J. Math. Biology (1982) 13: 305–324

Journal of Mathematical Biology

Cycling in Simple Genetic Systems

Ethan Akin*

Mathematics Department, The City College, 137 Street and Convent Avenue, New York, NY 10031, USA

Abstract. We describe the – unexpected – occurrence of stable limit cycles in the two locus, two allele model. No frequency dependence is involved. The cycles are due to the interaction between recombination and natural selection.

Key words: Adaptive topography – Cycling – Fisher's theorem – Kimura's maximum principle – Lyapunov function – Mean fitness – Two locus model

1. Introduction

In general the genetic state of a Mendelian population is represented by a point in some high dimensional space, but imagine this space is a flat plane. Over this plane there lies a surface whose height at each point represents some measure of the "fitness" or "selective value" of the corresponding state. This is Sewall Wright's adaptive topography. Under the influence of evolutionary forces the population state is presumed to move in a direction of increasing fitness, approaching an equilibrium represented by a peak, or local maximum, of the surface.

The purpose of this paper is to consider how well the adaptive surface picture represents the effects of the forces of natural selection and recombination.

Let I be the (finite) list of gamete genotypes of the population. The state of the gene pool is described by a distribution vector p with p_i for $i \in I$ the ratio of the number of gametes of type i to the total number of gametes. Thus, p is an element of the simplex Δ, i.e. the state space of the population is:

$$\Delta = \{p \in \mathbb{R}^I : p_i \geqslant 0 \text{ for all } i \in I \text{ and } \textstyle\sum_i p_i = 1\}.$$

p is called an interior distribution if all of the gamete types occur in the population. So the set of interior distributions is:

$$\mathring{\Delta} = \{p \in \Delta : p_i > 0 \text{ for all } i \in I\}.$$

A zygote type consists of a pair of gametes ij with i and $j \in I$. The fitness or Malthusian parameter of zygote type ij is a real number m_{ij} representing the birth rate minus the death rate of the type. Since ij and ji represent the same zygote type the matrix is symmetric:

* This work received support from the National Science Foundation and the Research Foundation of the City University of New York

0303–6812/82/0013/0305/$04.00

$$m_{ij} = m_{ji} \qquad (i \text{ and } j \in I). \tag{1.1}$$

The effect of natural selection is then modeled by the system of differential equations (see Crow and Kimura, 1970, p. 191):

$$\frac{dp_i}{dt} = p_i(m_i - \bar{m}) \qquad (i \in I), \tag{1.2}$$

where $m_i = \sum_j p_j m_{ij}$ is the mean fitness of gamete type i and $\bar{m} = \sum_i p_i m_i = \sum_{i,j} p_i p_j m_{ij}$ is the mean fitness of the entire population. Notice that while the m_{ij}'s are constants m_i and $\bar{m}$ vary with the population state p. These equations assume the Hardy-Weinberg condition that the zygote distribution is the product of the gamete distributions.

Fisher's Fundamental Theorem of Natural Selection says that under the influence of equations (1.2) the mean fitness $\bar{m}$ is constantly increasing, except at equilibria where $\bar{m}$ remains constant with p. Kimura's Maximum Principle says, moreover, that the direction of change in (1.2) is the direction of greatest increase of $\bar{m}$.

The latter result is a bit puzzling. It suggests that the right side of (1.2) should consist of the components of the gradient of $\bar{m}$ which it does not. The key to the puzzle is Kimura's definition of direction, or vector of unit length, namely a vector of unit variance with respect to the distribution p at which the vector is based.

Shahshahani and Conley noted that the concept of gradient depends on the choice of a mathematical structure called a Riemannian metric. With respect to the metric implicit in Kimura's definition $p_i(m_i - \bar{m})$ is exactly the i component of the gradient of $\bar{m}/2$. So we will denote the vectorfield on Δ corresponding to equations (1.2) by $\bar{\nabla}(\bar{m}/2)$ where $\bar{\nabla}$ is the operator of gradient in the Shahshahani sense. Fisher's Theorem and Kimura's Principle are immediate consequences of this result (Shahshahani, 1979, p. 5). The geometry of Δ equipped with the Shahshahani metric turns out to be a natural tool for population genetics.

Thus, for selection modeled by (1.2) Fisher's Theorem justifies the adaptive surface picture. The fitness function on Δ is just mean fitness $\bar{m}$. Furthermore, selection moves the population up the gradient of fitness once the concept of gradient is interpreted properly.

However, for multilocus models (1.2) is incorrect. Its derivation is based on the assumption that the gametes produced by an ij zygote are all of type i or j, whereas new gamete types will be produced by independent assortment among the chromosomes and by crossovers within the chromosomes. This introduces additional terms on the right of (1.2) which are the components of the recombination vectorfield. For the two-locus-two-allele model the recombination field was derived by Kimura. The extension to multilocus models is a matter of introducing notation to keep tracking of the required book-keeping. This was done by Shahshahani. The combined effect of selection and recombination is modeled by a vectorfield on Δ: $\bar{\nabla}(\bar{m}/2) + R$, where R is the sum of new terms. The recombination field has the following properties (these are proved in Akin, 1979, Sec. III.2).

(1) Recombination does not affect the distribution of alleles at each locus. It only changes the linkage between loci. A distribution p is said to be in linkage equilibrium if it is the product of the distributions of alleles at each locus. The set of

distributions in linkage equilibrium is a submanifold Λ of Δ. Furthermore, for each $p \in \Delta$ there is a unique $\pi(p) \in \Lambda$ such that p and $\pi(p)$ have the same marginal distributions at each locus. R vanishes on Λ and if we look at recombination alone, i.e. solve $dp_i/dt = R(p)_i$ at $p \in \Delta$, then p approaches the equilibrium $\pi(p)$.

(2) Define the entropy of a distribution by $H(p) = -\sum p_i \log p_i$ and the normalized entropy by $\hat{H}(p) = H(p) - H(\pi(p))$. H and $\hat{H}$ are real valued functions on Δ. $\hat{H}$ is negative except on Λ where it vanishes. Under the influence of recombination alone $\hat{H}$ is constantly increasing, except at the equilibria in Λ. So $\hat{H}$ is a Lyapunov function for recombination. However, in contrast with the selection field, recombination is not the gradient of $\hat{H}$. In fact, the recombination field is not the Shahshahani gradient of any function. This fact is crucial in what follows.

What happens to mean fitness under the combined effects of selection and recombination? Ewens (1969a) proved that Fisher's Theorem still holds in the absence of epistasis, i.e. provided that fitness is additive between loci. However, Moran (1964) showed that $\bar{m}$ need not always increase. With the benefit of hindsight this result is easy to see. Suppose that the matrix m_{ij} is such that $\bar{m}$ has an isolated maximum at a point p_0 in $\mathring{\Delta}$ but $p_0 \notin \Lambda$ (this requires epistasis!). Now introduce a small recombination term (tight but incomplete linkage). Since $p_0 \notin \Lambda$, R does not vanish at p_0 and so p_0 is not an equilibrium for the combined field. Since R is a small perturbation of the selection field there is a new point p_1 near p_0 which is an equilibrium for the combined field. $\bar{m}$ is less at p_1 than at p_0. So if one begins at p_0 and approaches p_1 under the influence of the combined field, $\bar{m}$ must necessarily decrease on part of the path.

Moran's work left open the conjecture that the combined field admits some Lyapunov function other than $\bar{m}$ (see Ewens, 1969b, p. 96). Akin (1979) showed this conjecture to be false by demonstrating that the differential equation of selection plus recombination can admit periodic solutions. Thus, the problem is not in the choice of fitness function. It lies in the original intuition underlying the adaptive surface picture, namely the belief that these systems always tend to equilibrium.

