

## Species Packing and Competitive Equilibrium for Many Species\*

ROBERT MACARTHUR

*Department of Biology, Princeton University, Princeton, N. J. 08540*

Received September 12, 1969

## I. INTRODUCTION

It has always been interesting to some scientists to construct minimum principles for their science. These not only provide conceptual unification but also lead to some technical simplifications. Thus, for instance, in mechanics, a particle follows the path which will minimize the time integral of the difference between kinetic and potential energies (principle of least action) and the stable equilibrium positions of such a system are those for which the potential energy is at a minimum. The first example is a minimum principle for dynamics, the second only applies to the static, equilibrium, conditions.

Here I attempt an ecological counterpart of the second, weaker, kind of minimum principle. I show that for some kinds of competition equations a quadratic expression is minimized, and I use this result to interpret species packing and competitive equilibria.

The *competition equations* for  $n$  species are of the form

$$\frac{1}{X_i} \frac{dX_i}{dt} = C_i \left[ X_i - \sum_{j=1}^n \alpha_{ij} X_j \right] \quad i = 1, \dots, n, \quad (1)$$

where all  $X$  and  $C$  are positive and all  $\alpha$  are, at least, nonnegative. More realistically, the equations should include higher order terms and perhaps higher derivatives, the effects of seasonality and patchy environments, age distributions, time lags, and the like; but, ecologists have found it interesting to see what consequences even these simplified equations may have. Here I continue this approach. Lotka (1932) regarded the expressions in square brackets in these equations as the first terms in a Taylor series and, hence, probably realistic near equilibrium. Although the dynamics of competition usually involve nonequilibrium conditions and, hence, doubtless require more elaborate equations; Eqs. 1 seem very useful for equilibria.

\* The work was supported by grants from the National Science Foundation.

## II. GENERAL DISCUSSION

We begin by forming the quadratic expression  $R = \sum_{k,j} \alpha_{kj} X_k X_j$ , where the  $\alpha$  are those of Eqs. 1. Taking the derivative with respect to  $X_i$ , we get

$$\frac{\partial R}{\partial X_i} = \sum_j \alpha_{ij} X_j + \sum_k \alpha_{ki} X_k = \sum_j (\alpha_{ij} + \alpha_{ji}) X_j. \quad (2)$$

Now, if  $\alpha_{ij} = \alpha_{ji}$  for all  $i, j$ , then this becomes  $2 \sum_j \alpha_{ij} X_j$ , which is a prominent term on the right side of Eqs. 1. This prompts the formation of new quadratic form

$$Q = -2 \sum_k \mathcal{K}_k X_k + R = -2 \sum_k \mathcal{K}_k X_k + \sum_{k,j} \alpha_{kj} X_k X_j$$

whose derivative, with respect to  $X_i$ , is  $\partial Q / \partial X_i = -2[\mathcal{K}_i - \sum_j \alpha_{ij} X_j]$ , whence Eqs. 1 take the form

$$\frac{dX_i}{dt} = -\frac{C_i X_i}{2} \frac{\partial Q}{\partial X_i}$$

and

$$\frac{dQ}{dt} = \sum_i \frac{\partial Q}{\partial X_i} \frac{dX_i}{dt} = -\frac{1}{2} \sum_i C_i X_i \left( \frac{\partial Q}{\partial X_i} \right)^2 < 0. \quad (3)$$

In other words,  $Q$  is minimized by competition and, at this point, there is a stable equilibrium. All this assumes only that  $\alpha_{ij} = \alpha_{ji}$ , i.e., that the  $\alpha$  matrix is symmetric. Under this condition, Eqs. 1 imply that  $Q$  is minimized. Conversely, suppose we know some  $Q$  is minimized at the equilibrium of Eqs. 1; then, since every quadratic form can be written as a symmetrical one, we are led to a symmetric matrix in  $Q$ . Hence, given Eqs. 1 at equilibrium, we find a quadratic form minimized if, and only if, the  $\alpha$  matrix is symmetrical. Of course, some other function may be minimized even if the  $\alpha$  matrix is not symmetrical, but it won't be a quadratic form.

