NUMERICAL ANALYSIS OF RANDOM DRIFT IN A CLINE†

THOMAS NAGYLAKI

*Dept. of Biophysics and Theoretical Biology, The University of Chicago,
Chicago, Illinois 60637*

AND

BRADLEY LUCIER

*Dept. of Mathematics, The University of Chicago,
Chicago, Illinois 60637*

ABSTRACT

The equilibrium state of a diffusion model for random genetic drift in a cline is analyzed numerically. The monoecious organism occupies an unbounded linear habitat with constant, uniform population density. Migration is homogeneous, symmetric and independent of genotype. A single diallelic locus with a step environment is investigated in the absence of dominance and mutation. The flattening of the expected cline due to random drift is very slight in natural populations. The ratio of the variance of either gene frequency to the product of the expected gene frequencies decreases monotonically to a nonzero constant. The correlation between the gene frequencies at two points decreases monotonically to zero as the separation is increased with the average position fixed; the decrease is asymptotically exponential. The correlation decreases monotonically to a positive constant depending on the separation as the average position increasingly deviates from the center of the cline with the separation fixed. The correlation also decreases monotonically to zero if one of the points is fixed and the other is moved outward in the habitat, the ultimate decrease again being exponential. Some asymptotic formulae are derived analytically.——The loss of an allele favored in an environmental pocket is investigated by simulating a chain of demes exchanging migrants, the other assumptions being the same as above. For most natural populations, provided the allele would be maintained in the population deterministically, this process is too slow to have evolutionary importance.

IN a recent paper (NAGYLAKI 1978a), the biological importance of random genetic drift in a cline was briefly discussed, previous work was reviewed and a diffusion model was developed. The diffusion treatment requires that the selection coefficient, s, the variance of the migrational displacement, σ^2, and the reciprocal of the population density, $1/\rho$, all be small and of the same order of magnitude (NAGYLAKI 1978b), and yields the following results (NAGYLAKI 1978a). If both alleles are favored in very large regions, for most selection gradients, s, σ^2 and ρ appear only combined in the single dimensionless parameter β, the ratio of the characteristic length for migration and random drift to

the natural distance for migration and selection. Random drift is highly significant if $\beta \ll 1$; it causes only relatively small variations of order $1/\beta$ around the deterministic cline if $\beta \gg 1$. With $\beta \gg 1$, the correlation between the gene frequencies at any two points is very nearly independent of β, and hence is parameter-free for particular forms of the selection gradient. For a very steep selection gradient, $\beta = 2\sqrt{2s}\rho\sigma$, i.e., essentially the product of the square root of the selection intensity and the neighborhood size. Thus, with weak selection, a large neighborhood size is required for large β.

Since the derivation of detailed analytic results appears to be quite difficult, we must resort to numerical methods. Such treatments of random drift in linear stepping-stone clines of finite length have been given by FELSENSTEIN (1975) and SLATKIN and MARUYAMA (1975). The strong-selection limit of the diffusion model (pertinent to FELSENSTEIN's work) was studied numerically by LUSKIN and NAGYLAKI (1979), who also obtained some asymptotic results analytically. The numerical investigation of the complete diffusion model in Section I of this paper relates to the computations of SLATKIN and MARUYAMA (1975), already discussed in NAGYLAKI (1978a). In Section II, we shall deduce analytically some aspects of the long-distance behavior of the diffusion approximation.

By combining parameters into the single quantity β, the diffusion approximation enormously increases the range of parameters accessible to numerical analysis. All migration patterns with the same variance can be treated simultaneously, and a single calculation with a fixed value of β covers all possible values of s, σ^2, and ρ yielding that β. Furthermore, without the diffusion approximation, numerical investigation of parameter sets with extremely weak migration and selection and high population density is not practicable.

Deterministic theory (FLEMING 1975; NAGYLAKI 1975) shows that an allele favored in an environmental pocket will be maintained in the population if and only if k, the ratio of the width of the pocket to the natural length of the cline corresponding to the step environment, exceeds a critical value, usually of order unity. With finite population density, however, random drift causes certain loss of this allele in a finite time. In Section III, we shall examine the effectiveness of spatial heterogeneity in preserving genetic diversity by studying the distribution of this extinction time. Although lack of a detailed analytic theory will force us to employ direct simulation, we shall find that the results may be simply and informatively interpreted in terms of the diffusion parameters k and β.

