GENETIC STRUCTURE OF A POPULATION OCCUPYING A CIRCULAR HABITAT[1]

THOMAS NAGYLAKI

Department of Medical Genetics, University of Wisconsin, Madison, Wisconsin 53706

Manuscript received December 19, 1973
Revised copy received April 23, 1974

ABSTRACT

The geographical structure of a finite population distributed continuously and homogeneously along a circular habit is explored. Selection is supposed to be absent, and the analysis is restricted to a single locus with discrete, nonoverlapping generations. Assuming every mutant is new to the population, the rate of decay of genetic variability is obtained, and the probability that two homologous genes separated by a given distance are different alleles is calculated. If moments of the migration function higher than second are neglected, the eigenvalue equation is shown to be a simple trigonometric one, and the Fourier series giving the transient and stationary probabilities of allelism are summed in terms of elementary functions. The proportion of homozygotes, the effective number of alleles maintained in the population, and the amount of local differentiation of gene frequencies are discussed.

I. INTRODUCTION

THE rate of approach to homozygosity and the amount of genetic variability maintained by mutation in a finite population are important classical problems in population genetics and evolutionary theory. Taking into account geographical structure makes these questions richer in interest and more applicable to natural populations. A good review of the history and applications of this subject is presented by MARUYAMA (1972).

MARUYAMA (1970a, 1970b, 1971a) has analyzed stepping-stone models of finite populations in one and two dimensions. He has also employed perturbation theory to calculate the asymptotic rate of decay of heterozygosity and the probability of allelism for continuously and homogeneously distributed finite populations in circular and linear habitats in the limiting cases of large populations and large habitats (MARUYAMA 1971b). For a torus-like two-dimensional space, he has given a Fourier expansion of the equilibrium probability of allelism, and has inferred the algebraic form of the dominant eigenvalue controlling the rate of decay of genetic variability without mutation for the "torus" and square from numerical calculations (MARUYAMA 1972).

The model treated in this paper is that of MALÉCOT (1967; 1969, pp. 64–69), with the formulation of MARUYAMA (1971b, 1972). We shall suppose that diploid

---

[1] Paper no. 1716 from the Genetics Laboratory, University of Wisconsin, Madison, Wisconsin 53706. Supported by the National Institutes of Health (Grant GM-15422).

monoecious individuals are distributed continuously and homogeneously along a circular habitat. They migrate in such a manner that the population density remains constant and uniform. Generations are discrete, and at the end of each generation every individual is replaced by a new individual formed by the random union of gametes from the immediately surrounding region. We shall restrict ourselves to a single locus without selection, and assume every mutation is new to the population. We use the term "circle" to designate any closed curve which is demographically homogeneous due to the uniformity of the population density and the migration pattern. Although the primary aim of the analysis is the conceptual understanding of the genetic structure of a population occupying what is (mathematically, at least) the simplest finite habitat, specific biological applications may be possible. As examples, one may think of individuals distributed around a mountain, lake, or the shore of an island. Amphibians or shallow-water organisms in a large, deep lake or around an island would also reside in a "circular" habitat.

In Section II, we shall expound the model in detail and separate the probabilities of allelism in the nonequilibrium and equilibrium cases. In Section III, we shall derive equations for the eigenvalues controlling the rate of decay of genetic variability and for the eigenfunctions which describe the genetic structure of the population. The equilibrium probability that two homologous genes separated by a given distance are the same allele will be deduced in Section IV. If moments of the migration function higher than second are neglected, we shall show that the eigenvalue equation becomes a simple trigonometric one, and the Fourier series for the transient and stationary probabilities of allelism will be summed in terms of elementary functions.

II. FORMULATION

Let individuals be distributed along a "circle" of length L with constant linear density ρ. From any point on the circle, we measure the distance x along the circle, positively in one direction (say, counterclockwise) and negatively in the other. We denote the probability that an individual migrates a distance between x and $x + \Delta x$ $(-\infty < x < \infty)$ in one generation by $m(x)\Delta x$. We do not assume m is symmetric. The homogeneity of the problem permits the use of a migration function depending only on displacement, rather than initial and final positions. Let $r(x)\Delta x$ designate the probability that the separation between two individuals changes by a distance between x and $x + \Delta x$ $(-\infty < x < \infty)$ in one generation. Then

$$r(x) = \int_{-\infty}^{\infty} m(y)m(x+y)dy\,. \tag{1}$$

(If $m(-x) = m(x)$, then (1) agrees with MARUYAMA's (1971b, Eq. (2–1)) definition.) The substitution $z = y - x$ in

$$r(-x) = \int_{-\infty}^{\infty} m(y)m(-x+y)dy\,. \tag{2}$$

shows that r is symmetric:

$$r(-x) = r(x) \tag{3}$$

Let $f(t,x)$ represent the probability that two homologous genes separated by a distance x at time t are the same allele. Then $f(t,0)$ is the proportion of homozygotes. Evidently, $f(t,x)$ is periodic in x with period L:

$$f(t, x + jL) = f(t,x) \ , \ j = \pm 1, \pm 2, \dots \ . \tag{4}$$

Considering f on the interval $-L/2 \leq x \leq L/2$, (3) implies f is even in x:

$$f(t,-x) = f(t,x) \ . \tag{5}$$

Together, (4) and (5) permit us to restrict our attention to the interval $0 \leq x \leq L/2$, and extend our solution by evenness and periodicity. We assume selection is absent, the mutation rate per generation for every gene is u, and every mutant is new to the population. We define

$$\tilde{\delta}(x) = \delta(x|\mathrm{mod}\, L) = \sum_{j=-\infty}^{\infty} \delta(x - jL) \ , \tag{6}$$

where δ is the Dirac delta function. The unidimensional version of MARUYAMA's basic equation (3–1) (1972) reads

$$f(t+1, x) = (1-u)^2 \int_{-\infty}^{\infty} r(y) \{ f(t, x-y) + \left[\frac{1-f(t,0)}{2\rho}\right] \tilde{\delta}(x-y) \} dy \ . \tag{7}$$

Equation (7) may be understood as follows (see also MALÉCOT 1967). For two genes in generation $t+1$ to be identical in state, they must be descended from the same gene or identical genes in generation t, and no mutation must occur. Hence, (7) must have the factor $(1-u)^2$. The first term of the integral is the probability of descent from identical genes separated by a distance between $x-y$ and $x-y-\Delta y$, taking into account the probability that the separation changes to x. The probability that the same gene is sampled twice in an interval of length Δx is $(2\rho\Delta x)^{-1}$, and the separation at time t must be between jL and $jL + \Delta x$, where $j = 0, \pm 1, \pm 2, \dots$. If the two genes are descended from the same gene, the probability of identity at time t is unity, not $f(t,0)$. Therefore, we must correct the first term in the integral in (7) by adding

$$\sum_{j=-\infty}^{\infty} \left[\frac{1-f(t,0)}{2\rho\Delta x}\right] r(x-jL) \Delta x \ .$$

Noting that the Δx cancels out in this expression and recalling (6), we see at once that we have derived the second term in (7).

The probability of nonallelism is

$$h(t,x) = 1 - f(t,x) \ . \tag{8}$$

We put

$$h(t,x) = H(t,x) + h_\infty(x) \ , \tag{9}$$

where $H(t,x) \to 0$ as $t \to \infty$. Thus, $H(t,x)$ and $h_\infty(x)$ are the transient and stationary parts of the probability that two homologous genes separated by a distance x are nonidentical alleles. Clearly, if $u = 0$, the population tends to homozygosity, which implies $h_\infty(x) = 0$. Substituting (8) and (9) into (7), and employing the obvious fact (easily proved from (1) and (10) with r replaced by m) that

$$\int_{-\infty}^{\infty} r(y) dy = 1 \ , \tag{10}$$

leads to

$$H(t+1,x)+h_{\infty}(x)=1-(1-u)^{2}+(1-u)^{2}\int_{-\infty}^{\infty} r(y)\cdot\{H(t,x-y)+h_{\infty}(x-y)\}\cdot$$

$$\cdot\{1-[\tilde{\delta}(x-y)/(2\rho)]\}dy\,. \tag{11}$$

As $t\rightarrow\infty$, we get from (11) the integral equation describing the stationary behavior of the population:

$$h_{\infty}(x)=1-(1-u)^{2}+(1-u)^{2}\int_{-\infty}^{\infty} r(y)h_{\infty}(x-y)\cdot\{1-[\tilde{\delta}(x-y)/(2\rho)]\}dy. \tag{12}$$

Substracting (12) from (11) yields the difference-integral equation for the transient function H:

$$H(t+1,x)=(1-u)^{2}\int_{-\infty}^{\infty} r(y)H(t,x-y)\cdot\{1-[\tilde{\delta}(x-y)/(2\rho)]\}dy\,. \tag{13}$$

Equation (13) will be analyzed in the following section; Eq. (12) in Section IV.

III. THE DECAY OF GENETIC VARIABILITY

The eigenvalues λ and eigenfunctions $\psi(x)$ which determine the transient behavior of the population are defined by the decomposition

$$H(t,x)=\lambda^{t}\psi(x)\,. \tag{14}$$

The general solution is an arbitrary linear combination of terms of the type (14). The expansion coefficients multiplying the various eigenfunctions may be calculated from the initial value $H(0,x)$ of $H(t,x)$. Writing

$$\lambda - (1-u)^{2}\kappa\,, \tag{15}$$

and substituting (14) and (15) into (13) yields

$$\kappa\psi(x)=\int_{-\infty}^{\infty} r(y)\psi(x-y)\{1-[\tilde{\delta}(x-y)/(2\rho)]\}dy\,. \tag{16}$$

Equations (15) and (16) show at once that the effect of mutation is merely to reduce all the eigenvalues of the mutation-free problem by the same factor $(1-u)^{2}$. For geographically unstructured populations, if the individuals are monoecious, this result follows directly from a simple modification of the work of KIMURA and CROW (1964); for dioecious organisms the statement is approximately valid for low mutation rates, and is proved by MALÉCOT (1969, p. 40).