The case of a skew-symmetric matrix,  $\alpha_{ij} = -\alpha_{ji}$ , is of special interest since this situation occurs in classical predator-prey equations. Here,

$$\frac{\partial R}{\partial X_i} = \sum_j (\alpha_{ij} + \alpha_{ji}) X_j = 0$$

so that  $R$  is constant. In other words, trajectories lie on surfaces  $R = \text{const.}$

One more general consideration is of interest here. If the form  $Q$  has a lower bound, say zero, then we can talk about communities of competitors which cannot be invaded. For if  $Q$  is already zero, then it cannot be reduced further. More exactly, quadratic forms like  $Q$  can always be diagonalized to a sum of

squares. When all the squares are positive in this diagonal form, then  $Q$  may be regarded as a distance squared which has a minimum of zero. We see this explicitly in what follows.

## III. COMPETITION FOR NONINTERACTING RESOURCES

Rather than begin with Eqs. 1, it is better to derive Eqs. 1 and, at the same time, get more informative formulas for the  $\mathcal{K}$  and  $\alpha$  terms. This section follows MacArthur (1969) closely. I begin with consumer-resource equations

$$\frac{1}{X_i} \frac{dX_i}{dt} = C_i \left[ \sum_k a_{ik} w_k R_k - T_i \right], \quad (4a)$$

where  $a_{ik}$  is the probability that an individual of species  $i$  (of abundance  $X_i$ ) encounters and eats an item of resource  $k$  (of abundance  $R_k$ ) in a unit of time;  $w_k$  is the weight of this item of resource in grams, and  $T_i$  is the threshold number of grams of resource captured to maintain the population, and  $C_i$  is a constant of proportionality governing the biochemical conversion of grams of resource  $R$  into grams of  $X_i$ . The resources have equations for their own renewal:

$$\frac{1}{R_k} \frac{dR_k}{dt} = \frac{r_k}{K_k} \{K_k - R_k\} - \sum_j a_{jk} X_j = \frac{r_k}{K_k} \left[ K_k - \sum_j a_{jk} \frac{K_k}{r_k} X_j - R_k \right]. \quad (4b)$$

The term in braces is a logistic self-inhibition of resource  $k$  by itself; since we are temporarily assuming that the resources do not interact, other  $R$ 's do not affect  $R_k$ , and there is no other  $R$  term. The  $a$  term appeared in one of Eqs. 4a because, when an  $X$  eats an  $R$  (Eq. 4a), an  $R$  disappears (Eq. 4b).

To understand the equilibrium abundances of the competing species, we need only solve the simultaneous equations  $dX_i/dt = 0$ ,  $dR_k/dt = 0$  for all  $i$  and  $k$ , which means we equate to zero the square brackets in (4a) and (4b). Actually, it is more informative to solve (4b) for  $R_k$  at equilibrium,

$$R_k = K_k - \sum_j a_{jk} \frac{K_k}{r_k} X_j$$

and to substitute this for  $R_k$  in the full Eq. 4a:

$$\frac{1}{X_i} \frac{dX_i}{dt} = C_i \left[ \left\{ \sum_k a_{ik} w_k K_k - T_i \right\} - \sum_j \left\{ \sum_k a_{ik} a_{jk} \frac{w_k K_k}{r_k} \right\} X_j \right]. \quad (5)$$

Here, I have gathered terms so that the expressions in braces correspond to the  $\mathcal{K}$  and  $\alpha$  terms in Eqs. 1. In other words, in Eqs. 1:

$$\mathcal{K}_i = \sum_k a_{ik} w_k K_k - T_i \quad \text{and} \quad \alpha_{ij} = \sum_k a_{ik} a_{jk} \frac{w_k K_k}{r_k}. \quad (6)$$