I. NUMERICAL ANAYLSIS OF THE INFINITE CLINE

Our diffusion model is the limit of the following discrete scheme. A diploid monoecious population occupies an infinite chain of equally spaced colonies. We suppose migration is homogeneous, symmetric and independent of genotype, and consider a single diallelic locus with alleles A_1 and A_2 in the absence of dominance and mutation. The generation starts with the same very large number of zygotes in each colony. Selection and migration, occurring in succession, may alter the subpopulation sizes, but the latter must remain extremely

large. Any biologically reasonable variation of the subpopulation numbers under selection, in particular, no change (soft selection) and change determined by the mean fitness in each colony (hard selection), leads to the same diffusion limit (NAGYLAKI 1978b). Random genetic drift acts through population regulation, which truncates each colony to the same finite number of adults. Each of these adults produces many gametes without fertility differences. These gametes fuse at random to form the next generation of zygotes, in Hardy-Weinberg proportions in each colony.

We confine ourselves to the step environment, positing $1 + s$ sgn x, 1, $1 - s$ sgn x $(s > 0$; sgn $x = x/|x|$, $x \neq 0$) for the fitnesses of A_1A_1, A_1A_2, A_2A_2 at position x $(-\infty < x < \infty)$. This choice represents precisely cases of abrupt environmental change, such as sudden shifts in soil or vegetation, and approximates fairly accurately situations in which the characteristic length of the environmental variation is less than $l = \sigma/\sqrt{2s}$ (SLATKIN 1973), the natural length for the cline corresponding to the step environment. Many of our general conclusions will concern long-distance behavior; these should hold qualitatively even for slower environmental variation (FISHER 1950; FELSENSTEIN 1975; MAY, ENDLER and McMURTRIE 1975; NAGYLAKI 1975, 1978c; SLATKIN and MARUYAMA 1975). Notice that for simplicity we have assumed equal selection intensities in the two environments.

We investigate only the equilibrium state of the population. For the step environment, the characteristic convergence time to equilibrium is $1/s$ (NAGYLAKI 1978a). Our treatment will be restricted to the expected gene frequency in the cline and the covariance between the gene frequencies at two arbitrary points. Let us denote the mean allelic frequency at x by $\bar{p}(\xi)$, where $\xi = x/l$ is a convenient, dimensionless spatial coordinate. We express the covariance between the gene frequencies at x and $y = l\eta$ as a function, $v(z,w)$, of the dimensionless average position $z = (\xi + \eta)/2$ and separation $w = (\xi - \eta)/2$.

Taking into account the symmetries $\bar{p}(-z) = 1 - \bar{p}(z)$ and $v(-z,w) = v(z,-w) = v(z,w)$, we have the following boundary value problem in the quarter-plane $z \geq 0$, $w \geq 0$ (NAGYLAKI 1978a):

$$\bar{p}_{zz} + \bar{p}(z)[1 - \bar{p}(z)] - v(z,0) = 0, \tag{1a}$$

$$v_{zz} + v_{ww} + 4\{1 - \bar{p}(z+w) - \bar{p}(|z-w|)\} v = 0, \tag{1b}$$

$$v_w(z,0+) + \beta^{-1}\{\bar{p}(z)[1 - \bar{p}(z)] - v(z,0)\} = 0, \tag{1c}$$

$$\bar{p}(0) = 1/2, \tag{1d}$$

$$\bar{p}(z) \to 1 \text{ as } z \to \infty, \tag{1e}$$

$$v_z(0,w) = 0, \tag{1f}$$

$$v(z,w) \to 0 \text{ as } z + w \to \infty, \tag{1g}$$

in which the subscripts represent partial derivatives (e.g., $v_{zz} = \partial^2 v/\partial z^2$).

The only parameter in (1) is $\beta = 2\sqrt{2s}\rho\sigma$. Having ignored mutation, we confine ourselves to selection intensities exceeding typical mutation rates: $s \geq 10^{-5}$. Since the neighborhood size $\rho\sigma$ always seems to exceed about 30, and is usually at least a few hundred (WRIGHT 1978, Chapter 2), the model is of evolutionary interest only if $\beta \gtrsim 1/4$. We shall solve the system (1) for $\beta = 1/4$, 1 and 4. As

noted above, for $\beta \gg 1$, random drift causes relatively small variations of order $1/\beta$ around the deterministic cline. Hence, if $\beta \gtrsim 4$, we expect the weak-random drift computations of LUSKIN and NAGYLAKI (1979) to be an adequate approximation. Our results will show that this is, indeed, the case.

We expect all the results in this section to deviate at most one percent from the (unknown) exact solution of (1). In fact, the agreement with the asymptotic behavior derived in Section II is much closer than that.