Equations (4) and (5) permit us to expand ψ in a Fourier cosine series:

$$\psi(x)=\sum_{n=0}^{\infty}\varepsilon_{n}a_{n}\cos(2n\pi x/L)\,, \tag{17}$$

where $\varepsilon_{0}=1/2$; $\varepsilon_{n}=1$, $n=1, 2, \ldots$;

$$a_{n}=\frac{2}{L}\int_{-L/2}^{L/2}\psi(x)\cos(2n\pi x/L)dx\,. \tag{18}$$

In the first term of (16), we employ (17) to expand $\psi(x-y)$, observing that, due to (3), the sines do not contribute. The delta function term in (16) may be reduced by recalling (4) and (6). We find from (16) and (18), then,

$$a_n = \frac{\psi(0)}{\rho L} \frac{R_n}{R_n - \kappa}, \tag{19}$$

with

$$R_n = \int_{-\infty}^{\infty} r(x) \cos(2n\pi x/L)\,dx\,. \tag{20}$$

But (10) gives $R_0 = 1$. Hence, (17) and (19) give the solution

$$\psi(x) = \frac{\psi(0)}{\rho L} \left[\frac{1}{2\mu} + \sum_{n=1}^{\infty} \frac{R_n \cos(2n\pi x/L)}{R_n - \kappa} \right], \tag{21}$$

where $\mu = 1 - \kappa$.

For a nontrivial solution, $\psi(0) \neq 0$. Putting $x = 0$ in (21), we obtain the eigenvalue equation

$$\rho L = \frac{1}{2\mu} + \sum_{n=1}^{\infty} \frac{R_n}{R_n - \kappa}\,. \tag{22}$$

For each eigenvalue κ_i computed from (22), we may calculate the corresponding eigenfunction $\psi_i(x)$ from (21). We shall order the eigenvalues so that $|\kappa_{i+1}| < |\kappa_i|$, $i = 0, 1, 2, \ldots$. We proceed now to discuss some limiting cases.

1. *Dense population*

In (22), on the right-hand side only κ depends on ρ. Therefore, as $\rho \to \infty$ with fixed L, the eigenvalues approach the values required to produce singularities in (22), and the corresponding nearly singular terms in (21) approximate the eigenfunctions. For the leading eigenvalue and eigenfunction, this argument yields

$$\mu_0 \approx 1/(2\rho L) = 1/(2N), \quad \psi_0(x) \approx \psi_0(0), \tag{23a}$$

where N is the population size. This agrees with the perturbation calculation of MARUYAMA (1971b). The other eigenvalues are close to the Fourier components (20). Since we shall usually have $|R_{i+1}| < |R_i|$, this means for $i = 1, 2, \ldots$

$$\kappa_i \approx R_i, \quad \psi_i(x) \approx \psi_i(0) \cos(2i\pi x/L). \tag{23b}$$

Notice that from (20) we have

$$|R_n| \leq \int_{-\infty}^{\infty} r(x) |\cos(2n\pi x/L)|\,dx \leq \int_{-\infty}^{\infty} r(x)\,dx = 1\,.$$

Thus, (15) and (23) show that $|\lambda_i| < 1$, $i = 0, 1, 2, \ldots$. This means that $H(t,x) \to 0$ as $t \to \infty$, as required, and (23) is an acceptable approximation for the eigenvalues.

2. *Panmixia*

Let σ^2 be the variance of the probability density r. (From (1) it is straightforward to prove that σ^2 is twice the variance of m.) The dimensionless parameter $k = \pi\sigma/L$ will be very useful. If $k \gtrsim 1$, (20) leads us to expect (due to the rapid oscillation of the cosine) $|R_n| << 1 < \rho L$ ($\rho L \geq 2$, $n \neq 0$). Therefore, we arrive at (23a), as we must. For the higher eigenvalues and eigenfunctions, observe that the sum in (22) is negligible unless κ_i is fairly close to R_i for $i = 1, 2, \ldots$. But then $|R_i| << 1$ implies that the higher eigenfunctions may be neglected after a few generations. In the limit $k \to \infty$, $R_i \to 0$, and only the expected panmictic solution (23a) survives.

