a gene-for-gene system. The development of this model and the preliminary attempts at analysis of it were discussed in Chapter 3. We discovered that Leonard's (1977) model did not adequately model the biological relationship that Leonard had had in mind. The fact that the pathogen population changed its genetic composition with each generation during the growing season and in turn the host population changed its genetic composition as a result of the genetic composition of the pathogen at the end of the growing season, led to the sequential type of model proposed in Chapter 3. This new model produced inconclusive results for the linear stability analysis of the internal singular point, but computer simulations supported the existence of either a locally stable spiral point

or stable limit cycles. In either case, a polymorphism in both the pathogen population and the host population would be preserved.

Nevertheless, when a so-called analogous continuous time model was proposed by Fleming (1980) and shown to produce a center, an analytic technique was needed to show that the discrete time model of Leonard (1977) was not as easily analyzed as the so-called continuous time analog and that in fact it did not produce a center. Such a technique was developed by adapting a method for discerning a center from a focus that Poincaré (1885) had developed for continuous time systems. We have also discussed why this technique was one of the relatively few techniques developed for continuous time systems that was applicable to discrete time systems.

The results of the long and arduous calculations required for the application of this new method for difference equations were in one respect, inconclusive. A completely general analytic result was not feasible since it was necessary to describe what portions of the space would produce either a positive or negative value for a sum of over two hundred terms. The expression was therefore calculated for various values of the parameters by using a computer. In all such cases, except for those where a biologically stable polymorphism was not probable, it was shown that the internal equilibrium point was a stable focus and not a center. Thus it appeared that Van der Plank's concept of "stabilizing selection" is a viable explanation of what was happening in the natural coevolution of a gene-for-gene host-pathogen system.

The only disturbing aspect of this analysis was that we had once again resorted to the computer and had apparently gained relatively

little, in this instance, from this new technique. It did, however, bring up a question of increasingly greater importance in the analysis of models of biological systems. That is, when should any attempt at an analytic answer be given up in favor of a numerical result? Had we needlessly wasted time seeking an analytical result? These questions should always be kept in mind whenever studying any model. The benefits, in terms of a general result, to be gained must be weighed against the cost, in terms of effort and time, of gaining an analytical result. We had originally decided that since the preliminary linear analysis was inconclusive and the numerical analysis did allow a conclusion, then an attempt at any further analysis was too costly. However, when Fleming (1980) published his so-called analog, the possible benefits of an analytic result were heightened until they outweighed the cost; that is, until it became necessary to analyze an expression with an discouragingly large number of terms. Thus, in our opinion, the decision of whether to do an analytical or numerical analysis, depends entirely on a cost-benefit analysis of the question.

Once we had studied this model of the coevolution in nature of a gene-for-gene host-pathogen system, we needed to study the genetic evolution of a pathogen population on a mixed host. Only then could we develop a model to compare the relative merits of a multi-line cropping system with a single variety, pure line cropping system. To this end, the classic model of Levene (1953) for a multiple niche polymorphism and its analog for haploids by Gliddon and Strobeck (1972) were studied. The applicability of the Glidden and Strobeck model to the evolution of the pathogen on the multiline crop was

discussed and a stability analysis of its internal equilibrium was conducted. It was concluded that a multiline cropping system could produce a stable polymorphism in the pathogen population or in other words, that a stable internal equilibrium point for gene frequencies was possible.

Although this model showed that a multipline crop could be used to stabilize gene frequencies at some internal equilibrium point, it told us relatively little about its effectiveness in terms of slowing down or stopping an epidemic. The fact that gene frequencies were stabilized is inconsequential if the pathogen number is increasing rapidly and ends up infecting a large part of the crop. Thus, in order to compare the effectiveness of a multiline crop with the effectiveness of a single variety monoculture that is changed each time the resistance breaks down, it was necessary to develop a model to analyze effectiveness. It was decided that a variety would remain effective until a certain maximum number of lesions were on that variety. Thus instead of using yield or some other quality to measure the loss in effectiveness, we used lesion number and thus developed a model that counted lesion number.

The model was simply a population growth model with the number of new lesions being limited by the available or uninfected plant leaf area. The model was analyzed by computer simulation because of its complexity. It was discovered that in almost all instances, a single variety monoculture, that is replaced each time an unacceptable number of lesions occur, is more effective at prolonging the cumulative lifetime of the varieties than is a multiline made up of the same varieties. The only time a multiline crop seems to be

better is when it slows down the rate of increase of the pathogen enough so that the pathogen cannot overcome the overwintering rate. Thus, if the overwintering rate is very low, the multiline crop may slow the pathogen reproduction enough so that it cannot recover and hence the pathogen would be eradicated. Of course, these results were contingent on the assumptions made in the model.

APPENDICES

Chapter 8. APPENDIX

8.1. Derivation of equation (3.1).

$$n' = \frac{nW_V}{nW_V + (1-n)W_V} = \frac{n\{(1-p)^2(1-k) + [1-(1-p)^2]\,(1-k+a)\}}{n\{(1-p)^2(1-k) + [1-(1-p)^2]\,(1-k+a)\} + (1-n)\,\{(1-p)^2 + [1-(1-p)^2](1-t)\}}$$

$$= \frac{n\{[1-(2p-p^2)](1-k) + (2p-p^2)(1-k+a)\}}{n\{[1-(2p-p^2)]\,[(1-k)-1] + (2p-p^2)[1-k+a-(1-t)]\} + [1-(2p-p^2)(1-1+t)]}$$

$$= \frac{n\{(1-k) + (2p-p^2)a\}}{n\{(2p-p^2)(a+t)-k\} + [1-(2p-p^2)\,t]}$$

$$= \frac{n\{(2p-p^2)a - k + 1\} + n(t-t)(2p-p^2)}{n\{(2p-p^2)(a+t)- k\} + [1-(2p-p^2)t]}$$

$$= \frac{n\{(2p-p^2)(a+t)-k\} + n[1-(2p-p^2)t]}{n\{(2p-p^2)(a+t)-k\} + [1-(2p-p^2)t]}$$

$$p' = \frac{p^2W_{RR} + pqW_{Rr}}{p^2W_{RR} + 2pqW_{Rr} + q^2W_{rr}}$$

$$= \frac{[p^2 + p(1-p)]\{(1-n)[1-c-s(1-t)] + n[1-c-s(1-k+a)]\}}{[p^2 + 2p(1-p)]\{(1-n)[1-c-s(1-t)] + n[1-c-s(1-k+a)]\} + (1-p)^2\{(1-n)(1-s) + n[1-s(1-k)]\}}$$

$$= \frac{p\{[1-c-s(1-t)]-n\ s[t+a-k]}{(2p-p^2)\{[1-c-s(1-t)] - ns[t+a-k]\} + [1-(2p-p^2)][1-s+nsk]}$$

$$= \frac{p\{1-s+nsk\} + p\{-c+st-ns(a+t)\}}{\{1-s+nsk\} + (2p-p^2)\{-c+st-ns(a+t)\}}$$

