UNSOLVED PROBLEMS IN EVOLUTIONARY THEORY

P. A. P. MORAN
AUSTRALIAN NATIONAL UNIVERSITY

1. Introduction

The theory of evolution is a field of research in which many mathematical investigations have been made and in which many unsolved problems remain. Whatever one's opinions are about the ultimate causes of evolutionary progress (a still debatable problem), it is now clear that an overwhelming part in determining the direction of evolution is played by selection and mutation, and furthermore, that the only empirically testable theory of evolution that has at the present time any plausibility is that evolution is the result of these two factors.

In the *Proceedings of the Second Berkeley Symposium*, Professor W. Feller gave a lecture on "Diffusion processes in genetics" in which he drew attention to the very interesting stochastic processes constructed and studied by Wright and Fisher which attempt to describe and explain what happens to the genetical structure of populations when we take account of the fact that the population has a finite size so that what happens at birth, mating, and death is not determinate but has a random character. This lecture has stimulated a good deal of further research on such random processes (see Moran [14] and more recent papers of Karlin, McGregor, and Ewens).

So long as the population size is effectively finite, use of the theory of random processes is essential. In this way we can study such problems as the rate at which a population becomes homozygous at some locus when there is no mutation (a phenomenon called "Drift" by Sewall Wright and which he holds, I think incorrectly, to be of importance in evolution), the stationary distributions of gene frequencies when there is mutation, and the probability of survival of new mutants.

Although these problems must remain of great interest to students of stochastic processes, I believe that the study of the deterministic processes which effectively describe what happens when the population size is large are more important from the evolutionary point of view. This is due to the fact that the theory of stochastic genetic processes shows that the influence of the random element is of the same order as the reciprocal of the population size. Thus, for example, for populations of size greater than 10^4, deterministic theory is sufficient to answer nearly all problems concerning selection in which the selection coefficients are, say, greater than 10^{-3}. Similarly, in studying the effect of mutation

we can use deterministic theory provided the population size is substantially greater than the reciprocal of the mutation rate (except, of course, for the special problem of finding the probability of survival of a single new mutant).

Most natural populations probably satisfy these conditions, as is illustrated by the extreme case of most species of insects, although occasionally there are some small isolated populations whose size is sufficiently small to require consideration of the stochastic element. This is illustrated, for example, by laboratory populations or by the populations of the grasshopper, Moraba Scurra, in the graveyards of Australian country towns, which have been extensively studied by M. J. D. White.

It is of some importance to note that for the stochastic element to be important from the evolutionary point of view, such populations must be completely isolated. It can be shown (Moran [13]) that conclusions about the stochastic behavior of large populations are unaffected if these populations are divided into many smaller populations, provided there is a small but nonzero amount of migration between them. The effective population size is then practically equal to the whole population.

Our aim, therefore, in this paper is to consider the theory of evolution in populations large enough for stochastic variations to be ignored. Since we are then concerned with average numbers or proportions of individuals of a specific type, the subject can still be rightly regarded as a part of statistical analysis.

We define the state of such a population at a given time, or in the case of non-overlapping generations in a given generation, by a finite set of quantities $x_1, \cdots, x_k$. It is often convenient to write this set in vector form, $\underset{\sim}{x} = (x_1, \cdots, x_k)$, although it does not have the algebraic properties of a vector. We attach a suffix, t, so that $\underset{\sim}{x}_t$ is the state in the tth generation, or if we are considering a model which evolves continuously in time, we write $\underset{\sim}{x}(t)$ where t is a continuous variable.

The assumed properties of the population then enable us to write down a relation between the values of $\underset{\sim}{x}$ for one generation (the parent generation) and the corresponding values in the next generation. We therefore begin by considering the properties of recurrence relations of the form

$$\underset{\sim}{x}_{t+1} = \underset{\sim}{f}(\underset{\sim}{x}_t) = (f_1(\underset{\sim}{x}_t), \cdots, f_k(\underset{\sim}{x}_t)). \tag{1}$$

In most cases the components of $\underset{\sim}{x}$ will be the relative proportions of genetic individuals of various types (genes or zygotes) and will therefore add to unity and be nonnegative. Equality (1) is then the mapping of the interior of a $(k - 1)$ dimensional simplex onto itself. This is the case considered in this paper.

2. Recurrence relations and their related differential equations

We first look for the stationary states of the system which will be given by the relevant solutions of the equation obtained from (1) by putting $\underset{\sim}{x} = \underset{\sim}{x}_t = \underset{\sim}{x}_{t+1}$, which is thus

$$\underline{x} = \underline{f}(\underline{x}). \tag{2}$$

It is possible for the solutions of this equation to consist of isolated points, lines in the space of the (x_i), or the whole region in which $\underline{x}$ is defined. Any solution of (2) will be said to be a "stationary point" if $\sum x_i = 1$, all $x_i \geqq 0$, and if we have further that all the x_i are greater than zero, we shall say it is an "internal" stationary point. In most cases we will be concerned with the number and position of isolated stationary internal points.

We then have to consider the local stability of the system for small variations about the stationary values. Write $\underline{x}_t = \underline{x} + \underline{y}_t$ where $\underline{x}$ is a solution of (2) and $\underline{y}_t$ is a vector $(y_{1t}, \cdots, y_{kt})$ of small quantities adding to zero. Then under suitable conditions on the function $\underline{f}(\underline{x})$,

$$x_i + y_{i,t+1} = f_i(\underline{x}) + \sum_j y_{j,t} \frac{\partial f_i(\underline{x})}{\partial x_j} + O(\sum_j y_{j,t}^2). \tag{3}$$

It is clear that a sufficient condition for the point $\underline{x}$ to be locally stable is that all the roots of the matrix,

$$(a_{ij}) = \left(\frac{\partial f_i(\underline{x})}{\partial x_j}\right), \tag{4}$$

should have moduli less than unity, and a sufficient condition for instability is that at least one of the roots has modulus greater than unity. When all the moduli are equal to unity the situation becomes more difficult. As an example of this, if we take $w_1 = 1$, $w_2 = w_3 = 2$ in the equation (9) discussed later, we obtain a recurrence relation of the form

$$x_{t+1} = f(x_t) = x_t \frac{2 - x_t}{2 - x_t^2}, \tag{5}$$

where $0 \leqq x_t \leqq 1$. Even though $x_t = 0$ is a stationary point, $f'(0) = 1$, and the above criterion does not apply, although this is clearly stable since $x_{t+1} < x_t$ for $0 < x_t < 1$.

It is sometimes easier to consider only the behavior of the population in the neighborhood of the boundary of the region. Starting from a boundary point and introducing a perturbation of an allowed kind (for example, all $x_i \geqq 0$), it is often fairly easy to show that all boundary points are unstable. This makes it likely that there is a stable stationary point in the interior, although it does not prove it since there remains the theoretical possibility of cyclic behavior. This approach has been extensively developed by Bodmer [3] (see also Bodmer and Parsons [4], [5]), who was primarily interested in the initial progress and survival of new mutant genes, and I have used it elsewhere (Moran [17]) to discuss a model in which I was unable to prove directly the existence of a stable stationary interior state.

Associated with equation (1), there is a related theory for continuous time using differential equations. This is most simply approached if we suppose that the components x_i of $\underline{x}$ are the relative frequencies of elements of k different kinds in some population. We assume that in a small interval of time $(t, t + dt)$,

a fraction, dt, of the population dies and is replaced by dt new offspring of types whose relative frequency is given by the components of $\underline{f}(\underline{x}(t))$, where $\underline{x}(t)$ is the vector giving the structure of the population at time t, the notation being changed to denote that t is now continuous. We then have $\underline{x}(t + dt) = (1 - dt)\underline{x}(t) + dt\underline{f}(\underline{x}(t))$ from which we derive the system of differential equations

$$\frac{d\underline{x}(t)}{dt} = -\underline{x}(t) + \underline{f}(\underline{x}(t)). \tag{6}$$

The stationary states are again given by the solutions of (2). Contrary to what we might expect, we shall see later that (6) does not always correspond to the natural model of a genetic population evolving continuously in time unless the components $x_i(t)$ of $\underline{x}(t)$ are such as to give a complete description of the state of the system. Nevertheless, (6) can be used as an approximation to the behavior of (1), has the same stationary points, and has a closely related stability theory.

Let $\underline{x}$ be a solution of (2) and write $\underline{x}(t) = \underline{x} + \underline{y}(t)$ where $\underline{y}(t)$ is a vector of small components representing a "variation" or perturbation around the point $\underline{x}$. The asymptotic behavior of $\underline{y}(t)$ will provide us with a theory of the stability of the system (6) around the point $\underline{x}$ (for a general survey of the analytical theory, see Cesari [6]).

Under suitable conditions on $\underline{f}(\underline{x}(t))$, we can immediately derive the "variational" equations for $\underline{y}(t)$ which are

$$\frac{dy_i(t)}{dt} = -y_i(t) + \sum_j y_j(t) \frac{\partial f_i(\underline{x})}{\partial x_j}. \tag{7}$$

Then all $y_i(t) \to 0$ if the roots of the matrix (4) are such that their real parts are less than unity.