In Figure 1, we plot $\bar{p}(z)$. The first curve from the left is the deterministic solution (HALDANE 1948)

$$-(1/2) + (3/2) \tanh^2 [(z/2) + \tanh^{-1} \sqrt{2/3}]; \tag{2}$$

continuing toward the right, the others correspond to $\beta = 4$, 1 and 1/4. The effect of random drift on the expected gene frequency is evidently fairly small; it is quite negligible for $\beta \geq 4$. Figure 1 exhibits the flattening of the cline due to random drift, previously noted by SLATKIN and MARUYAMA (1975). Setting $\bar{q}(z) = 1 - \bar{p}(z)$ and plotting $-\ln \bar{q}(z)$ shows that $\bar{q}(z)$ ultimately tends to zero exponentially, in precise agreement with the asymptotic results (7), (12) and (13) in Section II. The position where $\bar{p}(z) = 1/2$ is fixed at the origin by symmetry; as indicated by the variances displayed in Figure 2 below, individual
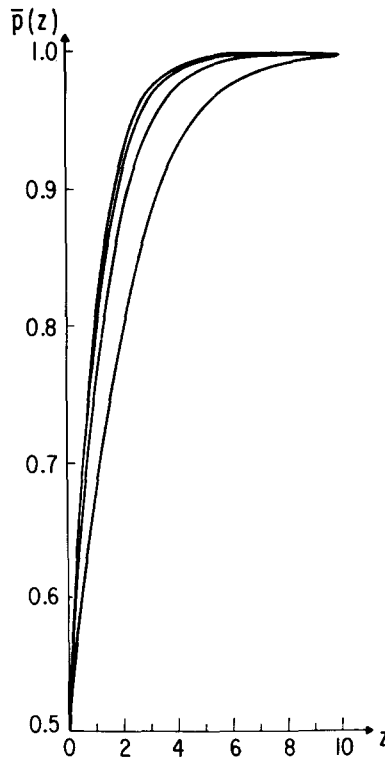

FIGURE 1.—The expected gene frequency as a function of position.

observations of the random variable $p(z)$ may be shifted considerably. Consult SLATKIN and MARUYAMA (1975) for further discussion.

Since the gene frequency has mean $\bar{p}(z)$ and must be in $[0,1]$, the maximum possible value of the variance $v(z,0)$ is $\bar{p}(z)[1-\bar{p}(z)]$. Hence, the ratio of these quantities,

$$F(z) = \frac{v(z,0)}{\bar{p}(z)[1-\bar{p}(z)]} \ , \tag{3}$$

shown in Figure 2, is a good measure of random variation. As expected, $F(z)$ decreases for each z as β increases. Although $F(z)$ becomes proportional to $1/\beta$ as $\beta \to \infty$ (NAGYLAKI 1978a), for finite β, $F(z)$ decreases more slowly than $1/\beta$. For $\beta = 4$, $\beta F(z)$ is already only about 15% less than its limiting value for $\beta = \infty$, as computed by LUSKIN and NAGYLAKI (1979). $F(z)$ is horizontal at $z = 0$ by dint of the symmetry condition $F(-z) = F(z)$; it decreases monotonically to a constant value (depending on β) as z increases; the limit $F(\infty)$ agrees exactly with the analytic expressions (10), (12) and (13) in Section II. Since $F(z)$ decreases as β increases, it is hardly surprising that $F(0) - F(\infty)$ decreases with β, showing that the curves are flattened; $\beta[F(0) - F(\infty)]$, however, increases with increasing β.

The correlation between the allelic frequencies at x and y reads

$$r(z,w) = v(z,w)[v(z+w,0)v(z-w,0)]^{-1/2} \ ; \tag{4}$$

we exhibit $-\ln r(z,w)$ in Figures 3 through 11. Far from the origin $(z = w = 0)$, the correlation increases if β increases with z and w fixed. Near the origin, in agreement with the results of SLATKIN and MARUYAMA (1975), the correlation
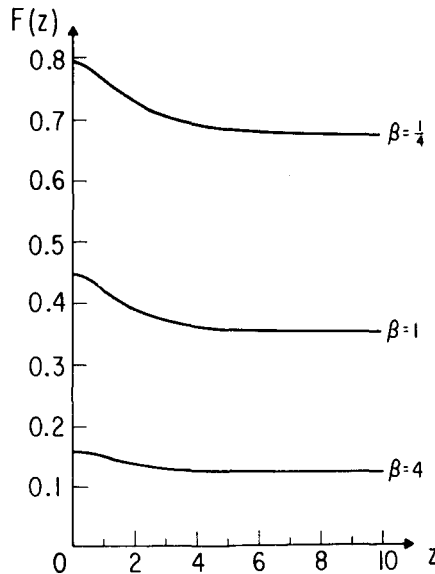

FIGURE 2.—The standardized variance as a function of position.

depends rather weakly on β, and decreases as β increases. As $\beta \to \infty$, $r(z,w)$ becomes independent of β (NAGYLAKI 1978a); for $\beta = 4$, $r(z,w)$ is already within about 10% of its limiting value for $\beta = \infty$, calculated by LUSKIN and NAGYLAKI (1979).