8.2. Derivation and Proof of equation (3.2).

Equation (3.1) is of the following form:

$$n' = \frac{nA + nB}{A + nB} \qquad p' = \frac{pC + pD}{c + (2p-p^2)D}$$

To show that equation (3.2) is the change in gene frequency when we have j generations of the pathogen per growing season, we shall use mathematical induction. The following is (3.2):

$$n' = n(A+B)^j/\{A^j + n[(A+B)^j - A^j]\} \qquad p' = p(C+D)/\{C+(2p-p^2)D\}$$

Notice for $j = 1$ (3.2) is equivalent to (3.1). Now suppose (3.2) was true if the pathogen had k generations in a growing season, then the frequency after the k^{th}-generation would be:

$$n^{(k)} = n(A+B)^k / \{A^k + n[(A+B)^k - A^k]\}$$

What would be the frequency after $k+1$ generations? From (3.1), we have:

$$n^{(k+1)} = \frac{n^{(k)}(A+B)}{A + n^{(k)}B} = \frac{\left\{\frac{n(A+B)^k}{A^k + n[(A+B)^k - A^k]}\right\}(A+B)}{A + \left\{\frac{n(A+B)^k}{A^k + n[(A+B)^k - A^k]}\right\}B}$$

$$= \frac{n(A+B)^{k+1}}{\{A^k + n[(A+B)^k - A^k]\}\left\{A + \frac{n(A+B)^k}{A^k + n[(A+B)^k - A^k]}B\right\}}$$

$$= \frac{n(A+B)^{k+1}}{\{A^{k+1} + nA[(A+B)^k - A^k]\} + n(A+B)^k B} = \frac{n(A+B)^{k+1}}{A^{k+1} + n[(A+B)^k (A+B) - A^{k+1}]}$$

This is just (3.2) for $j = k + 1$, so 3.2 holds by induction for any number of generations, j, of the pathogen in a growing season.

8.3. Calculation of eigenvalues of the Jacobian of (3.2) evaluated at the trivial singular points (0,0), (0,1), (1,0), (1,1).

For the point (0,0):

$A(0) = 1 \qquad B(0) = -k \qquad C(0) = 1-s \qquad D(0) = ts-c$

The values of the partial derivatives in (3.3) are:

$$\left.\frac{\partial f}{\partial n}\right|_{(0,0)} = \frac{(1-k)^j \cdot 1^j}{\{1^j + 0\}^2} = (1-k)^j$$

$$\left.\frac{\partial f}{\partial p}\right|_{(0,0)} = 0$$

$$\left.\frac{\partial g}{\partial n}\right|_{(0,0)} = 0$$

$$\left.\frac{\partial g}{\partial p}\right|_{(0,0)} = \frac{(1-s)[(1-s) + (ts-c)]}{(1-2)^2} = \frac{1-s+ts-c}{(1-s)} = 1 + \frac{ts-c}{1-s}$$

Thus to find the eigenvalues, we solve:

$$\begin{vmatrix} (1-k)^j - \lambda & 0 \\ 0 & 1 + \frac{ts-c}{1-s} - \lambda \end{vmatrix} = 0$$

So $\lambda_1 = (1-k)^j$ and $\lambda_2 = 1 + \frac{ts-c}{1-s}$. Since $ts > c$, if we are to have an internal equilibrium point, and since $s < 1$, then $|\lambda_2| > 1$. The point (0,0) is locally unstable.

For the point (0,1)

$A(1) = 1-t$ $B(1) = a+t-k$ $C(0) = 1-s$ $D(0) = ts-c$

The values of the partial derivatives in (3.3) are:

$$\frac{\partial f}{\partial n}\Big|_{(0,1)} = \frac{[(1-t) + (a+t-k)]^j (1-t)^j}{[(1-t)^j]^2} = \frac{(1 + a - k)^j}{(1-t)^j}$$

$$\frac{\partial f}{\partial p}\Big|_{(0,1)} = 0$$

$$\frac{\partial g}{\partial n}\Big|_{(0,1)} = \frac{-s(1)(1-1)^2[Dk + C(a+t)]}{[C + (2-1)\,D]^2} = 0$$

$$\frac{\partial g}{\partial p}_{(0,1)} = \frac{(C+D)^2}{(C+D)^2} = 1$$

Thus to find the eigenvalues we solve:

$$\begin{vmatrix} \frac{(1+a-k)^j}{(1-t)^j} - \lambda & 0 \\ 0 & 1-\lambda \end{vmatrix} = 0$$

So $\lambda_1 = \frac{(1+a-k)^j}{(1-t)^j}$ and $\lambda_2 = 1$. Since $t > k - a$ in Leonard's (1977) proposed ranges for the parameters, it follows that $|\lambda_1| > 1$. The point (0,1) is locally unstable.

For the point (1,0)

$A(0) = 1$ $B(0) = -k$ $C(1) = 1-s+ks$ $D(1) = ts-c-s(a+t)$

$$\frac{\partial f}{\partial n}\Big|_{(1,0)} = \frac{(1-k)^j}{(1-k)^{2j}} = \frac{1}{(1-k)^j}$$

$$\frac{\partial f}{\partial p}\Big|_{(1,0)} = 0 \quad \text{since } n = 1$$

$$\left.\frac{\partial g}{\partial n}\right|_{(1,0)} = 0 \qquad \text{since } p = 0$$

$$\left.\frac{\partial g}{\partial p}\right|_{(1,0)} = \frac{C(C+D)}{C^2} = \frac{1-s+ks+ts-c-s(a+t)}{1-s+ks} = 1 - \frac{c+sa}{1-s+ks}$$

Thus to find the eigenvalues we solve:

$$\begin{vmatrix} \frac{1}{(1-k)^j} - \lambda & 0 \\ 0 & 1 - \frac{c+sa}{1-s+ks} - \lambda \end{vmatrix} = 0$$

So $\lambda_1 = \frac{1}{(1-k)^j}$ and $\lambda_2 = 1 - \frac{c+sa}{1-s+ks}$. Since $k < 1$, then $|\lambda_1| > 1$. The point (1,0) is locally unstable.