Stability in the above sense is a "local" stability. Given a stationary point $\underline{x}$ we can associate with it the set of all points P, such that starting from P the population ultimately converges to $\underline{x}$. This may be called the domain of attraction of $\underline{x}$. If the domain of attraction of an internal point $\underline{x}$ contains all internal points of the simplex, we say that $\underline{x}$ is "globally" stable.

3. Selection dependent on a single locus with two alleles

Consider a model of a genetic population in which there are two alleles, A and a, at a single locus such that the "fitness" of a newly formed individual zygote is dependent on this locus alone. There are three possible types of zygote, AA, Aa, and aa, and we define their fitnesses to be nonnegative numbers w_1, w_2, w_3, proportional to the numbers of gametes that a zygote of these types, newly formed by union of gametes, will contribute to the formation of the next generation. We assume that successive generations are nonoverlapping and that mating is random, so that the probabilities of formation of zygotes AA, Aa, and aa will be $p(A)^2$, $2p(A)p(a)$, $p(a)^2$ respectively, where $p(A)$ and $p(a) = 1 - p(A)$ are the relative frequencies of the A and a gametes entering into the

formation of the zygotic population. Let $p = 1 - q$ be the frequency $p(A)$ in the first (parent) generation which we suppose to have been produced in this manner. Then the mean fitness of this population is

$$W = w_1p^2 + 2w_2pq + w_3q^2. \tag{8}$$

It is important to notice that in this discrete model the w's are not simply the relative fertilities of the adult zygotes, but they also depend on the probabilities that the newly formed zygotes survive to the age of reproduction.

If p' is the frequency of A in the next generation, we have

$$p' = p\frac{w_1p + w_2q}{W}, \tag{9}$$

an equation of type (1). This can also be written

$$p' - p = pq\frac{(w_1 - w_2)p + (w_2 - w_3)q}{W} = \frac{pq}{W}\frac{dW}{dp}, \tag{10}$$

where the derivative of W is calculated remembering that $q = 1 - p$. The stationary points are therefore given by equating the right-hand side of (10) to zero and thus, either $p = 0$, or $q = 0$, or the derivative of W is zero for some internal point satisfying $0 < p < 1$. It is further easily seen that for such an internal point to be locally stable, we must have

$$\frac{d^2W}{dp^2} < 0. \tag{11}$$

If this is so, $2w_2 - w_1 - w_2 > 0$ and also $w_2 > w_1$, $w_2 > w_3$. The internal point is then stable for any allowable deviations, and the points $p = 0, 1$ are unstable stationary points.

Thus if W attains a maximum at an internal stable stationary point, its value there is

$$W = \frac{w_2^2 - w_1w_3}{2w_2 - w_1 - w_3}, \tag{12}$$

and p is given by

$$p = \frac{w_2 - w_3}{2w_2 - w_1 - w_3}. \tag{13}$$

Thus the population moves to maximize the mean fitness W. Notice that the position and stability of such an internal point is unaltered if we transform w_i into another set of values by the transformation $w_i \to \lambda w_i + \mu$, where $\lambda > 0$, $\mu \geqq -\lambda \min w_i$, but the rate of convergence of the population to its equilibrium value is altered unless $\mu = 0$. It is also not hard to show that in this simple case the gene frequency changes monotonically from its initial value to its final value and does not overshoot the latter.

Now consider the effect of mutation on a stationary state of the population. Suppose first that in the absence of mutation there is a stationary stable internal state (that is, one with $0 < p < 1$). Let A mutate to a at a rate α_1 per gamete per generation, and a to A at a rate α_2. The rates α_1 and α_2 will usually not be

larger than 10^{-4}. Let $p_0 = 1 - q_0$ be the gene frequency of A in the absence of mutation so that p_0 is given by (13). Let selection in the production of the gametes occur first and then mutation. If p is the equilibrium gene frequency, we therefore have $Wp = p(1 - \alpha_1)(w_1p + w_2q) + q\alpha_2(w_2p + w_3q)$, from which we get, to a sufficiently good approximation,

$$p = \frac{w_3 - w_2}{2w_2 - w_1 - w_3} - \frac{w_1p_0 + w_2q_0}{p_0q_0(2w_2 - w_1 - w_3)}(\alpha_1p_0 - \alpha_2q_0) \tag{14}$$

$$= p_0 - \frac{W(\alpha_1p_0 - \alpha_2q_0)}{p_0q_0(2w_2 - w_1 - w_3)}.$$

The effect on the fitness, W, is therefore to lower it by an amount (ignoring terms of higher order) equal to

$$\frac{W^2(\alpha_1p_0 - \alpha_2q_0)^2}{2p_0^2q_0^2(2w_2 - w_1 - w_3)^2}\frac{d^2W}{dp^2} = \frac{2(w_2^2 - w_1w_3)^2(\alpha_1p_0 - \alpha_2q_0)^2}{p_0^2q_0^2(2w_2 - w_1 - w_3)^3}. \tag{15}$$

This is of the second order in the α's, and thus mutation has little effect on fitness dependent on a single locus at which there is a stable internal polymorphism.

The situation is different when there is no internal stable point. Consider the particular case $w_1 < w_2 < w_3$. Then the only stable stationary case is given by $p = 0$. The mean fitness W is a maximum for this value, but because it occurs at the end of the range, its derivative is not in general zero. The mutation of A to a will have little effect, so we put $\alpha_1 = 0$ and write α for α_2. The stationary points are then the solutions of $Wp = p(w_1p + w_2q) + q\alpha(w_2p + w_3q)$. One solution is $p = 1$, and the other is very close to $p = (w_3/w_3 - w_2)\alpha$, if we assume α small compared with $w_3 - w_2$. Since

$$\frac{dW}{dp} = 2(w_1p + w_2(q - p) - w_3q), \tag{16}$$

which is equal to $2(w_2 - w_3)$ at $p = 0$, the decrease in fitness is about $2w_3\alpha$ which is thus of the same order as the mutation rate.

It follows that if we consider a single locus, decrease in fitness due to mutation will only be important in the absence of a selectively balanced polymorphism. We shall see later that when fitness is controlled by more than one locus, the effect of mutation may be of the first instead of the second order at a stable stationary internal point, because fitness dependent on more than one locus is not in general a maximum at such a point.

4. Selection dependent on a single locus with more than two alleles

The above theory has been known for many years. More recently attention has been devoted to selection dependent on a single locus at which there are more than two alleles, and here the situation is not quite so simple. We assume again that we have nonoverlapping generations of diploid individuals who mate at random so that the frequencies of the various classes of zygotes in the next

generation are proportional to the products of the frequencies of the gametes. Let there be k alleles $A_1, \cdots, A_k$ at the given locus. Then we represent the zygotes by the symbols $(A_iA_j)(i, j = 1, \cdots, k)$. There are clearly $(1/2)n(n + 1)$ different pairs, but it is convenient algebraically to draw a distinction between (A_iA_j) and (A_jA_i) when $i \neq j$, while keeping their frequencies equal. This could be done, for example, by conventionally supposing that the first of the two alleles is derived from the father and the second from the mother. We shall not, however, consider models in which the gene frequencies or selection coefficients are different in the two sexes, although such models are of interest in some special cases.

Let the gene frequency of A_i among the gametes which formed the first generation be $p_i(\sum p_i = 1)$. Then using the above convention and the fact that mating is random, the frequency of newly formed (A_iA_j) in the first generation will be p_ip_j. The fitness of (A_iA_j) is defined to be w_{ij}, a nonnegative number proportional to the number of gametes this newly formed zygote will contribute to the next generation. Then the mean fitness of the first generation is

$$W = \sum_{ij} w_{ij}p_ip_j, \tag{17}$$

where the sum is taken over all values of i and j.

The numbers of A_i in the gametic product of this generation will be proportional to

$$\sum_j w_{ij}p_ip_j + \sum_k w_{ki}p_kp_i = 2p_i \sum_j w_{ij}p_j. \tag{18}$$

The sum of all such expressions is $2W$, and if we write p_i' for the frequency of A_i in the next generation, we have

$$p_i' = p_iW^{-1} \sum_j w_{ij}p_j. \tag{19}$$

The mean fitness, W', of the next generation is

$$W' = \sum w_{ij}p_i'p_j' = W^{-2} \sum_{ijkl} w_{ij}w_{ik}w_{lj}p_ip_jp_kp_l. \tag{20}$$

It is a remarkable fact that $W' \geqq W$, so that the mean fitness does not decrease from generation to generation. This was first proved by Scheuer and Mandel [22], and by Mulholland and Smith [18]. Later proofs were given by Atkinson, Watterson, and Moran [1], and by Kingman [9], [10]. The last proof is particularly simple and also makes clear the conditions under which $W' = W$, so that we give it here.