Notice  $\alpha_{ij} = \alpha_{ji}$ , so that the theorems of Section II apply. A word is in order about the validity of this procedure. I have put the equilibrium values of  $R$  into the nonequilibrium equations for  $X$  (although actually I shall mainly use the resulting equations at equilibrium). This means that I assume the population changes are taking place slowly enough that the  $R$  can track their equilibrium values accurately. Even without this assumption, we can proceed to the remarkable results of Section II.  $Q$ , apart from an additive constant, takes the explicit form

$$Q = \sum_k \frac{w_k K_k}{r_k} \left[ r_k - \sum_j a_{jk} X_j \right]^2 + 2 \sum_j T_j X_j. \quad (7)$$

This  $Q$  is not easy to interpret biologically; however, with two additional assumptions, we can get a more interpretable  $Q$  which is similarly minimized: Suppose all species have the same threshold resource requirements,  $T_i = T$ , and that all have the same total harvesting ability,  $\sum_j a_{ij} = a$ , independent of  $i$ . Then, we can form the new expression

$$Q = \sum_k \frac{w_k K_k}{r_k} \left[ \left( r_k - \frac{Tr_k}{aw_k K_k} \right) - \sum_j a_{jk} X_j \right]^2 \quad (8)$$

which is a weighted (by  $wK/r$  terms) squared deviation of *available production* ( $r_k - Tr_k/aw_k K_k$ ) from the species' harvesting abilities ( $\sum_j a_{jk} X_j$ ). As before, if we take the derivatives with respect to  $X_i$ , we get back the equilibrium conditions, and Eqs. 3 hold so that this  $Q$  is also minimized by competition.

#### IV. APPLICATION TO SPECIES PACKING

The fact that competition for noninteracting resources (and, as we shall see, many other kinds of competition also) leads to minimization of the expression  $Q$  of Eq. 8, makes possible a neat interpretation of how many species an environment will hold. Thus, suppose a continuum of resources,  $j$ , is arranged linearly (by size or height above the ground, etc.) and the available production,  $r_j - Tr_j/aw_j K_j$ , is plotted along this coordinate, as in any of the upper curves of Figs. 1 or 2. We can also plot the species' utilization curves,  $a_{ij}$ , along the same coordinate as in the lower curves, where three species,  $a_{1j}$ ,  $a_{2j}$ ,  $a_{3j}$ , are plotted. The result of the last section then says that those species will be present, and in such numbers that a positive linear combination of the lower curves (the utilizations) gives the best weighted least squares fit to the upper (production) curves. In the figures, the flatter production curves are exactly  $a_{1j} + a_{3j}$  so that the fit is perfect and any of species 2 would make the fit worse. Similarly, the more peaked production curves are exactly  $2a_{2j}$  so that any addition of  $a_{1j}$  or  $a_{3j}$  would make

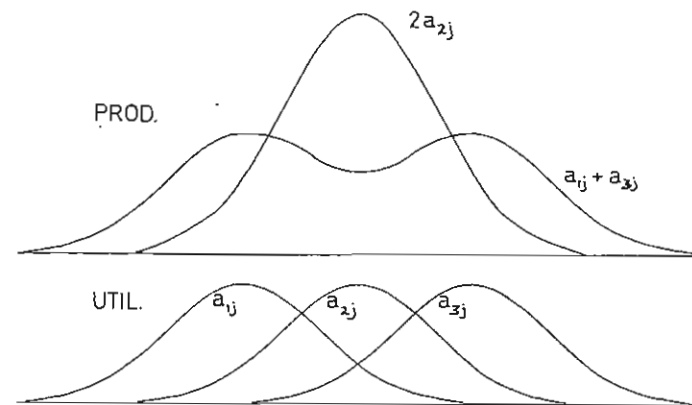


FIG. 1. The available production curves (PROD) illustrate the limits of environmental variation within which the three illustrated species,  $a_{1j}$ ,  $a_{2j}$ , and  $a_{3j}$  will coexist. See text for details.