<u>For the point (1,1)</u>

$A(1) = 1-t \qquad B(1) = a+t-k \qquad C(1) = 1-s+ks \qquad D(1) = ts-c-s(a+t)$

$$\left.\frac{\partial f}{\partial n}\right|_{(1,1)} = \frac{(A+B)^j A^j}{(A+B)^{2j}} = \frac{A^j}{(A+B)^j}$$

$$\left.\frac{\partial f}{\partial p}\right|_{(1,1)} = 0$$

$$\left.\frac{\partial g}{\partial n}\right|_{(1,1)} = 0$$

$$\left.\frac{\partial g}{\partial p}\right|_{(1,1)} = \frac{(C+D)^2}{(C+D)^2} = 1$$

Thus to find the eigenvalues we solve:

$$\begin{vmatrix} \frac{A^j}{(A+B)^j} - \lambda & 0 \\ 0 & 1-\lambda \end{vmatrix} = 0$$

So $\lambda_1 = \frac{A^j}{(A+B)^j} = \frac{(1-t)^j}{(1+a-k)^j}$ and $\lambda_2 = 1$. Since $t > k - a$, according to the proposed parameter values given by Leonard (1977), it follows that $|\lambda_1| < 1$.

We cannot tell from the values of the eigenvalues if the point (1,1) is stable or not. However, we can get an idea by appealing to the original system. If we take a point arbitrarily close to the point (1,1), we can determine whether it is locally stable or not by determining whether the systems trajectory moves toward or away from the point (1,1). So take the point $n = 1$ and $p = 1-\varepsilon$, with ε

arbitrarily small. We then have the following from system (3.2):

$$n' = \frac{n(A+B)^j}{\{A^j + n[(A+B)^j - A^j]\}} = \frac{1\ (A+B)^j}{\{A^j + [(A+B)^j - A^j]\}} = \frac{(A+B)^j}{(A+B)^j} = 1 .$$

Thus if $n = 1$, then $n' = 1$. For $p = 1-\varepsilon$, we have the following:

$$p' = \frac{(1-\varepsilon)[C+D]}{\{C + [2(1-\varepsilon) - (1-\varepsilon)^2]\ D\}} = \frac{(1-\varepsilon)\ [C+D]}{\{C + [2-2\varepsilon-(1-2\varepsilon+\varepsilon^2)]\ D\}}$$

$$= \frac{(1-\varepsilon)\ [C+D]}{[C + (1-\varepsilon^2)D]} = \frac{(1-\varepsilon)\ [(1-s) - (c+sa)]}{\{(1-s) - (1-\varepsilon^2)(c+sa)\}}$$

For values of $c + sa < 1 - s$, we see that $p' < 1 - \varepsilon$ and hence the point (1,1) is unstable. If this isn't the case and $c + sa > 1 - s$, then the virulent pathogen is very highly favored, and we would expect fixation at the stable point (1,1). In other words, for these parameter values, we would not expect, from a biological point of view, a stable internal equilibrium.

8.4. Calculation of the eigenvalues of the Jacobian of (3.5) evaluated at the internal equilibrium point (n^*,p^*).

$$A(p^*) = 1 - \frac{kt}{a+t}$$

$$B(p^*) = 0$$

$$C(n^*) = 1 - s + \frac{(ts-c)k}{a+t}$$

and

$$D(n^*) = 0 \ .$$

$$\left.\frac{\partial f(n,p)}{\partial n}\right|_{(n^*,p^*)} = \frac{(A+B)^j \ A^j}{A^j + n^*[(A+B)^j - A^j]} = \frac{A^{2j}}{A^{2j}}$$

$$\left.\frac{\partial f(n,p)}{\partial p}\right|_{(n^*,p^*)} = \frac{j\ 2(1-p^*)n^*(1-n^*)(A+B)^{j-1}\ A^{j-1}\ \{aA + t(A+B)\}}{\{A^j + n^*[(A+B)^j - A^j]\}^2}$$

$$= \frac{j\ 2(1-p^*)n^*(1-p^*)\ A^{2j-1}\ (a+t)}{A^{2j}} \stackrel{d}{=} jK_1$$

$$\left.\frac{\partial g(n',p)}{\partial n}\right|_{(n^*,p^*)} = \frac{\partial f}{\partial n} \cdot \frac{-sp^*(1-p^*)^2\ [Dk + C(a+t)]}{\{C + (2p^*-p^{*2})\ D\}^2} = 1 \cdot \frac{-sp^*(1-p^*)^2\ D(a+t)}{C^2} \stackrel{d}{=} L_1$$

$$\left.\frac{\partial g(n',p)}{\partial p}\right|_{(n^*,p^*)} = \frac{(C + p^{*2}D)(C+D)}{\{C + (2p^*-p^{*2})D\}^2} + \frac{\partial f}{\partial p} L_1 = \frac{C^2}{C^2} + jK_1L_1$$

Thus to find the eigenvalues, we must solve:

$$\begin{vmatrix} 1 - \lambda & jK_1 \\ L_1 & 1 + jK_1L_1 - \lambda \end{vmatrix} = 0$$

or

$$(1-\lambda)\ (1 + jK_1L_1-\lambda) - jK_1L_1 = 0 \qquad \lambda^2 - (2 + jK_1L_1)\lambda + 1 = 0$$

8.5. Calculation of eigenvalues of the Jacobian of (3.5) evaluated at the trivial singular points (0,0), (0,1), (1,0), (1,1) .

