A GENERAL METHOD FOR INVESTIGATING THE EQUILIBRIUM OF GENE FREQUENCY IN A POPULATION*

R. C. LEWONTIN

Genetics Faculty, North Carolina State College, Raleigh, N. C.

IT is a well known fact that nearly all, if not all, natural populations are in some degree polymorphic. On a phenotypic level, this simply means that there exist a number of different morphological or physiological types in the population. On a genotypic level, this often, although not always, means that there exists in the population more than one allelomorph at a given locus. The fact that alternate states of a gene persist over long periods in a population has been one of great interest to students of evolution.

In a recent paper Li (1955) discussed various cases of equilibrium and demonstrated that the adaptive value or average fitness of a population is maximized at the equilibrium value of the gene frequency, $\hat{q}$, when this equilibrium is stable, and minimized when it is unstable. Unfortunately, in order to make this generalization, it was necessary in most cases to define average fitness in a way which is not intuitively obvious. In addition, for the case of an equilibrium due to a balance of selection and mutation, the adaptive value of the population is not quite maximized. Despite these reservations, however, the principle of maximization of adaptive value does, in some sense, tie together the diverse conditions of equilibrium. It is interesting to note the parallel between Li's approach and the conditions for equilibrium of a physical particle subjected to various forces. For such a particle the condition of stable equilibrium is that of minimum potential energy, while unstable equilibrium is characterized by maximum potential energy.

A further, and extremely important, contribution to the general theory of genetic equilibria has been made by KIMURA (1956) in which that writer has developed general conditions for the stability of equilibria when more than two alleles exist at a locus. The restriction on these general conditions is that the selective forces be constant, irrespective of gene frequency, and that mating be at random.

In this paper I will propose a further extension of the general treatment of gene frequency equilibria in a different direction than did KIMURA. While he held selection and mating system constant and allowed the number of alleles to be n, I shall deal with only two alleles, but allow the factors acting on the gene frequency to have the widest possible scope. The methodology developed will

apply irrespective of the nature of selection, the mating system, the presence of mutation or migration, or the normality of Mendelian segregation.

To this end, a slightly different approach to finding the equilibrium value, $\hat{q}$, will be used, and it will be shown that with this method of attack, the general conditions for stability take an extremely simple form.

To complete the deterministic theory of gene frequency equilibria. it remains to draw together the n-dimensional generalization of KIMURA, with this treatment, a problem which, to this moment, has not been solved.

The method of weights

Let us assume that there are two alleles at a locus, a and A, and that in any particular generation they are in the relative proportions q_0 and $p_0 = (1-q_0)$. In the next generation the three genotypes AA, Aa, and aa will be present in the relative frequencies:

	AA	Aa	aa
(1)	$U_{11} p_0^2$	$U_{12} 2 p_0 q_0$	$U_{13} q_0^2$

This will be true *irrespective of the mating system,* mode of selection, mutation, migration, or any other force acting on the population. The U's are not to be confused with the adaptive value, W, used by LI. The U's are simply an arbitrary set of weights chosen in order to make expression (1) true. Thus, despite the superficial resemblance of (1) to the distribution of genotypes in a random mating population under selection, it is actually a perfectly general statement. These weights will generally be a function not only of the selection coefficients, mutation rates, migration, and inbreeding coefficient of the population, but also of q itself.

Now the gene frequency of allele a in this next generation is:

$$(2) \qquad q_1 = \frac{p_0 q_0 \, U_{12} + q_0^2 \, U_{22}}{p_0^2 \, U_{11} + 2 \, U_{12} \, p_0 q_0 + U_{22} \, q_0^2}$$

and the change in gene frequency from the initial generation is

$$(3) \qquad q_1 - q_0 = \Delta q = \frac{p_0 q_0 \, [(U_{12}-U_{11}) + q_0 \, (U_{11} + U_{22} - 2 \, U_{12})]}{p_0^2 \, U_{11} + 2 \, U_{12} \, p_0 q_0 + U_{22} \, q_0^2}$$

By equilibrium of gene frequency is meant simply the value or values of q for which $\Delta q = 0$. Setting $\Delta q = 0$ in equation (3) there are three possible solutions. Two of these ($\hat{q} = 1, \hat{q} = 0$) are trivial. We are not concerned with loss or fixation of the allele. The third solution is given by

$$(4) \qquad \hat{q} = \frac{U_{11} - U_{12}}{U_{11} + U_{22} - 2 U_{12}}$$

This expression (4) is analogous to LI's equation (4) which is simply a special case when the weights are the adaptive values of the three genotypes.

By substitution of the appropriate weights in (4), the value or values of q for which there is an equilibrium can be immediately found.

It was stated at the outset that only the case of two alleles at a locus was being considered here. In the last section of this paper the most general expression for the stability of equilibria involving n alleles will be given. It still remains to translate this general condition into restrictions upon the weights themselves.

The stability of equilibria

By analogy to the equilibrium of mass points in physics, we may classify equilibria as stable, unstable, or neutral. In a stable equilibrium, the gene frequency will return to the equilibrium point, $\hat{q}$, if it is displaced slightly in the positive or negative direction. The stipulation of a *slight* disturbance is necessary, for if several equilibria exist a large disturbance from one equilibrium point may cause the frequency to change toward another equilibrium. Mathematically we may say that the equilibrium is stable if $\dfrac{d \Delta q}{d q}\bigg|_{\hat{q}}$ is negative.

