MARCUS W. FELDMAN

Mathematical Evolutionary Theory

PRINCETON LEGACY LIBRARY

Mathematical Evolutionary Theory

MATHEMATICAL EVOLUTIONARY THEORY

Edited by Marcus W. Feldman

PRINCETON UNIVERSITY PRESS

PRINCETON, NEW JERSEY

Published by Princeton University Press, 41 William Street,
Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press,
Guildford, Surrey

Library of Congress Cataloging-in-Publication Data

Mathematical evolutionary theory / edited by Marcus W. Feldman.
p. cm.
Collection of papers presented in honor of Samuel Karlin.
Includes bibliographies and indexes.
ISBN 0-691-08502-1 (alk. paper) ISBN 0-691-08503-X
(pbk. : alk. paper)
1. Evolution—Mathematics. 2. Karlin, Samuel, 1923–
I. Feldman, Marcus W. II. Karlin, Samuel, 1923–
QH371.M295 1988 575.01'51—dc19 88-15591

This book has been composed in Lasercomp Times Roman
by Syntax International Private Limited

Clothbound editions of Princeton University Press books
are printed on acid-free paper, and binding materials are
chosen for strength and durability. Paperbacks, although satisfactory
for personal collections, are not usually suitable for library rebinding

Printed in the United States of America by Princeton University Press,
Princeton, New Jersey

Designed by Laury A. Egan

CONTENTS

DEDICATION

The papers in this volume celebrate Samuel Karlin's contributions to mathematical evolutionary theory. His earliest contributions dealt with genetic drift, the stochastic phenomena induced by the finiteness of a population's size. In the late 1960s and 1970s his work addressed the interaction between linkage, selection, and the mating system in a deterministic context. This was succeeded by papers advancing the statistical analysis of data for gene frequencies and familial aggregation. Karlin's most recent studies in the mathematical evolutionary theory have concerned the evolution of behavior and the development of numerical algorithms for comparing and interpreting DNA sequences.

The authors of the papers collected here have all been influenced by Sam Karlin, either as a mentor, collaborator, constructive critic, or through extended stays at Stanford. The papers span the wide range of topics in evolutionary theory to which Sam has contributed. This is our opportunity to acknowledge the profound effect that he has had on the direction and quality of interdisciplinary study in mathematical biology, and on our own research.

PREFACE

Mathematical theory has been central to the study of evolutionary biology since the rediscovery of Mendel's work. Will Provine's masterful biography of Sewall Wright* amply documents the profound effect that this theory has had on the leading natural historians of this century. The mathematical theory of evolution was established by Wright, R. A. Fisher, and J.B.S. Haldane in order to explain observations about variability within and between populations and species. The need to explain the origin and maintenance of this variability led to the dynamical theories of population genetics and, to some extent, to ecology. The need to describe the extant patterns of variation led to far-reaching developments in mathematical statistics.

In the past ten years, application of advances in biotechnology to the study of populations has resulted in the exposure of previously unexpected extensive genetic variability in nature. At the same time, more advanced mathematical technology has been applied to the models designed to describe the origin and maintenance of this variability. Samuel Karlin has, to a great extent, orchestrated this mathematical advance.

Mathematical evolutionary theory spans a range of subjects similar to that covered by population biology as a whole. What are the roles of population size and subdivision on the rate of evolution? Does it make a difference to our picture of the evolutionary process if genes are linked or unlinked? Can the evolutionary effect of departures from random mating such as inbreeding, assortative mating, and sexual selection be quantified? What are the advantages of sex, recombination, and dispersal? Is there a natural framework within which to simultaneously study behaviors and genes that affect these behaviors? What aspects of the environment are most important for the evolution of life-history patterns? What kinds of inference about organismal evolution can legitimately be made from molecular genetic variability? The papers here contribute to our understanding of some of these issues by suggesting new mathematical models, by more extensive analysis of standard models, or by comparing data and theory.

The papers have been divided, somewhat arbitrarily, into two groups. Part I addresses general problems in evolutionary genetics; these do not refer to a specific biological situation, organism, or behavior but take a

* W. B. Provine, *Sewall Wright and Evolutionary Biology* (Chicago: University of Chicago Press, 1986).

general perspective on the mathematical consequences of small population size, mutation, gene linkage, population subdivision, and gene flow. In Part II the theory is developed for specific biological situations, some of which have been associated with general theory, as is the case with kin selection. Both approaches to theory have been important to the modern study of evolution.

Stanford, California
December 1987

Mathematical Evolutionary Theory

PART 1

Stochastic and Deterministic Genetic Theory

The foundation stone of population genetic theory with finite population size is the Wright-Fisher sampling model. It is described in the paper by Ewens and underlies much of the discussion in the papers by Gillespie, Watterson, and Kaplan and Hudson. It also forms the background for part of Tavaré's paper. The binomial (or multinomial) sampling scheme produces a Markov chain that describes the change in the genetic constitution of the population over time. The eigenvectors of this Markov chain have not been found in a useful form, and Wright in 1931 used a diffusion approximation to the discrete time stochastic process in order to obtain information about the transient properties of the gene frequencies

The eigenvalues of the Wright-Fisher Markov chain determine the rate of evolution of the system. The leading nonunit eigenvalue in the simplest case, for example, with N diploid individuals, is $1 - 1/2N$, and this is the rate of approach to homozygosity. When the population is partitioned into two sexes, or undergoes inbreeding, the rate-determining eigenvalue is more complicated, and is often expressed in terms of N_e, the effective population size. Other definitions of effective population size have involved analogies with different properties of the stochastic model.

In his paper (Chapter 1), Ewens discusses the various definitions of effective population size when a population is subdivided into demes. At every generation some of these demes become extinct and are recolonized. The manner of recolonization, fast or gradual, is shown greatly to affect the various effective population sizes and their relative magnitudes.

Much of the huge volume of population genetic data accumulated during the past twenty years has been interpreted in terms of the interaction between drift and mutation. One of the most widely studied genetic models with finite population size includes a constant mutation rate, μ, to novel alleles. This is the infinite alleles model about which so much has been written in the context of the selection-neutrality controversy. The parameter $\theta = 4N\mu$, with N the population size, has been known for many years to govern the distribution of $\{\beta_i\}$, the number of alleles having i representatives $(i = 1, 2, \ldots, n)$ is a sample of size n from such a population.