We display the monotonic decrease of the correlation for fixed average position, z, and variable separation, w, in Figures 3 through 5. The curves are never too far from exponential; as $w \to \infty$ they clearly become exponential, the exponential approximation being quite accurate in all cases for $w > 4$. The decay rates in the exponentials are monotone decreasing in both z and β (i.e., the correlations decrease more slowly for large z and large β); they range from 0.93 for $z = 5$ and $\beta = 4$ to 1.42 for $z = 0$ and $\beta = 1/4$. The decay rates for $\beta = 4$ are within a few percent of the ones obtained by LUSKIN and NAGYLAKI (1979) for $\beta = \infty$. The correlations for $z = 0$ and $z = 1$ are quite close because $r(-z,w) = r(z,w)$ implies $\partial r/\partial z(0,w) = 0$.

Figures 6 through 8 exhibit the monotonic decrease of the correlation as a function of the average position, z, for various values of the separation, w. The curves are horizontal at $z = 0$ because $r(-z,w) = r(z,w)$. The constant limits as $z \to \infty$ agree with (11), (12) and (13) in Section II. The curves show greater variation for large w in the sense that $r(0,w)/r(\infty,w)$ increases as w increases.

Suppose that we now fix one point at η and vary the position of the other point, ξ, moving it outward from η. In Figures 9 through 11 we display $r(w+\eta,\eta)$ for various values of η. As noted by SLATKIN and MARUYAMA (1975) for $\eta = 0$, the monotone decrease is very close to exponential throughout. For $\eta = 3$ the results are already extremely close to the limiting form (11), (12), (13) derived in Section II, which applies as $\eta \to \infty$. The figures show that the correlation decreases monotonically to this limit as η increases.

II. ASYMPTOTIC BEHAVIOR OF THE INFINITE CLINE

Since in this section we shall be concerned only with long-distance properties of the cline, we can generalize the fitness pattern of Section I to $1 + sg(x)$, 1, $1 - sg(x)$, with the assumption that $g(x) \to 1$ as $x \to \infty$. If both points at which the gene frequency is measured are on the right side of the cline and far away, i.e., $z - w >> 1$, we can still deduce the asymptotic behavior from (1). In fact, (1e) and (1g) permit us to linearize (1) in $\bar{q}$ and v; as $z - w \to \infty$, these satisfy the boundary value problem

$$\bar{q}_{zz} - \bar{q}(z) + v(z,0) = 0, \tag{5a}$$

$$v_{zz} + v_{ww} - 4v = 0, \tag{5b}$$

$$v_w(z,0+) + \beta^{-1} [\bar{q}(z) - v(z,0)] = 0, \tag{5c}$$

$$\bar{q}(z) \to 0 \quad \text{as } z \to \infty, \tag{5d}$$

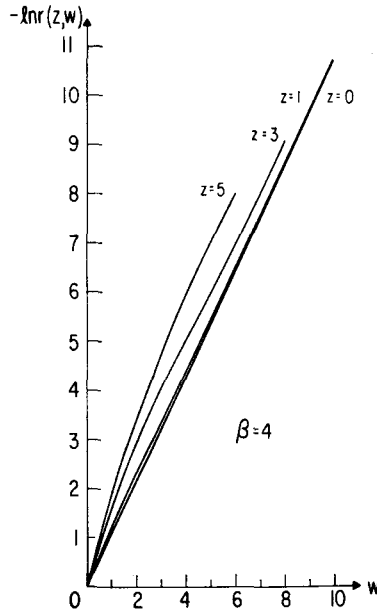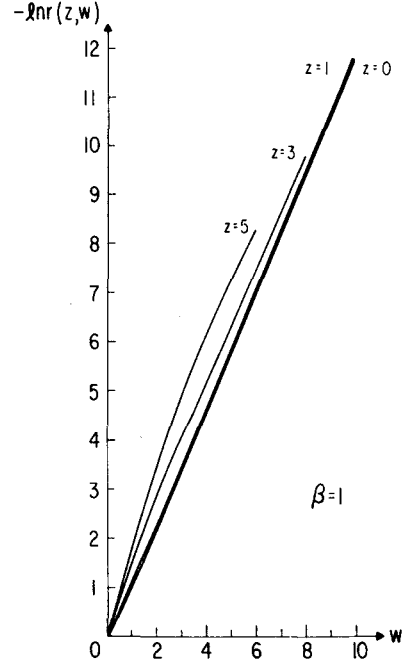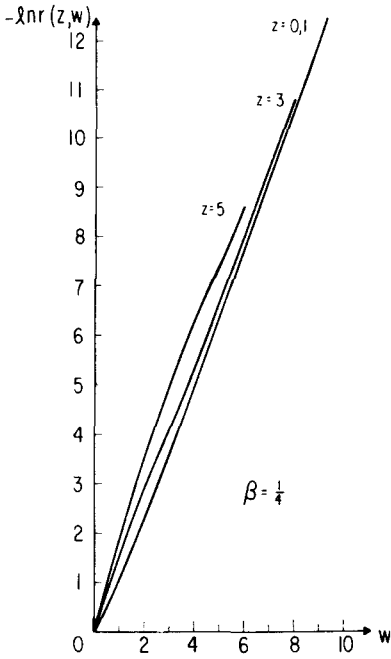$$v(z,w) \to 0 \quad \text{as } z + w \to \infty. \tag{5e}$$