Conversely, the equilibrium is unstable if the gene frequency continues to move away from the equilibrium point following a slight displacement. This may be symbolized as $\dfrac{d \Delta q}{d q}\bigg|_{\hat{q}}$ is positive.

Finally, the equilibrium is neutral if the gene frequency moves neither toward nor away from equilibrium following a small displacement, but remains at the point to which it has been displaced. This means that $\dfrac{d \Delta q}{d q}\bigg|_{\hat{q}}$ is zero.

It may happen that $\dfrac{d \Delta q}{d q}$ does not exist at the point $\hat{q}$. There are a great many functions for which this may be true. When such is the case, a slightly more general set of criteria may be useful. Let δq be an increment of q taken to be arbitrarily small. Then the conditions are:

(5a)   stable: $\Delta (\hat{q} + \delta q) \;<\; 0 \;<\; \Delta (\hat{q} - \delta q)$
(5b)   unstable: $\Delta (\hat{q} + \delta q) \;>\; 0 \;>\; \Delta (\hat{q} - \delta q)$
(5c)   neutral: $\Delta (\hat{q} + \delta q) \;=\; 0 \;=\; \Delta (\hat{q} - \delta q)$

In certain complex cases it may be most convenient, simply to graph Δq in the region of $\hat{q}$ and see which of these relations holds   .

It has been called to my attention by Dr. A. R. G. Owen of Cambridge, that the criterion of stability, involving only the sign of $\dfrac{d \Delta q}{d q}\bigg|_{\hat{q}}$ as it does, might not suffice since the absolute magnitude of this derivative also seems to affect the nature of the equilibrium. As the following demonstration proves, the absolute

magnitude of $\left.\dfrac{d\,\Delta\,q}{d\,q}\right|_{\hat{q}}$ in no way complicates the problem.

Let us assume that at a given instant the gene frequency is at the value of q^*.
Then the various possibilities for $\left.\dfrac{d\,\Delta\,q}{d\,q}\right|_{q^*}$ may be written as:

(6a) $\qquad 0 < \left.\dfrac{d\,\Delta\,q}{d\,q}\right|_{q^*}$ $\qquad$ (6d) $\quad -2 < \left.\dfrac{d\,\Delta\,q}{d\,q}\right|_{q^*} < -1$

(6b) $\qquad 0 = \left.\dfrac{d\,\Delta\,q}{d\,q}\right|_{q^*}$ $\qquad$ (6e) $\quad -2 = \left.\dfrac{d\,\Delta\,q}{d\,q}\right|_{q^*}$

(6c) $\quad -1 \le \left.\dfrac{d\,\Delta\,q}{d\,q}\right|_{q^*} < 0$ $\qquad$ (6f) $\qquad \left.\dfrac{d\,\Delta\,q}{d\,q}\right|_{q^*} < -2$

Condition 6a represents *no equilibrium* (instability) for after a perturbation of q away from q^* the gene frequency will continue to move away from q^* in the same direction as the initial change.

Condition 6b is a state of *neutrality* since there is no further change of q either away from or toward q^* after the initial perturbation.

Condition 6c states that if q is moved away from q^* it will return toward q^* but not pass it. q^* is then a point of *stable equilibrium*.

Condition 6d represents a state of *damped oscillation*. After a change of q, the gene frequency will return toward q^* but will pass it, coming to rest at a new value of q which deviates less in absolute magnitude from q^* than did the initially perturbed value of q. The result of such a process is an oscillating return to q^* by an infinite series of positive and negative steps. Strictly speaking q^* in such a case is a point of *stable equilibrium*.

Conditions 6e and 6f represent, respectively, *uniform* and *increasing* oscillation around q^*. Then at q^*, $\Delta\,q$ is not zero because q will pass through q^* and move to a new position deviating more than, or as widely from q^* as did the initially perturbed value. But $\hat{q}$, the equilibrium value, has been defined as the value of q for which $\Delta\,q$ is zero. Than any q^* for which $\left.\dfrac{d\,\Delta\,q}{d\,q}\right|_{q^*} \le -2$ is not $\hat{q}$. Conversely, if there is any value of q for which $\Delta\,q = 0$ neither conditions (6e) or (6f) can apply to it. The only possible conditions on the derivative at $\hat{q}$ are, thus, the first four, which accord completely with the sign criterion of stability.

Having established that the condition $\left.\dfrac{d\,\Delta\,q}{d\,q}\right|_{\hat{q}} < 0$ is a necessary and sufficient criterion of stability, we may now convert this inequality into restrictions on the weights themselves.

Restrictions on the weights

Defining

(7a) $\quad a = U_{11} - U_{12}$

(7b) $\quad b = U_{22} - U_{12}$

(7c) $\quad \bar{U} = p^2 U_{11} + 2pq\, U_{12} + q^2\, U_{22}$

then from (3)

(8) $\quad \Delta q = \dfrac{pq\, [-a + q\,(a+b)]}{\bar{U}}$.