The magnitude of the compound parameter θ has been a source of controversy between selectionists and neutralists since the mid-1960s. Watterson (Chapter 2) studies the distribution of $\{\beta_i\}$ when θ is not constant over time, analytically and by numerical simulation. In one simulation the mutation rate is held constant and the population is allowed to

cycle from one size to another. In particular, it is shown that the simulated sample mean of the homozygosity agrees well with the expected population homozygosity derived earlier by Maruyama and Fuerst (1985b; see Chapter 2). The variance of the sample mean, however, can be uncomfortably high.

Tavaré (Chapter 3) studies a birth and death process subject to immigration at the time points of a Poisson process. Starting from an initial immigrant propagule, the new arrivals initiate lineages that evolve independently from one another. The stochastic process of interest keeps track of the sizes of the families in the order of their appearance. When conditioned on the population size, the age-ordered family sizes have a joint distribution that is the size-biased version of that derived by Ewens (1972; see Chapter 2) for the number of alleles with i representatives in the infinitely-many neutral alleles model (the β_i of Watterson's paper). The special case of the birth process with immigration provides a simple way to study the genealogical structure of the (age-ordered) stationary infinite alleles model. For example, the asymptotic fractions of the population accounted for by the different families have, when written in decreasing order, the Poisson-Dirichlet distribution that plays so central a role in the mathematical theory of neutral evolution.

The next paper, by Gillespie (Chapter 4), further examines the Wright diffusion approximation. For Wright's diffusion approximation to be mathematically legitimate, parameters such as the rates of mutation, rates of migration, and fitness differences between genotypes must be of the same order of magnitude as the reciprocal of the population size, N, as N increases. In his paper, Gillespie surveys what is known about the properties of limiting processes that emerge when other assumptions are made about the relationship between these key parameters of evolution and the population size.

In the remaining articles of this section the evolutionary forces under study are genotypic selection, recombination, and migration. The interaction of these has been a major focus for population genetic theory since 1970. Linkage disequilibrium is the widely used term for the degree of association between the alleles at a pair of loci. With multiple loci an adequate description of the gamete frequency distribution must involve higher-order associations, that is, associations among three or more genes. Christiansen (Chapter 5) suggests that useful measures of association can be generated from the powers of a single matrix that has proven effective in earlier studies of selection on multiple linked loci (see his formula 4). He calls these *linear interactions*, and shows that they possess desirable properties under iteration in the absence of selection.

In earlier studies of subdivided populations, the linkage disequilibrium between a pair of genes had been shown to depend on the variation in

gene frequencies among the subpopulations. Christiansen extends this principle to multiple loci and to some simple migration arrangements among the subpopulations. It is shown, for example, that the measure of association among a given set of genes depends only on the frequencies of alleles at that set of genes.

Weir and Cockerham (Chapter 6) have been concerned for a number of years with the statistical and stochastic sampling properties of linkage disequilibrium in populations that are sampled with respect to the genotype at two or more loci. They use "descent measures," measures of the probability that specific alleles (or combinations of alleles) in the gametes that unite to form a zygote are identical by descent. In the present paper, they express the frequencies of the ten possible genotypes at two diallelic loci in terms of the two gene frequencies and a set of disequilibrium measures. The latter are defined as specific sums of the genotype frequencies and can be expressed in terms of the descent measures. Some of these disequilibrium measures depend on three or four positions in the two-locus genotype and may not have the same time-dependent behavior as the usual linkage disequilibrium.

With the assumption of multinomial sampling of individuals from a population, maximum likelihood estimates of the various disequilibria and variances of these estimates are computed, although some of the algebra is daunting. These estimates are then used to construct statistics for testing whether specific disequilibria are zero. The test statistics have chi-square asymptotic distributions.

The validity of this procedure was tested numerically with 10,000 replicate samples of size 100. When all disequilibria were zero, the fit of the test statistic to the chi-square distribution was very good. The presence of disequilibria in the underlying population did not have enough of an effect on the testing procedure to invalidate it.

Liberman, Feldman, and Holsinger (Chapter 7) address the notion of evolutionary optimality. Their approach extends the notion of Evolutionary Genetic Stability, originally used in the study of the sex ratio, to neutral genetic modifiers of the rate of migration between two populations. The genetic model is diploid and consists of two loci, one of which has two alleles and is called "major" because it is subject to viability selection on the genotypes. The population is divided into two habitats with arbitrary viability regimes in each. The second locus, the modifier, has an arbitrary number of alleles, and its function is to control genetically the level of migration between the demes. The linkage between the genes is arbitrary.

The first part of the paper develops a class of equilibria at which the migration-modifying alleles have frequencies equivalent to those that would emerge from a one-locus viability model with the migration rates

playing the roles of the genotypic viabilities. It is then shown that at such an equilibrium only those alleles that initially reduce the average migration rate in the population can invade. This suggests that *zero* migration has the property of EGS.

In the Appendix to Chapter 7, internal stability of the equilibrium is analyzed numerically. In view of the authors' earlier work on mutation modification, the equilibrium structure here is surprisingly complicated. In particular, the observation of several instances of cyclic behavior of the genotype frequencies poses interesting questions about the mathematical structure of migration models in population genetics.

CHAPTER ONE

The Effective Population Sizes in the Presence of Catastrophes

Warren J. Ewens

Introduction

My long association with Sam Karlin through our work in mathematical population genetics has been a most memorable and enjoyable experience for me. It started in 1964 when I was a postdoctoral student at Stanford, has continued to this day, and my aim in this paper is to continue it even further. A simplified description of one aspect of our association is as follows: I would become interested in some problem and partially develop its mathematical properties, but would eventually be defeated by some aspect of the mathematical analysis, or not see the full generality of the theory, whereupon I would take the problem to Sam, who would solve the mathematical problem, or generalize the theory (or do both). I do not wish to imply that more than a small fraction of Sam's work in population genetics was initiated in this way, but much of mine was finished along these lines. I remember problems relating to the distribution of the number of alleles maintained in the infinitely many alleles model (Ewens 1963, 1964; Karlin and McGregor 1967), the fundamental theorem of natural selection (Ewens 1969a,b; Karlin and Feldman 1970a and subsequent papers), the unexpected complexities of two-locus systems (Ewens 1968; Karlin and Feldman 1970b and subsequent papers), the evolution of dominance (Ewens 1967; Feldman and Karlin 1971), and in particular the eigenvalues (Ewens and Kirby 1975; Karlin and Avni 1975) and sampling properties (Ewens 1972; Karlin and McGregor 1972) of the infinitely many alleles model, as well as problems of heterozygosity and the effective population size (all the above references). I can never thank Sam enough for the inspiration of this association, and I here present him with a set of further problems, evocative of the those just described, and again concerning eigenvalues, sampling theory, heterozygosity, and effective population sizes, as well as other matters recognizable to the initiated.