Now the occurrence of cycling even under selection alone is well known for "frequency dependent" models, i.e. where the m_{ij}'s depend on p. Perhaps the basic example with three gamete types is given by Kimura (1958), p. 154:

$$\frac{dp_1}{dt} = p_1(p_2 - p_3),$$

$$\frac{dp_2}{dt} = p_2(p_3 - p_1),$$

$$\frac{dp_3}{dt} = p_3(p_1 - p_2). \tag{1.3}$$

Thus, under (1.3) type 1 increases where $p_2 > p_3$, type 2 increases when $p_3 > p_1$, etc.

These equations also arise in the study of evolutionary games (see Taylor and Jonker, 1978; Zeeman, 1979; Akin, 1980). The models are identical with (1.2) but the assumption that m_{ij} be symmetric, i.e. (1.1), is dropped. In fact, the matrix for (1.3) is the skew-symmetric matrix:

$$\begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}.$$

This is the so-called "paper-rock-scissors" game.

To see the cycles, introduce polar coordinates on $\mathring{\Delta}$ centered at $p^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, defining:

$$r = -\frac{1}{3}\sum_i \log 3p_i,$$

$$\theta = \tan^{-1}\frac{p_2 - p_1}{p_3 - p_1}. \tag{1.4}$$

The radial coordinate r is just a linear adjustment of the log of the product $p_1p_2p_3$. $r \geqslant 0$ by convexity of the log function, with $r = 0$ at p^*. The angular coordinate θ is undefined at the origin, p^*, and is defined up to addition of 2π.

It is easy to check that $dr/dt = 0$, i.e. under the influence of (1.3) r and hence $p_1p_2p_3$ remain constant. On the other hand, on $\mathring{\Delta} - p^*$ the vectorfield of (1.3) turns out to be the positive function $(p_2 - p_1)^2 + (p_3 - p_1)^2$ times the Shahshahani gradient of θ. Notice that while θ is not a well-defined real valued function its gradient is well-defined. Thus, starting at any point of $\mathring{\Delta} - p^*$ one remains on a closed curve defined by $p_1p_2p_3 =$ constant and moves so that θ continually increases.

In frequency dependent models like (1.3) the occurrence of cycling is often intuitively clear from the nature of the selection pressures. However, in the constant fitness case the cycles are completely unexpected.

The abstract nature of the proof in Akin (1979) leaves open several questions about the cycles.

(1) Position Effects: The absence of position effects amounts to the assumption of a number of additional symmetry conditions on the matrix m_{ij} in addition to (1.1) (see Akin, 1979, p. 106). Can the cycles occur for matrices with no position effects?

(2) Robustness: Is the cycling behavior robust, i.e. maintained despite small perturbations of the model? For example, Zeeman has shown that in evolutionary game models with three strategies cycles are never robust.

(3) Stability: Can stable cycles, i.e. limit cycles, occur?

To study these questions it is useful to look at the two-locus-two-allele (TLTA) case as this is the simplest model where selection and recombination can interact. Akin (1981) answers these questions affirmatively by constructing examples in the TLTA case without position effects admitting robust limit cycles. Furthermore, a program is there developed which generates all such examples in the TLTA case.

The cycles arise as a consequence of the Hopf Bifurcation Theorem. In the next section we review the phenomenon of Hopf bifurcation. In the following sections we summarize Akin (1981) by outlining the program and reporting a class of numerical examples.

Because the cycling is robust it will continue to occur in related models of selection and recombination. At the end of Sec. 3 we will discuss the model of Nagylaki and Crow (1974) which does without the mathematically convenient, but

biologically unrealistic assumption of Hardy-Weinberg proportions. We will also look at discrete-time models and the recent work of Hastings.

Because the cycling may be stable, it is, in principle, observable. Initial positions near such an attractive cycle tend to be pulled in toward it. However, the cycles appear to be rather low frequency, i.e. long period cycles. Such behavior, were it to occur in real data, would not be recognized as cycling. It would more likely be mistaken for drift or a "moving equilibrium" tracking slow environmental change. At the end of Sec. 4 we will return to this frequency question.

2. Hopf Bifurcation

The state space for the TLTA case is

$$\Delta = \left\{ p \in \mathbb{R}^4 : p_i \geqslant 0 \text{ for } i = 1, \ldots, 4 \text{ and } \sum_{i=1}^{4} p_i = 1 \right\}.$$

The vectors tangent to Δ form a three-dimensional subspace of $\mathbb{R}^4$:

$$\mathbb{R}_0^4 = \left\{ v \in \mathbb{R}^4 : \sum_{i=1}^{4} v_i = 0 \right\}.$$

Consider any vectorfield X on Δ, i.e. X is a function from Δ to $\mathbb{R}_0^4$. Associated with X is the differential equation:

$$\frac{dp}{dt} = X(p) \qquad (p \in \Delta). \tag{2.1}$$

Thus, a solution of (2.1) is a path in Δ which is everywhere tangent to X. The condition that $X(p)$ lie in $\mathbb{R}_0^4$, i.e. $\sum_i X(p)_i = 0$, comes from the fact that $\sum_i p_i$ remains constant for p in Δ.

A point, p_0, is an equilibrium for the equation (2.1) if X vanishes at that point, i.e. if $X(p_0) = 0$.

To study the behavior of (2.1) near the equilibrium p_0 we linearize, that is, change variables by $p = p_0 + v$ (with $v \in \mathbb{R}_0^4$) and differentiate X at p_0 to get a linear map $L: \mathbb{R}_0^4 \to \mathbb{R}_0^4$. The linearized equation is then:

$$\frac{dv}{dt} = L(v) \qquad (v \in \mathbb{R}_0^4). \tag{2.2}$$

The behavior of solutions of (2.2) depend on the eigenvalues of L which are the roots of the characteristic equation $\det(\sigma I - L) = 0$ or

$$\sigma^3 - T\sigma^2 + M\sigma - D = 0, \tag{2.3}$$

where T is the trace of L (= the sum of the eigenvalues) and D is the determinant of L (= the product of the eigenvalues). M is the sum of the 2×2 diagonal minors of L in any matrix representation.

The extent to which one can use (2.2) to study (2.1) near p_0 depends on these eigenvalues. p_0 is called a *nondegenerate equilibrium* if 0 is not an eigenvalue or equivalently if the determinant D does not vanish. It follows from the inverse function theorem that a nondegenerate equilibrium is isolated and if the original

vectorfield is perturbed slightly then there is a unique equilibrium of the perturbed field near p_0. We used this condition in discussing Moran's paper.

p_0 is called a *hyperbolic equilibrium* if no eigenvalue has real part zero or equivalently if the set of eigenvalues avoids the entire imaginary axis and not just 0. A hyperbolic equilibrium is not only isolated, but a theorem of Hartman says that the behavior of solutions of (2.1) near p_0 is similar to the behavior of solutions of (2.2) near 0. In other words, for a hyperbolic equilibrium the linearized equation really does describe the original one near equilibrium.

A classic result of Lyapunov says that if the real parts are all negative then p_0 is an attractor, i.e. solutions beginning near p_0 return toward p_0 and approach it in the limit as t approaches infinity. The imaginary part of the eigenvalues determines the nature of the returning path. Solutions associated with real negative eigenvalues approach equilibrium tangent to some ray. A nonzero imaginary part induces a rotation so that solutions associated with complex eigenvalues spiral in toward equilibrium.