To do this we use repeatedly the inequality

$$\frac{\sum a_ib_i^k}{\sum a_i} \geqq \left(\frac{\sum a_ib_i}{\sum a_i}\right)^k, \tag{21}$$

where $a_i \geqq 0$, $\sum a_i > 0$, $b_i \geqq 0$, $k \geqq 1$. We have, on renumbering the suffices,

$$\sum_{ijkl} w_{ij}w_{ik}w_{lj}p_ip_jp_kp_l = \frac{1}{2}\sum_{ijk} w_{ij}w_{ik}(\sum_l w_{lj}p_l + \sum_m w_{mk}p_m)p_ip_jp_k \geqq \sum_{ijk} w_{ij}w_{ik}(\sum_{lm} w_{lj}w_{mk}p_lp_m)^{1/2}p_ip_jp_k, \tag{22}$$

on using the inequality between arithmetic and geometric means. This in turn is equal to

$$\sum_i p_i\{\sum_j w_{ij}p_j(\sum_l w_{lj}p_l)^{1/2}\}^2 \geqq \{\sum_i p_i \sum_j w_{ij}p_j(\sum_l w_{lj}p_l)^{1/2}\}^2, \quad \text{by (20).} \tag{23}$$

This can be written as $\{\sum_j p_j(\sum_l w_{lj}p_l)^{3/2}\}^2$, and using (21) again, this is not less than

$$\{(\sum_{lj} w_{lj}p_lp_j)^{3/2}\}^2 = (\sum w_{ij}p_ip_j)^3. \tag{24}$$

The sequence of inequalities is then equivalent to $W' \geqq W$, and equality only occurs at an internal point (all $p_i > 0$) if

$$\sum_j w_{lj}p_j = \sum_{ij} w_{ij}p_ip_j \qquad \text{for all } l. \tag{25}$$

This last equation is equivalent to the condition for stationarity that we get when we put $p_i' = p_i > 0$ in (19). Thus the mean fitness must increase if the initial point is an internal point which is not stationary, and a stable stationary internal point must be a local maximum of W.

We can look at this in another way. Consider the stationary points of W subject to the condition that $\sum p_i = 1$. Since W is a quadratic form in the p_i, finding its stationary points under this condition is equivalent to finding the stationary values of $\Phi = W/(\sum p_i)^2$ when the p_i have arbitrary nonnegative values. Then

$$\frac{\partial \Phi}{\partial p_i} = (\sum p_j)^{-3}\left\{\sum_j p_j \frac{\partial W}{\partial p_i} - 2W\right\}. \tag{26}$$

Putting this equal to zero, we again obtain the conditions $\sum_j w_{ij}p_j = W$, for all i, which, by (19), must hold at any internal stationary point. The fact that W is a quadratic form implies that if there is an isolated stationary point, there is only one such point.

The conditions under which such an internal stationary point is stable have been discussed by Mandel [12], Kimura [7], Penrose, Smith, and Sprott [21], and Kingman [10].

We shall see later that the principle that the mean fitness of a population tends to increase until it reaches a maximum holds in general only for fitness dependent on a single locus and is not a general principle of evolutionary theory.

5. The single locus problem with continuous time

Associated with equations (19), we naturally have a continuous time model in $p_i(t)$, the frequency of A_i at time t, satisfying the differential equations

$$\frac{dp_i(t)}{dt} = -p_i(t) + W(t)^{-1}p_i(t)\sum_j w_{ij}p_j(t), \tag{27}$$

where $W(t) = \sum_{ij} w_{ij}p_i(t)p_j(t)$.

If the $p_i(t)$ satisfy this equation, it is now much easier to prove that $W(t)$ cannot decrease, for we have

$$\begin{aligned}\frac{1}{2}\frac{dW(t)}{dt} &= \sum_{ij} w_{ij}p_i(t)\frac{dp_j(t)}{dt} \\ &= \sum_{ij} w_{ij}p_i(t)\{-p_j(t) + W(t)^{-1}p_j(t)\sum_k w_{jk}p_k(t)\} \\ &= W(t)^{-1}\{\sum_{ij} w_{ij}p_i(t)p_j(t)\sum_k w_{jk}p_k(t) - W(t)^2\} \\ &= W(t)^{-1}\{\sum_j p_j(t)\,(\sum_k w_{jk}p_k(t))^2 - [\sum_j p_j(t)\sum_k w_{jk}p_k(t)]^2\} \geqq 0,\end{aligned} \tag{28}$$

because of the inequality $\sum p_i a_i^2 \geqq (\sum p_i a_i)^2$ for $p_i \geqq 0$, $\sum p_i = 1$, $a_i \geqq 0$. The mean fitness of those zygotes which contain A_i is $\sum_k w_{jk}p_k(t)$. Thus the rate of change of $W(t)$ on this model is proportional to the variance of these quantities. This is Fisher's so called "Fundamental Theorem of Natural Selection."

Unfortunately, however, equations (27) do not correspond to a natural model of an evolving population and can be regarded only as an approximation useful when the w_{ij} are nearly all equal. The reason for this is that although mating remains always "at random," the frequencies of zygotes (A_iA_j) do not remain equal to the products of the gene frequencies, even if they start that way.

To see this, consider a population of zygotes (A_iA_j) whose frequencies are $p_{ij}(t) = p_{ji}(t)$, and are such that initially (for example, at $t = 0$) they are equal to $p_i(t)\,p_j(t)$. In any small interval of time $(t, t + dt)$ we suppose that a fraction, dt, of the population of zygotes dies, and is replaced by dt offspring produced by random mating among the parents existing at time t, due accord being taken of the selective differences, w_{ij}, which now have to be regarded as relative fertilities since the chance of death of an individual in $(t, t + dt)$ must be independent of its age for this model. Then

$$p_{ij}(t + dt) = (1 - dt)p_{ij}(t) + dt\,p_i(t)p_j(t)W(t)^{-2}\sum_k w_{ik}p_k(t)\sum_l w_{lj}p_l(t). \tag{29}$$

The appropriate differential equation is then

$$\frac{dp_{ij}(t)}{dt} = -p_{ij}(t) + p_i(t)p_j(t)W(t)^{-2}\sum_k w_{ik}p_k(t)\sum_l w_{lj}p_l(t). \tag{30}$$

At $t = 0$, $p_{ij}(t) = p_i(t)p_j(t)$, and if this is to remain true for $t > 0$, we must have

$$\begin{aligned}\frac{dp_{ij}(t)}{dt} &= p_i(t)\frac{dp_j(t)}{dt} + p_j(t)\frac{dp_i(t)}{dt} \\ &= p_i(t)p_j(t)\{-2 + W(t)^{-1}\sum_k w_{ik}p_k(t) + W(t)^{-1}\sum_l w_{lj}p_l(t)\}.\end{aligned} \tag{31}$$

Equating this to (30), replacing $p_{ij}(t)$ by $p_i(t)p_j(t)$, and assuming the latter to be nonzero, we get

$$(W(t) - \sum_k w_{ik}p_k(t))(W - \sum_l w_{lj}p_l(t)) = 0, \tag{32}$$

which implies that either $dp_i(t)/dt = 0$, or $dp_j(t)/dt = 0$, which is not in general true unless all the w_{ij} are equal, or the population is stationary. The population will tend to the same stationary point as the two previous models, and the $p_i(t)$ will have the values obtained by maximizing $\sum w_{ij}p_i(t)p_j(t)$. The true mean fitness of the population is $\sum w_{ij}p_{ij}(t)$. At the stationary point this will have the value previously given, but in the course of approaching this point, it may increase or decrease. This is very simply seen to be the case if we consider a two allele situation in which the zygotes (AA), (Aa), and (aa) have (say) the fitnesses 1, 2, 1, and suppose that the population starts in a state in which all individuals are (Aa). The true mean fitness must then decrease. Notice also that in this continuous time model "random mating" does not imply that the frequency of a zygote is equal to the product of the frequencies of the gametes which entered into its formation, except when the state of the population is stationary.

Before leaving the subject of selection dependent on a single locus we mention two further generalizations. We have so far assumed that the w_{ij} are constants independent of the gene frequencies. It is easy to imagine plausible situations in which this is not so. The mean fitness may then decrease, and the theory is naturally more complicated. Some examples are given in Moran [14].

Another restriction that may be relaxed is that of random mating. One alternative is "assortative" mating. This occurs when there is a tendency for some types of mating to have a greater or lower probability than the product of the frequencies of the types involved. When individuals have an increased probability of mating with other individuals of the same genotype, the assortment is said to be positive, and in the reverse case negative. In deterministic models the resulting theory is complicated but not too difficult. The effect of negative assortative mating is quite different from that of positive assortative mating in that the latter does not tend to alter the gene frequencies whereas the former has the effect of forcing the gene frequencies (in the two allele case) towards the value 1/2. Thus, negative assortative mating results in a pseudo-selective heterozygote advantage which can produce a permanent polymorphism. This is really the process which keeps the sex ratio at conception near unity. In stochastic models assortative mating, especially when negative, is much more difficult to deal with (Watterson [23]).