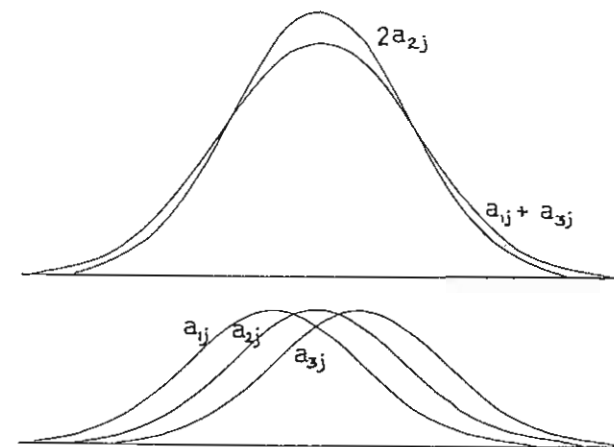


FIG. 2. The same as Fig. 1, except that the species are now more similar and a much smaller environmental variation is now tolerated. See text for details.

the fit worse. In other words, if in a given environment the production is as concentrated as the peaked curves, then only species 2 will be present, while if the production is as low in the middle as the flatter curves, only species 1 and 3 will be present. If the environmental production is a convex linear combination of the two curves pictured (i.e., is a fraction  $p$  times one curve plus  $(1 - p)$  times the other and hence is in between them), then all three species will coexist. Hence, the two production curves in each figure represent the limits of environ-

mental fluctuation in production within which all three species will coexist. Although in a constant environment there is almost no limit to the number of species which can improve the fit and hence be packed in (unless the previous species are already an excellent fit or the number of resources is less than the number of species), in a fluctuating environment the closer the species are packed the more susceptible they are to environmental fluctuations. Thus, in Fig. 1, the environment can fluctuate greatly before one of the species is eliminated, while in Fig. 2, with its closely packed species, only a small environmental fluctuation is tolerated.

If the resources come in two or more dimensions, say food size and food height in the trees, the same conclusions hold but one new additional result is plausible. Suppose one of the coordinates has a fairly constant resource distribution, but the other is more seasonal. For example, the distribution of food over height might be constant while the distribution of food size might vary with the age of the food and, hence, with season. Then competing species in subdividing this environment can subdivide the constant coordinate rather finely, but not the varying one. In the example, the species would be feeding height specialists but not food size specialists.

The proof of these results is rather inadequate, being from a graphical example, and it would be of great interest to know exactly what measure of environmental fluctuation is easy to relate to the distance between  $a_{ij}$  curves.

The range of validity of this case is not as restricted as it might seem, for spatially distinct resources have little or no interaction.

This interpretation of species packing also clarifies the effect of environmental productivity on species diversity, at least in a nonseasonal environment. For, picture a resource production curve and the total utilization curve of several consumer species. The utilizations of these consumers are adjusted to be a *best fit* to the available production and will hence be in some places slightly in excess of the available production and in other places, slightly below the production (unless by accident they should exactly match it). Where the utilizations fall slightly below the production, a new species can enter if the addition of one pair (or a *propagule*) will make the total utilization even closer to the available production. A small multiple of some appropriate species will always improve the fit somewhere so that some new species, in very small quantity, could always invade. However, if only, say, .3 of a pair is appropriate to improve the fit, only that many of the new species will be present—it will actually be absent. If, however, the total productivity were tenfold greater, the whole picture would be roughly blown up tenfold and now 3 pairs instead of .3 pairs of the new species will be appropriate. This time the species will be present. For this reason more productive environments will hold more species.