Equation (2.3) has three roots and complex roots come in conjugate pairs. Hence, there are either three real eigenvalues or else one real one and one complex conjugate pair. In the latter case we can write the roots as $\varepsilon \pm i\lambda$ and μ with $\lambda > 0$. The only way a nondegenerate equilibrium can fail to be hyperbolic is with a pair of conjugate imaginary roots and a nonzero real root, i.e. $\varepsilon = 0$ and the roots are $\pm i\lambda$, μ with $\lambda > 0$ and $\mu \neq 0$. In that case we will call p_0 a *Hopf equilibrium*. So p_0 is a Hopf equilibrium if L has an imaginary eigenvalue and the determinant of L is nonzero. These conditions can be detected from the coefficients of the characteristic equation.

Lemma 1. *Let* $L\colon \mathbb{R}^4_0 \to \mathbb{R}^4_0$ *be a linear map. L has an imaginary eigenvalue iff the following conditions hold*:

$$D = M \cdot T \tag{2.4}$$

and

$$M > 0. \tag{2.5}$$

In this case the real eigenvalue is T which has the same sign as D.

Proof. Equation (2.4) says that T is a root of (2.3). In that case, the other two roots sum to zero and M is their product. So $M \leqslant 0$ if these two are real and $M > 0$ if they are conjugate imaginaries. Q.E.D.

Because a Hopf equilibrium fails to be hyperbolic the behavior of solutions near it cannot be predicted from the linearization alone. For example, if $\mu < 0$ it may be true that the equilibrium is an attractor but if so this depends on higher order terms ignored by the linearization. In a heroic feat of calculation Marsden and McCracken have written down a formula for a number which depends on the second and third derivatives of X at p_0 and which determines the stability of the equilibrium in certain cases. Suppose that μ is negative. If this number, which I dub MARMC, is also negative then the Hopf equilibrium is an attractor and we will call it a *Hopf attractor*. If MARMC is positive then p_0 is not an attractor. Finally, if MARMC $= 0$ the decision depends on still higher terms.

The computation of MARMC depends on the use of a special coordinate system, or basis, analogous to a basis of eigenvectors in the real eigenvalue case.

Lemma 2. *Suppose an orientation and inner product* $(\ ,\ )$ *are chosen for* $\mathbb{R}_0^4$ *and suppose that the eigenvalues of* $L\colon \mathbb{R}_0^4 \to \mathbb{R}_0^4$ *are* $\varepsilon \pm i\lambda$ $(\lambda > 0)$ *and* $\varepsilon + \mu$ *(writing the real eigenvalue as* $\varepsilon + \mu$ *instead of* μ *is a matter of notational convenience). Then there exists a positively oriented orthonormal basis* $\{u, v, w\}$ *for* $\mathbb{R}_0^4$ *and a positive number* k *such that the matrix of* L *with respect to* $\{u, v, w\}$ *is given by*

$$\begin{pmatrix} \varepsilon & \lambda k & \gamma_1 \\ -\lambda k^{-1} & \varepsilon & \gamma_2 \\ 0 & 0 & \varepsilon + \mu \end{pmatrix}. \tag{2.6}$$

Proof. Associated with the complex eigenvalue $\varepsilon + i\lambda$ is a complex eigenvector $u_0 + iv_0$. Multiplying by the complex number $a + ib$ gives another complex eigenvector $u_1 + iv_1$ with

$$u_1 = au_0 - bv_0 \qquad \text{and} \qquad v_1 = av_0 + bu_0.$$

Computing the inner product of u_1 with v_1 we get:

$$(u_1, v_1) = (a^2 - b^2)(u_0, v_0) + ab[(u_0, u_0) - (v_0, v_0)].$$

Now if $(u_0, v_0) = 0$ then choose $a = 1$ and $b = 0$, i.e. $u_1 + iv_1 = u_0 + iv_0$. If $(u_0, v_0) \neq 0$ then choose $a, b > 0$ such that $r = a/b$ is the positive root of the quadratic equation:

$$r - r^{-1} = [(v_0, v_0) - (u_0, u_0)]/(u_0, v_0).$$

So in any case, $u_1 + iv_1$ is a complex eigenvector with $(u_1, v_1) = 0$. Now define:

$$u = u_1/\|u_1\| = u_1/(u_1, u_1)^{1/2},$$
$$v = v_1/\|v_1\| = v_1/(v_1, v_1)^{1/2},$$
$$k = \|u_1\|/\|v_1\|.$$

Here one needs to notice that $\{u_1, v_1\}$ are linearly independent since $u_1 + iv_1 \neq 0$ and $\lambda \neq 0$. In particular, neither vector equals 0.

Finally, since $\mathbb{R}_0^4$ is three dimensional there is a unique line through the origin perpendicular to the plane spanned by u_1, v_1 or equivalently by u, v. Choose w to be the unit vector on this line such that $\{u, v, w\}$ is positively oriented. $\{u, v, w\}$ is an orthonormal basis and it is fairly straightforward to check that the matrix of L is given by (2.6). Q.E.D.

We will call $\{u, v, w\}$ an *eigenframe* for the linear map L and k the *skewness* of L.

I recommend that the reader digest all this by ruminating on the following example:

$$\frac{dx}{dt} = \varepsilon x + \lambda y + \delta(x^2 + y^2)x,$$
$$\frac{dy}{dt} = -\lambda x + \varepsilon y + \delta(x^2 + y^2)y,$$
$$\frac{dz}{dt} = (\varepsilon + \mu)z. \tag{2.7}$$

The origin is an equilibrium whose linearization is obtained by setting $\delta = 0$. The eigenvalues are $\varepsilon \pm i\lambda$ and $\varepsilon + \mu$. The eigenframe is the standard basis and the skewness is 1.

Replacing the rectangular coordinates by cylindrical coordinates with $r^2 = x^2 + y^2$ and $\theta = \tan^{-1}(y/x)$ we get the equations:

$$\frac{dr}{dt} = \varepsilon r + \delta r^3,$$

$$\frac{d\theta}{dt} = -\lambda,$$

$$\frac{dz}{dt} = (\varepsilon + \mu)z. \tag{2.8}$$

If ε and $\varepsilon + \mu$ are not zero then the origin is a hyperbolic equilibrium, attracting if both are negative.

If $\varepsilon = 0$ and $\mu \neq 0$ then the origin is a Hopf equilibrium and MARMC is δ. If μ and δ are negative then the origin is still an attractor. However, if $\delta = 0$ so that the original equation is linear then the plane $z = 0$ is filled with concentric circular orbits. Compare system (1.3).

There is an important difference between a hyperbolic attractor and a Hopf attractor which is illustrated by this system.

Perturbing the equations slightly does not change the attracting character of a hyperbolic attractor. This is because the eigenvalues are perturbed slightly and so remain to the left of the imaginary axis. In other words, a hyperbolic attractor is a robust attractor. This is not true of a Hopf attractor. Consider what happens on the $z = 0$ plane as ε changes sign from negative to positive. The results depend on the sign of MARMC $= \delta$:

$\delta = 0$: When $\varepsilon < 0$ all of the orbits spiral in toward the origin and when $\varepsilon > 0$ all of the orbits spiral out away from the origin. When $\varepsilon = 0$ the nonequilibrium orbits are all cycles about the origin.