For the point (0,0):

$$A(0) = 1 \qquad B(0) = -k \qquad C(0) = 1-s \qquad D(0) = ts-c$$

$$\left.\frac{\partial f(n,p)}{\partial n}\right|_{(0,0)} = \frac{(A+B)^j\ A^j}{(A^j)^2} = \left(\frac{A+B}{A}\right)^j = \frac{(1-k)^j}{1}$$

$$\left.\frac{\partial f(n,p)}{\partial p}\right|_{(0,0)} = \frac{0 - 0}{A^{2j}} = 0$$

$$\left.\frac{\partial g(n',p)}{\partial n}\right|_{(0,0)} = \frac{\partial f}{\partial n} \cdot 0 \,/\, C^2 = 0$$

$$\left.\frac{\partial g(n',p)}{\partial p}\right|_{(0,0)} = \frac{C(C+D) + D}{C^2} = \frac{C+D}{C}$$

The results are the same as those in Appendix 8.3 for point (0,0).

For the point (0,1):

$A(1) = 1-t \qquad B(1) = a+t-k \qquad C(0) = 1-s \qquad D(0) = ts-c$

$$\left.\frac{\partial f(n,p)}{\partial n}\right|_{(0,1)} = \frac{(A+B)^j\, A^j}{A^{2j}} = \left(\frac{A+B}{A}\right)^j = \left(\frac{1+a-k}{1-t}\right)^j$$

$$\left.\frac{\partial f(n,p)}{\partial p}\right|_{(0,1)} = 0$$

$$\left.\frac{\partial g(n'p)}{\partial n}\right|_{(0,1)} = 0$$

$$\left.\frac{\partial g(n',p)}{\partial p}\right|_{(0,1)} = \frac{(C+D)(C+D)}{(C+D)^2} = 1$$

The results are the same as those in Appendix 8.3 for the point (0,1).

For the point (1,0):

$A(0) = 1 \qquad B(0) = -k \qquad C(1) = 1-s+ks \qquad D(1) = -(c+sa)$

$$\left.\frac{\partial f(n,p)}{\partial n}\right|_{(1,0)} = \frac{(A+B)^j A^j}{(A+B)^{2j}} = \frac{A^j}{(A+B)^j} = \frac{1}{(1-k)^j}$$

$$\left.\frac{\partial g(n',p)}{\partial \hat{p}}\right|_{(1,0)} = 0$$

$$\left.\frac{\partial g(n',p)}{\partial n}\right|_{(1,0)} = 0$$

$$\left.\frac{\partial g(n',p)}{\partial p}\right|_{(1,0)} = \frac{C(C+D) + 0}{C^2} = \frac{(C+D)}{C}$$

The results are the same as in Appendix 8.3 for point (1,0).

For the point (1,1):

$$A(1) = 1-t \qquad B(1) = a+t-k \qquad C(1) = 1 - s + ks \qquad D(1) = -(c+sa)$$

$$\left.\frac{\partial f(n,p)}{\partial n}\right|_{(1,1)} = \frac{A^j(A+B)^j}{(A+B)^{2j}} = \left(\frac{A}{A+B}\right)^j = \left(\frac{1-t}{1+a-k}\right)^j$$

$$\left.\frac{\partial f(n,p)}{\partial p}\right|_{(1,1)} = 0$$

$$\left.\frac{\partial g(n',p)}{\partial n}\right|_{(1,1)} = 0$$

$$\left.\frac{\partial g(n',p)}{\partial p}\right|_{(1,1)} = \frac{(C+D)^2}{(C+D)^2} = 1$$

The results are the same as in Appendix 8.3 for point (1,1).

8.6. Poincare's Method Applied to Leonard's Model.

The Taylor expansion of Leonard's model truncated after the 3rd order can be written as follows:

$$x' = x + K_1y + K_2xy + K_3y^2 + K_4x^2y + K_5xy^2 + K_6y^3$$

$$y' = L_1x + (1 + L_1K_1)y + (L_1K_2 + L_2)xy + (L_1K_3 + L_2K_1)y^2 + (L_3K_1 + L_1K_5 + L_4 + L_2K_2)xy^2$$

$$+ (L_2K_3 + L_1K_6 + \frac{L_1K_1^2}{C_1} + L_4K_1)y^3 \qquad (8.61)$$

$$f_2' = a_2(x + k_1y)^2 + b_2(x + k_1y)\,[(1 + L_1K_1)y + L_1x] + c_1[L_1x + (1 + L_1K_1)y]^2$$

$$- [a_2x^2 + b_2xy + c_2y^2] = 0$$

For the above to be true for all x and y, the following must hold:

$$x^2[a_2 + b_2L_1 + c_1L_2^2 - a_2] = 0$$

$$xy[a_2 2K_1 + b_2[L_1K_1 + (1 + L_1K_1)] + c_1[2L_1(1 + L_1K_1)] - b_2] = 0$$

$$y^2[a_2(K_1^2) + b_2[K_1(1 + L_1K_1)] + c_2(1 + L_1K_1)^2 - c_2] = 0$$

Then, we want a_2, b_2, and c_2 which satisfy the following matrix equation

$$A_2 \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad (8.61)$$

where

$$A_2 = \begin{bmatrix} 0 & L_1 & L_1^2 \\ 2K_1 & (2L_1K_1) & (2L_1^2K_1 + 2L_1) \\ K_1^2 & K_1(1 + L_1K_1) & K_1(2L_1 + L_1^2K_1) \end{bmatrix}$$

If we pre-multiply both side of (8.61) by the following matrix:

$$M_2 = \begin{bmatrix} 1 & L_{1/2} & -L_{1/K_1} \\ -K_1 & 1/2 & 0 \\ 0 & -1/2 & 1/K_1 \end{bmatrix}$$

we get

$$M_2 A_2 = \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

where

$$M_2 A_2 = \begin{bmatrix} 0 & 0 & 0 \\ K_1 & 0 & L_1 \\ 0 & 1 & L_1 \end{bmatrix}$$

Hence, (8.61) is true if $a_2 = L_1$, $b_2 = L_1 K_1$, and $c_2 = -K_1$.

