THEORETICAL POPULATION BIOLOGY 3, 210–238 (1972)

Polymorphisms for Genetic and Ecological Systems with Weak Coupling*

SAMUEL KARLIN

Weizmann Institute of Science, Rehovot, Israel
and
Stanford University, Stanford, California

AND

JAMES MCGREGOR

Stanford University, Stanford, California

Received October 4, 1971

1. INTRODUCTION

Recent experimental and observational studies have established conclusively that natural populations are often genetically polymorphic at a large proportion of loci. Specifically, the discovery of vast numbers of isozyme polymorphisms for plant, insect and animal organisms (Lewontin and Hubby, 1966; Harris, 1966; Brown and Allard, 1969; Marshall and Allard, 1969; Kojima and Tobari, 1969a, b, and others) provides incontestible evidence of the considerable variability present in wild populations.

Moreover, a number of animal populations which appear to be practicing partial assortative mating and plant populations which are almost obligate inbreeders, but which do outcross occasionally, exhibit a high degree of polymorphism. A noteworthy fact for many of these highly selfing plant populations where considerable polymorphism is present, is that the frequency of the heterozygotes is small while the homozygotes of the different alleles exist in approximately equal proportions [see Allard, Jain and Workman (1968), Marshall and Allard (1969) and references].

It is perhaps useful to review various known criteria which may lead to stable polymorphism.

* Work was supported in part by National Institute of Health Grant USPRS 10452-09.

© 1972 by Academic Press, Inc.

(i) *Heterozygote advantage.* The most prominent and classical case of selection balance involving two alleles at a single locus.

(ii) *Mutation selection balance.* Selection favoring one genotype and recurrent mutation producing the deleterious allele (Haldane, 1924).[1]

(iii) *Migration selection balance.* Selection benefiting different genotypes in distinct niches with some intermigration mixing the two populations (H. Levine, 1953; Moran, 1959).

(iv) *Temporal variation in selection.* Selection varying in successive generations alternately favoring different genotypes (Haldane and Jayakar, 1963a).

(v) *Haploid–diploid selection balance.* Selection tending to favor different genes during the haploid vs diploid phase (Scudo, 1967).

(vi) *Selection varying between sexes.* Selective advantage conferred on different homozygotes in the two sexes (Owens, 1953; Bodmer, 1965).

(vii) *Frequency-dependent selection.* For example, favoring the rare genotype (Haldane and Jayakar, 1963b; Kojima, 1971; Sheppard, 1958; O'Donald, 1969; and Clarke, 1964).

(viii) *General selection balance.* Polymorphic equilibria resulting from balance of fertility, viability and segregation distortion forces [e.g., see Hartle, Hiraizumi and Crow (1967), Feldman, Nabholz, Bodmer (1969), and Karlin (1968a)].

(ix) *Two-(or multi-) loci selection balance.* Epistasis and recombination effects creating polymorphisms [e.g., Lewontin and Kojima (1960), Lewontin (1964a, b), Karlin and Feldman (1970)].

(x) *Multilocus mutation selection balance.* Combined effects of mutation, recombination and selection (Karlin and McGregor, 1971).

All the above produce cases of polymorphism due more or less to opposing tendencies among the forces of natural selection. The following criteria describe some causes of stable polymorphic situations due to balance in sexual selection (e.g., mating pattern and/or mating behavior) as distinguished from natural selection.

(xi) *Negative assortative mating.* [e.g., Workman (1964), Scudo and Karlin (1969)]. Each phenotype is endowed with a tendency (usually genetically controlled) to mate more frequently with the opposite phenotype.

[1] We have assigned references to these criteria without intending to attribute priorities for their discoveries. The list of indicated references, of course, does not constitute a complete set.

(xii) *Incompatibility and self sterility mechanisms.* See, for example, Finney (1952), Wright (1939, 1969), Workman (1964), and Karlin and Feldman (1968a, b). This includes factors for sex determination, etc.

In almost all mathematical modeling of a genetic system producing a stable polymorphism, it is customarily postulated that there is some clear mechanism operating to the advantage of one of the types and another mechanism favoring an alternative type. The fact that opposing pressures maintain the equilibrium usually means that if one of these pressures is removed the other force effective alone will direct the population to an obvious fixation.

In contrast to cases producing balanced polymorphism we cite several established criteria usually bringing the population to fixation.

(a) *Directional selection in favor of a specific allele.* For example, where the viability of the homozygous genotype A_1A_1 exceeds or equals that of A_1A_2, whose viability exceeds that of A_2A_2. Manifestly, fixation on the A_1A_1 genotype results.

(b) *Disruptive selection.* (The viability of the heterozygote is smaller than the viabilities of both homozygotes.) It is known that the analytic consequences of disruptive selection yield ultimate fixation and which homozygote is fixed depends on the initial composition of the population.

(c) *Assortative mating for dominant traits.* (Preferences or tendencies for mating between similar genotypes) It was found by mathematical analysis [see Scudo and Karlin (1969), and Karlin (1968)] of several genetic models of partial assortative mating for dominant autosomal traits (even with effects of natural selection superimposed) produce consequences similar to those of disruptive selection or those of directed selection for a dominant gene, precluding the existence of stable polymorphism in these situations.

(d) *Selection for two-locus haploid populations.* Feldman (1971) established that in any haploid two-locus population practicing random mating, viability selection pressures alone cannot produce a stable polymorphism.

(e) *Regular inbreeding systems.* Including sibmating, parent–offspring mating, imprinting, consanguinous mating, etc., mostly lead to fixation.

The theoretical bases of fixation indicated in criteria (b)–(d) are not concordant with much observational data. Specifically, in situations of heterogeneous environments, wild populations are regularly exposed to some measure of disruptive selection for many characters; yet a considerable expression of polymorphism persists. In his discussion of the possible consequences of disruptive selection, Mather proposed that in certain circumstances disruptive selection could give rise to polymorphism and in others to isolation. The intuitive argument for this supposes that different optimum phenotypes when

dependent upon one another produce a polymorphic situation and otherwise, fixation or isolation evolves. This problem is taken up in detail in Karlin and McGregor (1972).

Consider next the situation of partial assortative mating [case (c)]. As pointed out above, most situations of populations involving partial assortative mating induced by a dominant trait and, more generally, most inbreeding systems of mating, produce pure populations [e.g., see Crow and Kimura (1970), Karlin (1968), and references therein]. Nonetheless, many animal populations in the presence of partial assortment and many plant populations practicing largely self-fertilization (over 95%) with little outcrossing, exhibit a high degree of polymorphism with no clear selective differences (Allard, Jain and Workman, 1968 and references). What is even more striking in many highly selfing plant populations (e.g., wild oats, barley) where polymorphism is widely present (especially for isozyme characters), the frequency of the heterzygotes is small and a preponderance of homozygotes of all the different alleles are represented in approximately equal proportions. This kind of polymorphism manifested for both isozyme and morphological characters definitely cannot be explained on the basis of heterosis. It has been suggested that frequency-dependent selection may be operating but this rationale seems difficult to justify and is not consistent with data manifesting large numbers of homozygotes. However, polymorphic phenomena for partial assortative mating populations can be based on a multiniche version of the same model (see Section 2).

In case (d) of a two-locus haploid population, it was stated previously that polymorphism cannot be maintained by selection balance alone. Thus, if polymorphism obtains some other influence apart from selection pressures should be operating. Along these lines, it is implicit in the work of Raper and others that random mating is not applicable to a number of haploid models of fungus populations (e.g., Schizophyllum Commune) and indeed, certain incompatibility mechanisms are in force. For such a two-locus haploid population, Feldman showed that a stable polymorphism can be obtained attributable to the incompatibility mating pattern alone. We show in Karlin and McGregor (1972) under wide conditions on the selection parameters, that a polymorphism is possible for the multiniche version of the same two-locus haploid population.

A general tenet will emerge from our analysis to the effect that almost all selection models bear possibilities for stable polymorphism by spreading spatially the selection effects (i.e., to several niches) or temporally (i.e., to several generations) and keeping the interaction between niches or generations slight.

Kimura (1968, 1969a, b), King and Jukes (1969) and some others, in attempting to explicate the vast observed variability and the apparent fast rate of substitution of new alleles, stress the proposition that most allelomorphs are

neutral, existing in populations with small population number where genetic drift is the salient influence. The theory asserts that all loci are either fixed or in a transient state approaching fixation. There is still much active discussion in the literature on these last points.

Most attempts in population-genetics literature to explain polymorphism concentrate on a single or at most a combination of two influences. It is likely that the actual basis for the polymorphism may be attributable to a large variety of sources each contributing in a slight way. Along these lines Ford (1964, 1965) presents many cases of polymorphism inexplicable on the basis of overdominance. Furthermore, he underscores some cases where several contributing influences are clearly involved. Mayr (1963) points up other examples.

In this paper we highlight a principle (perhaps intuitive, maybe surprising, that needs care in its applications) that serves as a basis for generating polymorphism. Its elaboration will provide a quantitative approach to analyzing a mixture of genetic or ecological systems influenced by a variety of pressures and mating patterns. We will find that stable polymorphisms in the combined system can be maintained by one or more dominant effects in each system (which may differ between systems) and small effects due to slight interaction among all the systems. We describe its essential feature loosely now and later a precise formulation will be given.