Another subject which might repay mathematical analysis is the evolution of genetic populations which depend on each other by, for example, competition, predation, or parasitism. This would be of interest in the study of the evolution of disease.

6. Selection dependent on several loci

We now turn to the case of more than one locus. Suppose that there are $k > 1$ loci at each of which there are two alleles, A_i and a_i $(i = 1, \cdots, k)$. These loci

need not all be on the same chromosome, but for a general theory we may imagine them all on a single chromosome separated, if necessary, into groups so far apart that there is so much crossing over between them that they behave as if they were on different chromosomes.

The restriction to two alleles at each locus is not important, for two loci at which there are alleles A_1, a_1, and A_2, a_2, respectively, might be so close together that there is no crossing over between them. They then behave as a single locus with four alleles which can be denoted by A_1A_2, A_1a_2, a_1A_2, and a_1a_2. Similarly, we could obtain a 'locus' with 2^s ($s > 2$) effective alleles, and then, by identification, a locus with any number of alleles less than 2^s.

We denote each gamete by a vector $(x_i) = (x_1, \cdots, x_k)$, where $x_i = 0, 1$ according to whether the gamete contains A_i or a_i. Similarly, we denote a zygote by an ordered pair of such vectors $(x_i|x_i') = (x_1, \cdots, x_k|x_1', \cdots, x_k')$ where it is again algebraically convenient to distinguish such a zygote from $(x_i'|x_i) = (x_1', \cdots, x_k'|x_1, \cdots, x_k)$ by some conventional distinction (such as the origin of the constituent gametes from father or mother) which ensures that the frequencies of such pairs of zygotes are equal.

The fitnesses of the zygotes are again defined by nonnegative numbers $w(x_i|x_i')$ where not only do we assume $w(x_i'|x_i) = w(x_i|x_i')$, but we shall also assume that the w's do not depend on how the alleles are distributed among the two components; that is, we shall suppose, for example, that $w(11|00)$ is equal to $w(10|01)$. This is equivalent to assuming that the w's are functions of $(z_i) = (x_i + x_i')$ only. We shall, however, have to distinguish (x_i) and (x_i') when we find the gametic output of each zygote (some evolutionary problems have been considered in which the w's depend on the (x_i) and (x_i') separately, but we shall not consider this "position effect").

Let the generations be nonoverlapping, and suppose that in the gametic output which formed the first generation the frequency of gametes of type $(x_i) = (x_1, \cdots, x_k)$ is $p(x_i) = p(x_1, \cdots, x_k)$. There are 2^k such gametes and

$$\sum_{x_i} p(x_1, \cdots, x_k) = 1. \tag{33}$$

These form zygotes by random mating so that the frequency of the zygote $(x_i|x_i')$ can be written as

$$p(x_i|x_i') = p(x_i)p(x_i'). \tag{34}$$

There are 2^{2k} such zygotes, 2^k of which are such that $x_i = x_i'$ for all i. These are the "homozygotes," and the remainder are "heterozygotic" in respect of one or more loci, and are identifiable in pairs with the same frequencies.

We now have to consider the manner in which the zygotes produce the gametes which form the next generation. The number of gametes they produce will be proportional to the corresponding w's, but now that we have more than one locus, we have to take account of the possibility of "recombination" due to "crossing over." Suppose that the loci are arranged in order on the chromosome at points $P_1, \cdots, P_k$. Recombination will occur between P_1 and P_2 if an odd

number of crossovers occur between them. Since there are $k - 1$ gaps between the loci, recombination can occur in 2^{k-1} ways (including the case of no recombination). The probabilities with which these occur will depend on the positions of the loci and their distances apart. An elaborate theory (Bailey [2]) exists for this dependence, but is not relevant here.

Suppose that an odd number of crossovers occur in each of the gaps $j_1, j_2, \cdots, j_s$ ($j_i = 1, \cdots, k - 1$), and zero or an even number in the remaining gaps. For the probability of this occurring we shall write

$$2R(0, 0, 0, \cdots, 0, 1, 1, 1, 0, \cdots) = 2R(1, 1, 1, \cdots, 1, 0, 0, 0, 1, \cdots), \tag{35}$$

there being k symbols equal to 0 or 1, and a change in the sequence occurring between the sth and $(s + 1)$th symbol if an odd number of crossovers occurs in the sth gap. We write this as $2R(\delta_1, \cdots, \delta_k)$ where $\delta_i = 0, 1$, and we have made this equal to $2R(1 - \delta_1, \cdots, 1 - \delta_k)$ in order to ensure symmetry in the following equations. We then have

$$\sum_{\delta_i} R(\delta_1, \cdots, \delta_k) = 1, \tag{36}$$

where the summation is taken over all sequences of k zeros and ones.

If a recombination of type $(\delta_1, \cdots, \delta_k)$ occurs in the production of gametes from a zygote of type $(x_1, \cdots, x_k|x'_1, \cdots, x'_k)$, the gametic output will consist of equal numbers of gametes of the form

$$((1 - \delta_1)x_1 + \delta_1 x'_1, \cdots, (1 - \delta_k)x_k + x'_k), \tag{37}$$

and of the form

$$(\delta_1 x_1 + (1 - \delta_1)x'_1, \cdots, \delta_k x_k + (1 - \delta_k)x'_k). \tag{38}$$

Thus, zygotes of type $(x_1, \cdots, x_k|x'_1, \cdots, x'_k)$ will have a total gametic output such that the numbers of gametes of type (38) will be proportional to

$$w(x_1, \cdots, x_k|x'_1, \cdots, x'_k)R(\delta_1, \cdots, \delta_k). \tag{39}$$

The frequency of such zygotes is

$$p(x_1, \cdots, x_k|x'_1, \cdots, x'_k) = p(x_1, \cdots, x_k)p(x'_1, \cdots, x'_k), \tag{40}$$

which we write as $p(x_i|x'_i) = p(x_i)p(x'_i)$. Thus, finally the total number of gametes of type $(y_1, \cdots, y_k)$ contributing to the next generation will be proportional to

$$\sum_{\delta_i} \sum_{x_i}{}' \sum_{x'_i}{}' R(\delta_1, \cdots, \delta_k)w(x_i|x'_i)p(x_i)p(x'_i), \tag{41}$$

the first sum being over all sequences $(\delta_1, \cdots, \delta_k)$, and the sums over x_i, x'_i being taken over all values such that $\delta_i x_i + (1 - \delta_i)x'_i = y_i$, the fact that this is a conditional sum being denoted by a dash. If we sum (41) over all possible sets $(y_1, \cdots, y_k)$, we get

$$W = \sum_{x_i} \sum_{x'_i} w(x_i|x'_i)p(x_i)p(x'_i), \tag{42}$$

the summation over the x_i and x_i' being now unrestricted. This is the mean fitness of the parent population.

Denote the frequency of the gamete $(y_1, \cdots, y_k)$ entering into the formation of the next generation by $p'(y_1, \cdots, y_k)$. We now have

$$p'(y_1, \cdots, y_k) = W^{-1} \sum_{\delta_i} \sum_{x_i}' \sum_{x_i'}' R(\delta_1, \cdots, \delta_k) w(x_i|x_i') p(x_i) p(x_i'). \tag{43}$$

These are the fundamental equations of evolutionary theory assuming random mating and fitnesses independent of gene frequencies.

Two extreme cases are of particular interest. First, we may suppose that the loci are so closely linked that no recombination at all is possible. We then have $R(0, \cdots, 0) = R(1, \cdots, 1) = 1/2$, and all the other values of R are zero. Equality (43) becomes

$$p'(y_1, \cdots, y_k) = W^{-1} \sum_{x_i} w(x_i|y_i) p(x_i) p(y_i) = \frac{1}{2} p(y_i) W^{-1} \frac{\partial W}{\partial p(y_i)}. \tag{44}$$

These equations are now the same as those considered in section 4, as is otherwise obvious from the fact that we may consider the population to be equivalent to one in which fitness is determined by a single locus at which there are 2^k alleles. Thus, in this case, the population moves to the unique state at which fitness is maximized, if the stability conditions referred to before are satisfied.

The other extreme case is that of independent assortment. Here we put $R(\delta_1, \cdots, \delta_k) = 2^{-k}$, so that

$$p'(y_1, \cdots, y_k) = 2^{-k} W^{-1} \sum_{x_i, x_i'}^{*} w(x_i|x_i') p(x_i) p(x_i'), \tag{45}$$

where the sum $\sum^*$ is taken over all sets (x_i), (x_i') such that (y_i) is obtained by a specified sequence of choices of the pairs x_i and x_i'. The theory of these equations is much less simple than that of (44).

7. Selection dependent on two loci

In an attempt to understand these equations better, we now confine ourselves to the case where selection is dependent on only two loci. Furthermore, since the above used notation is clumsy for only two loci, we change the notation to agree with that of a previous paper (Moran [16]) and which is similar to that used by other writers (for example, Bodmer and Parsons [5]).