Thus,

$$f_2 = L_1 x^2 + L_1 K_1 xy - K_1 y^2 .$$

$$f_2(x', y') = L_1 (x')^2 + L_1 K_1 (x')(y') - K_1 (y')^2$$

where x' and y' are defined in (8.61). The results are as follows:

$$x^2y[2L_1K_2 - L_1L_2K_1]$$

$$xy^2[2L_1K_3 + L_1K_1K_2 - 2L_2K_1 + L_1^2K_1^2K_2]$$

$$y^3[L_1K_1K_3 - L_1L_2K_1^3 - 2L_2K_1^2]$$

$$x^3y[2L_1K_4 + L_1K_1(L_1K_4)]$$

$$x^2y^2[L_1K_1^2 - L_2^2K_1 - L_1L_2K_1K_2]$$

$$xy^3\{2L_1K_2K_3 - L_1L_2K_1^2K_2 - L_1L_2K_1K_3 - 2L_2^2K_1^2 + L_1[K_6^2 + 2K_1K_5 + L_1K_1[(L_2K_3 + L_1K_6 + \frac{L_1K_1^2}{c_1} + L_4K_1)$$

$$+ K_1(L_3K_1 + L_1K_5 + L_4 + L_2K_2)] - K_1[2L_1(L_2K_3 + L_1K_6 + \frac{L_1K_1^2}{c_1} + L_4L_1) + 2(1 + L_1K_1)(L_3K_1$$

$$+ L_1K_5 + L_4 + L_2K_2)]\}$$

$$y^4\{L_1K_3^2 - L_1L_2K_1^2K_3 - L_2^2K_1^3 + L_1(2K_6K_1) + L_1K_1[K_1(L_2K_3 + L_1K_6 + \frac{L_1K_1^2}{c_1} + L_1K_1)$$

$$+ (1 + L_1K_1)K_6] - K_1[2(1 + L_1K_1)(L_2K_3 + L_1K_6 + \frac{L_1K_1^2}{c_1} + L_4K_1)]\}$$

Thus,

$$f_{23}(x,y) = (2L_1K_2 - L_1L_2K_1)x^2y + [2L_1K_3 + L_1K_1K_2 - 2L_2K_1 + L_1^2K_1^2K_2]xy^2 + [L_1K_1K_3 - L_1L_2K_1^3 - 2L_2K_1^2]y^3$$

Now $f_3' = 0$ for all values of x and y only if the following is true:

$$x^3\{b_3L_1 + c_3L_1^2 + d_3L_1^3\} = 0$$

$$x^2y\{3a_3K_1 + 3b_3L_1K_1 + (2L_1 + 3L_1^2K_1)L_3 + (2L_1K_2 - L_1L_2K_1) + d_3(3L_1^2 + 3L_1^3K_1)\} = 0$$

$$xy^2\{3a_3K_1^2 + (2K_1 + 3L_1K_1^2)b_3 + (4L_1K_1 + 3L_1^2K_1^2)c_3 + (3L_1 + 6L_1^2K_1 + 3L_1^3K_1^2)d_3$$

$$+ [2L_1K_3 + L_1K_1K_2 - 2L_2K_1 + L_1^2K_1^2K_2]\} = 0$$

$$y^3\{a_3K_1^3 + (K_1^2 + L_1K_1^3)b_3 + (K_1 + 2L_1k_1^2 + L_1^2K_1^3d_3 + (3L_1K_1 + 3L_1^2K_1^2 + L_1^3K_1^3)d_3$$

$$+ [L_1K_1K_3 - L_1L_2K_1^3 - 2L_1K_1^2]\} = 0$$

Then, we want a_3, b_3, c_3, and d_3 which satisfy the following matrix equation:

$$A_3\begin{pmatrix} a_3 \\ b_3 \\ c_3 \end{pmatrix} + \begin{matrix} 0 \\ G_1 \\ G_2 \\ G_3 \end{matrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad (8.62)$$

where

$$A_3 = \begin{bmatrix} 0 & L_1 & L_1^2 & L_1^3 \\ 3K_1 & 3L_1K_1 & (2L_1 + 3L_1^2K_1) & (3L_1^2 + 3L_1^3K_1) \\ 3K_1^2 & (2K_1 + 3L_1K_1^2 & (4L_1K_1 + 3L_1^2K_1^2) & (3L_1 + 6L_1^2K_1 + 3L_1^3K_1) \\ K_1^3 & (K_1^2 + L_1K_1^3) & (K_1 + 2L_1K_1^2 + L_1^2K_1^3) & (3L_1K_1 + 3L_1^2K_1^2 + L_1^3K_1^3) \end{bmatrix}$$

and G_1, G_2, and G_3 are as given on page 4.11 of Chapter 4. If we pre-multiply both sides of (8.62) by the following matrix:

$$M_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ (-3K_1) & 1 & 0 & 0 \\ \left(\frac{-2K_1}{L_1}\right) & (-K_1) & 1 & 0 \\ 2K_1^3 & (-K_1)^2 & 0 & 1 \end{bmatrix}$$

we get

$$M_3A_3 = \begin{pmatrix} a_3 \\ b_3 \\ c_3 \\ d_3 \end{pmatrix} + M_3 \begin{matrix} 0 \\ G_1 \\ G_2 \\ G_3 \end{matrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

where

$$M_3A_3 = \begin{bmatrix} 0 & L_1 & L_1^2 & L_1^3 \\ 3K_1 & 0 & 2L_1 & 3L_1^2 \\ 0 & 0 & 0 & L_1(3 + L_1K_1) \\ -2K_1^3 & K_1^2 & K_1 & 3L_1K_1 \end{bmatrix}$$

Note that if $L_1K_1 = -3$, then, (8.62) cannot be solved with any choice of a_3, b_3, c_3, and d_3, since $G_2 - K_1G_1 \neq 0$. The method would then stop here, but otherwise the values given in Chapter 4 for a_3, b_3, c_3, and d_3 will solve (8.62).