Suppose there exists a number (say ρ) of ecological or genetic systems $\mathscr{P}_1, \mathscr{P}_2, \ldots, \mathscr{P}_\rho$, e.g., separate communities, niches, with a finite number (say r) of possible types $A_1, A_2, \ldots, A_r$ that may be represented in each system. We generally denote the frequencies of type $A_1, A_2, \ldots, A_r$ in population (or systems) $\mathscr{P}_\alpha$ by $\bar{p}_\alpha = (p_{\alpha 1}, \ldots, p_{\alpha r})$ and frequently subscript α is suppressed when no ambiguity of interpretation is possible. Suppose each system reproduces independently in some fashion such that the frequencies $\bar{p}_\alpha' = (p_{\alpha 1}', \ldots, p_{\alpha r}')$ in the next generation are determined by the relations

$$\bar{p}_{\alpha j}' = f_{\alpha, j}(p_{\alpha 1}, \ldots, p_{\alpha r}), \qquad j = 1, 2, \ldots, r. \tag{1.1}$$

We sometimes write (1.1) in vector notation taking the form

$$\bar{p}' = \bar{f}_\alpha(\bar{p}). \tag{1.2}$$

In most genetic models the transformation (1.1) is displayed as a ratio of two algebraic polynomials in the frequency variables. These transformations naturally reflect mating behavior, segregation pattern, selection, migration and mutation pressures, temporal and spatial (cyclical or other) changes, the influence of recombination when more than one locus is involved and other relevant factors of the process.

In each system $\mathscr{P}_\alpha$ there usually exist certain equilibria [invariant points under the transformation (1.1)] which are locally stable. The collection of

equilibria includes polymorphic and peripheral (i.e., boundary equilibrium) points where in the latter case some types are not represented. Local stability is to be understood in the following generalized sense. A frequency vector p^* is said to be locally stable if for any prescribed neighborhood U of p^* there exists another neighborhood V, $p^* \in V \subset U$ such that $f(\bar{V}) \subset V$ ($\bar{V}$ denotes the closure of V), and therefore the iterates of $\bar{f}_\alpha^{(n)}(\bar{p}) = \bar{f}_\alpha^{(n-1)}(\bar{f}_\alpha(\bar{p}))$ for any starting point $\bar{p} \in \bar{V}$ never depart from V. In most cases, local stability of an equilibrium p^* actually entails that if the initial frequency vector $\bar{p}$ is sufficiently close to p^*, the iterates $f_\alpha^{(n)}(\bar{p})$ indeed converge to p^*.

The notion of stability prescribed above makes no stipulations on the rate of convergence to the equilibrium. However, in most genetic systems when the equilibrium expresses a stable polymorphic balance, then convergence takes place at a geometric rate.[2] But on the other hand when the equilibrium is of the boundary kind (mostly a population of a single type) then convergence not uncommonly occurs at an algebraic rate.

Suppose now that the system $(\mathscr{P}_1, \ldots, \mathscr{P}_\rho)$ is coupled by some form of interaction. The coupling or interaction parameters can be, for example, in the form of *small* (i) mutation pressures, (ii) recombination or crossover probabilities, (iii) migration or outcrossing rates, (iv) seed load reflecting carry over from previous generations, and similarly.

To fix the ideas we will first consider the example when there are migration coefficients $m_{\alpha,\beta}$ with the interpretation that after reproduction a proportion $m_{\alpha,\beta}$ of individuals from $\mathscr{P}_\beta$ migrate to $\mathscr{P}_\alpha$. The recursion relations describing the evolution of type frequencies in the migration-coupled system are then

$$p'_{\alpha,i} = \sum_{\beta=1}^{r} m_{\alpha,\beta} f_{\beta,i}(p_{\beta,1}, \ldots, p_{\beta,r}), \qquad \begin{array}{l} \alpha = 1, \ldots, \rho \\ i = 1, \ldots, r. \end{array} \tag{1.3}$$

It is naturally assumed that $m_{\alpha,\beta} \geqslant 0$ and $\sum_{\beta=1}^{\rho} m_{\alpha,\beta} = 1$.

The migration specified by the ρ-square matrix $M = \| m_{\alpha,\beta} \|$ would be called *weak if M is sufficiently close to the identity matrix*, i.e., the $m_{\alpha,\beta}$ with $\alpha \neq \beta$ are all sufficiently small. In this event the flow between systems is slight. (All subsystems are presumed in the present theory to be of large size.) *If the uncoupled system has a locally stable equilibrium point p^** (i.e., set of equilibrium vectors $\bar{p}_\alpha{}^*$) *then we expect the coupled system to have a locally stable equilibrium point q^* near p^* provided the migration coupling is sufficiently weak.*

The above statement is a special case of a general principle for weakly coupled systems which appears to have a wide variety of important applications to genetic and ecological systems. We formulate in this paper precise conditions

[2] A sequence $\bar{q}_n$ converges geometrically fast to $\bar{q}$ at rate λ ($0 < \lambda < 1$) if the distance between $\bar{q}_n$ and $\bar{q}$ is of the order λ^n. Algebraic convergence means an order of difference like $1/n^\alpha$ for some positive α, usually $\alpha = 1$.

under which the conclusion is valid. It should be emphasized that the conclusion may be violated for strongly coupled systems, in which equilibria behavior of the complex system may be dominated by tendencies in one or more of the subsystems. We do not mean to imply that the only possible situations of polymorphism are those corresponding to distinctive dominant effects in separate niches accompanied by slight interaction among the niches. When migration parameters are large, or more generally, even when the interaction is moderate, the influences operating on the total system become sufficiently mixed such that the evolutionary dynamics and equilibrium behavior of the process is exceedingly difficult to analyze except in a few oversimplified cases. There, obviously, occur in the wild, cases involving strong interaction among subniches, some producing polymorphism and others resulting in fixation.

For specific application of the theory in a number of cases we refer to Karlin and McGregor (1971, 1972b). In the first of these papers a small mutation pressure plays the role of furnishing the requisite mixing. In (1972a) the small parameter is served by the migration rates. In (1972b) small recombination parameters produce the desired interaction effects. See also Eshel (1972) for additional applications of this method of small parameters.

Usually one can parametrize the strength of coupling by one or more small parameters (in the specific example cited previously, by the $m_{\alpha,\beta}$ with $\alpha \neq \beta$), and show that as the parameters go to zero there is a unique equilibrium point q^* which approaches p^*.

We say a *full* polymorphism is attainable in the *uncoupled* system $(\mathscr{P}_1, \mathscr{P}_2, \ldots, \mathscr{P}_\rho)$ if there is a set of equilibrium frequency vectors $p_\alpha{}^*$, $\alpha = 1, 2, \ldots, \rho$ where $p_\alpha{}^*$ is a *locally stable* solution of

$$p_\alpha{}^* = f_\alpha(p_\alpha{}^*), \qquad \alpha = 1, 2, \ldots, \rho$$

such that for each i, $1 \leqslant i \leqslant r$, $p^*_{\alpha,i}$ is positive for at least one α (α may depend on i). *The set of equilibrium frequency vectors p^* comprise a fixed point $p_\alpha{}^*$ of the uncoupled system in which every possible type is represented in at least one subsystem.* Now if q^* is the nearby equilibrium point of the system under weak coupling then every type will still be represented in at least one subsystem provided the coupling is sufficiently weak. In the migration-coupling example there is a simple condition which will guarantee that at the equilibrium state q^* each possible type is actually represented in every subsystem. The matrix $M = \| m_{\alpha,\beta} \|$ is called irreducible if there is a power M^k whose elements are all strictly positive. This means that a k-th generation descendent of an individual from any subsystem has a positive probability to be in any other subsystem.

PRINCIPLE I. *If a full polymorphism (defined in paragraph above) is attainable in the uncoupled system and if the migration matrix is irreducible and sufficiently weak (i.e., M close to I = identity matrix), then the coupled system has a locally*

stable equilibrium state in which actually every possible type is represented in every subsystem.

Probably the crux of this assertion is intuitive (although not always valid) to some theoretical biologists and undoubtedly has been expressed in some qualitative terms by others (consult the discussion on this point).

As a simple application of the principle stated above, consider the case of two alleles A_1 and A_2 at a single locus with two niches with selection coefficients $1, 1 - \sigma, 1$ for the genotypes A_1A_1, A_1A_2, and A_2A_2, respectively, or more generally, we can suppose that the fitness coefficients of the genotypes are s_1, s_2, s_3 in niche 1, $\sigma_1, \sigma_2, \sigma_3$ in niche 2 where $s_2 < \min(s_1, s_3)$ and $\sigma_2 < \min(\sigma_1, \sigma_3)$. Thus, in each separate niche disruptive selection operates and ordinarily the population would be fixed. However, invoking the conclusion of our Principle I we infer if a small fraction (necessarily small) of the population of each niche migrates to the other niche then there are possible sets of stable polymorphisms with both alleles represented in each population. Of course, the possibility of stable global fixations also occur.