$\delta < 0$: When $\varepsilon \leqslant 0$ all of the orbits spiral in toward the origin. When $\varepsilon > 0$ then orbits beginning far away spiral in and orbits near the origin spiral out. These two regimes are separated by a unique cycle which is a circular orbit of radius $(\varepsilon/-\delta)^{1/2}$. If, in addition, $\mu < 0$ then this circular orbit is a limit cycle, i.e. it is an attractor. As ε increases the radius of the cycle increases proportionally to $\varepsilon^{1/2}$.

$\delta > 0$: This case is the reverse of the previous one. Unique cycles appear with ε negative as circular orbits of radius $(-\varepsilon/\delta)^{1/2}$. Now, however, the cycles are repelling separating a regime of spirals into the hyperbolic attractor at the origin and a regime of spirals outward to infinity.

The occurrence of cycles associated with Hopf equilibria is the phenomenon of Hopf bifurcation.

Suppose that p_0 is a Hopf equilibrium for a vectorfield X and that there is a family of vectorfields depending on a real parameter ρ with X corresponding to $\rho = 0$. Since p_0 is nondegenerate the perturbed equilibria and their eigenvalues are functions of ρ. Assume that the real part $\varepsilon(\rho)$ of the complex eigenvalues satisfies:

$$\frac{d\varepsilon(\rho)}{d\rho} \neq 0 \qquad \text{at} \quad \rho = 0 \tag{2.9}$$

This means that as ρ changes sign so does $\varepsilon(\rho)$ and the complex eigenvalues cross the imaginary axis. Furthermore, ε is a monotone function of ρ with $\varepsilon = 0$ when $\rho = 0$. Inverting to get the function $\rho(\varepsilon)$ and substituting has the effect of replacing the parameter ρ by ε itself.

Thus, for ε near zero we have a one parameter family of vectorfields X^ε with $X^0 = X$ and equilibria p_ε of X^ε depending smoothly on ε. The eigenvalues of X^ε at p_ε are given by $\varepsilon \pm i\lambda(\varepsilon)$ and $\mu(\varepsilon)$ with $\lambda(\varepsilon) > 0$ and $\mu(0) \neq 0$. So by shrinking the domain of ε we can assume $\mu(\varepsilon)$ has the same sign as $\mu(0)$ for all ε.

The Hopf Bifurcation Theorem (Marsden and McCracken, 1976, Theorem 3.15) says that if p_0 is a Hopf attractor for X, i.e. $\mu(0) < 0$ and MARMC < 0, then for $\varepsilon > 0$ small enough there is a unique periodic solution for X^ε cycling around p_ε and attracting nearby orbits. Furthermore, these cyclic attractors are robust in that any small enough perturbation of X^ε merely distorts the cycle.

Actually, the Hopf Bifurcation Theorem describes the occurrence of cycles for any family of vectorfields perturbing a Hopf equilibrium and satisfying (2.9). However, robust limit cycles are predicted only in the Hopf attractor case.

The program of the following section describes the Hopf equilibria for the TLTA model, shows how to decide which are Hopf attractors and provides a perturbation technique to get the cycles.

3. The Hopf Variety of the TLTA System

In the TLTA case, we denote by A, a and B, b the two alternate alleles at the two loci. The four gamete genotypes are numbered by:

$$1 = AB, \qquad 2 = Ab, \qquad 3 = aB, \qquad 4 = ab.$$

If the population is in state $p \in \Delta$ then the marginal distributions, i.e. the gene distributions at each locus are defined by:

$$p_A = p_1 + p_2, \qquad p_B = p_1 + p_3,$$
$$p_a = p_3 + p_4, \qquad p_b = p_2 + p_4.$$

On Δ the difference measure of linkage disequilibrium is defined by:

$$d = p_1 p_4 - p_2 p_3. \tag{3.1}$$

The name arises from the following easily checked equations:

$$\begin{aligned} p_1 &= p_A p_B + d, \\ p_2 &= p_A p_b - d, \\ p_3 &= p_a p_B - d, \\ p_4 &= p_a p_b + d. \end{aligned} \tag{3.2}$$

Define the vector ξ of $\mathbb{R}^4_0$ to be $(1, -1, -1, 1)$. Then (3.2) says that $p - \pi(p) = d\xi$. Recall that $\pi(p)$ is the distribution in Δ having the same marginals as p but in linkage equilibrium.

Table 1

	aa	*Aa*	*AA*
BB	m_{33}	$m_{13} = m_{31}$	m_{11}
Bb	$m_{34} = m_{43}$	$m_{14} = m_{41}$ = = $m_{23} = m_{32}$	$m_{12} = m_{21}$
bb	m_{44}	$m_{24} = m_{42}$	m_{22}

A selection matrix (m_{ij}) is the symmetric 4×4 matrix of zygotic fitnesses. We assume the absence of position effects which means that the two double heterozygotes have the same fitness. In addition to the symmetry condition, (1.1), this means we assume:

$$m_{14} = m_{23}. \tag{3.3}$$

Any selection matrix can be displayed by a 3×3 table (see Table 1).

Adding a common constant to all of the m_{ij}'s has no effect on the selection differential equation. This is because the magnitude of $\bar{m}$ describes the rate at which the entire population is growing while the rate of increase of p_i depends only on the difference $m_i - \bar{m}$. Alternatively, the addition of the constant to the function $\bar{m}$ does not affect its gradient on Δ. Thus, we can assume that the central element of Table 1 is zero. In other words, we define a *normalized selection matrix* to be a symmetric 4×4 matrix satisfying

$$m_{14} = m_{23} = 0. \tag{3.4}$$

The set M of normalized selection matrices is an eight dimensional vectorspace.

For the TLTA model the recombination vectorfield R is equal to $-rb\,d\xi$ (Crow and Kimura, 1970, p. 197), i.e. the differential equations for selection plus recombination are given by:

$$\frac{dp_i}{dt} = p_i(m_i - \bar{m}) - rb\,d\xi_i \qquad (i = 1, \ldots, 4). \tag{3.5}$$

Here r is the recombination rate in crossover per birth and $b = b_{14} = b_{23}$ is the birthrate for the double heterozygotes. Thus, rb is a nonnegative constant.

The behavior of (3.5) depends only on the ratio of rb to the fitness numbers m_{ij} provided that the former is positive. This is because multiplying the matrix (m_{ij}) and the scalar rb by a positive number s has the same effect as the time change of replacing the variable t by t/s. Thus, multiplication by a positive constant does not affect what happens only the rate at which it happens. We normalize (3.5) by assuming

$$rb = 1, \tag{3.6}$$

and define for every normalized selection matrix the vectorfield X^m on Δ by

$$X^m(p)_i = p_i(m_i - \bar{m}) - d\xi_i \qquad (i = 1, \ldots, 4). \tag{3.7}$$

Thus, everything depends on the ratio of fitness to recombination and the selection

matrices which our program generates are really ratios. One then can get an example with any positive rate of recombination rb one likes by multiplying by the chosen rb value. Compare the discussion in Akin (1979), Sec. I.9.

The vectorfields of (3.7) form an eight dimensional family parametrized by M. We can write this as a function X: $M \times \Delta \to \mathbb{R}_0^4$ defined by $X(m,p) = X^m(p)$. We define the *equilibrium manifold* Σ to be the set of pairs (m,p) such that p is an equilibrium for X^m, i.e.

$$\Sigma = \{(m,p) \in M \times \Delta \colon X(m,p) = 0\}.$$

Σ is an eight dimensional submanifold of $M \times \Delta$. In fact, the projection of Σ onto M is a local diffeomorphism at $(m,p) \in \Sigma$ iff p is a nondegenerate equilibrium for X^m. These results follow from the fact that X is a submersion of $M \times \Delta$ onto $\mathbb{R}_0^4$.