Noting that at equilibrium $q\,(a+b) - a = 0$ and that $\bar{U}$ and pq are always positive quantities, the requirement that $\dfrac{d\,\Delta q}{d\,q}\bigg|_{\hat{q}} < 0$ is equivalent [from (8)] to the condition

(9) $\quad \dfrac{d}{d\,q}\,[q\,(a+b) - a] < 0$.

This expression is identical with the requirement that $\dfrac{d^2\,\bar{U}}{d\,q^2}$ be negative when the weights are constant since the expression in brackets is the first derivative of $\bar{U}$ for constant selection. For selection which is functionally dependent upon gene frequency, this bracketed expression is easily shown to be equivalent to

(10) $\quad \dfrac{d\,\overline{w}}{d\,q} - \overline{\left(\dfrac{d\,w}{d\,q}\right)}$

which is Li's expression (17).

On simplification (9) becomes

(11) $\quad 0 > (\hat{a} + b)\left[1 - \dfrac{d}{d\,q}\left(\dfrac{a}{a+b}\right)\right]_{\hat{q}}$

The stability of the equilibrium, then, depends upon the sign of the two quantities $(\hat{a} + b)$ and $1 - \dfrac{d}{d\,q}\left(\dfrac{a}{a+b}\right)\bigg|_{\hat{q}}$. Ignoring for the moment the differentiated function, the condition on the sign of $(\hat{a} + b)$ is exactly that given by KIMURA and LI for constant selective values. Since $\hat{q}$ must be a positive quantity in the interval 0, 1 for a nontrivial solution, $\hat{a}$ and b must be of the same sign. Still ignoring the differentiated quantity, the condition for stability would then be:

(12) $\quad \hat{a} = U_{11}\,(\hat{q}) - U_{12}\,(\hat{q}) < 0$

$\qquad\quad b = U_{22}\,(\hat{q}) - U_{12}\,(\hat{q}) < 0$

That is, at equilibrium the heterozygote must have a higher fitness than either homozygote.

For the general case in which the weights are functionally related to q, however, the differential quantity cannot be ignored and its sign and magnitude become important. The most general conditions for a stable equilibrium are then (i) The heterozygote is greater in fitness than either homozygote at equilibrium

in which case the slope of the line $f(q) = q - \dfrac{a}{a+b}$ must be positive at $\hat{q}$

   or else

(ii) The heterozygote is lower in fitness than either homozygote in which case the slope of the line $f(q) = q - \dfrac{a}{a+b}$ must be negative at $\hat{q}$.

It should again be noted that in deriving (11) it was assumed that $\dfrac{d \Delta q}{d q}$ existed at $\hat{q}$. There is a set of functions for which this is not true so that recourse must be had to the more primitive conditions 5a, 5b and 5c. A case of this sort will be reviewed in the next section.

Two simple examples will show that an unstable equilibrium may exist even when the fitness of the heterozygote is greater than either homozygote at equilibrium, and that a stable equilibrium may result despite the inferiority of the heterozygote. To demonstrate the first case, consider the weights:

$$U_{11} = 1 \qquad\qquad U_{12} = 2 q^2 + 1 \qquad\qquad U_{22} = 4 q^2;$$

from (4)

(13a) $\hat{q} = \dfrac{\hat{a}}{\hat{a}+b} = 2 \hat{q}^2$

(13b) $\hat{q} = \dfrac{1}{2}$

At equilibrium, then, the fitnesses of the genotypes are:

AA	Aa	aa
1.0	1.5	1.0

but

(14) $1 - \dfrac{d}{dq}\left(\dfrac{a}{a+b}\right)\Big|_{\hat{q}} = 1 - 4\hat{q} = -1$

so that

(15) $(\hat{a}+b)\left(1 - \dfrac{d}{dq}\dfrac{a}{a+b}\right)\Big|_{\hat{q}} = (-1)(-1) = 1$

The equilibrium is thus an *unstable* one.

The second case can be illustrated by the weights

AA	Aa	aa
1	$1-2 q^2$	$2-4 q^2$.

Again $\hat{q}$ is ½ so that at equilibrium the fitnesses of the genotypes are:

AA	Aa	aa
1.0	0.5	1.0

and the heterozygote is inferior. The fact that $(\hat{a}+b)$ is equal to unity but $1 - \dfrac{d}{dq}\dfrac{a}{a+b}\Big|_{\hat{q}}$ is again equal to -1 so that $(\hat{a}+b)\left(1 - \dfrac{d}{dq}\dfrac{a}{a+b}\Big|_{\hat{q}}\right) = (1)(-1) = -1$, shows the equilibrium to be *stable*.

Determination of the weights

The determination of the weights is generally a very simple matter and will be illustrated for some cases which demonstrate the generality of the method. The cases chosen are those which have already been attacked by other methods and whose results are known, providing some sort of check on the procedure.

(a) *Inbred population with constant selection:* Under certain conditions a population may reach an equilibrium under inbreeding different from complete homozygosis. For example, a population of plants, a proportion, K, of which are outpollinated and a proportion $(1-K)$ of which are selfed will reach an equilibrium value of F, the coefficient of inbreeding, such that $F = \dfrac{1-K}{1+K}$.