The polymorphic equilibria of this example have the property that a preponderance of one homozygote occurs in one niche while a preponderance of the alternative homozygote is maintained in the second niche. It is reasonable to speculate that some forms of habitat selection confer an advantage, say on A_2A_2 in niche 2, on A_1A_1 in niche 1, while the heterozygote (or hybrid type) bears marked disadvantage to both homozygotes in each of the niches yet a global balance is preserved (consistent with data of Allard on certain traits of plant populations).

What evolves depends crucially on the initial composition of all the subpopulations. Thus, whether fixation transpires or polymorphism is attained could be a function of founders and random fluctuation effects determining the initial conditions. Small colonies of different homozygotes could inhabit neighboring localities with selection favoring both homozygotes over the heterozygotes in each locality. Subsequently, population size grows and presumably some slight gene flow binds the two localities. ["Perfect geographic isolation of two populations is most usually impossible" (Darlington, 1957).] A suitable application of the principle then points to a polymorphism with most existing types being homozygotes [see also Karlin and McGregor (1972a) for a more detailed quantitative analysis of this two-niche model].

The extension to the case of three alleles is as follows. Consider a three-allele model involving alleles A_1, A_2, A_3 with viabilities of A_iA_j specified by the matrix with the obvious interpretation

$$\begin{pmatrix} 1 + \alpha_1 & 1 - \epsilon_1 & 1 - \epsilon_2 \\ 1 - \epsilon_1 & 1 + \alpha_2 & 1 - \epsilon_3 \\ 1 - \epsilon_2 & 1 - \epsilon_3 & 1 + \alpha_3 \end{pmatrix} \qquad \epsilon_i, \alpha_i > 0, \qquad i = 1, 2, 3,$$

so that each homozygote is favored. Consider two replicate systems of the above structure with slight migration between them. It can be proved that there exists no stable polymorphism with all genotypes represented. (The proof involves the converse version of Principle II, stated as Principle III below.) However, if the above 3-allele genetic population is replicated in *three* systems with slight gene flow between them, then a stable polymorphism is possible involving all types. (The proof is accomplished by application of Principle I.)

We now state the principle in more general form.

PRINCIPLE II. *If a system of transformations acting on a certain set (in finite-dimensional space) has a "stable" fixed point then a slight perturbation of the system maintains a stable fixed point nearby.*

The theme of this principle is quite intuitive although care in its application and interpretation is vital. Its validity does not require the stability hypothesis to apply in a geometric sense. In fact, for numerous important genetic models the stability of the relevant equilibrium is manifested only in an algebraic sense (e.g., this happens in the case of assortative mating).

The precise mathematical statement of Principles II and III and refinements and extensions are elaborated in Sections 3 and 4. In Section 2 the theorem is applied to a multiniche model of partial assortative mating. Other applications are set forth in the companion paper of Karlin and McGregor. The final section discusses some implications and limitations of these principles.

A converse proposition to Principle II of considerable value in ascertaining *all* possible equilibria is stated in rough form now. For a precise mathematical statement the reader should consult the discussion of Theorem 5 in Section 4.

PRINCIPLE III. *If $f(x)$ is a differentiable transformation acting on a certain set S (in finite-dimensional space) having a finite number of fixed points, say $y_1, y_2, \ldots, y_r$, with the property that the linear approximation to $f(x)$ in the neighborhood of each fixed point has no eigenvalue of absolute value one, then a slight differentiable perturbation of $f(x)$ maintains at most a single fixed point $z_i \in S$ in the neighborhood of each y_i. Moreover, z_i is locally stable iff y_i is locally stable.*

It is worth noting that some fixed points of $f(x)$ (but none of the stable ones) may disappear under small perturbations.

We refer to Karlin and McGregor (1972a) for further application of Principle III.

2. Application of Principle II to Multiniche Models of Partial Assortative Mating

Many organisms have polymorphic characters associated with preferences for mating between similar individuals (assortative mating) or preference for mating between dissimilar individuals (disassortment). Such assorting is known to occur in certain bird, mammal (man included), insect, and plant populations.[3] Known examples of a Mendelian factor affecting mating frequencies include those associated with pigment color or pattern (Mainardi, 1968; Parsons, 1962).

Scudo and Karlin (1969), Karlin and Scudo (1969), and Karlin (1968) investigated through a series of genetic models, the effects of partial assortative mating on the equilibrium nature of populations incorporating certain ecological factors and more of the mechanics of assortment. It was found in these works that partial assortment for dominant autosomal genes will almost always have an effect similar to disruptive selection or that of directed selection for a dominant gene, and it will never produce stable polymorphism. In all these cases, convergence to the pure recessive state occurs at a geometric rate while that to a pure dominant state occurs more slowly at an algebraic rate. We review briefly the results for two of the simpler models.

Let X, Y, Z denote the frequencies of AA, Aa and aa, respectively, and assume A is dominant to a. Degrees of assortment in the phenotypes are measured by two parameters: $\alpha(0 \leqslant \alpha < 1)$ will be the fraction of dominants preferring to mate with their kinds, $\beta(0 \leqslant \beta < 1)$ that of recessive females preferring their kind. The frequencies of the mating types are listed in the array below (Table I).

TABLE I

Frequencies of Mating Types

Mating type	Model I	Model II	Random mating
	Assorting		
$AA \times AA$	$\alpha X^2/(X+Y)$	αX^2	$(1-\alpha)X^2$
$AA \times Aa$	$2\alpha XY/(X+Y)$	$2\alpha XY$	$2(1-\alpha)XY$
$Aa \times Aa$	$\alpha Y^2/(X+Y)$	αY^2	$(1-\alpha)Y^2$
$AA \times aa$			$(2-\alpha-\beta)XZ$
$Aa \times aa$			$(2-\alpha-\beta)YZ$
$aa \times aa$	βZ	βZ^2	$(1-\beta)Z^2$

[3] Assortative mating occurs in plant populations by the nature of the vector carrier of pollen, attracted by different pigment color of flower or shape, etc.

An underlying postulate in Model I is that all females are fertilized while in the second model there is a loss in fertility accompanying the effort of assortative mating. For justification, relevance of these formulations, and further discussion of other partial assortative models incorporating further ecological and genetic factors, the reader is directed to the papers referred to previously.

The recurrence formulas connecting frequencies in successive generations reduce to

Model I

$$p' = p(1 + \frac{\alpha - \beta}{2} Z), \qquad \text{where} \quad p = X + \frac{Y}{2}, \quad p' = Z' + \frac{Y'}{2}. \tag{2.1}$$

Model II

$$\begin{aligned} p' &= p \frac{(1 - aZ)}{1 - 2aZ(1 - Z)} \qquad p = X + Y/2, \\ Z' &= \frac{(1 - p)^2 - 2a(1 - p)Z + 2aZ^2}{1 - 2aZ(1 - Z)} \qquad (a = \alpha + \beta). \end{aligned} \tag{2.2}$$

It follows on the basis of (2.1) that

(i) Where $\alpha > \beta$, the only stable state for the population is that with only dominants ($X = 1$, $Y = Z = 0$), convergence taking place at an algebraic rate.

(ii) When $\alpha < \beta$, a pure recessive population is the only stable state and it is approached geometrically fast.

For Model II there exists a unique interior equilibrium (locally unstable)

$$\hat{Z} = \frac{1}{2}, \qquad \hat{p} = 1 - \frac{a + \sqrt{[(2 - a)(1 - a)]}}{2}.$$

If initially $Z_0 < \frac{1}{2}$, $p_0 > \hat{p}$ fixation of A takes place at an algebraic rate, while if $Z_0 > \frac{1}{2}$, $p_0 < \hat{p}$ fixation of a occurs.

Thus in both these models (and in the others as well), except for exceptional relations among the parameters, partial assortment based on a dominant effects the population dynamics by causing either gene substitution or has effects analogous to disruptive selection. On the other hand, there is much evidence that polymorphic populations practicing partial assortment occur widely in nature.

The general Principle II provides a basis for polymorphism with partial assortment. Indeed, consider at least two replica populations of Model II or two populations of Model I reproducing in two niches with parameters $\alpha_1, \beta_1 (\alpha_1 > \beta_1)$ relevant at the first niche and $\alpha_2, \beta_2 (\alpha_2 < \beta_2)$ at the second

niche. Thus in the last example, assortment for the dominant is stronger in one niche while that for the recessive is stronger in the second niche. Assume slight but positive migration between the two populations. Then a stable polymorphism can be attained. The validation of this assertion invokes Principle II. (Note that only algebraic convergence takes place at some of the pure equilibria in a single population.)

3. Formal Theorems and Proofs

The next two sections are technical and record a mathematical expression of the content of Principles I–II enunciated in Section 1. We develop the proofs here to have them available. The biologically oriented reader may pass over these sections (but perhaps a cursory look at Section 3 is advisable) and consult the discussion as well as the papers devoted to examples, implications and applications of these theorems; see Karlin and McGregor (1971, 1972a).