So, play it again, Sam: What do you make of this one?

Subdivided Populations and Catastrophes

One of the main concerns of theoretical population genetics is the analysis of the degree and nature of genetic variation in natural populations. The degree of variation is, of course, also of interest in practice to those who are concerned with the loss of genetic variation through random genetic drift, particularly in specific or unusual situations. One circumstance, discussed at length at a recent conference on minimum viable population sizes (Soulé 1987), is that of a large population divided into comparatively small subpopulations that are liable to complete extinction through catastrophic events, with the niche previously occupied by an extinct subpopulation being taken over by migrants from another subpopulation. We discuss here the rate of loss of genetic variation in such cases and the amount of genetic variation maintained when mutation is present, extending the work of Maruyama and Kimura (1980). We find that accepted ideas in this situation do not necessarily hold. This occurs largely because the very subdivision of the large population has an effect on genetic variation often counterbalancing the effects caused by the extinction process.

Many reasonable models of a "catastrophe and recolonization" process may be formed, and any conclusion drawn should not be largely an artefact of the particular model chosen. We consider two models here but it is clear that others are possible. Further, we sometimes reach different conclusions in the two models we do consider. A complete investigation of the process of modeling for this problem is needed and is far from straightforward.

The rate of loss of genetic variation in a population is often measured by calculating one or another concept of the effective population size. These are defined relative to the simple Wright-Fisher model and we therefore start by introducing this model and discussing the effective population sizes to which it leads.

The Wright-Fisher Model and the Effective Population Sizes

Although most populations of interest to us are diploid, our analysis here is more easily carried out, and the main features are not lost, in the haploid case, which we (and others, e.g., Maruyama and Kimura 1980) therefore use throughout. The "simple" Wright-Fisher model (i.e., allowing no selection, mutation, population subdivision, etc.) considers a population of M individuals (or genes) in any generation, each of which is of allelic type A_1 or A_2. It is assumed that random sampling of individuals with re-

placement occurs in choosing the parents forming any daughter generation, so that if there are i A_1 genes in any generation, the probability p_{ij} that in the next generation there will be j A_1 genes is given by

$$p_{ij} = \binom{M}{j}\left(\frac{i}{M}\right)^j\left(1 - \frac{i}{M}\right)^{M-j}, \qquad i, j = 0, 1, 2, \ldots, M. \tag{1.1}$$

Clearly $\mathbf{P} = \{p_{ij}\}$ is the transition matrix of a Markov chain with absorbing states at 0 and M. Many properties of this model are known, in particular

1. $\lambda_{\max}$ = largest nonunit eigenvalue of $\mathbf{P} = 1 - M^{-1}$, (1.2)
2. π = prob (two individuals taken at random have same parent) $= M^{-1}$, (1.3)
3. $\operatorname{var}(x_{t-1}|x_t) = x_t(1 - x_t)M^{-1}$, where x_t is the fraction of individuals in generation t who are A_1. (1.4)

Thus in this model,

$$M = (1 - \lambda_{\max})^{-1}, \tag{1.5}$$

$$M = \pi^{-1}, \tag{1.6}$$

$$M = x_t(1 - x_t)/\operatorname{var}(X_{t+1}|x_t). \tag{1.7}$$

Suppose in a more complicated model, allowing, say, for geographical subdivision, the largest nonunit eigenvalue of the appropriate transition matrix is λ^*. Then the (eigenvalue) effective population size for this model is defined, using (1.5), as

$$N_e^{(e)} = (1 - \lambda^*)^{-1}, \tag{1.8}$$

and the interpretation of this is that insofar as the leading eigenvalue is concerned, the complicated model behaves as a simple Wright-Fisher model of size $N_e^{(e)}$. Similarly, if π^* is the probability that two individuals have the same parent, the (inbreeding) effective population size is defined as

$$N_e^{(i)} = (\pi^*)^{-1}. \tag{1.9}$$

Finally, the variance effective population size $N_e^{(v)}$ (if it exists) is defined, using (1.7), as

$$N_e^{(v)} = x_t(1 - x_t)/\operatorname{var}(x_{t+1}|x_t). \tag{1.10}$$

These three quantities are not necessarily equal, so the expression "effective population size" should not be used without a further adjective describing which of the three is intended. Further, in complicated models they can be very difficult to calculate, and indeed $N_e^{(v)}$ might not even be

well defined, in that there might be no scalar Markovian variable x such that the right-hand side in (1.10) is a constant independent of x_t.

Finally, a fourth definition of effective population size, which we call here the mutation effective population size $N_e^{(m)}$, has been introduced by Maruyama and Kimura (1980), and this new concept should be of considerable value in describing the likely degree of genetic variation in populations subject to mutation. If genes mutate, at rate u, in such a way that every mutant is of an entirely novel allelic type, then genetic variation is maintained in the population. We now quickly summarize the standard theory (Kimura and Crow 1964) assessing this degree of variation for the simple Wright-Fisher model for purposes of comparison with later calculations. If P_{t+1} is the probability that two individuals chosen at random in generation $t+1$ are of the same allelic type, then neither can be a mutant [probability $(1-u)^2$] and they are either both descended from the same parent (probability M^{-1}) or different parents who are of the same allelic type [probability $(1-M^{-1})P_t$]. Thus

$$P_{t+1} = (1-u)^2[M^{-1} + (1-M^{-1})P_t]. \tag{1.11}$$

The stationary probability P is then, exactly,

$$P = (1-u)^2[M - (M-1)(1-u)^2]^{-1}. \tag{1.12}$$

If, in a more complicated model, the stationary value of this probability is P^*, we use (1.12) to define the mutation effective population size exactly as the solution for y of the equation

$$P^* = (1-u)^2[y - (y-1)(1-u)^2]^{-1}. \tag{1.13}$$

$N_e^{(m)}$ is not necessarily equal to $N_e^{(e)}$, $N_e^{(i)}$, or $N_e^{(v)}$, and is to be interpreted simply as the size of a simple Wright-Fisher population having the same stationary value of P as the more complicated model. Unfortunately, $N_e^{(m)}$ is not necessarily independent of the mutation rate u, although as we note later, it will often be rather insensitive to the value of u and also have a well-defined limit as $u \to 0$. It is therefore a potentially valuable parameter in describing an important aspect of the complicated model.