## V. EXTENSION TO OTHER FORMS OF COMPETITION

### (a) *The Action of the Self-Limitation Term*

So far I have assumed that the species have no self-limitation term but that their resources do. If neither had self-limitation but each was limited solely by the other, the situation would have been different; the right sides of (4a) and (4b) would have been independent. There can only be  $X$  solutions of the equilibrium equations if there are at least as many unknowns as equations, i.e., at least as many  $X$  as  $R$ ; similarly, the  $R$  can only come to equilibrium if there are at least as many  $R$  as  $X$ . So the number of equilibrium  $X$ 's must be equal to the number of resources. From convexity arguments we can also determine whether these equilibrium solutions are all positive. Other methods (see MacArthur, 1968) apply to this situation, but no quadratic  $Q$  is minimized by Section II since  $\alpha_{ij} = -\alpha_{ji}$ . Thus, some self-limitation term seems necessary for the existence of a  $Q$  and, in fact, for the derivation of Eqs. 1 from 2.

Suppose next that both consumers and resources have a self-limitation. Then Eqs. 4a have a  $-b_i X_i$  inside the square bracket and when the value of  $R_k$  from (4b) is substituted, we get Eq. 5 with an additional term  $-b_i X_i$  in the square bracket. Expressions (7) and (8) are then simply modified by subtracting  $b_i$  from  $a_{ii}$  so  $\alpha$  is symmetric and the rest proceeds as in Section II; the  $Q$  of Eq. 8 also is directly constructed.

### (b) *Interacting Self-Regulation*

If, in addition to an  $R_k$  term, Eq. 4b has terms in the other  $R$ 's, more complications result; but, a perfectly good  $Q$  is sometimes possible: The square bracket in (4b) becomes

$$K_k - \sum_j a_{jk} \frac{K_k}{r_k} X_j - \sum_l b_{kl} R_l.$$

If the  $R$  are weakly interacting, then only  $b_{ii}$  are significant and we can solve directly; otherwise, to solve for  $R_l$  we must invert the  $b_{kl}$  matrix. Now, the  $b$ , like the  $\alpha$  of Eq. 1, are competition coefficients and so can usually be decomposed as a product of a matrix with its transpose; hence, the inverse matrix of  $b_{kl}$  can often be written in the form  $\sum_n C_{mn} C_{kn}$  so that  $\sum_k \sum_n C_{mn} C_{kn} b_{kl} = \delta_{ml}$  (i.e.,  $= 0$  if  $m \neq l$  and  $= 1$  if  $m = l$ ). Then, on multiplying by  $\sum_n C_{mn} C_{kn}$  and summing over  $k$

$$\sum_{k,n} C_{mn} C_{kn} K_k - \sum_{j,k,n} C_{mn} C_{kn} a_{jk} \frac{K_k}{r_k} X_j = \sum_l \delta_{ml} R_l = R_m.$$

Substituting this into (4a), which says that  $\sum_m a_{im} w_m R_m = T_i$  at equilibrium,

$$\begin{aligned} 0 &= \left\{ \sum_{m,k,n} a_{im} C_{mn} C_{kn} w_k K_k - T_i \right\} - \sum_j \left\{ \sum_{m,k,n} a_{im} C_{mn} C_{kn} a_{jk} w_k \frac{K_k}{r_k} \right\} X_j \\ &= \left\{ \sum_{k,n} \left( \sum_m a_{im} C_{mn} \right) C_{kn} w_k K_k - T_i \right\} \\ &\quad - \sum_j \left\{ \sum_n \left( \sum_m a_{im} C_{mn} \right) \left( \sum_k a_{jk} C_{kn} \frac{w_k K_k}{r_k} \right) \right\} X_j. \end{aligned}$$

If we let  $g_{in} = \sum_m a_{im} C_{mn}$  and assume that  $w_k K_k / r_k = wK/r$ , a constant, this takes a form with symmetric  $\alpha$  matrix

$$\left\{ \sum_{k,n} g_{in} C_{kn} r - \frac{T_i r}{w_k} \right\} - \sum_j \left\{ \sum_n g_{in} g_{jn} \right\} X_j.$$

Now, a  $Q$ , even diagonalized like that of Eq. 8, can be written, assuming, as before, that  $\sum_j g_{jn} = g$

$$Q = \frac{w_k}{r} \sum_n \left[ \sum_j C_{jn} r - \frac{T_j r}{w_k g} - \sum_j g_{jn} X_j \right]^2.$$

When we can find a  $Q$ , we again match a production by a combination of consumptions, but this time both are smoothed by the operation of a  $C_{ij}$  matrix.