In this section we discuss the theorems relevant to the simple case when there are just two subsystems $\mathscr{P}_1$, $\mathscr{P}_2$ each with a population comprised of two types A, a and with simple migration coupling. Theorems 3.1 and 3.2 are prototypes of more general theorems proved in Section 4.

Let x_1, x_2 be the frequencies of A in $\mathscr{P}_1$, $\mathscr{P}_2$, respectively. For the uncoupled systems let the recursion relations giving the frequencies x_1', x_2' in the next generation be

$$\begin{aligned} x_1' &= f_1(x_1), \\ x_2' &= f_2(x_2). \end{aligned} \tag{3.1}$$

These functions are assumed to be continuous and to satisfy $0 \leqslant f_i(\xi) \leqslant 1$ when $0 \leqslant \xi \leqslant 1$, so that (3.1) defines a continuous mapping of the square

$$\Omega = \{(x_1, x_2); 0 \leqslant x_i \leqslant 1\}$$

into itself. We assume the uncoupled systems have equilibrium frequencies y_1, y_2, that is,

$$y_1 = f_1(y_1), \qquad y_2 = f_2(y_2),$$

so the mapping (3.1) has a fixed point (y_1, y_2) in the square Ω.

We now impose the condition of local stability for the fixed point (y_1, y_2), and at first the discussion will be restricted to the case when there is convergence at a geometric rate. For the applications of the simple kind discussed in this section it would be satisfactory to assume that f_1 and f_2 are differentiable at y_1 and y_2, respectively, and

$$|f_i'(y_i)| < 1, \qquad i = 1, 2. \tag{3.2}$$

However, there are multivariate examples in which the analogous assumption of differentiability is not satisfied. It is therefore preferable to require local stability in more general form as follows. We assume there are constants δ, C with $\delta > 0$, $C < 1$, such that

$$|f_i(x_i) - f_i(y_i)| \leqslant C\,|x_i - y_i| \quad \text{when} \quad |x_i - y_i| < \delta, \quad i = 1, 2. \tag{3.3}$$

This condition implies that (y_1, y_2) is a locally stable fixed point of (3.1) since if $\epsilon < \delta$ and the initial frequencies (x_1, x_2) satisfy $|x_1 - y_1| + |x_2 - y_2| < \epsilon$ then the frequencies (x_1', x_2') in the next generation will also satisfy $|x_1' - y_1| + |x_2' - y_2| < \epsilon$ as a result of (3.3).

Now suppose the two subsystems $\mathscr{P}_1, \mathscr{P}_2$ are migration coupled in such a way that the recursion relations for the resultant complex system are

$$\begin{aligned} x_1' &= (1 - m_1) f_1(x_1) + m_1 f_2(x_2) \equiv g_1(x_1, x_2), \\ x_2' &= m_2 f_1(x_1) + (1 - m_2) f_2(x_2) \equiv g_2(x_1, x_2), \end{aligned} \tag{3.4}$$

where $0 \leqslant m_i \leqslant 1$. Then (3.4) is a continuous mapping of the square Ω into itself and is nearly the same as the mapping (3.1) if m_1 and m_2 are small.

THEOREM 3.1. *Let the continuous mapping* (3.1) *of the square Ω into itself satisfy the local stability condition* (3.3). *Then for each positive ϵ with $\epsilon < \delta$ the set*

$$\Omega_\epsilon = \{(x_1, x_2); (x_1, x_2) \in \Omega, |x_1 - y_1| + |x_2 - y_2| \leqslant \epsilon\}$$

is sent into itself by the mapping (3.4) *if $m_1 + m_2 < (1 - C)\epsilon$.*

Proof. If $(x_1, x_2) \in \Omega_\epsilon$ then

$$\begin{aligned} &|g_1(x_1, x_2) - y_1| + |g_2(x_1, x_2) - y_2| \\ &\quad \leqslant (1 - m_1)\,|f_1(x_1) - y_1| + (1 - m_2)\,|f_2(x_2) - y_2| + m_1\,|f_2(x_2) - y_1| \\ &\qquad + m_2\,|f_1(x_1) - y_2| \\ &\quad \leqslant |f_1(x_1) - y_1| + |f_2(x_2) - y_2| + m_1 + m_2 \\ &\quad \leqslant C[|x_1 - y_1| + |x_2 - y_2|] + m_1 + m_2 \\ &\quad \leqslant C\epsilon + m_1 + m_2 \end{aligned}$$

which is $< \epsilon$ if $m_1 + m_2 < (1 - C)\epsilon$. Thus $(x_1', x_2') \in \Omega_\epsilon$ and the result is proved.

The theorem shows that when m_1 and m_2 are sufficiently small then the small domain Ω_ϵ is a "*domain of stability*" *for the coupled system. That there is actually at least one fixed point of the mapping* (3.4) *in Ω_ϵ is assured by Brouwer's fixed-point theorem*, which asserts, in particular, that in a finite-dimensional space a continuous map of a closed convex set into itself always has a fixed point.

In the above theorem the functions $g_i(x_1, x_2)$ depend linearly on the migration parameters m_1, m_2. It is not difficult to extend the conclusions to cases where this dependence is more complicated, and to cases where m_1 and m_2 are functions of the frequency variables x_1 and x_2. This and many other generalizations are encompassed by the results in Section 4.

If the functions $f_i(x)$ are continuously differentiable and the local stability condition (3.3) is replaced by the stronger assumption (3.2) then the above theorem can be improved. In this case the familiar implicit function theorem can be used to show that for all sufficiently small m_1, m_2 there is a neighborhood of (y_1, y_2) containing a fixed point of the mapping (3.4) which is unique and is locally stable. For more details see Section 4.

In many of the genetic models which appear to be weakly coupled complex systems, Theorem 3.1 (and its generalizations) is not applicable because condition (3.3) is not valid. In particular, this may occur if $f_i'(y_i) = 1$ for one or more i. Such equilibria, which are nevertheless locally stable, occur in models depicting the elimination of disadvantageous recessives, in assortative mating systems, incompatibility patterns of mating, etc. In these examples the equilibrium values y_i usually determine boundary points (pure states) and the stability is of a more delicate nature. These boundary equilibrium points do not represent polymorphism in the simple subsystems but may combine to give a full polymorphism for the complex system. Our aim will be to show that under suitable conditions the equilibrium point (y_1, y_2) has a small neighborhood V such that if the initial frequency vector (x_1, x_2) of the weakly coupled system is in V, then in all subsequent generations the frequency vector remains in V. In this way we establish the existence of what is, from the practical viewpoint, a balanced polymorphism.

The argument used in the proof for this more general and more delicate case involves the concept of a Liapunov function for a mapping. Let $\mathscr{E}$ be a bounded set in a finite-dimensional Euclidean space and consider a continuous mapping $P \to P' = f(P)$ of the set $\mathscr{E}$ into itself. A real-valued function $L(P)$ defined for P in $\mathscr{E}$ is called a Liapunov function for the mapping if $L(P)$ is nonnegative, continuous and $L(f(P)) \leqslant L(P)$ with equality only when $f(P) = P$. It is known that if the mapping has a set $\mathscr{S}$ of fixed points such that for each P in $\mathscr{E}$ the sequence of iterates $f_1(P) = f(P)$, $f_2(P) = f(f(P))$, $f_3(P) = f_2(f(P)), \ldots$ converges to some point in $\mathscr{S}$, then in the neighborhood of any locally stable fixed point the mapping usually has a local Liapunov function. As an example, there are some well-known simple genetic models which have a fitness function W which increases from one generation to the next. If then c is a suitably large constant, $L = c - W$ is a Liapunov function. As another example, consider the above system (3.1) under condition (3.3). The set

$$\mathscr{E} = \{(x_1, x_2); \mid x_i - y_i \mid \leqslant \delta, i = 1, 2\}$$

is mapped into itself by (3.1) and the function

$$L(x_1, x_2) = |x_1 - y_1| + |x_2 - y_2|$$

is a (local) Liapunov function. As a final example, consider (3.1) with the specific functions

$$f_1(x_1) = x_1/(1 + x_1), \qquad f_2(x_2) = 1/(2 - x_2).$$

In this case $y_1 = 0$; $y_2 = 1$ is a fixed point for the mapping of the square into itself. Condition (3.3) is not satisfied because $f_1'(0) = f_2'(1) = 1$. Nevertheless there is a (global) Liapunov function

$$L(x_1, x_2) = x_1 + 1 - x_2$$

and it is easily verified that (y_1, y_2) is locally stable.

The next theorem again deals with the continuous mappings (3.1) and (3.4) of the square Ω into itself. The fixed point (y_1, y_2) of (3.1) is assume to be locally stable in the sense that there exist neighborhoods $\mathscr{E}_1$ of y_1 and $\mathscr{E}_2$ of y_2 and Liapunov functions $L_1(x_1)$, $L_2(x_2)$ defined on $\mathscr{E}_1$ and $\mathscr{E}_2$, respectively, which are continuous and such that

$$\begin{gathered} L_i(y_i) = 0, \qquad L_i(x_i) > 0 \qquad \text{if} \quad x_i \neq y_i, \\ f_i(x_i) \text{ maps } \mathscr{E}_i \text{ into itself,} \\ L_i(f_i(x_i)) < L_i(x_i) \qquad \text{if} \quad x_i \neq y_i. \end{gathered} \tag{3.5}$$

Under these conditions $L(x_1, x_2) = L_1(x_1) + L_2(x_2)$ is a local Liapunov function for the mapping (3.2).