We now consider two models of catastrophes and the values of the various effective population sizes that they define.

Model 1: Total Replacement

Theory

We consider a total population of Mn individuals consisting of n subpopulations with M individuals in each. In each generation, k of the subpopulations become extinct, due to catastrophes, and the niche occupied

by an extinct subpopulation is refilled by randomly choosing one of the surviving subpopulations to produce (apart from its "normal" reproduction of M individuals) a further M to fill the niche. These choices are independent, so that a surviving subpopulation can fill more than one niche. Superimposed on this "demographic" process is a simple Wright-Fisher process within any subpopulation, as described above. Random changes in allele frequency thus arise at two levels, the demographic or "between subpopulations" changes (through the random choice of subpopulations to become extinct), and genetic or "within subpopulations" changes (through the Wright-Fisher process).

A fundamental parameter in this model is the probability q that two individuals chosen at random from distinct subpopulations have parents in the same subpopulation. We find

$$q = k(2n - k - 1)/[n(n - 1)(n - k)]. \tag{1.14}$$

To find the eigenvalue population size for this model, we must first find a Markovian variable having a transition matrix generalizing **P**. The appropriate variable is the vector $(a_1, a_2, \ldots, a_n)$, where $a_1, \ldots, a_n$ are the numbers of A_1 genes in the various subpopulations. Note that $a_1, \ldots, a_n$ are interchangeable: the two states $(a_1, a_2, \ldots, a_n)$ and $(a_2, a_1, \ldots, a_n)$, for example, are regarded as being identical. There are then $R = \binom{M+n}{M}$ states for the process, two of which $[(0, 0, \ldots, 0)$ and $(M, M, \ldots, M)]$ are absorbing, with the rest transient. Thus the matrix of transition probabilities allows two unit eigenvalues and $R - 2$ eigenvalues less than unity in modulus. The method described below for finding these is reminiscent of that used by Ewens and Kirby (1975); is there an approach reminiscent of Karlin and Avni (1975)?

We first write $a_1, \ldots, a_n$ in increasing order $a_1 \leq a_2 \leq a_3 \leq \cdots \leq a_n$, and then list the possible values in dictionary order, as in Table 1.1.

The first two states are now temporarily ignored, and if we also ignore all zeroes, we can define "sample configurations"

$$\{2\}, \{3\}, \ldots, \{M\}, \{1, 1\}, \{1, 2\}, \ldots, \{M, M, \ldots, M\} \tag{1.15}$$

in conformity with the vectors listed in Table 1.1. We write a typical sample configuration $\underset{\sim}{\mathbf{x}} = (x_1, \ldots, x_s)$ and seek the probability $P_{t+1}(\underset{\sim}{\mathbf{x}}) = P_{t+1}(x_1, \ldots, x_s)$ that in generation $t + 1$, a sample of x_1 genes from one subpopulation, x_2 from another, $\ldots$, x_s from an sth subpopulation, are all of the same allelic type. Since a sample of genes from s subpopulations can have parents in at most s subpopulations and since also a sample of x_i genes in any one subpopulation can have at most x_i parents (all, of course, in the same subpopulation), it follows by arguments parallel

Table 1.1
Values of the vector $(a_1, a_2, \ldots, a_n)$.

a_1	a_2	a_3	$\cdots$	a_{n-1}	a_n
0	0	0	$\cdots$	0	0
0	0	0	$\cdots$	0	1
0	0	0	$\cdots$	0	2
	.....				
0	0	0	$\cdots$	0	M
0	0	0	$\cdots$	1	1
0	0	0	$\cdots$	1	2
	.....				
M	M	M	$\cdots$	M	M

to those leading to (1.11) that, apart from an additive constant, $P_{t+1}(x_1, \ldots, x_s)$ is a linear combination of probabilities of the form $P_t(y_1, \ldots, y_u)$, where $(y_1, \ldots, y_u)$ precedes $(x_1, \ldots, x_s)$ in the ordering (1.15). Thus if we form a vector $\underset{\sim}{\mathbf{P}}_{t+1}$ from the $P_{t+1}(\underset{\sim}{\mathbf{x}})$ values, with the $\underset{\sim}{\mathbf{x}}$ values ordered as in (1.15),

$$\underset{\sim}{\mathbf{P}}_{t+1} = \mathbf{D}\underset{\sim}{\mathbf{P}}_t, \tag{1.16}$$

where $\mathbf{D}$ is a triangular matrix. The eigenvalues of $\mathbf{D}$ are the nonunit eigenvalues we seek, and these are its diagonal elements, that is, typically the coefficient of $P_t(\underset{\sim}{\mathbf{x}})$ in $P_{t+1}(\underset{\sim}{\mathbf{x}})$. This is easily seen to be

$$g(\underset{\sim}{\mathbf{x}}) = \left[\prod_{j=1}^{s} \{1 - M^{-1}\}\{1 - 2M^{-1}\}\{1 - 3M^{-1}\} \cdots \{1 - (x_j - 1)M^{-1}\}\right] \times \frac{\binom{n-k}{s}}{\binom{n}{s}} \sum_{j=0}^{s} \binom{s}{j} \frac{k!}{(k-s+j)!} \frac{1}{(n-k)^{s-j}}. \tag{1.17}$$

The latter term in this expression is the probability that s daughter subpopulations have s different parent subpopulations (i.e., is the "demographic" contribution to the eigenvalue), whereas the initial term comes from standard (Feller 1951) eigenvalues for the Wright-Fisher model (i.e., is the "genetic" contribution). The largest eigenvalue depends on the relative

contributions from these two sources, being

$$\lambda^* = \max(1 - M^{-1}, 1 - q), \tag{1.18}$$

and thus

$$N_e^{(e)} = \max(M, q^{-1}). \tag{1.19}$$

The ultimate rate of loss of genetic variation is thus decided entirely either by within population (genetic) factors, as measured by M, or by between population (demographic) factors as measured by q^{-1}, and not by any combination of the two.