We are looking for Hopf attractors and so we define the *Hopf variety*:

$$\Sigma_H = \{(m,p) \in \Sigma \colon \text{The linearization of } X^m \text{ at } p \text{ has an imaginary eigenvalue}\}.$$

It is possible to give an implicit description of Σ_H as the solution set of algebraic equations and inequalities in the variables p_i and m_{ij}. First, Σ is defined by the vanishing of the right side of (3.7). So Σ is described by four equations of which three are independent. Notice that these equations are linear in the m_{ij} variables. By choosing a coordinate system on Δ it is possible to write down explicit formulae for the coefficients of the characteristic polynomial of the linearization of X^m at p. This yields three functions T, M, D: $\Sigma \to \mathbb{R}$. Lemma 1 of the previous section then says that $(m,p) \in \Sigma$ lies in Σ_H iff $D(m,p) = T(m,p) \cdot M(m,p)$ and $M(m,p) > 0$. Removing the subset of degenerate equilibria where $D(m,p) = 0$ divides Σ_H into two open subsets of Hopf equilibria defined by the sign of the real eigenvalue. In principle, if one has a point (m,p) in Σ_H with $D(m,p) \neq 0$ it is possible to compute MARMC and so find the Hopf attractors.

The problem with this approach is the Eq. (2.4) is nonlinear in the variables p_i and m_{ij}. It is not clear how to solve it to get examples in Σ_H. Instead we look for a parametric description of Σ_H. This consists of a seven dimensional manifold S_H and a function σ_H of S_H onto Σ_H. Then by choosing parameter values, i.e. a point of S_H, the map σ_H gives an explicit formula for an element of Σ_H, i.e. a list of values for p_i and m_{ij} such that $(m,p) \in \Sigma_H$. Also the parametric description allows one to compute the eigenframe for the linearization (cf. Lemma 2 of the previous section) and so to compute MARMC fairly directly.

In order to get Hopf bifurcations from Hopf attractor examples we parametrize not just Σ_H but:

$$\Sigma_C = \{(m,p) \in \Sigma \colon \text{The linearization of } X^m \text{ at } p \text{ has a complex eigenvalue}\}.$$

Σ_C is an open subset of Σ and contains Σ_H.

As often happens in applied problems, it is important to choose the right system of coordinates. It is clear from the biology that two of the coordinates on Δ should describe the state of two loci separately, e.g. p_A and p_B. The proper choices turn out to be:

$$x = p_1 + p_2 - p_3 - p_4 = p_A - p_a = 2p_A - 1,$$
$$y = p_1 - p_2 + p_3 - p_4 = p_B - p_b = 2p_B - 1.$$

x and y vary between -1 and $+1$. The origin where $x = y = 0$ corresponds to $p_A = p_a = p_B = p_b = \frac{1}{2}$. This point, or rather the segment in Δ mapping to it, plays a special role in the theory as we will see below.

For the third coordinate there are three natural choices:

$$z = p_1 - p_2 - p_3 + p_4,$$
$$d = p_1 p_4 - p_2 p_3,$$
$$L = \log p_1 - \log p_2 - \log p_3 + \log p_4$$
$$= \log(p_1 p_4) - \log(p_2 p_3)$$
$$= \log(p_1 p_4 / p_2 p_3).$$

z has the advantage that like x and y it is linear in the p_i's. On the other hand it is not directly a measure of linkage disequilibrium. The condition of linkage equilibrium is described by the equation $z = xy$. d, the difference measure of disequilibrium, we have already met as it occurs in the formula for X^m. L, the log of the ratio measure of disequilibrium, is theoretically important but hard to use in computations. Its importance comes from the fact that the Shahshahani gradient of L, $\bar{\nabla} L$, is a nonvanishing vectorfield on Δ which is everywhere tangent to the segments of constant gene frequency at each locus. In other words as one moves tangent to $\bar{\nabla} L$, x and y remain constant. This implies that $\bar{\nabla} L$ is everywhere perpendicular to $\bar{\nabla} x$ and $\bar{\nabla} y$.

A key step in the parametric description is the introduction of the analogue of a cylindrical coordinate system at points of $\mathring{\Delta}$. We restrict attention to $\mathring{\Delta}$ because the Shahshahani metric is only defined at interior distributions.

The vertical coordinate is L. We normalize the gradient of L by dividing by its length to get the unit vectorfield:

$$\text{Ver} \equiv \bar{\nabla} L / \|\bar{\nabla} L\|. \tag{3.8}$$

The length of $\bar{\nabla} L$, and all other lengths, are computed using the Shahshahani metric.

Now define the angular coordinate:

$$\theta = \tan^{-1}(y/x).$$

As usual θ is not defined when $x = y = 0$ and is only defined up to multiples of 2π. The gradient of θ times $x^2 + y^2$ defines the vectorfield:

$$\text{Ang}_0 \equiv x\bar{\nabla} y - y\bar{\nabla} x. \tag{3.9}$$

Ang_0 vanishes when $x = y = 0$ but only vanishes there because $\bar{\nabla} x$ and $\bar{\nabla} y$ are linearly independent. So we can normalize to define:

$$\text{Ang} \equiv \text{Ang}_0 / \|\text{Ang}_0\| \qquad \text{(undefined when } x = y = 0\text{)}. \tag{3.10}$$

Because Ang_0 is a linear combination of $\bar{\nabla} x$ and $\bar{\nabla} y$ it is perpendicular to $\bar{\nabla} L$. By direct computation one can show that it is also perpendicular to the gradient of:

$$\begin{aligned} N &\equiv \log p_1 + \log p_2 + \log p_3 + \log p_4 \\ &= \log(p_1 p_4) + \log(p_2 p_3) \\ &= \log(p_1 p_2 p_3 p_4). \end{aligned}$$

This means that the functions $\log(p_1 p_4) \pm \log(p_2 p_3)$ are constants of motion for the vectorfield Ang_0. Consequently, $\log(p_1 p_4)$ and $\log(p_2 p_3)$ are constants of motion and hence so is d which is the difference of their antilogs. We conclude that Ang_0 and hence Ang are perpendicular to $\bar{\nabla} L$, Ver and $\bar{\nabla} d$.

Now define the vectorfield:

$$\begin{aligned} \mathrm{Rad}_0 &\equiv (x(1 - y^2) - 4dy)\,\bar{\nabla} x + (y(1 - x^2) - 4dx)\,\bar{\nabla} y \\ &= \bar{\nabla}\tfrac{1}{2}[x^2 + y^2 - x^2y^2 - 8xyd] + 4xy\,\bar{\nabla} d. \end{aligned} \tag{3.11}$$

Again Rad_0 vanishes only when $x = y = 0$ and is perpendicular to $\bar{\nabla} L$. So we can normalize to define

$$\mathrm{Rad} \equiv \mathrm{Rad}_0 / \|\mathrm{Rad}_0\| \qquad \text{(undefined when } x = y = 0\text{)}. \tag{3.12}$$

One can check that Rad_0 and Ang_0 are mutually perpendicular. This implies that we have a set of three mutually perpendicular unit vectorfields {Ver, Rad, Ang} defined at every interior distribution away from the origin segment where $x = y = 0$. In other words, we have an orthonormal basis or a frame defined at each point. We call it the *cylindrical frame*, $\mathfrak{B}_c$.