THEOREM 3.2. *Let the fixed point* (y_1, y_2) *of the continuous mapping* (3.1) *be locally stable in the sense that there are local Liapunov functions* L_1, L_2 *satisfying* (3.5). *Then, given any neighborhood* U *of* (y_1, y_2), *there is a neighborhood* V *of* (y_1, y_2) *with* $V \subset U$, *and a positive* $m > 0$ *such that if* $m_1 < m$, $m_2 < m$ *then the mapping* (3.4) *sends* V *into itself.*

Thus the arbitrarily small neighborhood U contains a smaller neighborhood V such that if m_1, m_2 are sufficiently small and if the initial frequency vector for system (3.4) is in V then the frequency vector in all subsequent generations remains in V.

Proof. Let U be a neighborhood of (y_1, y_2) and let

$$L(x_1, x_2) = L_1(x_1) + L_2(x_2).$$

For $\delta > 0$ define

$$V(\delta) = \{(x_1, x_2);\ x_1 \in \mathscr{E}_1,\ x_2 \in \mathscr{E}_2,\ L(x_1, x_2) \leqslant \delta\}.$$

We choose δ so small that $V(\delta) \subseteq U$. This is possible because L is continuous, nonnegative, and zero only at (y_1, y_2). Then $V(\delta)$ is a compact set which contains an open neighborhood of (y_1, y_2). Let δ' be the maximum value achieved on $V(\delta)$ by the continuous function $L(f_1(x_1), f_2(x_2))$. Since by (3.5)

$$L(f_1(x_1), f_2(x_2)) < L(x_1, x_2) \leqslant \delta$$

at each point (x_1, x_2) in $V(\delta)$ except (y_1, y_2), it follows that $\delta' < \delta$. Thus if $(x_1, x_2) \in V(\delta)$ then $(f_1(x_1), f_2(x_2)) \in V(\delta')$, where $\delta' < \delta$. It is clear that $V(\delta')$ is contained in the interior of $V(\delta)$ and since $V(\delta')$ is compact there is a positive $\epsilon > 0$ such that if $(z_1, z_2) \in V(\delta')$ and (z_1', z_2') satisfies

$$|z_1' - z_1| + |z_2' - z_2| < \epsilon$$

then $(z_1', z_2') \in V(\delta)$.

Let $m = \epsilon/4$ and suppose $m_1 < m$, $m_2 < m$. Then

$$\begin{aligned} &|g_1(x_1, x_2) - f_1(x_1)| + |g_2(x_1, x_2) - f_2(x_2)| \\ &\quad \leqslant m_1 |f_2(x_2) - f_1(x_1)| + m_2 |f_1(x_1) - f_2(x_2)| \\ &\quad \leqslant 2(m_1 + m_2) < 4m = \epsilon. \end{aligned}$$

Therefore $(g_1(x_1, x_2), g_2(x_1, x_2))$ is in $V(\delta)$ whenever $(f_1(x_1), f_2(x_2))$ is in $V(\delta')$, hence whenever (x_1, x_2) is in $V(\delta)$. With $V = V(\delta)$ the theorem is proved.

4. Theorems and Proofs (Continued)

In this section we present results which are generalizations of the theorems of Section 3 and discuss some auxiliary matters in greater detail. We require theorems analogous to Theorems 3.1 and 3.2 but applicable to the case of many subsystems, rather than just two subsystems. Moreover, it is necessary to treat cases where the individual subsystems may have greater complexity, so that the state of each subsystem must be described by a frequency vector rather than a single frequency. We also wish to consider more general kinds of coupling than the simple migration coupling and even in the case of migration coupling it is desirable to allow migration rates m_{ij} which are themselves functions of the frequency variables. In some problems one has a complex system in which the subsystems are already coupled in some manner, and one wishes to analyze the effects of an additional mode of coupling, or of a perturbation in the coupling.

With this variety of applications in mind we shall consider a domain Ω of n-dimensional Euclidean space R^n, and certain continuous mappings of Ω into itself. We assume that Ω is a compact convex set. Vectors in R^n will be denoted by x, y, z, etc., so that for example $x = (x_1, x_2, \ldots, x_n)$. Continuous functions which map Ω into itself will be denoted by $f(x)$, $g(x)$, etc. In more extended notation we should write

$$f(x) = (f_1(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n)).$$

We shall make use of a norm $\| x \|$ for vectors $x = (x_1, \ldots, x_n)$ in R^n, with the usual properties

(i) $\| x \| \geqslant 0$, $\| x \| = 0$ only if $x = 0$,
(ii) $\| \lambda x \| = | \lambda | \, \| x \|$ for any number λ,
(iii) $\| x + y \| \leqslant \| x \| + \| y \|$.

There are many choices for such a norm, and we shall choose one adapted to the particular problem at hand. In any case, $\| x \|$ is a continuous function of x and the sets $\{x; \| x \| \leqslant \epsilon\}$ are compact convex sets.

We shall consider two continuous mappings

$$x' = f(x) \tag{4.1}$$

and

$$x' = g(x) = f(x) + r(x) \tag{4.2}$$

of Ω into itself, where (4.1a) has a locally stable fixed point y in Ω, and where the *perturbation or coupling term* $r(x) = g(x) - f(x)$ is small in a sense to be specified.

It is convenient for applications and to expedite the analysis to slightly strengthen the concept of a locally stable fixed point as follows: *The fixed point y of f is said to be locally stable if for every y-neighborhood U there is a neighborhood V such that $y \in V \subset U$ and*

$$f(\bar{V}) \subseteq V \tag{*}$$

($\bar{V}$ = closure of V).

The usual definition of local stability asserts merely $f(V) \subset V$. A simple case where (*) is satisfied is in the following context.

If there exists a local Liapunov function for f at the point y (see also later Theorem 4.3) then y is a locally stable fixed point of f. In view of Lemma 4.2 we will see that this is the case provided there exists c, $0 \leqslant c < 1$ and $\epsilon > 0$ such that $\| x - y \| < \epsilon$ implies

$$\| f(x) - y \| \leqslant c \| x - y \|.$$

In the applications the coupling function $r(x)$ depends on one or more parameters $m_1, \ldots, m_k$ which describe migration rates, recombination fractions, outcrossing fractions, epistatic effects, etc., and $r(x)$ becomes small as the parameters approach suitable limiting values.

THEOREM 4.1. *Let* (4.1) *and* (4.2) *be continuous mappings of the set* Ω *into itself and let* y *be a fixed point of* (4.1). *Suppose that for some norm and some* $\delta > 0$ *the set*

$$V(\delta) = \{x; \| x - y \| \leqslant \delta\} \cap \Omega$$

satisfies the condition

$$x \in V(\delta) \qquad \textit{implies} \qquad \| f(x) - y \| \leqslant c \| x - y \|, \tag{4.3}$$

where $0 \leqslant c < 1$. *Then* (4.2) *maps* $V(\delta)$ *into itself and has a fixed point in* $V(\delta)$ *provided*

$$\max_{x \in V(\delta)} \| r(x) \| < (1 - c)\delta. \tag{4.4}$$

Remark. The condition (4.3), analogous to (3.3) in Section 3, asserts that the fixed point y of (4.1) is locally stable with geometric rate of convergence. However the hypotheses of the theorem are not strong enough to guarantee that any fixed point of (4.2) in $V(\delta)$ has the same property. On the other hand, it assures that the small region $V(\delta)$ is stable under the mapping (4.2). The value of the theorem lies in the fact that the weak hypotheses are more likely to be valid and, in any case, may be easier to verify. A result with stronger hypotheses and stronger conclusions follows below.

The "smallness condition" (4.4) is satisfied in the applications by imposing restrictions on the parameters.

Proof of Theorem 4.1. If $x \in V(\delta)$ then

$$\begin{aligned} \| g(x) - y \| &= \| [f(x) - y] + r(x) \| \\ &\leqslant \| f(x) - y \| + \| r(x) \| \\ &< c \| x - y \| + (1 - c)\delta \leqslant \delta. \end{aligned}$$

Consequently $g(x) \in V(\delta)$, that is, (4.2) maps $V(\delta)$ into itself. Since $V(\delta)$ is a compact convex set, the existence of a fixed point of (4.2) in $V(\delta)$ is insured by Brouwer's theorem. This completes the proof.

Condition (4.3) can only be satisfied if the norm is related to the mapping (4.1) in a suitable way. Fortunately (4.3) does not involve the perturbing term $r(x)$. We next indicate some criteria under which there will be a norm satisfying (4.3).

LEMMA 4.1. *Let A be an $n \times n$ real or complex matrix such that every root λ of the characteristic polynomial of A satisfies $|\lambda| \leqslant a$. Then, given a', $a < a'$, there is a norm such that for every vector x*

$$\| Ax \| \leqslant a' \| x \|.$$

Remark. This and the following lemma are quite familiar to most mathematicians concerned with dynamical systems.