We suppose there are two loci at which the possible alleles are A and a, B and b, respectively. The possible genotypes of the zygotes can then be arranged in the array

$$\begin{array}{ccc} AABB & AaBB & aaBB \\ AABb & AaBb & aaBb \\ AAbb & Aabb & aabb \end{array} \tag{46}$$

whose corresponding fitnesses we now denote by

$$\begin{matrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33}. \end{matrix} \tag{47}$$

Let c_1, c_2, c_3, c_4 be the relative frequencies of the gametes AB, aB, Ab, ab entering into the formation of the first (parental) generation. Then the frequencies of the above zygotes in this generation will be given by the array

$$\begin{matrix} c_1^2 & 2c_1c_2 & c_2^2 \\ 2c_1c_3 & 2c_1c_4 + 2c_2c_3 & 2c_2c_4 \\ c_3^2 & 2c_3c_4 & c_4^2, \end{matrix} \tag{48}$$

and the mean fitness W of this generation will be obtained by multiplying each of these frequencies by the corresponding w_{ij} and summing. Notice that $AaBb$ is ascribed the same fitness whether it is a "coupling" zygote (AB/ab), or a "repulsion" zygote (Ab/aB). When this is not true, the theory is less simple (Parsons [19]).

Let R be the probability of an odd number of crossovers between the two loci. Then $0 \leqq R \leqq 1$, and for the most part, we can also assume $R \leq 1/2$. Enumerating the terms of the sums in (43) we easily find the following equations for the frequencies (c_1', c_2', c_3', c_4') in the next generation:

$$\begin{aligned} c_1'W &= w_{11}c_1^2 + w_{12}c_1c_2 + w_{21}c_1c_3 + w_{22}c_1c_4 + Rw_{22}(c_2c_3 - c_1c_4) \\ &= \frac{1}{2}c_1\frac{\partial W}{\partial c_1} + Rw_{22}(c_2c_3 - c_1c_4); \end{aligned} \tag{49}$$

$$\begin{aligned} c_2'W &= w_{12}c_1c_2 + w_{13}c_2^2 + w_{22}c_2c_3 + w_{23}c_2c_4 - Rw_{22}(c_2c_3 - c_1c_4) \\ &= \frac{1}{2}c_2\frac{\partial W}{\partial c_2} - Rw_{22}(c_2c_3 - c_1c_4); \end{aligned} \tag{50}$$

$$\begin{aligned} c_3'W &= w_{21}c_1c_3 + w_{22}c_2c_3 + w_{31}c_3^2 + w_{32}c_3c_4 - Rw_{22}(c_2c_3 - c_1c_4) \\ &= \frac{1}{2}c_3\frac{\partial W}{\partial c_3} - Rw_{22}(c_2c_3 - c_1c_4); \end{aligned} \tag{51}$$

$$\begin{aligned} c_4'W &= w_{22}c_1c_4 + w_{23}c_2c_4 + w_{32}c_3c_4 + w_{33}c_4^2 + Rw_{22}(c_2c_3 - c_1c_4) \\ &= \frac{1}{2}c_4\frac{\partial W}{\partial c_4} + Rw_{22}(c_2c_3 - c_1c_4). \end{aligned} \tag{52}$$

The quantities c_1, c_2, c_3, c_4 enter into these equations in a symmetrical manner. Notice that this symmetry is not described by the full permutation group on four symbols but by a permutation group of order 8.

In these equations it is convenient to remember that W is a quadratic form in the four variables (c_1, c_2, c_3, c_4) with matrix

$$
\begin{array}{cccc}
w_{11} & w_{12} & w_{21} & w_{22} \\
w_{12} & w_{13} & w_{22} & w_{23} \\
w_{21} & w_{22} & w_{31} & w_{32} \\
w_{22} & w_{23} & w_{32} & w_{33}.
\end{array} \tag{53}
$$

Equations (49)–(52) were derived by Lewontin and Kojima [11] (see also Kimura [8]).

If $R > 0$, the maximization of W subject to $\sum c_i = 1$ does not give equations equivalent to those obtained for internal stationary points by putting $c_i' = c_i$ in (49)–(52). In fact, to maximize W with $\sum c_i = 1$ is equivalent to maximizing $W/(\sum c_i)^2$ for unrestricted nonnegative c_i, and then rescaling the c_i if necessary. This leads to the equations

$$\frac{\partial W}{\partial c_i} = 2W, \qquad i = 1, \cdots, 4 \tag{54}$$

for an internal point. If $R > 0$, these equations are not consistent with (49)–(52) unless $c_2c_3 = c_1c_4$, which is not in general true. Thus the stationary points of the system defined by (49)–(52) are not found by maximizing W, and it can be shown (Moran [16]) that W can decrease from one generation to the next even when each is produced by random mating.

Lewontin and Kojima call $(c_2c_3 - c_1c_4)$ the "coefficient of linkage disequilibrium," but this is misleading since it implies that the fact $c_2c_3 \neq c_1c_4$ is dependent on the existence of linkage between the loci (namely, $R \neq 1/2$). This is not true, and $c_2c_3 - c_1c_4$ might be better called a "coefficient of epistatic disequilibrium."

Wright thought that evolutionary change could be described by saying that populations tend to move so as to maximize W, and his argument can be summarized as follows. Using nonoverlapping generations and the above model it is clear that the population at each generation satisfies the Hardy-Weinberg law for each locus separately; that is, the frequencies of the homozygotes are the squares of the corresponding gene frequencies. Consider what happens at the first locus and write p for the frequency of A and $q = 1 - p$ for the frequency of a (thus $p = c_1 + c_3$, $q = c_2 + c_4$). Then the gene frequency p' in the next generation will be given by (10) where

$$W = w_1p^2 + 2w_2pq + w_3q^2, \tag{55}$$

and

$$
\begin{aligned}
w_1 &= \frac{1}{(c_1 + c_3)^2}\{w_{11}c_1^2 + 2w_{21}c_1c_3 + w_{31}c_3^2\}, \\
w_2 &= \frac{1}{(c_1 + c_3)(c_2 + c_4)}\{w_{12}c_1c_2 + w_{22}(c_1c_4 + c_2c_3) + w_{32}c_3c_4\}, \\
w_3 &= \frac{1}{(c_2 + c_4)^2}\{w_{13}c_2^2 + 2w_{23}c_2c_4 + w_{33}c_4^2\}.
\end{aligned} \tag{56}
$$

In carrying out the differentiation in (10), the w_i have to be kept fixed. Wright thus concluded that for stationarity we must have

$$\frac{dW}{dp} = 0, \tag{57}$$

and this is true in the above sense when the w_i are kept fixed. Since W is not a function of the gene frequencies alone, this does not imply that the internal stationary states of the population correspond to stationary values of W. In fact, Wright assumed that W could be written in the form

$$\begin{aligned} W = {} & w_{11}p^2P^2 + 2w_{12}pqP^2 + w_{13}q^2P^2 \\ & + w_{21}p^2PQ + 4w_{22}pqPQ + 2w_{23}q^2PQ \\ & + w_{31}p^2Q^2 + 2w_{32}pqQ^2 + w_{33}q^2Q^2 \end{aligned} \tag{58}$$

where $P = 1 - Q$ is the frequency of B; that is, $P = c_1 + c_2$. This is clearly not correct. Plotting (58) as a function of p and P we obtain an "adaptive topography," but from what has been said above, this throws no light on the behavior of the population.

Using this incorrect theory, I attempted (Moran [15]) to discuss the problem of determining how many different isolated stationary internal points such a population could have, and how many of these could be stable for small perturbations. The condition that the points be isolated is necessary as it is trivial to show that stationary points can fill up a whole interval. The problem is then that of finding the stationary points of (58), and the answer is that there can be at most five such points, and at most three local maxima, both bounds being attainable.

This theory is incorrect, and the number of possible stationary and stable points of the system (49)–(52) is an open question. However, it is possible that five and three are the correct bounds, and this is suggested by the following argument.

In the single locus problem considered in sections 3 and 4, the position and stability of the unique internal stationary point was unaltered when the w_i underwent the transformation $W_i \to \lambda w_i + \mu$, where $\lambda > 0$, $\mu > -\lambda \min w_i$. On the other hand, for the system (49)–(52) the stationary solutions are, in general, only invariant for transformations of the form $w_{ij} \to \lambda w_{ij}$, $\lambda > 0$. However, consider the effect of a transformation of the type $w_{ij} \to 1 + \lambda(w_{ij} - 1)$. If $\lambda \to 0$, the new w_{ij} tend to unity. It can then be easily shown that $c_2c_3 - c_1c_4 \to 0$, and the stationary points tend in the limit to the stationary points of the function defined by (58). This suggests that for w_{ij} near unity the number of possible stationary points remains the same, and hence, this holds true generally. However, this has not been proved.