In an inbred population at equilibrium with respect to inbreeding, the zygotes forming the second generation will not be in the ratio $p_0{}^2 : 2\,p_0 q_0 : q_0{}^2$. As WRIGHT shows, the zygotic ratio may be put in the form

(16) $(1-F)p_0{}^2 + Fp_0$: $2\,p_0\,q_0\,(1-F)$: $(1-F)q_0{}^2 + F\,q_0$

where F is the equilibrium inbreeding coefficient. After selection, then, the three genotypes are in the relative frequencies

$$
\begin{array}{ccc}
AA & Aa & aa
\end{array}
$$

(17) $W_{11}\left[(1-F)p_0{}^2 + Fp_0\right]$ $W_{12}\,(1-F)\,2\,p_0\,q_0$ $W_{22}\left[(1-F)\,q_0{}^2 + Fq_0\right]$

Now by factoring out the appropriate quantities from each of these expressions, distribution (17) may be rewritten as:

$$
\begin{array}{ccc}
AA & Aa & aa
\end{array}
$$

(18) $W_{11}\left(1-F+\dfrac{F}{p_0}\right)p_0{}^2$ $W_{12}\,(1-F)\,2\,p_0 q_0$ $W_{22}\left(1-F+\dfrac{F}{q_0}\right)q_0{}^2$

These are in the form given by (1) and

$$U_{11} = W_{11}\left(1-F+\frac{F}{p_0}\right)$$

$$U_{12} = W_{12}\,(1-F)$$

$$U_{22} = W_{22}\left(1-F+\frac{F}{q_0}\right).$$

Substituting these values in the basic equation (4), and solving for $\hat{q}$, yields:

(19) $\hat{q} = \dfrac{(1-F)\,(W_{11}-W_{12}) + F\,(W_{11}-W_{22})}{(1-F)\left[(W_{11}-W_{12}) + (W_{22}-W_{12})\right]}$

which is identical with LI's solution (21), in his 1955 paper.

Expansion of $\left.\dfrac{d}{dq}\,\dfrac{a}{a+b}\right|_{\hat{q}}$ shows that it is always a number smaller than unity, thus assuring that $1 - \left.\dfrac{d}{dq}\,\dfrac{a}{a+b}\right|_{\hat{q}}$ is always positive.

The conditions for stability then become:

(20) $0 > W_{11}\left(1-F+\dfrac{F}{\hat{p}}\right) - W_{12}\,(1-F)$

and

(21) $\qquad 0 > W_{22} \left(1 - F + \dfrac{F}{\hat{q}}\right) - W_{12}\,(1 - F)$

or their equivalents:

(22) $\qquad 0 > (W_{11} - W_{12})\,(1 - F) + \dfrac{F\,W_{11}}{\hat{p}}$

(23) $\qquad 0 > (W_{22} - W_{12})\,(1 - F) + \dfrac{F\,W_{22}}{\hat{q}}$

In order that relations (22) and (23) hold it is necessary that the heterozygote be superior to both homozygotes. This is not sufficient however. Rearrangement of (22) and (23) yields

(24) $\qquad \hat{p} > \dfrac{F}{1 - F} \left(\dfrac{W_{11}}{W_{12} - W_{11}}\right)$

(25) $\qquad \hat{q} > \dfrac{F}{1 - F} \left(\dfrac{W_{22}}{W_{12} - W_{22}}\right)$

which by addition results in

(26) $\qquad \dfrac{1 - F}{F} > \dfrac{W_{11}}{W_{12} - W_{11}} + \dfrac{W_{22}}{W_{12} - W_{22}}$

This last requirement, together with that of overdominance defines stability. (b) *Compensation:* An excellent illustration of variable selection coefficients is provided by the phenomenon of *compensation*. By compensation is meant the production of a relatively greater number of offspring by certain matings. This is not the same as simple selection since the differential fecundity is a function of the mating combination, rather than of the genotypes of the parents alone. Lɪ (1953) dealt with the case of selection against heterozygotes, accompanied by compensation in those matings which would produce heterozygotes. Homozygous recessive females when mated either to heterozygous or homozygous dominant males, produce $(1 + t)$ offspring for each offspring produced by other matings. Here t is the compensation coefficient. The assumption is made that for the selection coefficients $W_{11} = W_{22} = 1$ and $W_{12} = (1 - s)$. Lɪ then shows that the relative frequencies of the three genotypes following selection and compensation are:

$$\begin{array}{ccc} AA & Aa & aa \end{array}$$

(27) $\qquad p_0^{\,2} \qquad\qquad \left(1 - \dfrac{s\,q_0}{2}\right) 2\,p_0\,q_0 \qquad (1 + t\,p_0\,q_0)\,q_0^{\,2}$

On substituting these weights into the basic equation (4) we find that

(28) $\qquad \hat{q} = \dfrac{t + s - \sqrt{t^2 + s^2}}{2\,t}$

which is identical with Lɪ's solution. By inspection of the weights, it is clear that for all $q \neq 0$

(29a) $\qquad U_{11} > U_{12} < U_{22}$

In addition

(29b) $$\left. \frac{d}{dq} \frac{a}{a+b} \right|_{\hat{q}} = \frac{2st}{(s+t+\sqrt{t^2+s^2})^2}$$

which is smaller than unity for all s, t. The equilibrium is then, *unstable*, as concluded by Li (1953).