Proof. It is clearly sufficient to consider the case when the underlying vector space is the n-dimensional complex space C^n, because a norm on C^n induces a norm on the embedded R^n. Over C^n the matrix A is similar to a matrix in Jordan canonical form. That is, there is a basis for C^n consisting of vectors $z^{(1)},\ldots, z^{(n)}$ such that

$$Az^{(i)} = \lambda_i z^{(i)} + \delta_i z^{(i-1)}, \qquad i = 1,\ldots, n,$$

where the λ_i are eigenvalues of A and each δ_i is either 0 or 1 and $\delta_1 = 0$. Any vector x has a unique representation

$$x = \sum_{i=1}^{n} \xi_i z^{(i)}.$$

With $\rho > 1$ we define the norm

$$\| x \| = \sum_{i=1}^{n} | \xi_i | \rho^i.$$

Now for any x

$$Ax = \sum_{i=1}^{n} \lambda_i \xi_i z^{(i)} + \sum_{i=1}^{n-1} \delta_{i+1} \xi_{i+1} z^{(i)}$$

so

$$\begin{aligned} \| Ax \| &\leqq \sum_{i=1}^{n} | \lambda_i | \cdot | \xi_i | \rho^i + \sum_{i=1}^{n-1} \frac{| \xi_{i+1} |}{\rho} \rho^{i+1} \\ &\leqslant a \| x \| + \frac{1}{\rho} \| x \| = \left(a + \frac{1}{\rho}\right) \| x \|. \end{aligned}$$

The conclusion follows if $\rho > (a' - a)^{-1}$.

If a mapping $f(x)$ has first-order partial derivatives, then the matrix of first-order partial derivatives

$$f'(x) = \left[\frac{\partial f_i(x)}{\partial x_j}\right]_{i,j=1,\ldots,n}$$

is called the *gradient* of $f(x)$.

LEMMA 4.2. *If the mapping $f(x)$ has, in some neighborhood of the fixed point y, continuous first-order partial derivatives, and if all eigenvalues λ of the gradient matrix $f'(y)$ satisfy $|\lambda| < 1$, then there is a norm and numbers $\delta > 0, 0 \leqslant C < 1$ such that condition* (4.3) *holds.*

Proof. By Lemma 4.1 there is a norm and a number $a', 0 \leqslant a' < 1$ satisfying

$$\|f'(y) \cdot u\| \leqslant a' \|u\|$$

for every vector u. Choose C so $a' < C < 1$. Then by continuity of the gradient and of the norm there is a number $\delta > 0$ such that

$$\|f'(z) \cdot u\| \leqslant C \|u\| \qquad \text{if} \quad \|z - y\| \leqslant \delta. \tag{4.5}$$

If $x \in \Omega$ and $\|x - y\| \leqslant \delta$ we can integrate both members of the identity

$$(d/dt) f(y + t(x - y)) = f'(y + t(x - y)) \cdot (x - y)$$

over $0 \leqslant t \leqslant 1$ (since the line segment joining y to x lies in the convex set Ω). This yields

$$f(x) - f(y) = \int_0^1 f'(y + t(x - y)) \cdot (x - y)\, dt$$

and hence by (4.5)

$$\|f(x) - y\| \leqslant \int_0^1 C \|x - y\|\, dt = C \|x - y\|,$$

which completes the proof.

Theorem 4.1 asserts the existence of at least one fixed point of (4.2) in $V(\delta)$. For many cases of practical interest the implicit function theorem of advanced calculus can be used to show that the fixed point is unique. Consider a family of continuous mappings

$$x' = f(x, \theta) \tag{4.6}$$

depending on a parameter θ. The parameter θ is assumed to range over some set H in a k-dimensional Euclidean space with a limit point $\theta_0 \in H$ such that

$$f(x, \theta_0) = f(x),$$

that is, (4.6) reduces to (4.1) for $\theta = \theta_0$.

THEOREM 4.2. *Suppose that for each $\theta \in H$ the mapping* (4.6) *is a continuous mapping of the compact convex set Ω into itself. Let $y \in \Omega$ be a solution of $y = f(y, \theta_0)$. Assume the family of mappings* (4.6) *are defined and have partial*

derivatives with respect to the variables $x_1, \ldots, x_n$ for each $\theta \in H$ and for all x in a set $\mathscr{E}$ in R^n, where $\mathscr{E}$ contains y as an interior point. The partial derivatives are assumed to be jointly continuous in the variables x and θ. If all eigenvalues λ of the gradient $f'(y)$ satisfy $|\lambda| < 1$, then the conclusions of Theorem 4.1 *hold, the fixed point $y(\theta)$ in $V(\delta)$ will be unique when θ is sufficiently close to θ_0, and all eigenvalues λ of the gradient $f'(y(\theta), \theta)$ will satisfy $|\lambda| < 1$.*

Remark. When y is a boundary point of Ω the hypotheses of the theorem require that the mapping functions and their partial derivatives be defined in a full neighborhood of the fixed point y, and so at points outside of Ω. In the applications this is usually the case, in fact the mapping functions are frequently rational functions of the coordinates. However, it is necessary to verify carefully the continuity of the partial derivatives at y even in these applications.

Proof. Straightforward application of the implicit function theorem leads to the conclusion that when θ is sufficiently cose to θ_0 there is a small neighborhood of y containing a *unique* fixed point $y(\theta)$ and that the assertion concerning the gradient $f'(y(\theta), \theta)$ will hold. On the other hand the hypotheses of Theorem 4.1 hold so for all small $\delta > 0$ there is a fixed point in $V(\delta)$. This would lead to a contradiction unless $y(\theta) \in V(\delta) \subseteq \Omega$ and there is no other fixed point in $V(\delta)$.

The main perturbation theorem is next and involves local stability via the existence of local Liapunov functions.

THEOREM 4.3. *Let* (4.1) *and* (4.2) *be continuous maps of Ω into itself and let y be a fixed point of* (4.1). *Suppose there is a neighborhood $\mathscr{E}$ of y and a real-valued function $L(x)$ defined and continuous on $\mathscr{E} \cap \Omega$ such that*

$$\begin{gathered} L(y) = 0, \qquad L(x) > 0 \qquad \text{if} \quad x \neq y, \\ f(x) \text{ maps } \mathscr{E} \text{ into itself}, \\ L(f(x)) < L(x) \qquad \text{if} \quad x \neq y. \end{gathered} \tag{4.7}$$

Then, given any neighborhood U of y there is a neighborhood V of y with $V \subseteq U$, and a positive number ϵ, such that (4.2) *maps V into itself provided*

$$\max_{x \in V} \| r(x) \| < \epsilon. \tag{4.8}$$

(Here we use the usual Euclidean norm.)

Proof. For $\delta > 0$ define

$$V(\delta) = \{x, x \in \mathscr{E} \cap \Omega, L(x) \leqslant \delta\}.$$

With δ chosen so small that $V(\delta) \subseteq U$, $V(\delta)$ is compact and contains an open (relative) neighborhood of y. Let δ' be the maximum of the continuous function

$L(f(x))$ on the compact set $V(\delta)$. As in Section 3 we have $\delta' < \delta$. Choose $\epsilon > 0$ so that if $z \in V(\delta')$ and $\| z - z' \| < \epsilon$ then $z' \in V(\delta)$. Now if $x \in V(\delta)$ then $f(x) \in V(\delta')$ so $g(x) = f(x) + r(x)$ is in $V(\delta)$ provided (4.8) is satisfied.

It should be noted that if the sets $V(\delta)$ are also convex then (4.2) will have a fixed point in $V(\delta)$.

We close the developments with a more precise formulation of Principle III (cf. Section 1).

THEOREM 4.4. *Let $f(x)$ be a continuously differentiable mapping of a compact convex set $\Omega \subset E^n$ (Euclidean n space) into itself admitting a finite number of fixed points $\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_k$. Assume that $f(x)$ is actually defined in an open set V containing Ω. Suppose that all eigenvalues λ of the gradient matrix $f'(\underline{y}_i)$ at each y_i satisfy $|\lambda| \neq 1$. And suppose the mapping $g(x) = f(x) + r(x)$ sends Ω into itself, where $r(x)$ is continuously differentiable on V with $\max(\| r(x)\| + \| r'(x)\|) \leqslant \epsilon$ sufficiently small. Then $g(x)$ admits at most k fixed points $z_1, \ldots, z_k$ in Ω with each z_i in a neighborhood y_i. Moreover, if y_i is locally stable for $f(x)$ then z_i in Ω certainly exists in a neighborhood of y_i and is locally stable for $g(x)$. When y is locally unstable (i.e., if at least one eigenvalue λ of $f'(\underline{y}_j)$ has $|\lambda| > 1$) then if z_j exists in Ω, it is also locally unstable.*

Proof. Applying the implicit function theorem as in Theorem 4.2 we deduce the existence of exactly one fixed point z_i in the neighborhood of each y_i. (If y_i is locally unstable then z_i could move to $V - \Omega$, i.e., outside Ω.) On the other hand, if y_i is locally stable then the analysis of Theorem 4.1 establishes $z_i \in \Omega$. The concordant stability nature of z_i associated with y_i follows from elementary continuity considerations. It remains to show that $g(x)$ admits no other fixed points in Ω aside from $z_1, \ldots, z_k$. To this end, note that

$$\underline{f}(\underline{x}) - \underline{x} \neq 0 \qquad \text{for all} \quad x \neq y_i, \qquad i = 1, 2, \ldots, k.$$

Thus a sufficiently small neighborhood W of the fixed-point set can be chosen satisfying

$$W = W_1 \cup \cdots \cup W_r, \; W_i \text{ of } y_i$$

$\min_{x \notin W} \| \underline{f}(\underline{x}) - \underline{x} \| \geqslant \delta > 0$. The W_i can be further specified small enough so that $g(x)$ has only the fixed points $z_1, \ldots, z_k$ in W, with the inequality $\epsilon = \max_{x \in \Omega} \| r(x)\| < \delta$ fulfilled. Clearly $\min_{x \notin W} \| g(\underline{x}) - \underline{x} \| \geqslant \delta - \epsilon > 0$ and the proof of the theorem is complete.