It is interesting to confirm the eigenvalues (1.17) in the special case $M = 1$, $k = 1$. The model reduces in this case to the well-known Moran model (1958) of genetics where at unit time points a single individual dies and is replaced, with the dying individual not being a possible parent of the new individual. (This model first interested Karlin [Karlin and McGregor 1962] in genetics.) The largest nonunit eigenvalue in this model is well known to be $1 - 2/n(n-1)$, and this is what $N_e^{(e)}$ reduces to with $M = k = 1$.

It is usually accepted that a subdivided population subject to extinction of subpopulations will lose genetic variation more rapidly than an equally large random-mating population, or equivalently that it has a smaller eigenvalue effective population size. The above shows that this is not necessarily true: for example, when $M = 10^3$, $n = 10^4$, $k = 4$, the eigenvalue effective population size in the subdivided case is 1.25×10^7, whereas the actual population size is only 10^7. We will see later that when mutation exists, the subdivided population can maintain more genetic variation, on the average, than a random-mating population of the same size, again against accepted views.

We find $N_e^{(i)}$ by calculating the probability π^* that two different individuals chosen at random have the same parent. Elementary arguments show that

$$\pi^* = [M - 1 + qM(n-1)]/(Mn - 1)M],$$

so that

$$N_e^{(i)} = M(Mn - 1)/[M - 1 + qM(n-1)]. \tag{1.20}$$

It is interesting to compare the numerical value of $N_e^{(i)}$ with the actual population size Mn. Elementary calculations show that $N_e^{(i)} > Mn$ if and only if

$$qMn < 1. \tag{1.21}$$

This turns out to be a fundamental inequality, which we will return to several times later.

There appears to be no well-defined expression for $N_e^{(v)}$, since there is no scalar Markovian variable in this model. We therefore do not discuss $N_e^{(v)}$ further, other than to remark that the only hope for a reasonable definition of $N_e^{(v)}$ is through the quasi-Markovian variable concept of Norman (1975).

We turn finally to $N_e^{(m)}$, which can be calculated, using (1.13), once we have an expression for P^*. Now two individuals taken from the same subpopulation will be of the same allelic type with probability P given by (1.12). Two individuals taken from different subpopulations will have parents in the same subpopulation (possibly the same parent) with probability q. Thus, if P_d is the probability that two individuals from different subpopulations are of the same allelic type, the identity

$$P_d = (1-u)^2[(1-q)P_d + q\{M^{-1} + (1-M^{-1})P\}]$$

must hold. This gives

$$P_d = qP[1-(1-u)^2(1-q)]^{-1}, \tag{1.22}$$

and hence

$$P^* = [(M-1)P + M(n-1)P_d]/(Mn-1) \tag{1.23}$$

$$= P[M-1+M(n-1)q\{1-(1-u)^2(1-q)\}^{-1}]/(Mn-1). \tag{1.24}$$

This is an exact expression and from it we find, exactly,

$$N_e^{(m)} = M\beta + (\beta-1)/\alpha, \tag{1.25}$$

where

$$\alpha = (1-u)^{-2} - 1$$

$$\beta = (Mn-1)[M-1+M(n-1)q\{1-(1-u)^2(1-q)\}^{-1}]^{-1}.$$

It is necessary to use these exact calculations to develop various properties of $N_e^{(m)}$, which we now do.

First, $N_e^{(m)}$ is dependent on u, but as we note later, this dependence is often rather weak and $N_e^{(m)}$ thus provides a useful parameter for measuring the expected degree of variation in the subdivided population. The limit of $N_e^{(m)}$ as the mutation rate u approaches zero is well defined, being

$$\bar{N}_e^{(m)} = M + M(n-1)(1-q)/[q(Mn-1)]. \tag{1.26}$$

This is close to an expression given by Maruyama and Kimura (1980) in a similar model. Next, the behavior of $N_e^{(m)}$ as u approaches zero is of some interest. Careful handling of small-order terms shows that $N_e^{(m)}$ ap-

proaches its limiting value from *below* when (1.21) holds, and otherwise from *above*. We will return to this property later.

Finally, it is interesting to compare $N_e^{(m)}$ with $N_e^{(e)}$ and $N_e^{(i)}$. Elementary calculations show that $N_e^{(e)} > \bar{N}_e^{(m)}$ if and only if (1.21) holds. So far as the comparison of $N_e^{(m)}$ and $N_e^{(i)}$ is concerned, a simple (but incorrect) argument would make these identical. This argument claims that two randomly chosen individuals will be of identical genetic type if neither is a mutant [probability $(1-u)^2$], and they are either descended from the same parent [probability $(N_e^{(i)})^{-1}$], or different parents [probability $1-(N_e^{(i)})^{-1}$] who are of identical allelic type (probability P^*). This argument would lead to

$$P^* = (1-u)^2[(N_e^{(i)})^{-1} + \{1-(N_e^{(i)})^{-1}\}P^*], \tag{1.27}$$

so that, from (1.13), $N_e^{(i)} = N_e^{(m)}$. However, this cannot be the case, since $N_e^{(m)}$, unlike $N_e^{(i)}$, is a function of u, and the fallacy in the above argument (pointed out to me by R. C. Griffiths) is that the probability that the two (different) parents of two randomly chosen individuals are of the same allelic type is not the same as the corresponding probability for the individuals themselves. Note that this fallacy does not arise in the argument leading to (1.25) and (1.26). Thus the final P^* in (1.27) should be replaced by P^* (parents), the probability that the parents of two randomly chosen individuals who are themselves different individuals are of the same allelic type.

This fallacious argument is nevertheless of use, since it shows that $N_e^{(i)} < N_e^{(m)}$ if and only if $P^*(\text{parents}) < P^*$. Now

$$P^*(\text{parents}) = P\gamma + P_d(1-\gamma), \tag{1.28}$$

where $\gamma = (N-1)/(N_e^{(i)}-1)$ is the probability that two randomly chosen individuals have different parents who are nevertheless in the same subpopulation. Using (1.23) and (1.28), we eventually find that $N_e^{(i)} < N_e^{(m)}$ if and only if (1.21) holds. We have thus shown, by piecing together various of the above conclusions, that if (1.21) holds,

$$Mn < N_e^{(i)} < N_e^{(m)} < \bar{N}_e^{(m)} < N_e^{(e)}, \tag{1.29a}$$

while if (1.21) does not hold,

$$Mn > N_e^{(i)} > N_e^{(m)} > \bar{N}_e^{(m)} > N_e^{(e)}. \tag{1.29b}$$

When $qMn = 1$, all effective population sizes equal the actual population size. These inequalities highlight the importance of the parameter qMn in assessing the likely genetic properties of the subdivided population. They also raise a curious point: $N_e^{(i)}$ and the actual population size are non-genetic concepts, $N_e^{(m)}$ is a totally genetic concept and $N_e^{(e)}$ is a partially genetic concept. Despite this, the various effective population sizes can be

ordered, as in (1.28) and (1.29), according to the value of the single parameter qMn.