This frame does not come from a system of coordinates. To see this, look at the candidate for the radial coordinate. Define R^2 to be the bracketed expression in (3.11):

$$R^2 = x^2 + y^2 - x^2y^2 - 8xyd. \tag{3.13}$$

As it happens $R^2 = \|\mathrm{Ang}_0\|^2 + x^2y^2$ and so $R^2 \geqslant 0$ on $\mathring{\Delta}$ vanishing only at the origin segment.

Because Ang is perpendicular to Rad_0 and to $\bar{\nabla} d$ it follows from (3.11) that it is perpendicular to $\bar{\nabla}\frac{1}{2}R^2$, i.e. R, too, is a constant of motion for Ang. The equations $L =$ some constant and $R =$ some positive constant define a closed curve in $\mathring{\Delta}$ and Ang is everywhere tangent to it pointing in the direction of increasing θ. So the flow of Ang consists entirely of closed cycles at constant L levels. However, the cycles do not fit together to make vertical cylinders. Because of the $\bar{\nabla} d$ term in (3.11) $\bar{\nabla}\frac{1}{2}R^2$ is not perpendicular to Ver.

The concept of a moving frame was introduced by the geometer Elie Cartan to do just the job we want, namely to generalize the concept of coordinate system.

To see why the cylindrical frame is natural for the study of these cycles, we recall the proof of Akin (1979), Theorem IV.3, the existence theorem for the cycles.

If X is any vectorfield on $\mathring{\Delta}$ we can define the linearization of X at any point p (equilibrium or not). It is a linear map $L_p(X)\colon \mathbb{R}^4_0 \to \mathbb{R}^4_0$ obtained by differentiating X at p. Dual to $L_p(X)$ is a bilinear form called the *Hessian* $H_p(X)$. The Hessian operator is linear in X, i.e. if X and Y are vectorfields on $\mathring{\Delta}$ then $H_p(X + Y) = H_p(X) + H_p(Y)$ at every point p. Furthermore, a vectorfield X is a gradient (with respect to the Shahshahani metric) iff $H_p(X)$ is a symmetric form at every point p. In particular, the Hessian of the selection field $\bar{\nabla}(\bar{m}/2)$ is everywhere symmetric.

Furthermore, if we fix some point p then by choosing the selection matrix properly we can arrange that $H_p(\bar{V}(\bar{m}/2))$ be any symmetric form at all. On the other hand, the recombination field R is not a gradient. Hence, at some point p, $H_p(R)$ is not symmetric, i.e. it has a nonzero antisymmetric part. So by varying the selection matrix one can cause a Hopf bifurcation to occur at such a point.

In eigenvalue language, if $H_p(X)$ is symmetric then $L_p(X)$ can be represented by a symmetric matrix and so has only real eigenvalues. In other words, complex eigenvalues for $L_p(X)$ require an antisymmetric part for $H_p(X)$. In particular, for X^m this antisymmetric part must come from the recombination term. It turns out that $H_p(R)$ is symmetric for $p \in \mathring{\Delta}$ iff p is in linkage equilibrium, i.e. $d = 0$, or p lies on the origin segment, i.e. $x = y = 0$. Thus, $L_p(X^m)$ has only real eigenvalues if $d = 0$ or if $x = y = 0$. Notice that if p lies in the boundary of Δ and $d \neq 0$ then p cannot be an equilibrium point for X^m because selection points along the boundary and recombination points into the interior. As a consequence, if $(m, p) \in \Sigma_C$ then $p \in \Delta'$ where

$$\Delta' = \{p \in \mathring{\Delta} : d \neq 0 \text{ and } (x, y) \neq (0, 0)\}.$$

This explains the special role of the origin segment and also why we can ignore boundary points.

So everything depends on the antisymmetric part of $H_p(R)$. As it happens this bilinear form annihilates the vectorfield Ang. It follows that the cylindrical frame is an eigenframe for the antisymmetric part of the linearization of R.

We arrive at the parametric description by working backwards. Begin with $(m, p) \in \Sigma_C$. $p \in \Delta'$ and so the cylindrical frame is defined at p and induces an orientation on $\mathbb{R}^4_0$. By Lemma 2 of the previous section there exists an eigenframe $\mathfrak{B}_u = \{u, v, w\}$ for the linearization $L^{(m,p)}$ of X^m at p. The orientation of $\mathfrak{B}_u$ is chosen to agree with that of the cylindrical frame $\mathfrak{B}_c$. With respect to $\mathfrak{B}_u$ the matrix of $L^{(m,p)}$ is given by (2.6). In particular, the eigenvalues of $L^{(m,p)}$ are $\varepsilon \pm i\lambda$ ($\lambda > 0$) and $\varepsilon + \mu$ and the skewness is $k > 0$. The bases $\mathfrak{B}_u$ and $\mathfrak{B}_c$ are related by a 3×3 matrix O. Since the bases are orthonormal and of consistent orientation O is an orthogonal matrix of determinant 1, i.e. O is a member of the *special orthogonal group* $SO(3)$. The convention that λ is positive turns out to imply the condition

$$dO_{33} > 0, \tag{3.14}$$

i.e. O_{33} is nonzero and has the same sign as d. Notice that $d \neq 0$ because $p \in \Delta'$.

For the parametric description define

$$S_H = \{(p, O, k) \in \Delta \times SO(3) \times \mathbb{R} : p \in \Delta',\ k > 0 \text{ and } dO_{33} > 0\},$$

$$S_C = S_H \times \mathbb{R} = \{(p, O, k, \varepsilon) : p \in \Delta',\ k > 0 \text{ and } dO_{33} > 0\}.$$

Because $SO(3)$ is a three dimensional manifold, S_H and S_C are open manifolds of dimension seven and eight respectively.

On S_H there are defined two real valued functions Nm and Dn. Now define

$$\tilde{S}_H = \{(\mu, p, O, k) \in \mathbb{R} \times S_H : Nm - \mu\, Dn = 0\}$$

$$\tilde{S}_C = \tilde{S}_H \times \mathbb{R}.$$

To a point $(\mu, p, O, k, \varepsilon) \in \tilde{S}_C$ there corresponds a unique normalized selection matrix $m = (m_{ij})$ satisfying the following properties:

(1) p is an equilibrium point for X^m, i.e. $(m,p)\in\Sigma$.

(2) $(m,p)\in\Sigma_C$. In fact, if the frame $\mathfrak{B}_u$ is defined by the condition that it is related to $\mathfrak{B}_c$ by the matrix O then $\mathfrak{B}_u$ is an eigenframe for the linearization. With respect to $\mathfrak{B}_u$ the matrix of the linearization is given by (2.6) with ε, μ and k given by the point of $\tilde{S}_C$.

(3) The entries of m as well as $\lambda, \gamma_1, \gamma_2$ are given by explicit formulae in the variables μ, p_i, O_{ij}, k and ε.

In practice we use the open subsets:

$$\mathring{S}_H = \{(p,O,k)\colon Dn \neq 0\}, \qquad \mathring{S}_C = \mathring{S}_H \times \mathbb{R}.$$

These subsets map to $\tilde{S}_H$ and $\tilde{S}_C$ respectively by defining:

$$\mu = Nm/Dn. \tag{3.15}$$

The map $\sigma_C\colon \tilde{S}_C \to \Sigma_C$ defined by $\sigma_C(\mu,p,O,k,\varepsilon) = (m,p)$ and its restriction $\sigma_H\colon \tilde{S}_H \to \Sigma_H$ are smooth maps onto their ranges. The induced maps on $\mathring{S}_C$ and $\mathring{S}_H$ are onto an open dense subset of their ranges.