Lewontin (1953) investigated the equilibrium for the case of compensation and selection for a dominant gene. Using Li's notation the weights are found to be

(30) $U_{11} = (1-s) (1+t)$
$U_{12} = (1-s) (1+t)$
$U_{22} = 1 + t (1-q^2)$

Now, since $U_{11} = U_{12}$ for all q, the only possibility of equilibrium is when
(31) $U_{11} (\hat{q}) = U_{12} (\hat{q}) = U_{13} (\hat{q})$
The only value of q which satisfies (31) is

(32) $$\hat{q} = \sqrt{\frac{s (1+t)}{t}} .$$

Moreover, this is a stable equilibrium since q will increase if $U_{11}(q) = U_{12}(q) < U_{22}(q)$ while it will decrease if $U_{11}(q) = U_{12}(q) > U_{22}(q)$.

These are equivalent to

(33a) Increase of q : $q < \sqrt{\dfrac{s (1+t)}{t}}$

(33b) Decrease of q : $q > \sqrt{\dfrac{s (1+t)}{t}}$

This solution by inspection is simply a special application of condition (5a) above, and is necessary because $\left. \dfrac{d}{dq} \left(\dfrac{a}{a+b} \right) \right|_{\hat{q}}$ does not exist.

(c) *Selection and mutation:* The simplest form of balance between selection and mutation will be considered. It is assumed that the homozygous recessive genotype is selected against. Its adaptive value is $W_{22} = (1-s)$. At the same time A mutates to a at a rate μ per generation. If p_0 and q_0 are the gene frequencies in the initial generation, then the relative frequencies of the gametes will be $p_0 - \mu p_0$ and $q_0 + \mu p_0$ for A and a respectively.

Following random union of gametes and selection, the relative frequencies of the genotypes in the next generation will be:

(34)

AA	Aa	aa
$p_0^2 (1-\mu)^2$	$2 p_0 q_0 (1-\mu) [q_0 (1-\mu) + \mu]$	$[q_0 (1-\mu) + \mu]^2 W_{22}$

On factoring out the appropriate terms in each entry we have:

(35)

AA	Aa	aa
$p_0^2 (1-\mu)^2$	$2 p_0 q_0 \left[(1-\mu)\left(1-\mu+\dfrac{\mu}{q_0}\right) \right]$	$q_0^2 \left[1-\mu+\dfrac{\mu}{q_0} \right]^2 (1-s)$

Substituting the weights given by (35) into equation (4) and solving for $\hat{q}$ yields

$$(36) \qquad \hat{q} = \frac{1}{1-\mu} \left(\sqrt{\frac{\mu}{s}} - \mu \right)$$

This solution differs slightly from that given by Lı for two reasons. The model we have used is slightly more exact since Lı assumes selection and mutation to act simultaneously on the same gene frequency q, while we assume selection to act on the altered gene frequencies arising from mutation. In addition Lı rejects higher powers of μ and assumes $(1 \pm \mu)$ to be not different from unity. For the usual mutation rates postulated of the order of 10^{-5} these assumptions are quite reasonable. Applying them to equation (36) we have

$$(37) \qquad q \cong \sqrt{\frac{\mu}{s}}$$

which is identical with that given by Lı, and is the form usually accepted. Now it is clear that for all values of q,

$$(38) \qquad 0 > U_{11} - U_{12}$$

moreover

$$(39) \qquad 0 > U_{22} - U_{12}$$

if $s > \mu$. In order for there to be any nontrivial equilibrium the selection against the gene must be greater than the mutation rate.

Moreover this equilibrium is stable since, for $\quad s > \mu, \quad \hat{a} + b < 0 \quad$ and

$$\frac{d}{dq} \left. \frac{a}{a+b} \right|_{\hat{q}} = \frac{-(1-\mu)^2 \, \mu \, (s-\mu)}{[\mu \, (1-\mu) + (s-\mu) \, (q-\mu q+\mu)]^2} < 0$$

(d) *Abnormal segregation ratios:* To complete the picture of the generality of equation (4) as a means of finding equilibrium conditions, it is interesting to apply the method of weights to the case of abnormal segregation of genes. Dunn (1953) has described a highly aberrant condition in mice in which males heterozygous for a recessive allele, t, do not produce t and $+$ gametes in a normal 1:1 ratio. In an accompanying article Prout (1953) demonstrated that this aberrant segregation ratio could lead to an equilibrium when the various genotypes were differentially selected. Prout's result applies only when the segregation abnormality is present in both sexes. Let

$\qquad q_0 =$ frequency of t alleles in the gametic pool after segregation
$\qquad 1 - r = W_{11}$
$\qquad 1 - s = W_{22}$

and $m =$ proportion of t alleles among the gametes produced by heterozygotes.