From equations (49)–(52) we can easily derive, as in section 5, a set of four differential equations to describe the evolution of the population in a continuous time model. However, these are incorrect for the same reasons as those considered in section 5, and a full description of such a population requires the frequencies

of each of the zygotes. There are nine of these, adding to unity, and their values do not, in general, remain equal to the values $c_1^2, 2c_1c_2, \cdots$ given in section 7. The differential equations derived directly from (49)–(52) may, however, be regarded as a useful approximation when the w_{ij} are nearly equal.

Since the maximization of W does not provide a method for determining the stationary points, we may ask if there is any other function whose extreme values would serve. Relatively trivial examples are easy to construct, for example, by summing the squares of the differences of the right-hand sides of (49)–(52) from the values Wc_i, $(i = 1, \cdots, 4)$. This is not a quadratic form in the c_i, and it seems probable that there is no such quadratic form whose extreme or stationary values correspond to stationary states of the population. Furthermore, if a suitable function could be found, it would be desirable that it should be nondecreasing from generation to generation. One might hope in this way to construct a theory of the equations of evolutionary change analogous to the general theories of dynamics.

I have not succeeded in constructing such a function directly from equations (49)–(52). Professor S. Karlin has, however, shown that if we assume that iteration of these recurrence relations always results in the state of the system converging to a set of points S, then there exists a continuous function of the c_i which is nondecreasing. I am indebted to him for the following proof.

Let $\underset{\sim}{x}$ be a vector in a bounded space with a distance function $\rho(\underset{\sim}{x}, \underset{\sim}{y})$, and let $\underset{\sim}{x}' = T\underset{\sim}{x}$ be a transformation which has the property that the nth iterate, $T^n\underset{\sim}{x}$, converges to a subset S; that is, $\rho(T^n\underset{\sim}{x}, S) \to 0$. Then there exists a continuous function, $u(\underset{\sim}{x})$, such that $u(T\underset{\sim}{x}) \geqq u(\underset{\sim}{x})$ for all $\underset{\sim}{x}$. To prove this, define a continuous function

$$\begin{aligned} u_j(\underset{\sim}{x}) &> 0, \quad \text{if} \quad \rho(x, S) > j^{-1}, \\ u_j(\underset{\sim}{x}) &= 0, \quad \text{if} \quad \rho(x, S) \leqq j^{-1}, \end{aligned} \tag{59}$$

and then define

$$u(\underset{\sim}{x}) = \sum_{j=1}^{\infty} 2^{-j} \left\{ 1 + \sum_{n=0}^{\infty} u_j(T^n\underset{\sim}{x}) \right\}^{-1}. \tag{60}$$

The inner sum is really finite since $T^n\underset{\sim}{x} \to S$, and the outer sum is convergent. It is then obvious that $u(T\underset{\sim}{x}) \geqq u(\underset{\sim}{x})$.

This result assumes that we know that the iterates converge to the set S. This seems highly probable in the genetic case but is not known for certainty. What is needed is to construct a function like $u(\underset{\sim}{x})$ directly from the w_{ij} and c_i and use this to deduce the asymptotic behavior of $T^n\underset{\sim}{x}$.

8. The stability of equilibria and the effects of mutation

Suppose that the c_i are the values of a stationary internal point so that $0 < c_i < 1$. Then equations (49)–(52) must hold with c_i' replaced by c_i. As in Moran [16] (see also Kimura [8]), we investigate the effect of a small perturbation around this point by replacing the c_i with $c_i + \delta_i$, where the δ_i are small.

In the next generation we will get values c_i' which we put equal to $c_i + \delta_i'$, where the δ_i' are also small. Then expanding the recurrence relation around the point (c_1, c_2, c_3, c_4) and taking the linear terms only, we get

$$\delta_i' = \sum_j a_{ij}\,\delta_j \tag{61}$$

where (a_{ij}) is a matrix such that

$$a_{ii} = W^{-1}\left\{\frac{1}{2}\frac{\partial W}{\partial c_i} + \frac{1}{2}c_i\frac{\partial^2 W}{\partial c_i^2} + Rw_{22}k(i)\frac{\partial D}{\partial c_i} - c_i\frac{\partial W}{\partial c_i}\right\}, \tag{62}$$

and if $i \neq j$,

$$a_{ij} = W^{-1}\left\{\frac{1}{2}c_i\frac{\partial^2 W}{\partial c_i\,\partial c_j} + Rw_{22}k(i)\frac{\partial D}{\partial c_j} - c_i\frac{\partial W}{\partial c_j}\right\} \tag{63}$$

where $k(i) = +1$ if $i = 1, 4$, and where $k(i) = -1$ if $i = 2, 3$, and D is defined as $c_2c_3 - c_1c_4$.

It is easy to verify that $\sum_i a_{ij} = 0$, and therefore $\sum_i \delta_i' = 0$, as we expect. The matrix, therefore, has one root equal to zero, and the state will be stable if the other three roots are less than unity in modulus.

Since in general there are neighboring points to a stationary point (c_i) at which W is both greater and less than its value at (c_i), we might expect that the effect of mutation, which will shift the stationary point somewhat, could either increase or decrease the fitness. Consider the effect of such mutation at a stable point.

Suppose that A mutates to a at a rate α_1 per locus per generation, and a to A, B to b, b to B similarly at rates α_2, β_1, and β_2, respectively. If the (c_i) are the initial frequencies of the gametes, their frequencies after mutation will be

$$c_1 + \epsilon_1 = (1-\alpha_1)(1-\beta_1)c_1 + \alpha_2(1-\beta_1)c_2 + (1-\alpha_1)\beta_2c_3 + \alpha_2\beta_2c_4, \tag{64}$$

$$c_2 + \epsilon_2 = \alpha_1(1-\beta_1)c_1 + (1-\alpha_2)(1-\beta_1)c_2 + \alpha_1\beta_2c_3 + (1-\alpha_2)\beta_2c_4, \tag{65}$$

$$c_3 + \epsilon_3 = (1-\alpha_1)\beta_1c_1 + \alpha_2\beta_1c_2 + (1-\alpha_1)(1-\beta_2)c_3 + \alpha_2(1-\beta_2)c_4, \tag{66}$$

$$c_4 + \epsilon_4 = \alpha_1\beta_1c_1 + (1-\alpha_2)\beta_1c_2 + \alpha_1(1-\beta_2)c_3 + (1-\alpha_2)(1-\beta_2)c_4. \tag{67}$$

The α_i and β_i being very small, we may neglect terms of higher order than the first and replace the above equations by

$$\epsilon_1 = -(\alpha_1+\beta_1)c_1 + \alpha_2c_2 + \beta_2c_3, \tag{68}$$

$$\epsilon_2 = \alpha_1c_1 - (\alpha_2+\beta_1)c_2 + \beta_2c_4, \tag{69}$$

$$\epsilon_3 = \beta_1c_1 - (\alpha_1+\beta_2)c_3 + \alpha_2c_4, \tag{70}$$

$$\epsilon_4 = \beta_1c_2 + \alpha_1c_3 - (\alpha_2+\beta_2)c_4. \tag{71}$$

To find the new stationary position we use (49)–(52) with $c_1', \cdots, c_4'$ replaced by $c_1, \cdots, c_4$ on the left-hand sides, and $c_1, \cdots, c_4$ on the right-hand sides replaced by $c_1 + \epsilon_1, \cdots, c_4 + \epsilon_4$, and then solve for the c_i. The new values of the c_i will differ from the previous stationary values by quantities $\eta_1, \eta_2, \eta_3, \eta_4$, say, and the effect on the fitness is to add to it an amount

$$\sum \eta_i \frac{\partial W}{\partial c_i} \tag{72}$$

where the derivatives are taken at the previous stationary stable point.

The η_i will in general be of the same order as the mutation rates, and if W is not at a stationary point, the effect on W will be of the same order as these mutation rates. This is in sharp contrast to the situation for a single locus where the effect of mutation is of the second order in the mutation rates if the original stationary point was internal, and of the first order only when the original position was at the boundaries. Furthermore, it is now no longer obvious that the effect of mutation is to decrease the fitness, for W is not in general a maximum value. If the transformation (64)–(67) were a perfectly general one, it seems that the c_i could be displaced in any direction (consistent with $\sum c_i = 1$), and W could therefore be decreased or increased. We would then have the intriguing conclusion that the mean fitness of a population could be increased by mutation. However, the transformation (64)–(67) is not a general one, since it depends on only four constants instead of twelve (twelve because $\sum \epsilon_i$ must equal zero). Thus it may be true that this special type of perturbation can only result in a decrease in W, but this has not been proved.

9. What is epistasis?

Epistasis may be said to occur in the joint effect of two loci if the effect of varying the genotype at one locus depends on the genotype at the other. This is an ambiguous definition, particularly when the effect concerned is the fitness of the joint genotype. Thus suppose first that w_{ij} can be written as $u_i + v_j$, where u_i, $(i = 1, 2, 3)$ depends on the state at the first locus and v_j on the state at the second. The effects of the two loci may then be described as "additive," and if this condition does not occur, we could say that there is epistasis between the two loci.