#### (c) Competition by Interference

Here an ambiguity makes precise results difficult, for we do not have a clear notion of how interference acts. If one species intimidates another, they presumably have to be within communicating distance (although bears or wolves, for instance, may leave signs which intimidate, even when their maker is far away), but how the intimidation affects the  $\mathcal{K}$  of Eqs. 1 is not at all clear. However, the  $\alpha$  terms are, at least, often of the symmetric form

$$\alpha_{ij} = \sum_k p_{ik} p_{jk},$$

where  $k$  runs through all localities and  $p_{ik}$  is the probability that a given individual of species,  $i$ , is in locality,  $k$ , at a given time.  $\alpha_{ij}$ , then, represents the probability that two individuals, one of each species, are in the same place simultaneously;  $\alpha_{ii}$  is the corresponding probability that two individuals of the same species,  $i$ , are together at the same time. For this kind of interference, a perfectly good  $Q$  exists, as in Section II.

Many other kinds of competition by interference do not lead to a  $Q$  which is

minimized. If, for instance, species  $i$  looks unlike species  $j$ , and the psychological effect of the intimidation is greater from different-looking individuals, then we could have  $\alpha_{ij} \neq \alpha_{ji}$ . Hutchinson (1947) pointed this out in considering tern species which were conjectured to gang up on one another. For such situations no  $Q$  is possible since the  $\alpha$  matrix is not symmetrical.

#### (d) Competition to Withstand Predators

There is a kind of duality between predator and prey in the Eqs. 4a, 4b. We solved the prey Eqs. 4b for  $R$  and substituted into the predator Eqs. 4a. If we had let the self-limitation term be in the predator equations (or both), we could have solved the predator equations for  $X$  in terms of  $R$  at equilibrium, and substituted this into the prey equations. The prey equations would then be competition equations in which the best competitor is the one which can withstand the greatest predator pressure. A completely analogous  $Q$  can be derived since  $\alpha$  is symmetrical and we can proceed as before.

Species which are simultaneously regulated by both predators and food supply will also have a  $Q$ , at least if they do for both predators and resources separately, for their  $\alpha$  matrix will be the sum of the two separate  $\alpha$  matrices and the sum of two symmetric matrices is still symmetric.

### V. REMARKS ON STABILITY

Although the bulk of this paper is devoted to the nature of the many-species equilibria, one aspect of stability follows so naturally that it would seem a pity not to include it. Basically, the rate of return to equilibrium depends upon the eigenvalues of the  $\alpha$  matrix, and the eigenvalues of symmetric matrices have special properties. Thus, if  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  are the eigenvalues (always real) of  $\alpha$ , and if we now add a new competitor to the system so that  $\alpha$  is augmented by the addition of a new row and column, then the eigenvalues

$$\mu_1 \leq \mu_2 \leq \dots \leq \mu_{n+1}$$

of the augmented matrix must be of a magnitude such that

$$\mu_1 \leq \lambda_1 \leq \mu_2 \leq \dots \leq \lambda_n \leq \mu_{n+1}.$$

(For a proof of this and other well-known properties of the eigenvalues of symmetric matrices, see Chap. 6 of Franklin, 1968). Although the rate of approach to equilibrium will vary according to the direction of the displacement, it will be slowest from the direction of the eigenvector corresponding to the smallest eigenvalue,  $\lambda_1$ , which is thus one kind of measure of stability: The larger the  $\lambda_1$  the more stable the system. Now the smallest eigenvalue of the augmented

system was  $\mu_1 \leq \lambda_1$ , so the augmented community cannot be more stable than the original. Thus, when an additional species is added to a community of competitors with symmetric  $\alpha$  matrix, the stability as measured by the smallest eigenvalue cannot increase and usually decreases.