Now that we have found $f_3(x,y)$, we can evaluate $f_3(x',y') = a_3(x')^3 + b_3(x')^2(y') + c_3(x')(y')^2 + d_3(y')^3$ to get fourth degree term $f_{34}(x,y)$ plus higher order terms if need be. The fourth degree terms are arrived at as follows:

$$a_3(x')^3 = a_3[(x + K_1y) + (xyK_2 + y^2K_3)]^3 \qquad \text{4th degree terms} \qquad 3a_3(x + yK_1)^2(xyK_2 + y^2K_3)$$

$$b_3(x')^2(y') = b_3[(x + K_1y) + (xyK_2 + y^2K_3)]^2\{[xL_1 + y(1 + L_1K_1)] + [xy(L_1K_1 + L_2) + y^2(L_1K_3 + L_2K_1)]\}$$

$$\text{4th degree terms} \qquad b_3\{(x + yK_1)^2[xy(L_1K_1 + L_2) + y^2(L_1K_3 + L_2K_1)]$$

$$+2(x+yK_1)(xyK_2 + y^2K_3)[xL_1 + y(1 + L_1K_1]$$

$$c_3(x')(y')^2 = c_3[(x + yK_1) + (xyK_2 + y^2K_3)][\{xL_1 + y(1 + L_1K_1)\} + \{xy(L_1K + L_2)$$

$$+ y^2(L_1K_3 + L_2K_1)\}]^2 \quad \text{4th degree terms} \quad c_3\{[(x + yK_1)2\{xL_1 + y(1 + L_1K_1)\}]\ [xy(L_1K_1 + L_2)$$

$$+ y^2(L_1K_3 + L_2K_1)] + (xyK_2 + y^2K_3)[xL_1 + y(1 + L_1K_1)]^2\}$$

$$d_3(y')^3 = d_3\{[xL_1 + y(1 + L_1K_1)] + [xy(L_1K_2 + L_2) + y^2(L_1K_3 + L_2K_1)]\}^3 \quad \text{4th degree terms}$$

$$d_3\{3[xL_1 + y(1 + L_1K_1)]^2\ [xy(yK_2 + L_2) + y^2(L_1K_3 + L_2K_1)]\} \quad .$$

From this, we get the following terms making up $f_{34}(x,y)$:

$$x^3y\{3a_3K_2 + b_3\ [(L_1K_1 + L_2) + 2(L_1K_2)] + c_3[2L_1(L_1K_1 + L_2) + L_1^2K_2] + d_3[3L_1^2(L_1K_2 + L_2]\}$$

$$x^2y^2\{a_3(6K_1K_2 + 3K_3) + b_3[(L_1K_3 + L_2K_1) + 2K_1(L_1K_1 + L_2) + 2\{K_3L_1 + K_1K_2L_1 + K_2(L + L_1K_1)\}]$$

$$+ c_3[2L_1(L_1K_3 + L_2K_L) + 2(1 + L_1K_1)(L_1k_1 + L_2) + 2K_1L_1(L_1K_1 + L_2) + 2L_1(1 + L_1K_1)K_2 + L_1^2K_3]$$

$$+ d_3[3L_1^2(L_1K_3 + L_2K_1) + 6(L_1 + L_1^2K_1)(L_1K_2 + L_2)]\}$$

$$xy^3\{a_3(6K_1K_3 + 3K_1^2K_2) + b_3[2K_1(L_1K_3 + L_2K_1) + K_1^2(L_1K_1 + L_2) + 2\{K_3(1 + L_1K_1) + K_1K_3L_1$$

$$+ K_1K_2(1 + L_1K_1)\}] + c_3[2\{(1 + L_1K_1)(L_1K_3 + L_2K_1) + L_1K_1(L_1K_3 + L_2K_1) + K_1(1 + L_1K_1)$$

$$(L_1K_1 + L_2)\} + K_2(1 + L_1K_1)^2 + 2L_1(1 + L_1K_1)K_3] + d_3[3\{2L_1\ (1 + L_1K_1)(L_1K_3 + L_2K_1) +$$

$$+ (1 + L_1K_1)^2\ (L_1K_2 + L_2)\}]\}$$

$$y^4\{a_3[3K_1^2K_3] + b_3[K_1^2(L_1K_3 + L_2K_1) + 2K_1K_3(1 + L_1K_1)] + c_3[2K_1(1 + L_1K_1)(L_1K_3 + L_2K_1)$$

$$+ K_3(1 + L_1K_1)^2] + d_3[3(1 + L_1K_1)^2\ (L_1K_3 + L_2K_1)]\}$$

We can now write the expressions for F_1, F_2, F_3, F_4 in the sum of the 4th order terms from f_2 and f_3. $f_{24}(x,y) + f_{34}(x,y) = F_1x^2y + F_2x^2y^2 + F_3xy^3 + F_4y^4$.

$$F_1 = [2L_1K_4 + L_1K_1(L_1K_4)] + \{3a_3K_2 + b_3[(L_1K_1 + L_2) + 2(L_1K_2)] + c_3[2L_1(L_1K_1 + L_2) +$$

$$+ L_1^2K_2] + d_3[3L_1^2(L_1K_2 + L_2)]\}$$

$$F_2 = [L_1K_1^2 - L_2^2K_1 - L_1L_2K_1K_2] + \{a_3(6K_1K_2 + 3K_3) + b_3[(L_1K_3 + L_2K_1) + 2K_1(L_1K_1 + L_2)$$

$$+ 2\{K_3L_1 + K_1K_2L_1 + K_2(1 + L_1K_1)\}] + c_3[2L_1(L_1K_3 + L_2K_1) + 2(1 + L_1K_1)(L_1K_1 + L_2)$$

$$2K_1L_1(L_1K_1 + L_2) + 2L_1(1 + L_1K_1)K_2 + L_1^2K_3] + d_3[3L_1^2(L_1K_3 + L_2K_1)$$

$$+ 6(L_1 + L_1^2K_1)(L_1K_2 + L_2)]\}$$

$$F_3 = \{2L_1K_2K_3 - L_1L_2K_1^2K_2 - L_1L_2K_1K_3 - 2L_2^2K_1^2 + L_1[2K_6 + 2K_LK_5] + L_1K_1[(L_2K_3 + L_1K_6$$

$$+ \frac{L_1K_1^2}{c_k} + L_4K_1) + K_1(L_3K_1 + L_1K_5 + L_4 + L_2K_2)] - K_1[2L_1(L_2K_3 + L_1K_6 + \frac{L_1K_1^2}{c_1} + L_4K_1)$$

$$+ 2(1 + L_1K_1)(L_3K_1 + L_1K_5 + L_4 + L_2K_2)]\} + \{a_3(6K_1K_3 + 3K_1^2K_2) + b_3[2K_1(L_1K_3 + L_2K_1)$$

$$+ K_1^2(L_1K_1 + L_2) + 2\{K_3(1 + L_1K_1) + K_1K_3L_1 + K_1K_2(1 + L_1K_1)\}] + c_3[2\{(1 + L_1K_1)(L_1K_3 + L_2K_1)$$