FIGURE 3.—The correlation for $\beta = \frac{1}{4}$ as a function of separation for various fixed values of the average position.

FIGURE 4.—The correlation for $\beta = 1$ as a function of separation for various fixed values of the average position.

FIGURE 5.—The correlation for $\beta = 4$ as a function of separation for various fixed values of the average position.
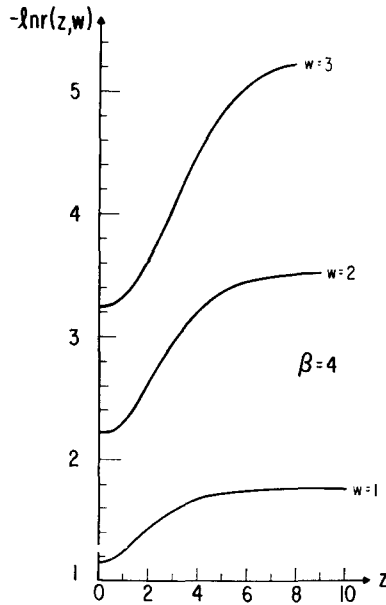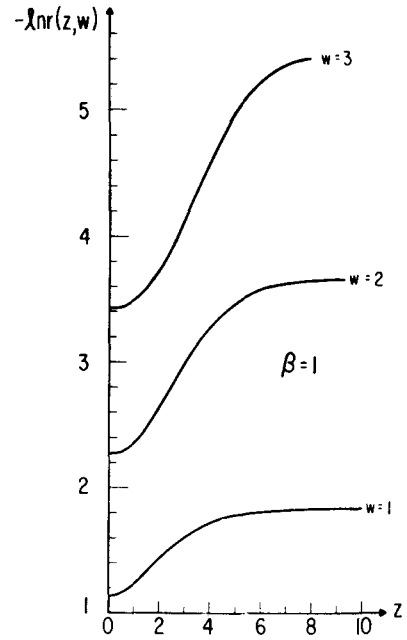
FIGURE 6.—The correlation for $\beta = \frac{1}{4}$ as a function of average position for various fixed values of the separation.
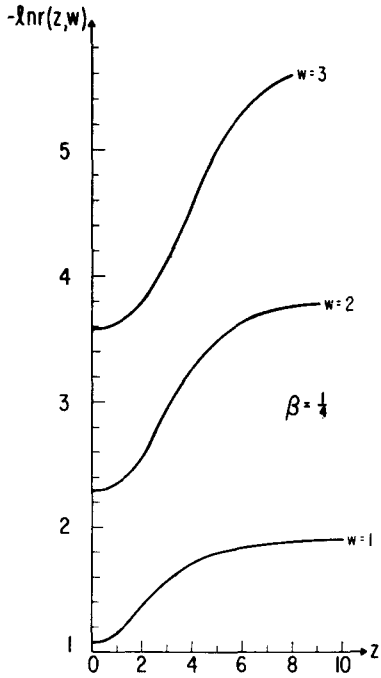
FIGURE 7.—The correlation for $\beta = 1$ as a function of average position for various fixed values of the separation.