## VI. DISCUSSION OF EVOLUTIONARY IMPLICATIONS

In many cases of interest, we have seen that competition acts to minimize a quadratic form,  $Q$ . Exceptions are some cases of interference, possibly complicated interaction among the resources, and the oversimplified predator-prey equations in which neither has self limitation. When  $Q$  exists, the stable equilibria are those which minimize  $Q$ , and the chief limitations on the number of coexisting species are environmental fluctuation and productivity. Using  $X_1, X_2, X_3 \dots$  as coordinates, we can plot  $Q$  as a landscape whose primary feature is a regular valley. A point on this landscape is a set of populations of  $X_i$  and our theorems (Eq. 7) say that as competition acts, the point moves down the landscape, the equilibrium point being the lowest point of the valley. If the bottom of the valley has some negative  $X$  coordinates, those species are eliminated. This is an exact counterpart of classical natural selection in which, provided there is no density dependence and a constant degree of inbreeding, the fitness,  $(1/X)(dX/dt)$ , always increases. Here, we allow for density dependence and find that, when a  $Q$  exists it always decreases. I have called such competitive selection among genes  $\mathcal{X}$  selection (MacArthur and Wilson, 1967) but have pointed out (MacArthur, 1962) that  $\mathcal{X}$  often, but by no means always, increases under such competition. Here, we have seen that in interspecies selection  $Q$  is always minimized, but before we know which  $Q$  to use, for I have defined several in this paper, we must know what kind of competition it is. Thus, we can speak of  $\mathcal{X}$  selection whenever there is competition; under certain conditions, we can go farther and get precise results in terms of a  $Q$ . The equations used here would only apply directly to gene competition of the genes contribute additively to the diploid genotype.

## VII. SUMMARY AND CONCLUSIONS

1. For a wide variety of competition equations there exists a quadratic form,  $Q$ , which is minimized by the action of the competition. A new species can enter a community only if it reduces  $Q$ , and species will be eliminated if  $Q$  is lower without them.

2. By the properties of approximations, a very large number of competing

species can coexist if the environmental production is constant, although more productive environments will hold more species. The more variable the environment, the fewer competitors it will hold.

3. When it exists, the form  $Q$  plays a role analogous to fitness in classical natural selection theory. Natural selection by interspecies competition will lead to low points on a  $Q$  landscape, and so will selection by intergene competition, at least if the genes are additive.

## ACKNOWLEDGMENTS

Richard Lewontin showed me how Section II could be developed and Motoo Kimura pointed out the validity of Eq. 3. To both I am immensely grateful. Henry Horn and Egbert Leigh made numerous useful suggestions.

## REFERENCES

- FRANKLIN, J. N. 1968. "Matrix Theory," Prentice Hall, Englewood Cliffs, N. J.
- HUTCHINSON, G. E. 1947. A note on the theory of competition between two social species, *Ecology* 28, 319-321.
- LOTKA, A. J. 1932. Growth of mixed populations, *J. Wash. Acad. Sci.* 22, 461-469.
- MACARTHUR, R. H. 1962. Some generalized theorems of natural selection, *Proc. Natl. Acad. Sci. U.S.* 48, 1893-1897.
- MACARTHUR, R. H. 1968. The theory of the niche, in "Population Biology and Evolution" (R. Lewontin, Ed.) Syracuse Univ. Press, Syracuse, N. Y.
- MACARTHUR, R. H. 1969. Species packing and what competition minimizes, *Proc. Natl. Acad. Sci. U.S.* (in press).
- MACARTHUR, R. H., AND WILSON E. O. 1967. "The Theory of Island Biogeography," Princeton Univ. Press, Princeton, N. J.