Then, following random mating and selection, the three genotypes will be in the proportions

	$+/+$	$+/t$	t/t
(40)	$p_0^2 \, (1-r)$	$2 \, p_0 \, q_0$	$q_0^2 \, (1-s)$

Following selection, segregation will take place. The gametes will then be in the proportion

$$
\begin{array}{cc}
+ & t \\
(41) \quad p_0^2(1-r) + 2\,p_0\,q_0\,(1-m) & 2\,p_0\,q_0\,m + q_0^2\,(1-s)
\end{array}
$$

Unlike the previous examples, the application of a complete cycle of selection and segregation, results in ratios of the gametes, not the zygotes. The weights of the zygotes are very simply determined, however, by finding what weights could have been applied to the zygotes in order to yield the gametic ratios (41) if segregation had been normal. This is done by adding and subtracting the quantity $4\,p_0 q_0\,m$ in the frequency of $+$ gametes. This yields

$$
\begin{array}{cc}
+ & t \\
(42) \quad p_0^2\,(1-r) + 2\,p_0\,q_0 + 2\,p_0\,q_0\,m - 4\,p_0\,q_0\,m & 2\,p_0\,q_0\,m + q_0^2\,(1-s)
\end{array}
$$

But this would be the gametic ratio resulting from normal segregation had the zygotes been in the frequency

$$
\begin{array}{ccc}
+/+ & +/t & t/t \\
(43) \quad p_0^2\,(1-r) + 2\,p_0\,q_0\,(1-2\,m) & 4\,p_0\,q_0\,m & q_0^2\,(1-s)
\end{array}
$$

Factoring out the appropriate terms in each entry the weights are seen to be

$$
U_{11} = 1 - r + 2\,\frac{q_0}{p_0}\,(1-2\,m)
$$
$$
U_{12} = 2\,m
$$
$$
U_{22} = 1 - s
$$

Substituting these weights in the basic formula (4) yields

$$
(44) \quad \hat{q} = \frac{r + 2\,m - 1}{s + r}
$$

as a solution. This is identical with PROUT's result.

Thus the method of weights is applicable not only irrespective of the mating system and the form of selection, but also of any assumption about the normality of the underlying Mendelian segregation.

The various combinations of s, r and m which satisfy the stability requirement are too numerous and complex to present. It is sufficient to observe that the numerical cases discussed by PROUT in his original paper all satisfy conditions (11) and so are stable.

Multiple alleles

Suppose that at any locus there are n alleles a_i each present in the the frequencies q_i. Then in a manner entirely analogous to the case of two alleles the genotype frequencies can be written as:

$$
\begin{array}{llll}
a_i\,a_i & : & U_{ii}\,q_i^2 & i = 1,2,3,\cdots n \\
a_i\,a_j & : & 2\,U_{ij}\,q_i\,q_j & i < j = 2,3,4\cdots n
\end{array}
$$

LI (1955) and WRIGHT (1949) show quite simply that

$$
(45) \quad \Delta\,q_i = \frac{q_i}{\bar{n}}\,(U_i. - \bar{U})
$$

where the U are constants but this applies equally well to variable weights. In expression 45

$$(46) \qquad U_{i\cdot} = \sum_{j=1}^{n} q_j \, U_{ij}$$

and

$$(47) \qquad \bar{U} = \sum_{i=1}^{n} q_i \, U_{i\cdot}$$

That is, $U_{i\cdot}$ is the average weight for allele a_i in all its combinations and $\bar{U}$ is the grand average weight of the population.

At equilibrium $\Delta \, q_i = 0$ for all i. There are $n-1$ independent equations of the form $U_{i\cdot} - \bar{U} = 0$ and the solution to these simultaneous equations can easily be shown to be the determinantal equations:

$$(48) \qquad \hat{q}_i = \frac{D_i}{D}$$

where
(49)

$$D_i = \begin{bmatrix} (U_{1,\,1}-U_{n,\,1}) & (U_{1,\,2}-U_{n,\,2}) & \cdot & \cdot & \cdot & (U_{1,\,n\text{-}1}-U_{n,\,n\text{-}1}) \\ (U_{2,\,1}-U_{n,\,1}) & (U_{2,\,2}-U_{n,\,2}) & \cdot & \cdot & \cdot & (U_{2,\,n}-U_{n,\,n\text{-}1}) \\ \cdot & \cdot & & & & \vdots \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ (U_{n\text{-}1,\,1}-U_{n,\,1}) & (U_{n\text{-}1,\,2}-U_{n,\,2}) & \cdot & \cdot & \cdot & (U_{n\text{-}1,\,n\text{-}1}-U_{n,\,n\text{-}1}) \end{bmatrix}$$

except that the ith column is replaced by the elements:

$$\begin{bmatrix} (U_{n,\,n}-U_{1,\,n}) \\ (U_{n,\,n}-U_{2,\,n}) \\ \cdot \\ \cdot \\ \cdot \\ (U_{n,\,n}-U_{n\text{-}1,\,n}) \end{bmatrix}$$

and

$$(50) \qquad D = \sum_{i=1}^{n} D_i$$

It should be remembered that $U_{ij} = U_{ji}$. This is the most general form of the determinatal equations which have been given by LI (1955) for the case of $n=3$ and constant selective values.