3. *Large habitat*

With ρ fixed and $L \to \infty$ (*i.e.*, $k << 1$), expanding the cosine in (20) in a Maclaurin series shows that we may neglect the contribution to R_n of moments of r higher than second. The resulting simplification of (21) will still be convergent. Further justification is provided by the numerical example at the end of this section, and by the agreement with the continuous time results of Section V. Since the cosine in (20) will oscillate more and more rapidly as n increases, generally, R_n will be a decreasing function of n (*e.g.*, if r is a normal distribution, $R_n = e^{-2n^2k^2}$). We have chosen the κ_i to decrease with increasing i. Consequently, the denominator of the sum in (21) informs us that for the smaller eigenvalues (larger i), more of the contribution comes from large n than for the leading ones. But expanding the cosine in (20) shows that higher moments of r are more significant for high n than for low n. Therefore, we conclude that our approximation, as is biologically desirable, is best for the leading eigenvalues and eigenfunctions. This will also be obvious *a posteriori*.

To second order, (20) yields

$$R_n \approx 1 - 2n^2k^2 . \tag{24}$$

Substituting (24) into (21), we obtain

$$\psi(x) \approx \frac{\psi(0)}{\rho L}\left[\frac{1}{2\mu} + \frac{1}{2k^2}\sum_{n=1}^{\infty}\frac{\cos(2n\pi x/L)}{c^2 - n^2}\right] , \tag{25}$$

where $c^2 = \mu/(2k^2)$. If c is not an integer, and $2j\pi \leq \xi \leq 2(j+1)\pi$ for some integer j,

$$\sum_{n=1}^{\infty}\frac{\cos n\xi}{n^2 - c^2} = \frac{1}{2c^2} - \frac{\pi}{2}\,\frac{\cos\{c[(2j+1)\pi - \xi]\}}{c \sin c\pi} \tag{26}$$

(Gradshteyn and Ryzhik 1965, p. 40). Here we have $\xi = 2\pi x/L$, $0 \leq x \leq L/2$, $j = 0$. Therefore, (25) simplifies to

$$\psi(x) \approx \frac{\psi(0)}{\alpha c}\,\frac{\cos\{c\pi[1 - (2x/L)]\}}{\sin c\pi} , \tag{27}$$

where $\alpha = 4k^2\rho L/\pi = 4\pi\rho\sigma^2/L$ is a dimensionless constant which, provided distance is measured in the natural units of habitat size, completely determines the solutions in this approximation. Putting $x = 0$ in (27), we deduce the simple eigenvalue equation

$$\alpha c \approx \cot c\pi . \tag{28}$$

The abscissae at the points of intersection in Figure 1 give the eigenvalues. From Figure 1 we see at once that $0 < c_i < i + 1/2$, and hence $0 < \mu_i < 2k^2\,(i + 1/2)^2$, $i = 0, 1, 2, \ldots$. Since $\mu_i < 1$, our approximation must fail for sufficiently large i and is most accurate, as stated above, for small i.

If there is no mutation, the asymptotic probability of nonallelism is proportional to $\psi_0(x)$. Therefore, with a normalizing constant of the appropriate sign

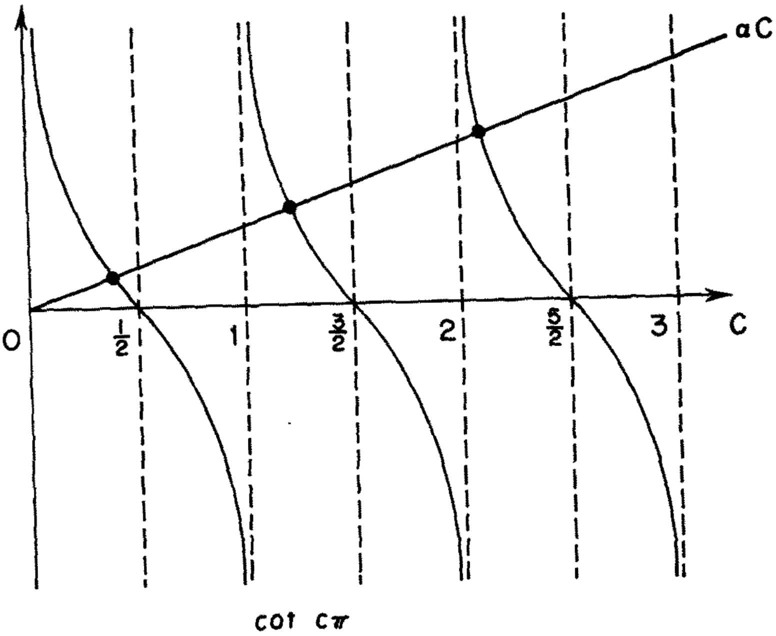

FIGURE 1.—The eigenvalues from Eq. (28).

(determined by the initial conditions, not by the homogeneous equation (27)), $\psi_0(x)$ must be positive definite. Since $0 < c_0 < ½$ and $0 \leq x \leq L/2$ in (27), this is indeed the case.