FIGURE 8.—The correlation for $\beta = 4$ as a function of average position for various fixed values of the separation.
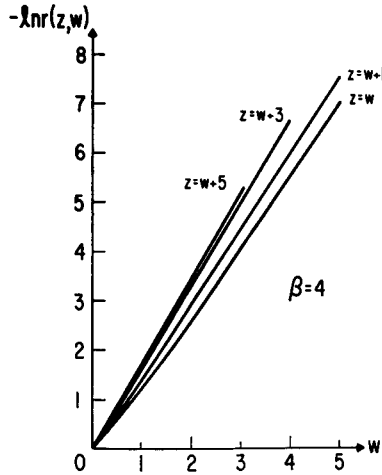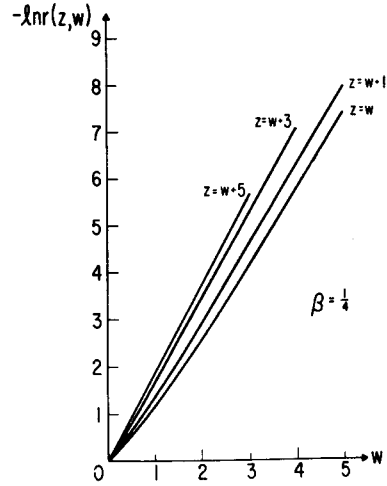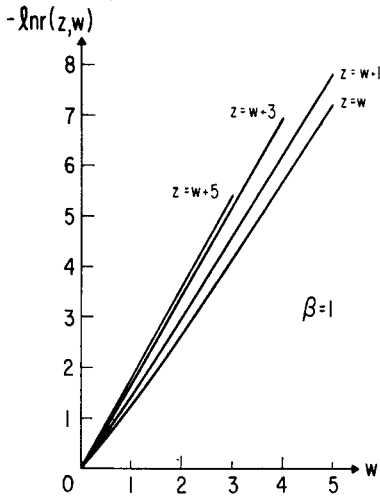
FIGURE 9.—The correlation for $\beta = \frac{1}{4}$ as a function of separation for various fixed locations of the point closer to the environmental transition.

FIGURE 10.—The correlation for $\beta = 1$ as a function of separation for various fixed locations of the point closer to the environmental transition.

FIGURE 11.—The correlation for $\beta = 4$ as a function of separation for various fixed locations of the point closer to the environmental transition.

Separation of variables shows that solutions of (5b) have the form

$$v(z,w) \sim Ae^{-\lambda z - \mu w}, \quad \lambda^2 + \mu^2 = 4. \tag{6}$$

In view of (5e) and the required positivity of v for all z and w, λ and μ must be real and positive. Inserting (6) into (5c) yields

$$\bar{q}(z) \sim A(1 + \beta\mu)e^{-\lambda z} \text{ as } z \to \infty, \tag{7}$$

the exponential decay agreeing with (5d). Substituting (6) and (7) into (5a), we find

$$\lambda^2 = \beta\mu/(1+\beta\mu), \tag{8}$$

and combining this with (6) yields the cubic

$$f(\mu) \equiv \beta\mu^3 + \mu^2 - 3\beta\mu - 4 = 0 \tag{9}$$

for $\mu = \mu(\beta)$. From (6) and (7) we obtain

$$F(z) \to F(\infty) = 1/(1+\beta\mu) \quad \text{as } z \to \infty, \tag{10}$$

$$r(z,w) \sim e^{-\mu w} \quad \text{as } z - w \to \infty. \tag{11}$$

Thus, although our simple method does not determine the constant A, the standardized variance and the correlation fortunately turn out to be independent of A.

To fully specify F and r, it remains to solve (9) for $\mu(\beta)$. By Descartes' Rule of Signs (see e.g., FIEDLER 1969, p. 1170), the number of positive roots of a polynomial equation is either equal to the number of changes of sign of the coefficients, or it is less by an even number. Since there is exactly one change of sign in $f(\mu)$, (9) has precisely one positive root, $\mu(\beta)$. Observing that $f(\sqrt{3}) = -1 < 0$ and $f(2) = 2\beta > 0$, we conclude that the root μ is in $(\sqrt{3}, 2)$.

The solution of (9) is standard (see, e.g., VILHELM 1969, pp. 77–79). Define the following functions of the coefficients:

$$b = -1 - (9\beta^2)^{-1}, \quad c = (27\beta^3)^{-1} - 3(2\beta)^{-1}, \tag{12a}$$

$$D = -b^3 - c^2 = 1 - 23(12\beta^2)^{-1} + 4(27\beta^4)^{-1}. \tag{12b}$$

If the discriminant $D \geq 0$, then (9) has three real roots; these are all distinct if $D > 0$. If $D < 0$, then (9) has one real and two complex conjugate roots. Now, D is a quadratic in β^{-2}, so it is easy to see that $D < 0$ if and only if $\beta_+ < \beta < \beta_-$, where

$$\beta_\pm = [(27/96)(23 \pm 11\sqrt{11/3})]^{-1/2}.$$

Hence, $\beta_+ = 0.284$ and $\beta_- = 1.355$. Which one of the three roots is the unique positive one depends on the sign of c; (12a) informs us that $c > 0$ if and only if $\beta < \beta_0 = \sqrt{2}/9 = 0.157 < \beta_+$. With

$$d = (\text{sgn } c)\sqrt{|b|}, \tag{12c}$$

we have the following cases:

(1) $\beta \leq \beta_+$ or $\beta \geq \beta_-$:
   Set
   $$\phi = \cos^{-1}(c/d^3). \tag{12d}$$

   (a) $\beta < \beta_0$:
   $$\mu = -(3\beta)^{-1} + 2d \cos[(\pi - \phi)/3]. \tag{13a}$$

(b) $\beta_0 < \beta \le \beta_+$ or $\beta \ge \beta_-$:

$$\mu = -(3\beta)^{-1} - 2d \cos(\phi/3).$$ (13b)

(2) $\beta_+ < \beta < \beta_-$:

Put

$$\theta = \cosh^{-1}(c/d^3) = \ln\{(c/d^3) + [(c/d^3)^2 - 1]^{1/2}\}.$$ (12e)

Then

$$\mu = -(3\beta)^{-1} - 2d \cosh(\theta/3).$$ (13c)

[The degenerate case $c = 0$, *i.e.*, $\beta = \beta_0$, is trivially solved (VILHELM 1969, pp. 77–79).]

We can derive approximations for $\mu(\beta)$ for small and large β. Substituting a power series in β into (9) and equating coefficients produces

$$\mu = 2 - (1/2)\beta + (17/16)\beta^2 - (5/2)\beta^3 + O(\beta^4)$$ (14a)

as $\beta \to 0$. A series in inverse powers of β yields

$$\mu = \sqrt{3} + (6\beta)^{-1} - 5\sqrt{3}(72\beta^2)^{-1} + 8(81\beta^3)^{-1} + O(\beta^{-4})$$ (14b)

as $\beta \to \infty$. Thus, as $\beta \to 0$, $\mu \to 2$ and $\lambda \to 0$; as $\beta \to \infty$, $\mu \to \sqrt{3}$ and $\lambda \to 1$. The latter limit recaptures the results $\beta F(\infty) = 1/\sqrt{3}$ and $r(z,w) \sim \exp(-\sqrt{3}w)$ of LUSKIN and NAGYLAKI (1979) for weak random drift.

A few qualitative conclusions, confirming some of the computations in Section II, follow without appealing to the explicit solution (12), (13).

Solving (9) for β as a function of μ, differentiating and using that $\sqrt{3} < \mu < 2$ shows that

$$\frac{d\mu}{d\beta} < 0, \quad \text{whence} \quad \frac{d\lambda}{d\beta} > 0.$$ (15)

Hence, as the amount of random drift is increased (*i.e.*, β decreased), the expected cline is broadened, and, provided $z - w$ is sufficiently large, the correlation is decreased.

By an easy calculation, with the aid of (9) and (10), we find

$$\frac{d}{d\mu}F(\infty) = 2\mu > 0;$$

combined with (15) this implies

$$\frac{d}{d\beta}F(\infty) < 0,$$ (16)

thereby establishing the decrease of the asymptotic standardized variance with weakening random drift.

A few numerical examples will relate the above theory to the computations in Section I and indicate the accuracy of the series (14). The exact values of

TABLE 1

Comparison of analysis, computation, and approximation

β	μ	Analysis λ	$F(\infty)$	μ	Computation λ	$F(\infty)$	Approximation μ
1/10	1.95863	0.40468	0.83622	—	—	—	1.95813
1/4	1.91729	0.56921	0.67599	1.90601	0.56896	0.67501	1.90234
1	1.83117	0.80425	0.35321	1.82631	0.80418	0.35180	1.87720
4	1.76747	0.93598	0.12392	1.76328	0.93557	0.12328	1.76774

μ, λ and $F(\infty)$ in Table 1 were calculated from (13), (6) and (10), respectively. We can estimate the computed value of μ most accurately from the correlations for $w = 1$ in Figures 6 through 8 because these yield the highest value of $z - w$; as expected from the figures, this procedure slightly underestimates the true value of μ. For an independent check on λ, we employ the values of the expected gene frequency plotted in Figure 1. The results displayed in Figure 2 supply the calculated values of $F(\infty)$. The agreement between all the exact and computed results is excellent. The last column in Table 1 is based on (14): we use (14a) for $\beta = 1/10$ and 1/4, and (14b) for $\beta = 1$ and 4. The inverse-power series (14b) must be utilized for $\beta = 1$ because, in contrast to the power series (14a), the coefficients in (14b) decrease fairly rapidly. These approximations are quite adequate; only for β fairly close to 1/2 is neither series acceptable.