$$+ K_1L_1(L_1K_3 + L_2K_1) + K_1(1 + L_1K_1)(L_1K_1 + L_2)\} + K_2(1 + L_1K_1)^2 + 2L_1(1 + L_1K_1)K_3]$$

$$+ d_3[3\{2L_1(1 + L_1K_1)(L_1K_3 + L_2K_1) + (1 + L_1K_1)^2(L_1K_2 + L_2)\}]\}$$

$$F_4 = \{L_1K_3^2 - L_1L_2K_1^2K_3 - L_1^2K_1^3 + L_1(2K_6K_1) + L_1K_1[K_1(L_2K_3 + L_1K_6 + \frac{L_1K_1^2}{c_1} + L_4K_1)$$

$$+ (1 + L_1K_1)K_6] - K_1[2(1 + L_1K_1)(L_2K_3 + L_1K_6 + \frac{L_1K_1^2}{c_1} + L_4K_1)]\} + \{a_3[3K_1^2K_3] + b_3[K_1^2(L_1K_3$$

$$+ L_2K_1) + 2K_1K_3(1 + L_1K_1)] + c_3[2K_1\ (1 + L_1K_1)(L_1K_3 + L_2K_1) + K_3(1 + L_1K_1)^2]$$

$$+ d_3[3(1 + L_1K_1)^2\ (L_1K_3 + L_2K_1)]\}$$

8.7. Expressions in terms of the original parameters for the coefficients of the Taylor Series expansion.

The values of the various coefficients were arrived at by taking the various partial derivatives of the functions f and g in equation (3.5) and evaluating them at the singular point, in this case the internal singular point (n*, p*). Once this was done and for $A = 1 - \frac{Kt}{a+t}$ and $C_1 = 1-s + \frac{(ts-c)K}{a+t}$, we have the following values for the K_i and L_i with respect to the original parameters.

$$K_1 = \frac{(2-2p^*)n^*(1-n^*)(a+t)}{A}$$

$$K_2 = \frac{(2-2p^*)(1-2n^*)(a+t)}{A}$$

$$K_3 = \frac{n^*(a+t)(n^*-1)}{A} - \frac{(2-2p^*)^2\ n^*[n^*(a+t)-t](a+t)(1-n^*)}{A^2}$$

$$K_4 = \frac{-(2-2p^*)(a+t)}{A}$$

$$K_5 = \frac{n^*(n^*-1)(a+t)^2\ (2-2p^*)^2 - (a+t)(1-2n^*)\ \{A + (2-2p^*)^2[n^*(a+t)-t]\}}{A^2}$$

$$K_6 = \frac{-(2-2p^*)[n^*(a-t)-t]\ (a+t)(n^*-n^*)^2\ \{2A + (2-2p^*)^2\ [n^*(a+t)-t]\}}{A}$$

$$L_1 = \frac{p^*s(k-a-t)}{C_1}$$

$$L_2 = \frac{p^*(2-2p^*)s(a+t) + s(K-a-t)}{C_1}$$

$$L_3 = \frac{2p^*s^2(K-a-t)(2-2p^*)(a+t)}{C_1^2}$$

$$L_4 = \frac{(2-2p^*)s(a+t) - p^*(a+t)s}{}$$

8.8. Proof that A_{1245} is a 5x4 matrix of Rank 4.

If we take the matrix consisting of rows 1, 2, 4, and 5 from matrix A, and display just the first four columns, we have the following matrix:

$$A_{1245} = \begin{bmatrix} 0 & L_1 & L_12 & L_1^3 \\ 4K_1 & 0 & 2L_1 & 3L_1^2 \\ 4K_1^3 & 3K_1^2 & (6L_1K_1^2 + 2K_1) & (6L_1K_1 + 9L_1^2K_1^2) \\ K_1^4 & K_1^3 & (K_1^2 + 2L_1K_1^3) & (K_1 + 3L_1K_1^2 + 3L_1^2K_1^3) \end{bmatrix}$$

By performing the following row operations on A_{1245} in order, we can show that A_{1245} is of rank 4:

1. Divide row 1 by L_1.
2. Divide row 4 by K_1.
3. Multiply row 2 by $-K_1^2$ and add it to row 3.
4. Multiply row 3 by -1/2 and add it to row 4.
5. Multiply row 1 by 2 and add it to row 2.
6. Multiply row 1 by $-3K_1^2$ and add it to row 3.
7. Divide row 3 by $(2K_1 + L_1K_1^2)$.
8. Multiply row 3 by $-L_1$ and add it to row 1.
9. Add row 1 to row 2.
10. Multiply row 2 by $-1/4\ K_1^2$ and add it to row 4.
11. Multiply row 4 by L_1^2 and add it to row 1.
12. Multiply row 1 by $4/(4-L_1K_1^2)$ and add it to row 2.
13. Multiply row 1 by $1/4\ K_1^2$ and add it to row 4.
14. Multiply row 4 by $-2L_1$ and add it to row 3.

14. Multiply row 4 by $-2L_1$ and add it to row 3.

15. Divide row 2 by $4K_1$.

After these row operations have been completed, matrix A_{1245} becomes the following row-equivalent matrix.

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Thus, since the first four columns of A_{1245} are nonsingular, then A_{1245} is of rank 4.

9. LITERATURE CITED

Barrett, J. A. 1978. A model of epidemic development in variety mixtures, in Plant Disease Epidemiology, ed. P. R. Scott and A. Bainbridge. Pages 129-137. Blackwell Scientific Publications, Oxford. 329 pp.

Barrett, J. A. and M. S. Wolfe. 1978. Multilines and super-races -- a reply. Phytopathology 68:1537.

Borlaug, N. E. 1953. New approach to the breeding of wheat varieties resistant to Puccinia graminis tritici. Phytopathology 43:467. (Abstract).

Browning, J. A. and K. J. Frey. 1969. Multiline cultivars as a means of disease control. Ann. Rev. Phytopathology 7:355-82.

Christianson, F. B. 1975. Hard and soft selection in a subdivided population. Amer. Natur. 109:11-16.

Coddington, E. A. and N. Levinson. 1955. Theory of Ordinary Differential Equations. McGraw-Hill, New York. 429 pp.

Crill, P. 1977. An assessment of stabilizing selection in crop variety development. Ann. Rev. Phytopathology 15:185-202.