The stability of equilibria involving more than two alleles where the U's are constant has been discussed, as I have already pointed out, by KIMURA. When the U's are functions of gene frequency, however, KIMURA's rules do not apply although a general expression can be derived using the same logic as KIMURA.

Let $\hat{Q}$ represent a point of equilibrium in n-1 dimensional space (corresponding to the n-1 independent gene frequencies) given by equations (48), and let $\delta \hat{Q}$ be a

vector representing a small displacement of gene frequencies away from $\hat{Q}$ to a new point Q^*. If $\hat{Q}$ is a point of stable equilibrium there will be a new change in gene frequencies from Q^* back toward $\hat{Q}$ which change can be represented by a vector ΔQ^*. What is meant by *back toward* $\hat{Q}$ is best illustrated in Figure 1. An *n*-1 dimensional sphere is drawn with $\hat{Q}$ at its center and $\delta\hat{Q}$ as its radius. Then if $\hat{Q}$ is a point of stable equilibrium ΔQ^* will carry the gene frequency to a point Q^{**} which lies *within the sphere*. In order for this to be true, two require-
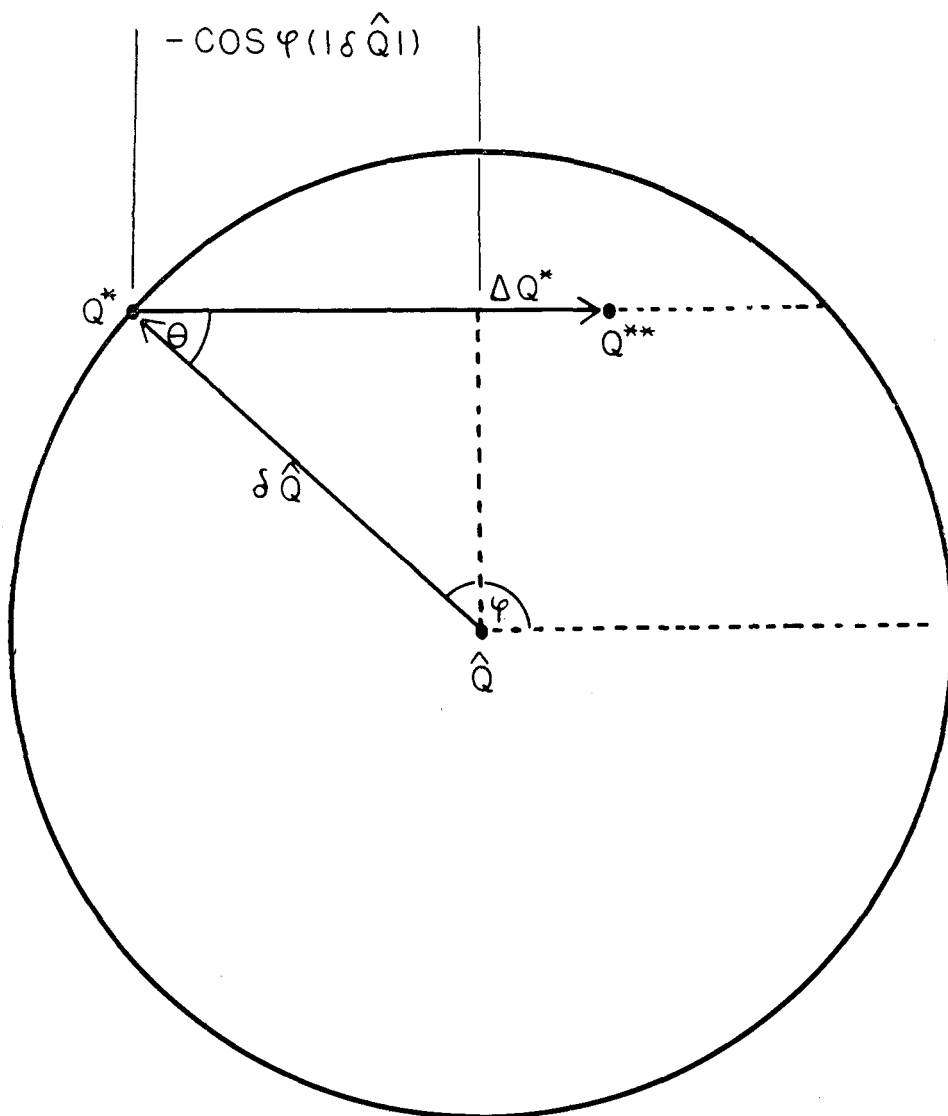

FIGURE 1.—Geometrical representation of the conditions for a stable equilibrium with more than two alleles at a locus. See text for explanation of symbols.

ments must be satisfied. First, the angle θ must be strictly less than 90° which is equivalent to requiring that the angle ϕ between the two vectors be strictly greater than 90°. Second, the vector ΔQ^* must not be too long or it will carry Q^{**} outside of or exactly on to the sphere, no matter what the angle θ.