Equations (27) and (28) reduce considerably in two important special cases:
a) $\alpha >> 1$:

This means either the population is large, or nearly panmictic, or both. From Figure 1, we see that

$$c_i = i + \eta_i\,, \tag{29}$$

with $0 < \eta_i << 1$. Substituting (29) into (28) leads to

$$\alpha(i + \eta_i) \approx 1/(\pi\eta_i)\,, \tag{30}$$

and hence

$$c_0 = \eta_0 \approx 1/\sqrt{\pi\alpha}\,, \tag{31}$$

$$\eta_i \approx 1/(i\pi\alpha), \quad i = 1, 2, \ldots. \tag{32}$$

Then (27) and (31) yield (23a), while from (27), (29), and (32), we obtain for $i = 1, 2, \ldots$

$$\mu_i \approx 2k^2 i^2\,, \tag{33}$$

$$\psi_i(x) \approx \psi_i(0)\cos(2i\pi x/L)\,. \tag{34}$$

Since (24) holds for $k << 1$, (33) and (34) agree with (23b).
b) $\alpha << 1$:

Now the habitat must be extremely large relative to the population density and the variance of the migration distribution. From Figure 1 and expanding (28) it is easy to prove

$$c_i \approx (i + ½)[1 - (\alpha/\pi)], \quad i = 0, 1, 2, \ldots, \tag{35}$$

so

$$\mu_i \approx 2k^2(i+\tfrac{1}{2})^2\,, \tag{36}$$

$$\psi_i(x) \approx \psi_i(0)\left\{\cos[(2i+1)\pi x/L] + \frac{\sin[(2i+1)\pi x/L]}{\alpha(i+\tfrac{1}{2})}\right\}\,. \tag{37}$$

For $i=0$, by an elegant semibiological argument, MARUYAMA (1971b) obtained (36). In (37), the second term dominates unless $(2i+1)\pi x/L$ is very close to $n\pi$, $n=0, 1, 2, \ldots$. Consequently, MARUYAMA's less precise formula, $\psi_0(x) \approx (\text{const.})\sin(\pi x/L)$, is adequate for $x \gtrsim \alpha L/(2\pi)$.

Finally, let us compare the largest eigenvalue computed from (28) to a numerical calculation of MARUYAMA (1971b). With $k=0.140$ and $\rho L=50$, for three different migration functions, he obtained $0.0072 \le \mu_0 \le 0.0076$ (MARUYAMA 1971b; Figure 1). Then $\alpha = 1.257$. We have selected this as an especially unfavorable example: the simplications (a) and (b) do not apply here, MARUYAMA's analytical methods also fail, and k is not sufficiently small compared to unity to expect the neglect of the higher moments of the migration function to make only an insignificant difference. From (28), we get $c=0.364$, and consequently $\mu_0 = 0.00523$, about 25% smaller than MARUYAMA's value.

IV. THE STEADY STATE

We begin by rewriting (12) in terms of the stationary probability of allelism,

$$f_\infty(x) = 1 - h_\infty(x)\,. \tag{38}$$

Using (10) and (38), (12) becomes

$$f_\infty(x) = (1-u)^2\int_{-\infty}^{\infty} r(y)\{f_\infty(x-y) + [1-f_\infty(0)]/(2\rho)\tilde{\delta}(x-y)\}dy\,. \tag{39}$$

The Fourier analysis employed in Section III gives

$$f_\infty(x) = \frac{(1-u)^2[1-f_\infty(0)]}{\rho L}\sum_{n=0}^{\infty}\frac{\varepsilon_n R_n\cos(2n\pi x/L)}{1-(1-u)^2R_n}\,, \tag{40}$$

as is obvious by making the replacements $\psi(0)\to -[1-f_\infty(0)]$ *and* $\kappa\to(1-u)^{-2}$ in (16) and (19). Equation (40) is just Eq. (3–9) of MARUYAMA (1972). Setting $x=0$ and solving for $f_\infty(0)$ in (40) yields

$$f_\infty(0) = \frac{(1-u)^2S}{\rho L + (1-u)^2S}\,, \tag{41}$$

where

$$S = \sum_{n=0}^{\infty}\frac{\varepsilon_n R_n}{1-(1-u)^2R_n}\,. \tag{42}$$

If ρ and L are fixed and we let the mutation rate $u\to 0$, we find from (42)

$$S \sim 1/(4u) + O(1)\,, \tag{43}$$

whence we obtain from (41)

$$1 - f_\infty(0) \sim 4\rho L u + O(u^2)\,, \tag{44}$$

so that (40) leads to

$$f_\infty(x) \sim 1 + O(u)\,. \tag{45}$$

Since the total population size, N, is finite, we know that if there is no mutation the population must eventually become homozygous. Equation (45) confirms this fact, and adds the information that for weak mutation the equilibrium probability of nonallelism is proportional to the mutation rate.