The program is used as follows: Begin with a point $(p,O,k)\in S_H$. Compute Nm and Dn. Check that $Dn \neq 0$ and that μ defined by (3.15) is negative. Compute the matrix (m_{ij}) and $\lambda, \gamma_1, \gamma_2$. Compute MARMC using the frame $\mathfrak{B}_u$. Check that MARMC is negative. At each stage if one of the check steps fails go back and begin with a new point in S_H. In this way one hunts for Hopf attractors. Actually the formulae of the program allow one to make asymptotic estimates which describe certain regions of the parameter space where the associated pair (m,p) is a Hopf attractor.

Once one has found $(p^*,O^*,k^*)\in\mathring{S}_H$ such that μ and MARMC are negative then one looks at the line $(p^*,O^*,k^*,\varepsilon)$ in $\mathring{S}_C$ and computes the selection matrix m^ε_{ij}. The entries turn out to be linear functions of ε. For each ε, $(m^\varepsilon,p^*)\in\Sigma_C$ and the family X^{m^ε} parametrized by ε execute a Hopf bifurcation at the common equilibrium point p^*. Since p^* is a Hopf attractor when $\varepsilon = 0$, it follows that robust limit cycles occur for all X^{m^ε} with ε small and positive.

The formulae for Nm, Dn, m_{ij}, etc. are all rational functions of the p_i, O_{ij} and k variables except for three expressions which are square roots of simple polynomials in the p_i's. This is important because while we need only determine the sign of MARMC our examples depend on landing exactly on points of the submanifold Σ_H. While the cycles are robust it is not clear a priori how robust, i.e. to what accuracy one can round off and still observe them. Thus, for the examples in the next section I give exact results, which are rather messy fractions and then round off the three place decimals so that the reader can get a sense of the shape of the selection pressures. However, the cycling is only certain for the fraction version.

We conclude this section by looking at related models.

In deriving (1.2) and (3.5) the zygote genotype distribution is computed from the gamete distribution by assuming Hardy-Weinberg proportions. The direct study of selection and recombination for zygote distributions was initiated by Nagylaki and Crow (1974) and carried on by Nagylaki and by Hoppensteadt. In particular, they made rigorous the belief that the older models are good approximations when the evolutionary forces are weak.

Following the notation of Nagylaki and Crow (1974) we define:

$$\begin{aligned} &\text{fertility:} && a_{ik,lj} = b + sb_{ik} + sb_{lj}, \\ &\text{mortality:} && d_{ij} = d + se_{ij}, \\ &\text{recombination:} && c = sr, \end{aligned}$$

where b_{ij} and e_{ij} satisfy (1.1) and (3.3), and $r > 0$. Fertility is assumed additive and s is a positive parameter measuring the strength of selection and recombination.

If P_{ij} is the frequency of the ij zygote then the departure from Hardy-Weinberg proportions is measured by Q_{ij}, i.e.

$$P_{ij} = p_i p_j + Q_{ij}. \tag{3.16}$$

When the correct equations for selection and recombination are written down it becomes apparent that Hardy-Weinberg proportions are not preserved, i.e. $\{Q_{ij} = 0\}$ is not an invariant manifold for the system. However, by applying the invariant manifold theorem of Hirsch, Pugh and Shub (1977) it follows that there is an attracting invariant manifold nearby, provided that s is sufficiently small. It can be defined by writing Q_{ij} as a function of p and s. Its precise shape depends on b_{ij}, e_{ij}, etc. but as we are holding these numbers fixed we suppress them as arguments of the functions Q_{ij}. Thus, at quasi-Hardy-Weinberg equilibrium the zygote genotypes are again functions of the gamete genotypes via (3.16). Furthermore, as s approaches zero the functions $Q_{ij}(p, s)$ approach zero together with all derivatives in the p variables. Finally, on the invariant manifold selection and recombination can be written as

$$\frac{dp_i}{dt} = s[p_i(m_i - \bar{m}) - rbd\xi_i + F(p, s)_i], \tag{3.17}$$

where $F(p, s)_i$ approaches zero with s and $m_{ij} = b_{ij} - e_{ij}$.

Now replace t by the time variable $\tau = st$ to get

$$\frac{dp_i}{d\tau} = p_i(m_i - \bar{m}) - rbd\xi_i + F(p, s)_i. \tag{3.18}$$

So on the invariant manifold of quasi-Hardy-Weinberg proportions the motion is a perturbation of (3.5). If (3.5) has a robust limit cycle of period T, (3.18) will have a limit cycle of period approximately T for all s sufficiently small. Hence, (3.17) will have a limit cycle of period approximately T/s for all s sufficiently small.

For discrete-time models the results also carry over if selection is weak. Define $m_{ij} = \log w_{ij}$ where w_{ij} is the positive matrix of multiplicative parameters of selection. For weak selection, i.e. $w_{ij} - \bar{w}$ small compared with $\bar{w}$:

$$m_{ij} = \log w_{ij} \approx \log \bar{w} + \frac{w_{ij} - \bar{w}}{\bar{w}}.$$

In particular, $\bar{m} \approx \bar{w}$ and $m_i - \bar{m} \approx (w_i - \bar{w})/\bar{w}$. Substituting in (3.5) and replacing dp_i/dt by $\Delta p_i/s$ we get the usual discrete model with a strength parameter s. Thus, under weak selection and tight linkage a numerical solution of (3.5) is a path for the discrete model. When the cycles are attracting they will be observable in a numerical solution.

The discrete-time model can be examined directly. Alan Hastings has been investigating equilibria of the discrete TLTA model using a linear programming approach. Recently, and independently, he has discovered stable cycles. He should soon have a comprehensive plot of his examples.

4. Cycling Examples

The actual programming of my "program" was done in collaboration with my colleague Stanley Ocken. In particular, he introduced me to the MACSYMA package at MIT which is nicely adapted to our algebraic needs.

The point p is chosen to be:

$$p = (\tfrac{13}{40}, \tfrac{9}{40}, \tfrac{9}{40}, \tfrac{9}{40}) = (0.325, 0.225, 0.225, 0.225)$$

at which

$$x = y = z = \tfrac{1}{10} = 0.1 \qquad \text{and} \qquad d = \tfrac{9}{400} = 0.0225.$$

The orthogonal matrix O is the 3×3 identity matrix. In particular, $O_{33} = 1$ which is positive as is d.

The parameter k as well as ε are left as free variables. Dn is not zero and $\mu = Nm/Dn$ is given by

$$\mu = -\frac{54}{325} + \frac{1}{100\sqrt{390}}\left(\frac{k^2-1}{k^2+1}\right) \sim -0.166.$$

In particular, μ is negative for all values of k.

$$\lambda = \frac{3}{5\sqrt{585}}\left(\frac{k}{k^2+1}\right) \sim 0.0248\left(\frac{k}{k^2+1}\right), \qquad \gamma_1 = \gamma_2 = 0.$$

These values describe the matrix (2.6) of the linearization of X^m at the point p with respect to the eigenframe. Since O is the identity the eigenframe is the cylindrical frame {Ver, Rad, Ang} at p.