A set of striking applications of Theorem 4.4 to multiniche selection models is set forth in Karlin and McGregor (1972a).

FURTHER EXTENSIONS. The conditions of Theorems 4.1–4.4 involve only estimates of the crudest kind on the remainder term $r(x)$. Consequently the

theorems and the proofs remain valid with only slight changes if the remainder term r is allowed to depend on the frequencies in the present generation and one or more of the preceding generations. Suppose the initial frequency in the complex system is $x^{(1)}$ and the frequency in the k-th generation is $x^{(k)}$. If the recursion law is of the form

$$x^{(k+1)} = f(x^{(k)}) + r(x^{(k)}, x^{(k-1)}) \tag{4.9}$$

then in Theorem 4.1 we must replace condition (4.4) by the condition

$$\max_{x \in V, z \in V} \| r(x, z) \| < (1 - c)\delta$$

and in Theorem 4.3 we must replace Condition (4.8) by the condition

$$\max_{x \in V, z \in V} \| r(x, z) \| < \epsilon.$$

Similar modifications can be made to cover the case when the remainder term depends on the entire past history of the system.

Finally, instead of dealing with fixed points and perturbations of the mapping in the neighborhood of these points the whole theory can be extended to treat locally stable *invariant curves* or manifolds. Then appropriate perturbations of the mapping f along the lines indicated previously assure the existence of nearby invariant manifolds for the mapping g.

5. Discussion

1. The abundance of polymorphic phenomena is well established in natural and experimental populations. Uncovering electrophoretic variants for several enzymes in different organisms has in recent years been a common laboratory practice.

In explaining polymorphisms and clines, emphasis is usually given to changes in selective factors between and within environments. It is also widely recognized that in many natural situations migration may play an important, even a dominant, role. The following theme recurs in many works concerned with population genetics: "The ongoing process of evolution probably requires adjustment to a constantly varying environment, and of the combination of characteristics from the different populations that survive. Given an environmental flux, a necessary and sufficient condition for genetic change is availability in the population of genetic variants." Important sources of variability in natural population can be genes and gene complexes transferred from other populations. Also widely recognized is that differentiated populations retain the ability

for exchange of genetic material. Put in a more descriptive language, spacial and temporal variation in environment is considered to be highly involved in the maintenance of genetic variation in population (Darlington, 1957; Wright, 1968, 1969; Dobzhansky, 1967).

We have attempted in this paper to partly provide a *quantitative basis* for these concepts by giving a methodology for analyzing a mixture of genetic systems and niches influenced by a variety of pressures and mating patterms. We find that stable polymorphisms in the total system (composed of several subsystems) can be maintained as a result of one or more main effects in each system with slight interaction and/or coupling persisting among all the subsystems. Of course, stable polymorphism is known to occur also in some cases where interaction among all the subsystems is quite strong and not slight.

2. In most studies of migration selection balance it is assumed that in one niche one homozygote carries a selective advantage while in a second niche an alternate homozygote is selectively superior [e.g., see Moran (1962, Chapter 8), Maynard Smith (1966), Clarke (1966)]. An application of Principle I [Karlin and McGregor (1972a)] demonstrates that two or more populations individually subject to disruptive selection pressures would separately fix but when slight gene flow connects the populations, then a stable polymorphism with all types persisting can be attained. Note that, for a single-niche situation subject to disruptive selection, only fixation is possible. On the other hand, a multiniche (at least two niches) version of the same model with disruptive selection in force in each niche now provides possibilities both of fixation (monomorphism) and polymorphism and which occurs depends on the nature of the initial compositions of the subpopulations or factors other than the forces of natural selection.

Determined from theoretical study, some genetic models of selfing and/or partial assortative mating for dominant autosomal traits bear consequences similar to those of disruptive selection or those of directed selection for a dominant gene, precluding the existence of stable polymorphism in these situations [see Scudo and Karlin (1969) or Karlin (1968)]. On the other hand, there is positive evidence that natural polymorphic populations (especialy certain bird populations) practice partial assortative mating [e.g., see Parsons (1962)]. Presumably, selfing or assortative mating arises in response to selection pressures for immediate fitness to special environments. We have proved in Section 2 by invoking Principle I, that two or more replica populations subject to the same partial assortative mating behavior with slight gene flow connecting the systems can maintain a stable polymorphism.

3. It is important to emphasize that the validity of these polymorphisms requires that the migration rates among the distinct systems be very small although some forms of migration selection (e.g., migration to certain niches

favoring certain genotypes) may be in force. When migration rates are not small then the various systems (or niches) blend and the ultimate effects of all the selection parameters are complex. On the other hand, when the migration rates are sufficiently small then the selection effects in each niche strongly dominate in their individual niche while the remaining niches help to keep a slight flow of alternative genotypes and thereby maintain a balance manifesting a global stable polymorphism.

4. Many of the stable polymorphisms established by virtue of the principles of Section 1 for multiniche selection (natural or sexual) balance have the property that a preponderance of one homozygote is maintained in one niche while a preponderance of an alternative homozygote exists in a second niche and so on for the other niches. It is reasonable to speculate that some forms of habitat selection confer an advantage, say on BB in niche 2, AA is favored in niche 1 while the heterozygote (or hybrid type) bears marked disadvantage to both homozygotes in each of the niches and yet a global balance may be preserved.

Aside from the kinds of polymorphisms described above, there also exists the possibility of stable global fixations. What evolves may depend decisively on the initial composition of the various subpopulations. Founder effects, sampling and statistical fluctuation, and changing environmental conditions may be important original forces causing a local population to be mostly of one type.

To sum up the above discussion we obtain the general dictum to the effect that almost all selection models bear possibilities of stable polymorphism by spreading spatially the selection effects (i.e., to several niches) and keeping the interaction between niches slight.

5. The well-studied example of variation of shell color and banding patterns in the land snails, *Cepaea nemoralis* and *Cepaea hortensis* is reviewed in Ford (1964). The genetics and selection forces for these traits are reasonably well understood. The distribution and data on these populations underscore the phenomena of "area effects" where two or more populations with different frequencies of color and banding patterns are found in an apparently uniform habitat. Moreover, relatively small numbers of intermediate gene frequencies are observed among adjacent populations. Explanations and interpretation of this phenomena have been put forward in terms of the concepts of adaptability and the founder principle. We call attention to Principles I-III which are clearly relevant. We have explained earlier how the founder's principle combined with forms of disruptive selection could bring evolution of coadapted gene complexes and an accompanying reduction of fitness and frequency of the heterozygote.

6. It is important to realize that the nature of local stability for the pure equilibria, for example, (algebraic versus) geometric local convergence in the separate niches, is irrelevant to the main conclusions of Principles I and II.

Examples treating multiniche selection models usually involve geometric convergence while for the single-niche assortative mating model of Section 2 the rate of fixation to the pure equilibria is algebraic. Generally, forms of sexual selection lead to algebraic rates of fixation while effects of natural selection bring geometric convergence. Our general principles apply in both situations. However, there is a difference of order of magnitude allowed for the migration parameters in the two cases yielding polymorphism. When algebraic rather than geometric convergence is operating in the corresponding one-niche model at a specified equilibrium, it manifestly takes longer to fix in such a population. In this circumstance, it can be established that the order of migration permitted assuring stable polymorphism is at least the square root of that where geometric convergence is operating. Therefore, more possibilities of polymorphism arise with algebraic convergence than in the case of geometric convergence. In particular, migration rates can be larger in the cases of assortative mating producing polymorphism than in the cases where disruptive selection is operating.

7. An interesting consequence of Principle I concerns the problem of the minimum number of niches required to guarantee the existence of p types in a stable polymorphic state. Assuming that sexual or natural selection operating in each niche fixes a different type, then a minimum of p niches with slight migration flow between them can maintain a global stable polymorphism with all p types represented in each subsystem (see Section 1).

When it is possible to have more than one type kept in a stable state in some or several niches then the minimum required number of niches to maintain all p types is diminished from p. There is some similarity between this result and the Gause exclusion principle of ecology.