The important limiting case of an infinite habitat is easy to deduce from (40). As $L \to \infty$, we put $s = 2n\pi/L$, and obtain from (20) and (40)

$$R(s) = \int_{-\infty}^{\infty} r(x) \cos sx dx\,, \tag{46}$$

the Fourier cosine transform of r, and

$$f_\infty(x) = \frac{(1-u)^2[1-f_\infty(0)]}{2\pi\rho} \int_0^\infty \frac{R(s)\cos sx}{1-(1-u)^2R(s)} ds. \tag{47}$$

Equation (47) is a more general version of the result of MALÉCOT (1955), in whose Eq. (3) the integral should be multiplied by 1/2. From (47) one obtains (MALÉCOT 1969, pp. 83, 84; MALÉCOT uses the variance of m) for $u << 1$

$$f_\infty(x) \sim \frac{e^{-\frac{2\sqrt{u}}{\sigma}x}}{1+4\rho\sigma\sqrt{u}}\,. \tag{48}$$

Let us return to (40) and assume that the habitat is sufficiently large ($k << 1$) for the approximation (24) to be permissible. Then (40) becomes

$$f_\infty(x) \approx \frac{1-f_\infty(0)}{2k^2\rho L}\left[\frac{1}{2b^2} + \sum_{n=1}^{\infty} \frac{\cos(2n\pi x/L)}{b^2+n^2}\right]\,, \tag{49}$$

where

$$b^2 = \frac{1-(1-u)^2}{2k^2(1-u)^2} \approx \frac{u}{k^2}\,, \tag{50}$$

the last relation being valid for $u << 1$. But (GRADSHTEYN and RYZHIK 1965, p. 40)

$$\sum_{n=1}^{\infty} \frac{\cos n\xi}{b^2+n^2} = \frac{\pi}{2b}\,\frac{\cosh[b(\pi-\xi)]}{\sinh b\pi} - \frac{1}{2b^2}\,, \quad 0 < \xi < 2\pi\,. \tag{51}$$

Substituting (51) into (49) with $\xi = 2\pi x/L$, we obtain for $0 \le x \le L/2$

$$f_\infty(x) \approx \frac{1-f_\infty(0)}{\alpha b}\,\frac{\cosh\{b\pi[1-(2x/L)]\}}{\sinh b\pi}\,. \tag{52}$$

Setting $x = 0$ in (52), we find

$$f_\infty(0) \approx \frac{\coth b\pi}{\alpha b + \coth b\pi}\,, \tag{53}$$

and hence (52) will read

$$f_\infty(x) \approx \frac{1}{\alpha b + \coth b\pi}\,\frac{\cosh\{b\pi[1-(2x/L)]\}}{\sinh b\pi} \tag{54}$$

$$= \frac{\cosh\{b\pi[1-(2x/L)]\}}{\alpha b \sinh b\pi + \cosh b\pi}\,, \quad 0 \le x \le L/2\,. \tag{55}$$

In a slightly different notation, and with the approximate form of b, (53) is stated by MARUYAMA (1970b). It is instructive to compare the decay eigenfunctions (27) to the stationary solution (55). The trigonometric functions for the decay problem correspond to the hyperbolic function in equilibrium. The dimensionless constant α completely determines the eigenfunctions since (28) allows the calculation of c as a function of α, but b is a known independent dimensionless parameter in the steady-state case. Note that according to (55), $f_\infty(x)$ is positive definite, as it must be.

As in Section III, we continue with the consideration of some biologically important limiting cases of our solution:

1. ρ, σ, L fixed; $u \to 0$:

Equation (50) shows that $b \to 0$, so (50) and (55) give immediately for low mutation rate

$$f_\infty(x) \sim \frac{1 + (u/2)(L/\sigma)^2[1 - (2x/L)]^2}{1 + 4Nu + (u/2)(L/\sigma)^2}, \tag{56a}$$

in agreement with (44) and (45).

2. ρ, u fixed; $k \to \infty$:

Equation (50) again implies $b \to 0$, so in the panmictic limit we obtain the result of MALÉCOT (1969, p. 76) and of KIMURA and CROW (1964),

$$f_\infty(x) \sim 1/(1 + 4Nu)\,. \tag{56b}$$

3. ρ, u fixed; $k \to 0$:

From (50), for a large habitat $b \to \infty$, and therefore we deduce from (55)

$$f_\infty(x) \sim \frac{e^{-2b\pi x/L} + e^{-2b\pi[1-(x/L)]}}{1 + \alpha b}\,. \tag{57}$$

Also,

$$\alpha b = 2\sqrt{2}\rho\sigma\sqrt{1-(1-u)^2}(1-u)^{-1} \approx 4\rho\sigma\sqrt{u}\,, \tag{58}$$

where the last formula holds for $u << 1$. For a very large habitat, $x/L < 1 - (x/L)$ and we may neglect the second exponental in (57). With the low mutation rate formulae (50) and (58), this leads to (48), as it must.