Crow, J. F. and M. Kimura. 1970. An Introduction to Population Genetics Theory. Harper and Row, New York.

Dempster, E. R. 1955. Maintenance of genetic heterogeneity. Cold Spring Harbor Symposium Quantitative Biology 20:25-32.

Denward, T. 1967. Differentiation in Phytophthora infestans. I. A comparative study of eight different biotypes. Hereditas 58:191-200.

Fleming, R. A. 1980. Selection pressures and plant pathogens: robustness of the model. Phytopathology 70:175-178.

Gliddon, C. and C. Strebeck. 1975. Necessary and sufficient conditions for multiple-niche polymorphism in haploids. Amer. Natur. 109:233-235.

Groth, J. V. 1976. Multilines and "super-races": a simple model. Phytopathology 66:937-39.

Groth, J. V. and C. O. Person. 1977. Genetic interdependence of host and parasite in epidemics. Ann. N. Y. Acad. Sci. 287:97-106.

Jayakar, S. C. 1970. A mathematical model for interaction of gene frequencies in a parasite and its host. Theor. Pop. Biol. 1:140-64.

Jensen, N. F. 1952. Intra-varietal diversification in oat breeding. Agron. J. 44:30-34.

Kiyosawa, S. 1972. Theoretical comparison between mixture and rotation cultivations of disease resistant varieties. Ann. Phytopathol. Soc. Japan 38:52-59.

Kiyosawa, S. and S. Yabuki. 1976. Modeling on the race frequency change in a host-pathogen system with genes for resistance and avirulence. Japan J. Breed. 26:237-46.

Kiyosawa, S. and S. Yabuki. 1976. Simulation of race frequency change in a host-pathogen system with true resistance and avirulence genes. Japan J. Breed. 26:247-55.

LaSalle, J. P. 1978. Stability of dynamical systems. Conference Board of the Mathematical Sciences Regional Conference in Mathematics, Mississippi State University, August, 1976. SIAM.

Leonard, K. J. 1969. Selection in heterogeneous populations of Puccinia graminis f. sp. avenae. Phytopathology 59:1851-57.

Leonard, K. J. 1977. Selection pressures and plant pathogens. Ann. N. Y. Acad. Sci. 287:207-22.

Leonard, K. J. and R. J. Czochor. 1978. In response to 'Selection pressures and plant pathogens: stability of equilibria'. Phytopathology 68:971-73.

Leonard, K. J. and Czochor, R. J. 1980. Theory of genetic interactions among populations of plants and their pathogens. Ann. Rev. Phytopathology 18:237-258.

Levene, H. 1953. Genetic equilibrium when more than one ecological niche is available. Amer. Natur. 87:331-333.

Marshall, D. R. and A. J. Prior. 1978. Multiline varieties and disease control. I. The "dirty crop" approach with each component carrying a unique single resistance gene. Theor. Appl. Genet. 51:177-84.

Marshall, D. R. and A. J. Prior. 1979. Multiline varieties and disease control. II. The "dirty crop" approach with components carrying two or more genes for resistance. Euphytica 28:145-59.

Marshall, D. R. and B. S. Weir. 1982. Multiline varieties and disease control. The "dirty crop" approach with complex mixtures of genotypes. (in preparation).

May, R. M. 1973. Stability and Complexity in Model Ecosystems. Princeton University Press. Princeton, New Jersey.

May, R. M. 1974. Biological populations with nonoverlapping generations: stable points, stable cycles, and chaos. Science 186:647-647.

Mode, C. J. 1958. A mathematical model for the co-evolution of obligate parasites and their hosts. Evolution 12:158-165.

Mode, C. J. 1960. A model of a host-pathogen system with particular reference to the rusts of cereals. In Biometrical Genetics, ed., Kempthorne, O., International Union Biological Sciences, Proc., Series B., No. 38, pp. 84-97.

Mode, C. J. 1961. A generalized model of a host-pathogen system. Biometrics. 17:386-404.

Nelson, R. R. 1972. Stabilizing racial populations of plant pathogens by use of resistance genes. J. Environ. Quality 1:220-27.

Nelson, R. R. 1973. Breeding Plants for Disease Resistance. Pennsylvania State University Press, University Park, Penn.

Nelson, R. R. 1978. Genetics of horizontal resistance to plant disease. Ann. Rev.Phytopathol. 16:359-78.

Nelson, R. R. 1979. The evolution of parasitic fitness. In Plant Disease, Volume IV. How Pathogens Induce Disease., ed. J. G. Horsfall, E. B. Cowling, Academic Press, New York., pages 23-45.

Parlevliet, J. E. and J. C. Zadoks. 1977. The integrated concept of disease resistance: a new view including horizontal and vertical resistance in plants. Euphytica 26:5-21.

Person, C., J. V. Groth, and O. M. Mylyk. 1976. Genetic change in host-parasite populations. Ann. Rev. Phytopathol. 14:177-88.

Poincaré, J. H. 1885. Sur les courbes définies par les équations differentialles. J. de Mathematiques pures et alliqueés. 4ne Serie, Tome premier. Pages 167-196.

Robinson, R. A. 1976. Plant Pathosystems. Springer-Verlag, New York.

Roughgarden, J. 1979. Theory of Population Genetics and Evolutionary Ecology: An Introduction. Macmillan Publishing Co., Inc., New York.

Sedcole, J. R. 1978. Selection pressures and plant pathogens: stability of equilibria. Phytopathology 68:967-70.

Strobeck, C. 1979. Haploid selection with n alleles in m niches. Amer. Natur. 113:439-444.

Trenbath, B. R. 1979. Interactions among diverse hosts and diverse parasites. Ann. N. Y. Acad. Sci. 287 124.150.

van der Plank, J. E. 1963. Plant Diseases: Epidemics and Control. Academic Press, New York.

van der Plank, J. E. 1968. Disease Resistance in Plants. Academic Press, New York.

van der Plank, J. E. 1975. Principles of Plant Infection. Academic Press, New York.

van der Plank, J. E. 1978. Genetic and Molecular Basis of Plant Pathogenesis. Springer-Verlag, New York.

van der Vaart, H. R. 1973. A comparative investigation of certain difference equations and related differential equations: implications for model-building. Bulletin of Mathematical Biology 35:195-211.