Alternatively, we might define the absence of epistasis by putting $w_{ij} = u_i v_j$ where u_i and v_j are defined in a similar way. We shall call this the "multiplicative" case. Since "fitness" is defined to be proportional to the expected number of gametes that a newly formed zygote of the specified type will contribute to the next generation, it appears at first sight that the behavior of the gene frequencies at the two loci in the multiplicative case should be independent of each other so that epistasis is best defined as the absence of the condition $w_{ij} = u_i v_j$.

This is not correct, as we shall now see in investigating some of the consequences of assuming that in (49)–(52) the w_{ij} are one of these two forms.

We suppose $u_1 < u_2 > u_3$, $v_1 < v_2 > v_3$, and consider separately the two cases $w_{ij} = u_i + v_j$ ("additive"), and $w_{ij} = u_i v_j$ ("multiplicative"). Write

$$P = 1 - Q = \frac{u_2 - u_3}{2u_2 - u_1 - u_3}, \tag{73}$$

$$p = 1 - q = \frac{v_2 - v_3}{2v_2 - v_1 - v_3}. \tag{74}$$

Then the set of values

$$(c_1 = Pp,\ c_2 = Pq,\ c_3 = Qp,\ c_4 = Qq) \tag{75}$$

is a solution of (49)–(52) for both the additive and multiplicative case. This is an internal point and for it $c_2c_3 - c_1c_4 = 0$, so that the value of R is not relevant. In the additive case, equations (49)–(52) can be written as follows:

$$c_1'W = c_1\{u_1c_1 + u_1c_2 + u_2c_3 + u_2c_4\} + c_1\{v_1c_1 + v_2c_2 + v_1c_3 + v_2c_4\} + R(u_2 + v_2)(c_2c_3 - c_1c_4), \tag{76}$$

$$c_2'W = c_2\{u_1c_1 + u_1c_2 + u_2c_3 + u_2c_4\} + c_2\{v_2c_1 + v_3c_2 + v_2c_3 + v_3c_4\} - R(u_2 + v_2)(c_2c_3 - c_1c_4), \tag{77}$$

$$c_3'W = c_3\{u_2c_1 + u_2c_2 + u_3c_3 + u_3c_4\} + c_3\{v_1c_1 + v_2c_2 + v_1c_3 + v_2c_4\} - R(u_2 + v_2)(c_2c_3 - c_1c_4), \tag{78}$$

$$c_4'W = c_4\{u_2c_1 + u_2c_2 + u_3c_3 + u_3c_4\} + c_3\{v_2c_1 + v_3c_2 + v_2c_3 + v_3c_4\} + R(u_2 + v_2)(c_2c_3 - c_1c_4), \tag{79}$$

where

$$W = u_1(c_1 + c_2)^2 + 2u_2(c_1 + c_2)(c_3 + c_4) + u_3(c_3 + c_4)^2 + v_1(c_1 + c_3)^2 + 2v_2(c_1 + c_3)(c_2 + c_4) + v_4(c_2 + c_4)^2. \tag{80}$$

From this we see that the equations for a stationary point with $R = 0$ involve only the marginal sums $c_1 + c_2$, $c_3 + c_4$, $c_1 + c_3$, $c_2 + c_4$, that is, the gene frequencies. It follows therefore that for $R = 0$, $(c_1 = Pp + x,\ c_2 = Pq - x,\ c_3 = Qp - x,\ c_4 = Qq + x)$ is also a solution for any value of x that makes all the c_i nonnegative. Thus none of these stationary states can be strictly stable.

Notice further that all these points are maxima of W as given by (80), although they are not strict local maxima. For $R = 0$, W is nondecreasing from one generation to the next as follows from the case of one locus with four alleles.

If $R > 0$, we can prove that there is only one internal point which is stationary, and this must be given by (75). To do this, put $c_i' = c_i$ in (76)–(79) and divide by c_1, c_2, c_3, c_4, respectively, since these are nonzero. Subtracting (77) and (78) from the sum of (76) and (79), we get

$$R(u_2 + v_2)(c_2c_3 - c_1c_4)(c_1^{-1} + c_2^{-1} + c_3^{-1} + c_4^{-1}) = 0, \tag{81}$$

from which it follows that $c_2c_3 = c_1c_4$. If the latter is true, we can put $c_1 = xy$, $c_2 = x(1 - y)$, $c_3 = (1 - x)y$, $c_4 = (1 - x)(1 - y)$, and the only values of x and y satisfying (76)–(79) will be found to be $x = P$, $y = p$.

It is not known whether this solution is globally stable, namely that all internal points correspond to populations which converge to (75), although local stability could be investigated by using the criterion of section 8. It might seem that a proof of stability would be easily carried out along the lines of the single locus case, but this is not true. It is in general not true that the expression $(c_2c_3 - c_1c_4)$ decreases in absolute value from generation to generation, nor that the gene

frequencies $c_1 + c_2$, $c_1 + c_3$ converge monotonically to their final values. These surprising facts can be illustrated by simple numerical examples.

The multiplicative case is even less simple. Equations (49)–(52) now become the following:

$$c_1'W = c_1\{c_1u_1v_1 + c_2u_1v_2 + c_3u_2v_1 + c_4u_2v_2\} + Ru_2v_2(c_2c_3 - c_1c_4), \tag{82}$$

$$c_2'W = c_2\{c_1u_1v_2 + c_2u_1v_3 + c_3u_2v_2 + c_4u_2v_3\} - Ru_2v_2(c_2c_3 - c_1c_4), \tag{83}$$

$$c_3'W = c_3\{c_1u_2v_1 + c_2u_2v_2 + c_3u_3v_1 + c_3u_3v_2\} - Ru_2v_2(c_2c_3 - c_1c_4), \tag{84}$$

$$c_4'W = c_4\{c_1u_2v_2 + c_2u_2v_3 + c_3u_3v_2 + c_3u_3v_3\} + Ru_2v_2(c_2c_3 - c_1c_4). \tag{85}$$

The set of values (75) is again a solution of these equations but may or may not be stable. We may investigate some of the possible forms of behavior by considering the simple symmetric scheme of fitnesses given by

$$\begin{matrix} 1 & 2 & 1 \\ 2 & a & 2 \\ 1 & 2 & 1 \end{matrix} \tag{86}$$

where $a > 0$. Lewontin and Kojima [11] and Parsons [20] have also studied special cases like this. The case $a = 3$ is "additive" and $a = 4$ is "multiplicative." Equations (49)–(52) then become

$$Wc_1' = c_1^2 + 2c_1c_2 + 2c_1c_3 + ac_1c_4 + Ra(c_2c_3 - c_1c_4), \tag{87}$$

$$Wc_2' = 2c_1c_2 + c_2^2 + ac_2c_3 + 2c_2c_4 - Ra(c_2c_3 - c_1c_4), \tag{88}$$

$$Wc_3' = 2c_1c_3 + ac_2c_3 + c_3^2 + 2c_3c_4 - Ra(c_2c_3 - c_1c_4), \tag{89}$$

$$Wc_4' = ac_1c_4 + 2c_2c_4 + 2c_3c_4 + c_4^2 + Ra(c_2c_3 - c_1c_4). \tag{90}$$

Suppose that we start from an initial state in which $c_1 = c_4$, $c_2 = c_3$. Then in the next generation $c_1' = c_4'$, $c_2' = c_3'$ by symmetry, and furthermore, $c_1 + c_2 = c_1 + c_3 = c_1' + c_2' = c_1' + c_3' = 1/2$. Simplifying the equations we find $c_1' - c_2' = F(c_1 - c_2)$ where

$$F = \frac{1 + a - 2Ra}{1 + a - 8(a - 3)c_1c_2}. \tag{91}$$

Suppose first that $R = 0$. If $a = 3$, $F = 1$ and $c_i' = c_i$. This is the neutral equilibrium case already considered above since if $a = 3$, the fitnesses are additive. If $a > 3$, $F > 1$ and $c_1 = c_4$ converges to $1/2$, and $c_2 = c_3$ to zero. If $a < 3$, then $F < 1$ and all c_i converge to $1/4$. Thus, if $a < 3$ the state $c_i = 1/4$ is probably globally stable (we have not actually proved this as we have only considered the situation where $c_1 = c_4$, $c_2 = c_3$), while if $a > 3$, the state $c_i = 1/4$ is certainly unstable.