Finally, a measure of the amount of genetic variability maintained in the population is the probability $\bar{f}_\infty$ that two homologous genes drawn at random are identical. It reciprocal is the effective number of alleles in the population (CROW and MARUYAMA 1971). By averaging over the positions of the two genes, one can prove the intuitively reasonable formula

$$\bar{f}_\infty = \frac{2}{L}\int_0^{L/2} f_\infty(x)\,dx\,. \tag{59}$$

Employing (55) to perform the integration, and recalling (53), we find

$$\bar{f}_\infty \approx \frac{[1 - f_\infty(0)](1-u)^2}{2N(2u - u^2)} \tag{60}$$

$$\approx [1 - f_\infty(0)]/(4Nu) \quad \text{if} \quad u << 1\,, \tag{61}$$

as demonstrated generally by CROW and MARUYAMA (1971).

V. RELATION TO THE DIFFERENTIAL EQUATION APPROACH

If we neglect moments of $r(y)$ higher than second in (7), we can easily derive a difference (in time)-differential (in space) equation for f. A better approximation is to rewrite (7) for an infinitesimal change Δt in t and to let $\Delta t \to 0$. Then (7) will read

$$f(t+\Delta t, x) = (1-u\Delta t)^2 \int_{-\infty}^{\infty} r(\Delta t, y)\{f(t, x-y) + [\Delta t/(2\rho)][1-f(t,0)]\tilde{\delta}(x-y)\}dy\,, \qquad (62)$$

where $r(\Delta t, y)$ is the migration function for time interval Δt. Following MARUYAMA (1971b) in taking the limit $\Delta t \to 0$, with the assumption that moments of $r(\Delta t, y)$ higher than second are $o(\Delta t)$ we find

$$\frac{\partial f}{\partial t} = -2uf + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2} + \left[\frac{1-f(t,0)}{2\rho}\right]\delta(x)\,, \quad -\frac{L}{2} \le x \le \frac{L}{2}\,. \qquad (63)$$

Applying the derivation between (8) and (13), we deduce from (63), instead of (13) and (39), the differential equations

$$\frac{\partial H}{\partial t} = -2uH + \frac{\sigma^2}{2}\frac{\partial^2 H}{\partial x^2} - \frac{H(t,0)}{2\rho}\delta(x)\,, \quad -\frac{L}{2} \le x \le \frac{L}{2}\,. \qquad (64)$$

$$2uf_\infty - \frac{\sigma^2}{2}\frac{d^2 f_\infty}{dx^2} - \left[\frac{1-f_\infty(0)}{2\rho}\right]\delta(x) = 0\,, \quad -\frac{L}{2} \le x \le \frac{L}{2}\,. \qquad (65)$$

Equation (64) was given by MARUYAMA (1971b) without mutation, while (65) is due to MALÉCOT (1967).

Let us write, in analogy with (14),

$$H(t,x) = e^{\beta t}\psi(x)\,, \quad -\mu = \beta + 2u\,. \qquad (66)$$

Then μ will correspond to the same quantity employed in Section III. Substituting (66) into (64), we obtain

$$-\mu\psi = \frac{\sigma^2}{2}\frac{d^2\psi}{dx^2} - \frac{\psi(0)}{2\rho}\delta(x)\,, \quad -\frac{L}{2} \le x \le \frac{L}{2}\,. \qquad (67)$$

As above, the solutions of (65) and (67) must be even and periodic. From (17) and (40) we infer the boundary conditions

$$\frac{d\psi}{dx}\left(\frac{L}{2}\right) = 0\,, \qquad (68)$$

$$\frac{df_\infty}{dx}\left(\frac{L}{2}\right) = 0\,. \qquad (69)$$

We integrate (67) between $-\varepsilon$ and η, $0 < \varepsilon,\eta << L$, and let $\varepsilon,\eta \to 0$. This yields

$$\frac{\sigma^2}{2}\left[\frac{d\psi}{dx}(0+) - \frac{d\psi}{dx}(0-)\right] = \frac{\psi(0)}{2\rho}\,, \qquad (70)$$

or, taking into account the evenness of ψ,

$$\sigma^2 \frac{d\psi}{dx}(0+) = \frac{\psi(0)}{2\rho}. \tag{71}$$

Similarly, (65) leads to

$$\sigma^2 \frac{df_\infty}{dx}(0+) = -\frac{1-f_\infty(0)}{2\rho}. \tag{72}$$

It is an easy matter to show that (27), (28) and (55) are the unique solutions of the boundary value problems (67), (68), (71), and (65), (69), (72), respectively.