Our principle of Section 1 has, perhaps, two striking general qualitative implications:

(a) Systems with simple equilibria when combined with slight interactions (e.g., gene flow) among them produce a complex system with a stable equilibria involving several types possible.

(b) Complex systems when combined with slight migration between them produce even more complex systems with more possibilities and representations of stable polymorphisms.

The last assertion is consistent with the "tenet" that complex systems tend to be more stable with the degree of stability increasing with the extent of complexity. However, there may be a point of diminishing return as to the number of niches needed to maintain the desired types.

Acknowledgments

It is a pleasure to acknowledge valuable criticisms of this manuscript by Professors W. Bodmer, W. Ewens, M. Feldman, W. Hill, E. Nevo, and G. Simchen.

References

Allard, R. W., Jain, S. K., and Workman, P. L. 1968. The genetics of inbreeding populations, *Advan. Genet.* **14**, 55–131.

Bodmer, W. F. 1965. Differential fertility in population genetic models, *Genetics* **51**, 411–424.

Brown, A. H. D., and Allard, R. W. 1969. Inheritance of isozyme difference among the inbred parents of a reciprocal recurrent selection populations of maize, *Crop Science* **9**, 72–75.

Clarke, B. C. 1964. Frequency-dependent selection for the dominance of rare polymorphic genes, *Evolution* **18**, 364–369.

Clarke, B. C. 1966. The evolution of morph-ratio clines, *Amer. Natur.* **100**, 389–402.

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

Darlington, P. J. 1957. "Zoogeography: The Geographical Distribution of Animals," John Wiley, New York.

Dobzhansky, Th. 1967. Genetic diversity and diversity of environments, *Proc. Fifth Berkeley Symp. Math. Stat. Prob.* **4**, 295–304.

Eshel, I. 1972. About neighbor effects in natural selection and the evolution of individually disadvantageous features, *Theor. Pop. Bio.*, to appear.

Feldman, M. W., Nabholz, M., and Bodmer, W. F. 1969. Evolution of the Rh polymorphism: A model for the interaction of incompatibility, reproductive compensation, and heterozygote advantage, *Amer. J. Hum. Genet.* **21**, 171–193.

Finney, D. J. 1952. The equilibrium of a self-incompatible polymorphic species, *Genetica* **26**, 33–64.

Ford, E. B. 1964. "Ecological Genetics," Methuen, London; John Wiley, New York.

Ford, E. B. 1965. "Genetic Polymorphism," Faber and Faber, London.

Haldane, J. B. S. 1924. A mathematical theory of natural and artificial selection. Part V. Selection and mutation, *Proc. Camb. Phil. Soc.* **28**, 838–844.

Haldana, J. B. S. 1932. "The Causes of Evolution," Harper and Row, New York.

Haldana, J. B. S. 1948. The theory of a cline, *J. Genet.* **48**, 277–284.

Haldane, J. B. S. and Jayakar, S. D. 1963a. Polymorphism due to selection of varying direction, *J. Genet.* **58**, 237–242.

Haldane, J. B. S. and Jayakar, S. D. 1963b. Polymorphism due to selection depending on the composition of a population, *J. Genet.* **58**, 318–323.

Harris, H. 1966. Enzyme polymorphisms in man, *Proc. Roy. Soc. B* **164**, 298–310.

Hartl, D. L., Hiraizumi, Y., and Crow, J. F. 1967. Evidence for sperm dysfunction as the mechanism of segregation distortion in *Drosophila melanogaster*, *Proc. Nat. Acad. Sci.* **58**, 2240–2245.

Karlin, S. 1968a. Equilibrium behavior of population genetic models with non-random mating. Part I. Preliminaries and special mating systems, *J. Appl. Prob.* **5**, 231–313.

Karlin, S. 1968b. Equilibrium behavior of population genetic models with non-random mating. Part II. Pedigrees, homozygosity, and stochastic models, *J. Appl. Prob.* **5**, 487–566.

KARLIN, S. AND FELDMAN, M. W. 1968a. Analysis of models with homozygote × heterozygote matings, *Genetics* **59**, 105–116.

KARLIN, S. AND FELDMAN, M. W. 1968b. Further analysis of negative assortative mating, *Genetics* **59**, 117–136.

KARLIN, S. AND FELDMAN, M. W. 1970. Two locus symmetric viability model, *Theor. Pop. Bio.* **1**, 39–71.

KARLIN, S. AND MCGREGOR, J. 1971. On mutation selection balance for two-locus haploid and diploid populations, *Theor. Pop. Bio.* **2**, 60–70.

KARLIN, S. AND MCGREGOR, J. 1972a. Applications of methods of small parameters to multi-niche population genetic models, *Theor. Pop. Bio.*

KARLIN, S. AND MCGREGOR, J. 1972b. Polymorphisms for multi-locus selection system with tight linkage, to be published.

KARLIN, S. AND SCUDO, F. M. 1969. Assortative mating based on phenotype. II. Two autosomal alleles without dominance, *Genetics* **63**, 499–510.

KIMURA, M. 1968. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles, *Genet. Res.* **11**, 247–269.

KIMURA, M. 1969a. The length of time required for a selectively neutral mutant to reach fixation through random frequency drift in a finite population, *Genet. Res.*

KIMURA, M. 1969b. The rate of molecular evolution considered from the standpoint of population genetics, *Proc. Nat. Acad. Sci.* **63**, 1181–1188.

KING, J. L. AND JUKES, T. H. 1969. Non Darwian evolution, *Science* **164**, 788–798.

KOJIMA, K. 1971. The distribution and comparison of "genetic loads" under heterotic selection and simple frequency-dependent selection in finite populations, *Theor. Pop. Bio.* **2**, 159–173.

KOJIMA, K. AND TOBARI, Y. N. 1969a. The pattern of viability changes associated with genotype frequency at the alcohol dehydrogenase locus in a population of *Drosophila melanogaster*, *Genetics* **61**, 201–209.

KOJIMA, K. AND TOBARI, Y. N. 1969b. Selective modes associated with karyotypes in *Drosophila ananassae*, II. Heterosis and frequency-dependent selection, *Genetics* **63**, 639–651.

LEVENE, H. 1953. Genetic equilibrium when more than one ecological niche is available, *Amer. Natur.* **87**, 311–313.

LEWONTIN, R. C. 1964a. The interaction of selection and linkage. I. General considerations of heterotic models, *Genetics* **49**, 49–67.

LEWONTIN, R. C. 1964b. The interaction of selection and linkage. II. Optimum models, *Genetics* **50**, 757–782.

LEWONTIN, R. C. AND HUBBY, J. L. 1966. A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*, *Genetics* **54**, 595–609.

LEWONTIN, R. C. AND KOJIMA, K. 1960. The evolutionary dynamics of complex polymorphisms, *Evolution* **14**, 458–472.

MARSHALL, D. R. AND ALLARD, R. W. 1969. The genetics of electrophoretic variants in Avena, *J. Heredity* **60**, 17–19.

MAINARDI, D. 1968. "La Scelta Sussuale nell Evoluzione della Specie," Borignhieri, Torino.

MAYNARD SMITH, J. 1966. Sympatric speciation, *Amer. Natur.* **100**, 637–650.

MAYR, E. 1963. "Animal Species and Evolution," Harvard Univ. Press, Cambridge, Mass.

MORAN, P. A. P. 1959. The theory of some genetical effects of population subdivision, *Aust. J. Biol. Sci.* 15, 109–116.

MORAN, P. A. P. 1962. "The Statistical Processes of Evolutionary Theory," The Clarendon Press, Oxford.

O'DONALD, P. 1968. Models of the evolution of dominance, *Proc. Roy. Soc. B* **171**, 127–143.

OWEN, A. R. G. 1953. A genetical system admitting of two distinct stable equilibria under natural selection, *Heredity* **7**, 97–102.

PARSONS, P. A. 1962. The initial increase of a new gene under positive assortative mating, *Heredity* **17**, 267–276.

PIMENTEL, D. G., SMITH, J. C., AND SOAMS, J. 1967. A population model of sympatric speciation, *Am. Nat.* **101**, 493–503.

SCUDO, F. M. 1967. Selection on both haplo and diplophase, *Genetics* **56**, 693–704.

SCUDO, F. M. AND KARLIN, S. 1969. Assortative mating based on phenotype. I. Two alleles with dominance, *Genetics* **63**, 479–498.

SHEPPARD, P. M. 1958. "Natural Selection and Heredity," Hutchinson and Co., London.

WORKMAN, P. L. 1964. The maintanance of heterozygosity by partial negative assortative mating, *Genetics* **50**, 1369–1382.

WRIGHT, S. 1939. The distribution of self-sterility alleles in populations, *Genetics* **24**, 538–552.

WRIGHT, S. 1940. Breeding structure of population in relation to speciation, *Amer. Natur.* **74**, 232–248.

WRIGHT, S. 1968. "Evolution and the Genetics of Populations. Vol. I. Genetic and Biometric Foundations," Univ. of Chicago Press, Chicago.

WRIGHT, S. 1969. "Evolution and the Genetics of Populations. Vol. II. The Theory of Gene Frequencies," University of Chicago Press, Chicago.