Numerical Example

Table 1.2 provides numerical examples of the values of the actual population size $N_a = Mn$, together with the various effective population sizes defined above, for various n and k combinations. There are many points of interest in the table, the most striking being the wide variation possible in the various values. Usually N_a and $N_e^{(i)}$ are quite close, as are $N_e^{(e)}$ and

Table 1.2

Values of the actual population size (N_a), the inbreeding effective size $N_e^{(i)}$, the mutation effective size for $u = 10^{-5}$(a) and 10^{-6}(b), and the limiting value $\bar{N}_e^{(m)}$, together with the eigenvalues effective size $N_e^{(e)}$, for various (n, k) combinations. ($M = 1000$)

Value of n		Value of k 1	10	100
10^3	N_a	10^6	10^6	10^6
	$N_e^{(i)}$	10^6	0.98×10^6	0.83×10^6
	$N_e^{(m)}$(a)	0.51×10^6	0.51×10^6	0.58×10^4
	$N_e^{(m)}$(b)	0.50×10^6	0.50×10^6	0.57×10^4
	$\bar{N}_e^{(m)}$	0.50×10^6	0.50×10^6	0.57×10^4
	$N_e^{(e)}$	0.50×10^6	0.50×10^6	0.47×10^4
10^4	N_a	10^7	10^7	10^7
	$N_e^{(i)}$	10^7	10^7	0.98×10^7
	$N_e^{(m)}$(a)	0.45×10^8	0.50×10^7	0.51×10^6
	$N_e^{(m)}$(b)	0.50×10^8	0.50×10^7	0.50×10^6
	$\bar{N}_e^{(m)}$	0.50×10^8	0.50×10^7	0.50×10^6
	$N_e^{(e)}$	0.50×10^8	0.50×10^7	0.50×10^6
10^5	N_a	10^8	10^8	10^8
	$N_e^{(i)}$	10^8	10^8	10^8
	$N_e^{(m)}$(a)	0.25×10^{10}	0.46×10^9	0.50×10^8
	$N_e^{(m)}$(b)	0.45×10^{10}	0.50×10^9	0.50×10^8
	$\bar{N}_e^{(m)}$	0.50×10^{10}	0.50×10^9	0.50×10^8
	$N_e^{(e)}$	0.50×10^{10}	0.50×10^9	0.50×10^8

$N_e^{(m)}$, but the former two vary from being some two hundred times larger than the latter two to being fifty times smaller. This shows that great care must be taken in assessing whether the subdivided population tends to maintain more variation, or preserve variation longer, than an undivided population. The reason why such large differences between the N_e's are possible (and this is largely an artefact of modeling of the situation) is that the very fact of subdivision tends to preserve genetic variation, and the loss of variation through the random loss of subpopulations might, or might not, be strong enough to offset this. What is perhaps needed is a catastrophe model that does not rely so strongly on subdivision into isolated populations. Migration is a factor that should also be taken into account.

Returning to the above model, the observation that the various effective population sizes can differ radically from each other, together with the two sets of inequalities (1.28) and (1.29) and the condition (1.21), highlights the importance of the parameter qMn in this model. This leads to the first of the problems this model suggests, namely, why is the parameter qMn of such significance, and what is important about the equation $qMn = 1$?

A second problem follows from this. It is well known (Ewens 1979) that for so-called exchangeable models (Cannings 1974), the equations

$$N_e^{(v)} = N_e^{(i)} = N_e^{(e)}$$

are true (assuming $N_e^{(v)}$ is well defined). Is it then true that the above model is exchangeable when $qMn = 1$? Is there a well-defined $N_e^{(v)}$ in this case?

A third problem is to construct a realistic catastrophe model that is not dependent, as is the above model, on population subdivision, so that the effects of catastrophes can be disentangled from the effects of subdivision. A possible model is to assume that some fixed number M^* of individuals dies in a catastrophe in each generation and that all individuals in the daughter generation are descended from the surviving individuals. (Simple models along these lines are exchangeable, so that the various effective population sizes are equal.)

As a final problem, it is unclear to what extent the sampling theory for random mating populations (Ewens 1972; Karlin and McGregor 1972), referred to in several papers in this volume, holds in the subdivided population model: the calculations above give no clues on this question.

Model 2: Gradual Replacement

This model has many of the features of the previous one. There exists a fixed number n of subpopulations, and in each generation k of this number become extinct because of catastrophes. However, the niche previously

occupied by an extinct subpopulation is now filled by one single offspring individual, drawn at random from the surviving individuals. At the same time, the number of individuals in each surviving subpopulation increases by unity. This model is very close to that analyzed by Maruyama and Kimura (1980).

We consider first the nongenetic properties of the model, focusing on the sizes of the various subpopulations. When $k = 1$, these sizes may be written in increasing order,

$$1 = M_1 < M_2 < M_3 < \cdots < M_n. \tag{1.30}$$

Apart from M_1, these sizes are random variables and it can be shown that, at stationarity,

$$\begin{aligned} P(M_i) = \binom{n-1}{i-1}\Bigg[\left(\frac{i-1}{n}\right)^{M_i-1} - \binom{i-1}{1}\left(\frac{i-2}{n}\right)^{M_i-1} \\ + \binom{i-1}{2}\left(\frac{i-3}{n}\right)^{M_i-1} \cdots \\ \pm \binom{i-1}{i-2}\left(\frac{1}{n}\right)^{M_i-1}\Bigg], \\ M_i = i, i+1, i+2, \ldots, \end{aligned} \tag{1.31}$$

$$\begin{aligned} P(M_1, \ldots, M_i) = \binom{n-1}{i-1}\Bigg[\left(\frac{1}{2}\right)^{M_2-1}\left(\frac{2}{3}\right)^{M_3-1}\left(\frac{3}{4}\right)^{M_4-1} \\ \times \left(\frac{i-2}{i-1}\right)^{M_{i-1}-1}\left(\frac{i-1}{n}\right)^{M_i-1}\Bigg], \\ 1 = M_1 < M_2 < M_3 < M_4 < \cdots < M_i, \end{aligned} \tag{1.32}$$

and in particular

$$\begin{aligned} P(M_1, \ldots, M_n) = \left[\left(\frac{1}{2}\right)^{M_2-1}\left(\frac{2}{3}\right)^{M_3-1} \cdots \left(\frac{n-1}{n}\right)^{M_n-1}\right], \\ 1 = M_1 < M_2 < M_3 < M_4 < \cdots < M_n. \end{aligned} \tag{1.33}$$