VI. DISCUSSION

Let us consider the implications of the equilibrium result (55) for the proportion of homozygotes, $f_\infty(0)$, the effective number of alleles maintained in the population, $n_e = 1/\bar{f}_\infty$ (CROW and MARUYAMA 1971), and the degree of local differentiation of gene frequencies. From (53),

$$f_\infty(0) = (1 + \alpha b \tanh b\pi)^{-1}, \tag{73}$$

while (53) and (61) yield

$$n_e = b\pi(\alpha b + \coth b\pi). \tag{74}$$

In discussing local differentiation, we shall follow the method of MARUYAMA (1974). If $\bar{f}_\infty$ is not too close to unity, we expect significant local differentiation whenever

$$\bar{f}_\infty / f_\infty(0) = \tanh b\pi/(b\pi) << 1, \tag{75}$$

whereas the population will be near panmixia if this ratio is close to 1. When $\bar{f}_\infty$, and hence $f_\infty(0) \geq \bar{f}_\infty$, are almost 1, (75) is a very insensitive measure of local differentiation. The ratio

$$\frac{1-\bar{f}_\infty}{1-f_\infty(0)} = (\pi\alpha b^2 + b\pi\coth b\pi - 1)/(\pi\alpha b^2), \tag{76}$$

however, will be close to or significantly greater than 1 according to whether population is nearly random mating or not. The assertions below follow directly from Eqs. (73) to (76).

If $\pi b = \sqrt{u}L/\sigma << 1$, then

$$1/f_\infty(0) \approx 1/\bar{f}_\infty = n_e \approx 4Nu + 1, \tag{77}$$

$$\frac{1-\bar{f}_\infty}{1-f_\infty(0)} \approx 1 + \frac{\pi}{3\alpha} = 1 + \frac{L}{12\rho\sigma^2}. \tag{78}$$

If, in addition, $4Nu >> 1$, (77) shows that the population will be quite heterogeneous (as measured by the effective number of alleles or the homozygosity) and close to panmixia (since $f_\infty(0) \approx \bar{f}_\infty << 1$). If $4Nu << 1$, however, the amount of genetic variability at equilibrium will be very small, and the population will be approximately panmictic only if $\alpha = 4\pi\rho\sigma^2/L >> 1$.

If $\pi b = \sqrt{u}L/\sigma >> 1$, then

$$1/f_\infty(0) \approx 1/(\pi b \bar{f}_\infty) \approx 1 + 4\rho\sigma\sqrt{u}, \tag{79}$$

$$n_e \approx (\sqrt{u}L/\sigma) + 4Nu. \tag{80}$$

Thus, if the habitat is sufficiently large ($L >> \sigma/\sqrt{u}$), there will always be a great deal of genetic variability and local differentiation of gene frequencies. If, also, $4\rho\sigma\sqrt{u} << 1$, most individuals will be homozygous, but for different alleles.

We shall comment briefly on the applicability of the methods employed in this paper to other types of habitats. If the habitat is inhomogeneous owing to the existence of boundaries (*e.g.*, "line" or square), at least in its present simple form, Fourier analysis does not yield a solution. For the "torus" (MARUYAMA 1972), an analogous Fourier expansion does hold, but higher moments of the migration function cannot be neglected because, if they are, the eigenvalue equation and the equation determining the local homozygosity in the steady state become divergent. This is probably the underlying reason for the failure of MARUYAMA's (1971b) large-habitat approximation in this case. Therefore, in two dimensions, he was compelled to resort to numerical methods (MARUYAMA 1972).

It is a pleasure to thank PROFESSOR WARREN EWENS for many suggestions which have improved the clarity and readability of the manuscript.

LITERATURE CITED

CROW, J. F. and T. MARUYAMA, 1971 The number of neutral alleles maintained in a finite, geographically structured population. Theoret. Pop. Biol. **2**: 437–453.

GRADSHTEYN, I. S. and I. M. RYZHIK, 1965 *Table of Integrals, Series, and Products.* Academic Press, New York.

KIMURA, M. and J. F. CROW, 1964 The number of alleles that can be maintained in a finite population. Genetics **49**: 725–738.

MALÉCOT, G., 1955 Decrease of relationship with distance. Cold Spring Harbor Symp. Quant. Biol. **20**: 52–53. ——, 1967 Identical loci and relationship. Proc. Fifth Berkeley Symp. Math. Stat. Prob. **4**: 317–332. ——, 1969 *The Mathematics of Heredity.* W. H. Freeman and Company, San Francisco.

MARUYAMA, T., 1970a On the rate of decrease of heterozygosity in circular stepping-stone models of populations. Theoret. Pop. Biol. **1**: 101–119. ——, 1970b Analysis of population structure. I. One-dimensional stepping-stone models of finite length. Ann. Hum. Genet. **34**: 201–219. ——, 1971a Analysis of population structure. II. Two-dimensional stepping-stone models of finite length and other geographically structured populations. Ann. Hum. Genet. **35**: 179–196. ——, 1971b The rate of decrease of heterozygosity in a population occupying a circular or a linear habitat. Genetics **67**: 437–454. ——, 1972 The rate of decrease of genetic variability in a two-dimensional continuous population of finite size. Genetics **70**: 639–651. ——, 1974 Allelism probability and genetic variability in a geographically structured finite population. In press.

Corresponding editor: W. J. EWENS