The cosine of an angle between two vectors is defined as the ratio of the inner product of those vectors to the product of their lengths. The requirement that ϕ be greater than 90° can then be written as:

$$(51) \quad \cos \phi = \frac{|\Delta Q^* \cdot \delta \bar{Q}|}{|\Delta Q^*| \cdot |\delta Q|} < 0$$

The requirement that the length of the vector ΔQ^* be short enough to insure that Q^{**} lies strictly within the sphere is equivalent to

$$(52) \quad \frac{|\Delta Q^*|}{|\delta \bar{Q}|} < -2 \cos \phi$$

by the simple geometry of Figure 1. Combining (51) and (52) gives the most general requirement for stability

$$(53) \quad 0 > \frac{|\Delta Q^*|^2}{|\Delta Q^* \cdot \delta \bar{Q}|} > -2$$

Condition (51) is given by Kimura and condition (52) appears to be his condition that no "cyclically winding paths" are involved. A cyclically winding path in n-1 dimensions is the equivalent of an oscillating equilibrium in one dimension and while we showed that oscillation around $\hat{q}$ is not possible by definition in the one dimensional case, this is not true for higher dimensions, so that the second requirement must be added. If expression 53 is reduced to the one dimensional case it becomes simply

$$(54) \quad 0 > \frac{d \Delta q}{dq}\bigg|_{\hat{q}} > -2$$

which was shown earlier to be precisely the condition for stability (6c and 6d).

Whereas condition 53 can be written in terms of the weights in any specific case, I have been unable to find a general restriction on the weights analogous to (11) so that for the present, at least, it is necessary to treat each case specially.

DISCUSSION

As was pointed out at the beginning, the methodology described here is designed as a constructive tool for investigating the equilibrium condition of a population under a wide variety of situations. The examples given merely serve to illustrate the power and generality of the method. It is becoming apparent that the adaptive values of genotypes in populations are functions, and often complex functions, of the gene frequency. This is not surprising since the environment of an organism includes the relative frequencies of other sorts of organisms in the population (See Levene, Pavlovsky and Dobzhansky 1954; Lewontin 1955). As various new situations arise, the method may be used in investigation of the equilibrium conditions.

Perhaps the most suggestive result of the mathematical analysis is the demonstration that stable equilibria may exist despite an inferiority of the heterozygote, provided that the adaptive values of the genotypes change properly with gene frequency. Whether evolution does in fact produce such balanced, nonheterotic, polymorphisms is problematical, but their existence is at least possible.

Whereas a general theory of equilibria encompassing adaptive values functionally related to gene frequency on the one hand and multiple allelism on the other is a step in the direction of increased reality of our models of population dynamics, there are a number of areas for theoretical study which are yet untouched or virtually so, but which are essential to a proper understanding of natural events. Among the most important are a study of general equilibrium conditions for many loci interrelated by linkage and epistasis, and the problem of environments which vary both in time and in space. This latter problem demands a revision of our static concept of equilibrium to include the idea of a dynamic and ever fluctuating genetic structure of a population. It is impossible to say whether such problems will ever be amenable to general mathematical analysis, but if population geneticists are to lay any claim to an understanding of the forces molding the genetic structure of natural populations, attempts upon these problems must be made.

SUMMARY

A basic method has been suggested for investigating the equilibrium condition of gene frequency in populations. The method is applicable to any population irrespective of the nature of the forces which change gene frequency. It applies equally well to random and nonrandom mating schemes provided an equilibrium condition of inbreeding can be specified, and irrespective of any assumptions about the regularity of Mendelian segregation at meiosis. It cannot be applied to more than one locus if epistasis is present nor is it useful if the various forces change over time in a way unrelated to gene frequency.

Within these limits it has been shown that superiority in fitness of the heterozygote is neither a necessary nor sufficient condition for equilibrium but that nonheterotic balanced polymorphic systems may exist. Conversely heterotic systems do not lead ineluctably to stable equilibria.

LITERATURE CITED

DUNN, L. C., 1953 Variations in the segregation ratio as causes of variations in gene frequency. Acta Genet. et Statist. Med. **4**: 139–147.

KIMURA, M., 1956 Rules for testing stability of a selective polymorphism. Proc. Natl. Acad. Sci. U.S. **42**: 336–340.

LEVENE, H., O. PAVLOVSKY, and TH. DOBZHANSKY, 1954 Interaction of the adaptive values in polymorphic experimental populations of *D. pseudoobscura*. Evolution **9**: 335–349.

LEWONTIN, R. C., 1953 The effect of compensation on populations subject to natural selection. Am. Naturalist **87**: 375–381.
1955 The effects of population density and composition on viability in *D. melanogaster*. Evolution **9**: 27–41.

Li, C. C., 1953   Is Rh facing a crossroad? A critique of the compensation effect. Am. Naturalist
     87: 257–261.
     1955   The stability of an equilibrium and the average fitness of a population. Am. Naturalist
     89: 281–296.

Prout, T., 1953   Some effects of the variations in segregation ratio and of selection on the
     frequency of alleles under random mating. Acta Genet. et Statist. Med. **4**: 148–151.

Wright, S., 1942   Statistical genetics and evolution. Bull. Am. Math. Soc. **48**: 223–246.
     1949   Adaptation and selection. pp. 365–389. *Genetics, Paleontology and Evolution.* Edited
     by G. L. Jepson. Princeton Univ. Press. N.J.