Any randomly chosen subpopulation has size given by the geometric distribution, so that for general k,

$$P(M) = (k/n)(1 - k/n)^{M-1}, \qquad M = 1, 2, 3, \ldots. \tag{1.34}$$

Thus

$$E(M) = n/k, \qquad \mathrm{Var}(M) = n(n-k)/k^2, \tag{1.35}$$

and the mean total population size is n^2/k. From this, we get an approximation to the probability that two individuals drawn at random are from the same subpopulation:

$$[2n^2(n-k)/k^2] \div (n^4/k^2) \simeq 2(n-k)/n^2. \tag{1.36}$$

This is approximately twice the value arising in the previous model and occurs because the two individuals are more likely to be chosen from the larger subpopulations. It is also possible to derive an expression for the probability distribution of the total population size. However, this calculation is complicated, and we note here only that the mean total population size is (as noted above) n^2/k, with variance (computed by G. A. Watterson) of

$$n^2(n-k)^2/\{k^2(2n-k-1)\}. \tag{1.37}$$

Note that this result is less than the sums of the variances of the individual population sizes, due to a correlation of $-(2n-k-1)^{-1}$ between individual population sizes.

We now add a genetical component to the model. First, suppose any individual is of allelic type A_1 or A_2, and the allelic type of any parent is passed without mutation to each offspring. The population will eventually consist of entirely A_1 or A_2 individuals, and the rate at which this occurs will be measured by the eigenvalue effective size of the model. To calculate this, we must first describe the state of the process (generalizing the vector $\underline{a}$ in Model 1) by a vector of pairs,

$$[(a_1, M_1), (a_2, M_2) \ldots, (a_n M_n)], \tag{1.38}$$

where a_i is the number (out of M_i) of individuals in subpopulation i who are A_1. We next impose a Wright-Fisher model in each subpopulation: if the subpopulation survives the catastrophe at any given time, then in an obvious notation,

$$P[(a_i, M_i) \longrightarrow (b_i, M_i + 1)] = \binom{M_i + 1}{b_i} \left(\frac{a_i}{M_i}\right)^{b_i} \left(1 - \frac{a_i}{M_i}\right)^{M_i + 1 - b_i}. \tag{1.39}$$

The eigenstructure of this process appears to be very difficult, and I can make no progress toward finding $N_e^{(e)}$. (There are, of course, infinitely many eigenvalues.)

An intuitive argument giving an approximation for the inbreeding effective population size is as follows. If at any time the subpopulation sizes are $M_1, M_2, \ldots, M_n$, with $\sum M_i = M$, the probability that two individuals taken at random are from a population size M_i is $M_i(M_i - 1)/M(M-1)$. The probability that they then have the same parent is $1/(M_i - 1)$, and

so the overall probability that two individuals have the same parent is

$$\sum_i M_i(M_i - 1)(M_i - 1)^{-1}/M(M - 1) = (M - 1)^{-1}.$$

The expected value of this quantity is approximately $(n^2/k)^{-1}$, and hence the inbreeding effective population size should be approximately the mean actual population size, n^2/k. It would be desirable to have a more precise argument leading to a more accurate value for $N_e^{(i)}$, using the complete distribution of population size.

To discuss the mutation-effective population size, we now allow mutation at rate u, with all mutants being of novel allelic type. We first calculate P_M, the probability that two individuals taken from a subpopulation of size M are of the same allelic type. Clearly $P_1 = 1$. For $M > 1$, an argument similar to that leading to (1.12) shows that

$$P_M = (1 - u)^2[(M - 1)^{-1} + (M - 2)(M - 1)^{-1}P_{M-1}]. \quad (1.40)$$

G. A. Watterson has pointed out to me that putting $Q_M = (M - 1)P_M$ leads to a very rapid solution of this recurrence relation:

$$P_M = Q[1 - (1 - u)^{2M-2}](M - 1)^{-1}, \qquad M = 2, 3, 4, \ldots, \quad (1.41)$$

where

$$Q = (1 - u)^2/[1 - (1 - u)^2].$$

Since any subpopulation size actually assumes the value M with probability given by (1.34), we may calculate the mean probability $\bar{P}$ that two individuals taken at random from the same subpopulation are of the same allelic type as

$$\bar{P} = \sum P_M(k/n)(1 - k/m)^{M-1} \simeq \alpha \log(1 + \alpha^{-1}), \quad (1.42)$$

where

$$\alpha = k/(2un). \quad (1.43)$$

A numerical check for (1.42) is available. Maruyama and Kimura (1980) considered a model almost identical to the above, the only difference being that catastrophes and population size increases occur in continuous time rather than at discrete time points (as here), with the probability of extinction of a subpopulation in time $(t, t + \delta t)$ being $\lambda\,\delta t$. To compare their simulations with (1.48), it is necessary to put $\lambda = k/n$ (and also their v equal to u). Table 1.3 compares their simulated values of $\bar{P}$ with those calculated from (1.48). We see that the two sets of values are quite close.

It seems much more difficult to arrive at an expression for $\bar{P}_d$, the probability that two individuals from different populations are of the same

Table 1.3
Simulation (Maruyama and Kimura 1980) and approximating theoretical values [from (42) and (45)] of $\bar{P}$ and $P_d/\bar{P}$.