The selection parameters m_{ij} are given by:

$$m_{11} = -\frac{4{,}194}{7{,}605} + \frac{18}{45\sqrt{390}}\left(\frac{k^2-1}{k^2+1}\right) + \frac{40}{13}\varepsilon,$$

$$m_{22} = m_{33} = -\frac{2{,}236}{7{,}605} + \frac{2}{45\sqrt{390}}\left(\frac{k^2-1}{k^2+1}\right) + \frac{40}{9}\varepsilon,$$

$$m_{44} = \frac{234}{7{,}605} - \frac{38}{45\sqrt{390}}\left(\frac{k^2-1}{k^2+1}\right) + \frac{40}{9}\varepsilon,$$

$$m_{12} = m_{21} = m_{13} = m_{31} = -\frac{15{,}561}{7{,}605} - \frac{18}{45\sqrt{390}}\left(\frac{k^2-1}{k^2+1}\right),$$

$$m_{24} = m_{42} = m_{34} = m_{43} = -\frac{18{,}187}{7{,}605} + \frac{14}{45\sqrt{390}}\left(\frac{k^2-1}{k^2+1}\right),$$

$$m_{14} = m_{41} = m_{23} = m_{32} = 0.$$

Table 2

$k = 1.1$	aa	Aa	AA
BB	$-0.294 + 4.44\varepsilon$	-2.05	$-0.550 + 3.08\varepsilon$
Bb	-2.39	0	-2.05
bb	$0.027 + 4.44\varepsilon$	-2.39	$-0.294 + 4.44\varepsilon$
$k = 7$	aa	Aa	AA
BB	$-0.292 + 4.44\varepsilon$	-2.07	$-0.532 + 3.08\varepsilon$
Bb	-2.38	0	-2.07
bb	$-0.010 + 4.44\varepsilon$	-2.38	$-0.292 + 4.44\varepsilon$

Setting $\varepsilon = 0$ to get the Hopf equilibria we compute MARMC and discover

$$\text{MARMC} < 0 \qquad \text{provided} \qquad k > 1.02.$$

Hence, for every value of $k > 1.02$ the above matrix with $\varepsilon = 0$ has a Hopf attractor at p. Hence, robust limit cycles occur for each such k provided that ε is a small enough positive number.

To see the pattern of the selection matrices we write them in the form of the genetic table for two representative values of k, rounding off to three decimal places (see Table 2).

When $\varepsilon = 0$ the maximum of mean fitness in the $k = 1.1$ case occurs at fixation at a and b, i.e. when $p_4 = 1$. This vertex is an attractor for selection alone and recombination increases the strength of attraction. However, in the $k = 7$ case the maximum of mean fitness occurs on the line between vertex 4 and vertex 1 (= fixation at AB). Recombination moves this stable equilibrium into the interior.

In both cases mean fitness has a saddle point in $\mathring{\Delta}$. So selection alone has a hyperbolic interior equilibrium. The simplest conjecture is that this equilibrium continues to the Hopf attractor p at unit recombination. If so this might yield a general picture of how these cycles arise, namely, it may be that increasing strength of recombination can have the effect of stabilizing an initially hyperbolic, i.e. saddle point, equilibrium in the interior. For weaker levels of recombination, i.e. before the critical level at which the equilibrium becomes a Hopf attractor and beyond which it is a hyperbolic attractor, cyclic attractors occur in a neighborhood of the equilibrium. The resolution of this conjecture requires qualitative analysis of the behavior of the examples.

This analysis should be made easier by the use of the symmetry which occurs in certain of the examples and is exhibited by the particular cases given above.

The symmetric viability model studied by a number of authors culminating in Karlin and Feldman's work (1970) is the class of models invariant under the involution $\pi_0 : \Delta \to \Delta$ defined by $\pi_0(p_1, p_2, p_3, p_4) = (p_4, p_3, p_2, p_1)$. It is related to the origin symmetry of the xy square defined by $\pi_0(x, y) = (-x, -y)$. For cycling examples the involution to look at is $\pi_+ : \Delta \to \Delta$ defined by $\pi_+(p_1, p_2, p_3, p_4) = (p_1, p_3, p_2, p_4)$ which is related to reflection across the diagonal in the xy square defined by $\pi_+(x, y) = (y, x)$. Notice that our examples are diagonal symmetric.

For examples which are π_+ symmetric, and this includes the ones above, the fixed point set of π_+, defined by $p_2 = p_3$ or equivalently $x = y$, is an invariant subset for the vectorfields. This fixed point set is a two dimensional cell in Δ and the cycles occur in it. Since two dimensional dynamical systems are rather tractable, it should be possible to give a complete description of the behavior of these symmetric examples, at least on and close to the fixed point set.

Notice that as k approaches ∞, μ and the selection parameters m_{ij} do not change much, but λ approaches 0. Since λ is the approximate frequency of the cycles, we get cycles of arbitrarily low frequency. Multiplying X^m by s, i.e. replacing m_{ij} and rb by sm_{ij} and srb, also multiplies λ by s. So as selection gets weaker and linkage tighter (s approaches zero) the pattern of behavior is the same but occurs at a slower rate. In particular, the frequency gets lower. Furthermore, one cannot hope to have λ approach ∞ while the m_{ij}'s remain bounded because an eigenvalue is bounded by the norm of the linear map. This means that there exists some constant $K > 0$ such that

$$\lambda \leqslant K \max(|m_{ij}|, rb).$$

All this suggests that low frequency, long period cycles may be the rule.

References

Akin, E.: The geometry of population genetics. Lecture notes in biomathematics. Vol. 31. Berlin-Heidelberg-New York: Springer 1979

Akin, E.: Domination or equilibrium. Math. Biosci. **50**, 239–250 (1980)

Akin, E.: Cycling in the two locus system (Manuscript 1981)

Crow, J. F., Kimura, M.: An introduction to population genetics theory. New York: Harper and Row Publishers, Inc. 1970

Ewens, W. J.: With additive fitness, the mean fitness increases. Nature **221**, 1076 (1969a)

Ewens, W. J.: Population genetics. London: Methuen and Co. 1969b

Hastings, A.: Disequilibrium, selection, and recombination: Limits in two-locus two-allele models. Genetics (in press)

Hastings, A.: Induced underdominance at a stable equilibrium (Two locus models) (to appear)

Hirsch, M. W., Pugh, C. C., Shub, M.: Invariant manifolds. Lecture notes in mathematics. Vol. 583. Berlin-Heidelberg-New York: Springer 1977

Hoppensteadt, F. C.: A slow selection analysis of two-locus, two-allele traits. Theoret. Population Biology **9**, 68–81 (1976)

Karlin, S., Feldman, M. W.: Linkage and selection: Two locus symmetric viability models. Theoret. Population Biology **1**, 39–71 (1970)

Kimura, M.: On the change of population fitness by natural selection. Heredity **12**, 145–167 (1958)

Marsden, J. E., McCracken, M.: The Hopf bifurcation and its applications. Berlin-Heidelberg-New York: Springer 1976

Moran, P. A. P.: On the nonexistence of adaptive topographies. Ann. Hum. Genet. **27**, 338–393 (1964)

Nagylaki, T.: The evolution of one- and two-locus systems: I. Genetics **83**, 583–600 (1976)

Nagylaki, T.: Selection in one- and two-locus systems. Lecture notes in biomathematics. Vol. 15. Berlin-Heidelberg-New York: Springer 1977

Nagylaki, T., Crow, J. F.: Continuous selection models. Theoret. Population Biology **5**, 257–283 (1974)

Shahshahani, S.: A new mathematical framework for the study of linkage and selection. Memoirs AMS **211**, (1979)

Taylor, P. D., Jonker, L. B.: Evolutionarily stable strategies and game dynamics. Math. Biosci. **40**, 145–156 (1978)
Wright, S.: Surfaces of selective value. PNAS **58**, 165–172 (1967)
Zeeman, E. C.: Population dynamics from game theory. In: Proc. Int. Conf. Global Theory of Dynamical Systems, Northwestern, Evanston, Illinois 1979

Received February 19/Revised July 16, 1981