Now suppose $R > 0$. If $a < 3$, we again get $F < 1$, and the state $c_i = 1/4$ is probably globally stable. If $a > 3$, the value of F depends on a, and $F < 1$ or > 1 according to whether $2Ra$ is greater or less than $8(a - 3)c_1c_2$. Near the point $c_i = 1/4$, c_1c_2 is nearly $1/16$ and thus, if $2Ra > 1/2(a - 3)$, $c_1 - c_2$ will tend to zero (in particular, this is true for $R = 1/2$), and the point $c_i = 1/4$ is probably

locally stable (this could be investigated using the criterion of section 8). On the other hand, if $2Ra < 1/2(a-3)$ (for instance, for $a = 4$, $R < 1/16$), $c_1 - c_2$ will increase until $c_1c_2 = R\{4(a-3)\}^{-1}$. Beyond this point $c_1 - c_2$ will decrease. This is therefore a stationary point and probably also a locally stable point. As an example, consider the fitnesses (86) with $a = 4$, $R = 3/64$. Then $(c_i) = (3/8, 1/8, 1/8, 3/8)$ and $c_i = (1/8, 3/8, 3/8, 1/8)$ are stationary points. Thus for sufficiently tight linkage, the point $c_i = 1/4$ is certainly unstable. For $a = 4$ and $R = 1/2$ there is no point for which $c_1c_2 = Ra\{4(a-3)\}^{-1}$, since $2Ra > 1/2(a-3)$ and c_1c_2 necessarily not greater than $1/16$ if $c_1 = c_4$, $c_2 = c_3$. This suggests that for $R = 1/2$ the point $c_i = 1/4$ is locally stable (as could be investigated) and even globally stable.

The above results suggest that for $R = 1/2$ the internal point (75) is stable, and even globally stable, in both the additive and multiplicative cases, but as stated above the c_i do not in the additive case always converge monotonically to their final values as they do for a single locus. In this case the expression $c_2c_3 - c_1c_4$ can also increase from an initial value of zero.

It is also curious to note that in the multiplicative case it is possible for $c_2c_3 - c_1c_4$ to increase in absolute value, but if it is initially zero, it remains zero. This can be shown by an interesting identity pointed out to me by Professor K. Mahler. Multiplying the equations (82)–(85) together in pairs and subtracting, we get

$$(92) \qquad (c_2'c_3' - c_1'c_4')W^2 = (c_2c_3 - c_1c_4)(\Omega - Ru_2v_2W),$$

where

$$(93) \qquad \begin{aligned}\Omega = {} & u_1u_2(v_1c_1 + v_2c_2)(v_2c_1 + v_3c_2) + u_2u_3(v_1c_3 + v_2c_4)(v_2c_3 + v_3c_4) \\ & + (u_1u_3 + u_2^2)v_1v_2c_1c_3 + (u_1u_3 + u_2^2)v_2v_3c_2c_4 \\ & + (u_1u_3v_1v_3 + u_2^2v_2^2)(c_1c_4 + c_2c_3).\end{aligned}$$

It is easy to verify from this that it is possible to have $\Omega - Ru_2v_2W > W^2$, so that $|c_2'c_3' - c_1'c_4'| > |c_2c_3 - c_1c_4|$. On the other hand, if $c_2c_3 - c_1c_4 = 0$, then $c_2'c_3' - c_1'c_4' = 0$.

10. Final remarks

It will be seen that much remains to be done in this subject. We need to understand much better the general theory of section 6 and the algebraic structure which lies behind it. The behavior of the relatively simple case of two loci each with two alleles also needs a great deal more study.

Applications of these equations to particular problems have been made. Thus, Kimura [8] has studied a system in which selection leads to closer linkage between two loci, and Ewens (personal communication) has used (49)–(52) to study the evolution of dominance.

Applications to human disease are, perhaps, even more interesting. Incompatibility between blood groups can cause the death of a foetus, and furthermore,

there is now evidence that a considerable proportion of fertilized eggs never result in a pregnancy, and that quite large selective differences in this respect exist. There is even evidence that existence of an *ABO* incompatibility can inhibit the effects of an *Rh* incompatibility. This type of phenomenon may well lead to an explanation of some of the puzzling features of the genetics of schizophrenia (Moran [17]). This is a mental disease of young adults for which there is now substantial evidence of a genetic bases, although its exact mode of inheritance is not known. The overt expression of the disease has undoubtedly a negative selective effect, but the underlying genetic polymorphism must be very stable since the frequency of the disease does not vary greatly between different countries (although there are large variations between different social classes). Calculations based on simple models suggest that if the polymorphism is maintained by selective advantages in heterozygotes, these will have to be at least of the order of 5%, and more probably 10%. The puzzle is to understand how so large a selective advantage can come about. If, however, the presence of a factor for schizophrenia can inhibit one of the incompatibilities, such as those for blood groups, which result in such a large loss of fertilized eggs, the polymorphism would be explained.

It will then be necessary to study the stability of systems like (49)–(52) with more than two alleles at each locus. However, further complications are likely to arise in this and similar problems of human genetics, for the definition of the fitnesses w_{ij} conceals strongly simplifying assumptions; it is probable that it will be necessary to define "fitness" as a function of both parental genotypes in a mating. This leads to a much more complicated theory.

Problems of a different kind arise when we attempt to fit a model of the type described by (49)–(52) to observed data. If we can observe in a large population the relative numbers of reproductive adults of the nine genotypes $AABB, \cdots,$ $aabb$, we obtain eight independent quantities. Supposing the polymorphism to be stationary, we have eleven quantities to estimate: the eight independent ratios of the w_{ij}, and three independent quantities out of c_1, c_2, c_3, c_4. These eleven quantities are all identifiable because the assumption of stationarity used in their estimation implies three independent relations between them (that is, three of the equations (49)–(52)). We could then go further and test the resulting system for stability by the criterion of section 8, but the effect of sampling variations on this is not known.

I am indebted to Professors S. Karlin and K. Mahler, and Drs. H. Cann, A. W. Edwards, and W. J. Ewens for some helpful criticisms and suggestions.

REFERENCES

[1] F. V. Atkinson, G. A. Watterson, and P. A. P. Moran, "A matrix inequality," *Quart. J. Math. Oxford Ser.*, Vol. 11 (1960), pp. 137–140.

[2] N. T. J. Bailey, *Introduction to the Mathematical Theory of Genetic Linkage*, Oxford, Clarendon Press, 1961.

[3] W. F. BODMER, "Discrete stochastic processes in population genetics," *J. Roy. Statist. Soc. Ser. B*, Vol. 22 (1960), pp. 218–236.
[4] W. F. BODMER and P. A. PARSONS, "The initial progress of new genes with various genetic systems," *Heredity*, Vol. 15 (1960), pp. 283–299.
[5] ———, "Linkage and recombination in evolution," *Advances in Genetics*, Vol. 11 (1962), pp. 1–100.
[6] L. CESARI, *Asymptotic Behaviour and Stability Problems in Ordinary Differential Equations*, Ergebnisse der Math., Vol. 16, Berlin, Springer, 1959.
[7] M. KIMURA, "Rules for testing stability of a selective polymorphism," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 42 (1946), pp. 336–340.
[8] ———, "A model of a genetic system which leads to closer linkage by natural selection," *Evolution*, Vol. 10 (1956), pp. 278–287.
[9] J. F. C. KINGMAN, "A matrix inequality," *Quart. J. Math.*, Vol. 12 (1961), pp. 78–80.
[10] ———, "A mathematical problem in population genetics," *Proc. Cambridge Philos. Soc.*, Vol. 57 (1961), pp. 574–582.
[11] R. C. LEWONTIN and K. KOJIMA, "The evolutionary dynamics of complex polymorphisms," *Evolution*, Vol. 14 (1960), pp. 458–472.
[12] S. P. H. MANDEL, "The stability of a multiple allelic system," *Heredity*, Vol. 13 (1959), pp. 289–302.
[13] P. A. P. MORAN, "The theory of some genetical effects of population subdivision," *Austral. J. Biol. Sci.*, Vol. 12 (1959), pp. 109–116.
[14] ———, *The Statistical Processes of Evolutionary Theory*, Oxford, Clarendon Press, 1962.
[15] ———, "Balanced polymorphisms with unlinked loci," *Austral. J. Biol. Sci.*, Vol. 16 (1963), pp. 1–5.
[16] ———, "The non-existence of adaptive topographies," *Ann. Human. Genet.*, Vol. 27 (1964), pp. 383–393.
[17] ———, "The schizophrenic polymorphism," *Ann. Human. Genet.*, Vol. 28 (1965), pp. 261–268.
[18] H. P. MULHOLLAND and C. A. B. SMITH, "An inequality arising in genetical theory," *Amer. Math. Monthly*, Vol. 66 (1959), pp. 673–683.
[19] P. A. PARSONS, "Complex polymorphisms where the coupling and repulsion double heterozygotes differ," *Heredity*, Vol. 18 (1963), pp. 369–374.
[20] ———, "Polymorphism and the balanced polygenic combination," *Evolution*, Vol. 17 (1963), pp. 564–574.
[21] L. S. PENROSE, S. M. SMITH, and D. A. SPROTT, "On the stability of allelic systems, with special reference of haemoglobins A, S, and C," *Ann. Human. Genet.*, Vol. 21 (1956), pp. 90–93.
[22] P. A. G. SCHEUER and S. P. H. MANDEL, "An inequality in population genetics," *Heredity*, Vol. 31 (1959), pp. 519–524.
[23] G. A. WATTERSON, "Non-random mating and its effect on the rate of approach to homozygosity," *Ann. Human. Genet.*, Vol. 23 (1959), pp. 204–220.