		$\bar{P}$		$P_d/\bar{P}$	
u	λ	Sim.	Theoret.	Sim.	Theoret.
0.05	0.1	0.599	0.693	0.154	0.166
0.02	0.1	0.599	0.841	0.295	0.333
0.01	0.1	0.872	0.912	0.507	0.500
0.001	0.1	0.987	0.990	0.853	0.909
0.002	0.05	0.958	0.962	0.735	0.714
0.005	0.05	0.907	0.912	0.506	0.500

allelic type. The argument leading to (1.22) does not appear to carry over to this model. We note from (1.22) that in Model 1, if q and u are both small,

$$P_d/P \approx 1/(1 + 2u/q) \approx 1/(1 + n^2u/k), \tag{1.44}$$

which for $n = 10$ (the value used in the Maruyama and Kimura simulations) gives

$$P_d/P \approx 1/(1 + 10u/\lambda). \tag{1.45}$$

We also present, in Table 1.3, the simulation values of P_d/P and the values calculated from (1.45). Again, "theoretical" and simulation values are close, although here we have little justification for using (1.45) as it derives from a model different from the present one. Despite this, and encouraged by the simulation results, we now form an approximation to P^*, the probability that two individuals drawn at random are of the same allelic type. We have from (1.36) that

$$P^* \approx 2n^{-1}\bar{P} + (n - 2)n^{-1}P_d,$$

and this is approximately

$$\alpha(\alpha + 1)(\alpha + \tfrac{1}{2}n)^{-1} \log(1 + \alpha^{-1}). \tag{1.46}$$

Letting $u \to 0$ and using (1.13) leads to a suggested approximate limiting mutation-effective population size of $n^2/2k$, about half the mean actual population size. In the case $n = 10^5$, $k = 100$, discussed by Muruyama

and Kimura (1980), this is 5×10^7, in exact agreement with the value they calculate, and half of the mean actual population size (10^8).

The expression (1.37) suggests that the standard deviation in the actual population size is about 2.2×10^5, so that the population size will seldom exceed 1.005×10^8. In other words, the actual population size and the mutation-effective population size will seldom differ by more than a factor of 2.01. The conjecture by Maruyama and Kimura that the actual population size might often be of order 10^{20} (so that there would be an enormous difference between actual and mutation-effective population sizes) seems to be without foundation.

Note that if the mutation-effective population size in this model is about half the actual population size, we reach conflicting conclusions between Models 1 and 2 on the relative values of these quantities when, in Model 1, $qMn < 1$. This emphasizes strongly a point made above and returned to below, that the conclusions we draw should not be artefacts of particular models and that we need a more general approach to the catastrophe problem.

It is clear that in Model 1, explicit expressions can be found for many quantities of interest. However, for Model 2 the best that has been found above, for most quantities of interest, is a set of approximations, sometimes based on tenuous arguments. Some specific problems for Model 2, some of which are very difficult, are the following. First, find the complete set of eigenvalues and hence the eigenvalue effective population size. Second, find the probability that two individuals have the same parent and then find the inbreeding population size. Third, find an exact expression for P^* and thus find the mutation-effective population size. Fourth, is there a parameter (analogous to qMn in Model 1) that determines the relative magnitudes of these effective population sizes and the mean actual population size? Fifth, is there ever a well-defined concept of a variance-effective population size? Sixth, is there a more realistic catastrophe model analogous to Model 2 but not so dependent on population structure? Seventh, can the whole concept of modeling the catastrophe process be considered so that the conclusions drawn are not artefacts of particular models? Finally, can one derive a reasonable sampling theory for Model 2, generalizing the one we found together many years ago for the random mating population case?

References

Cannings, C. C. 1974. The latent roots of certain Markov chains arising in genetics: A new approach, I. Haploid models. *Adv. Appl. Prob.* 6: 260–290.

Ewens, W. J. 1963. The diffusion equation and a pseudo-distribution in genetics. *J. Roy. Stat. Soc. B.* 25: 405–412.

Ewens, W. J. 1964. The maintenance of alleles by mutation. *Genetics* 50: 891–898.
Ewens, W. J. 1967. A note on the mathematical theory of the evolution of dominance. *Amer. Nat.* 101: 35–40.
Ewens, W. J. 1968. A genetic model having complex linkage behaviour. *Theoret. Appl. Genet.* 38: 140–143.
Ewens, W. J. 1969a. Mean fitness increases when fitnesses are additive. *Nature* 221: 1076.
Ewens, W. J. 1969b. A generalized fundamental theorem of natural selection. *Genetics* 63: 531–537.
Ewens, W. J. 1972. The sampling theory of selectively neutral alleles. *Theor. Pop. Biol.* 3: 87–112.
Ewens, W. J. 1979. *Mathematical Population Genetics.* Springer-Verlag, Berlin and NewYork.
Ewens, W. J., and Kirby, K. 1975. The eigenvalues of the neutral alleles process. *Theor. Pop. Biol.* 7: 212–220.
Feldman, M., and Karlin, S. 1971. The evolution of dominance: A direct approach through the theory of linkage and selection. *Theor. Pop. Biol.* 2: 482–492.
Feller, W. 1951. Diffusion processes in genetics. In *Proc. Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 227–246. Ed. J. Neyman. University of California Press, Berkeley.
Karlin, S., and Avni, H. 1975. Derivation of the eigenvalues of the configuration process induced by a labeled direct product branching process. *Theor. Pop. Biol.* 7: 221–228.
Karlin, S., and Feldman, M. 1970a. Convergence to equilibrium of the two-locus additive viability model. *J. Appl. Prob.* 7: 262–271.
Karlin, S., and Feldman, M. 1970b. Linkage and selection: Two-locus symmetric viability model. *Theor. Pop. Biol.* 1: 39–71.
Karlin, S., and McGregor, J. L. 1962. On a genetic model of Moran. *Proc. Camb. Phil. Soc.* 58: 299–311.
Karlin, S., and McGregor. J. L. 1967. The number of mutant forms maintained in a population. In *Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 415–438. Ed. L. Le Cam and J. Neyman. University of California Press, Berkeley.
Karlin, S., and McGregor, J. L. 1972. Addendum to a paper of W. Ewens. *Theor. Pop. Biol.* 3: 113–116.
Kimura, M., and Crow, J. F. 1964. The number of alleles that can be maintained in a finite population. *Genetics* 49: 725–738.
Maruyama, T., and Kimura, M. 1980. Genetics variability and effective population size when local extinction and recolonization of subpopulations are frequent. *Proc. Nat. Acad. Sci. USA* 77: 6710–6714.
Moran, P.A.P. 1958. Random processes in genetics. *Proc. Camb. Phil. Soc.* 54: 60–71.
Norman, M. F. 1975. Diffusion approximation of non-Markovian processes. *Ann. Prob.* 3: 358–364.
Soulé, M. E. 1987. *Viable Populations for Conservation.* Cambridge University Press, New York.