Before ending this section, we present the asymptotic behavior far to the left. Let us assume $g(x) \to -\alpha^2$ as $x \to -\infty$, but is otherwise arbitrary. By definition, $w > 0$. If the gene frequency is measured at two points far to the left, we may suppose that $z + w \to -\infty$. The asymptotic boundary value problem for $\bar{p}(z)$ and $v(z,w)$ differs only slightly from (5); by the techniques used above we find

$$\bar{p}(z) \sim B(1 + \alpha\beta\mu_0)e^{\alpha\lambda_0 z} \quad \text{as} \quad z \to -\infty, \tag{17a}$$

$$F(z) \to F(-\infty) = 1/(1 + \alpha\beta\mu_0) \quad \text{as} \quad z \to -\infty, \tag{17b}$$

$$r(z,w) \sim e^{-\alpha\mu_0 w} \quad \text{as} \quad z + w \to -\infty, \tag{17c}$$

where B is a constant, $\lambda_0^2 + \mu_0^2 = 4$, and μ_0 is given in terms of the unique positive root $\mu = \mu(\beta)$ of (9) by $\mu_0 = \mu(\alpha\beta)$. It follows at once that all qualitative properties are the same on the two sides of the cline ($\bar{p}$ and $\bar{q}$ being interchanged), as they must be.

III. LOSS OF AN ALLELE FAVORED IN AN ENVIRONMENTAL POCKET

In the previous sections we assumed that both alleles were favored in such large regions that extremely long absorption times render the study of equilibrium properties of the infinite cline biologically meaningful. If at least one allele is favored in a region not greatly exceeding the characteristic lengths in the problem, the absorption times may be sufficiently short to have biological relevance. Hence, to discover both the range of applicability of the previous sections and the efficacy of spatial diversity in maintaining genetic variability, we must investigate the extinction of an allele favored in an environmental pocket.

Lacking a detailed diffusion approximation for the absorption times, we must resort to numerical analysis of special cases of the Markov chain on which a diffusion approximation would be based. Both biology and the desire for valid, simple and informative diffusion interpretation compels us to choose parameters for which simulation is much more effective than matrix iteration.

The model described in the first paragraph of Section I is specialized as follows. Each deme has N individuals just before reproduction. Migration is confined to nearest neighbors, which exchange migrants at rate $m/2$. Since one obviously cannot simulate an infinitely long chain of colonies, we assume that there are i demes, A_1 being favored in j of them and deleterious in the remainder. With the discrete form of the step environment discussed in Section I, the selection coefficient s_n in deme n has the form ($s > 0$)

$$s_n = \begin{cases} s, & 1 \leq n \leq j, \\ -s, & j+1 \leq n \leq i. \end{cases} \tag{18}$$

Intuition and deterministic analyses (NAGYLAKI 1975, 1978d) suggest that placing the environmental pocket in the interior of the habitat rather than at one end would make no qualitative difference. In any case, the environment often *is* different near the end of a habitat, as it was in HALDANE's (1948) original application to *Peromyscus polionotus*.

If $p_n(t)$ is the frequency of A_1 in adults just prior to reproduction in colony n in generation $t (= 0,1,2,\ldots)$, p_n is also the gene frequency after reproduction, and after selection it becomes

$$p_n^* = \frac{p_n(1 + s_n p_n)}{1 + s_n(2p_n - 1)}, \quad 1 \leq n \leq i. \tag{19a}$$

Migration alters this to

$$p_1^{**} = [1 - (m/2)]p_1^* + (m/2)p_2^*,$$
$$p_n^{**} = (1 - m)p_n^* + (m/2)(p_{n-1}^* + p_{n+1}^*), \quad 1 < n < i, \tag{19b}$$
$$p_i^{**} = [1 - (m/2)]p_i^* + (m/2)p_{i-1}^*.$$

After population regulation, the gene frequencies in the colonies are independently distributed, the distribution of $p_n(t+1)$ begin binomial with parameter p_n^{**} and index $2N$.

To keep the number of parameters to a minimum and to model a common biological situation, we wish to choose i so large that there are no end effects. Therefore, we must use only such initial conditions that even with $i = \infty$ A_1 will be lost with probability one in a finite time. If A_1 is initially represented in an infinite number of colonies, however, then the probability of its being represented in a finite number of colonies in generation $t(< \infty)$ is zero, so its probability of extinction in a finite time is also zero. Hence, we suppose that A_1 initially occurs in a finite number of demes in the infinite chain we wish to approximate.

Let us posit that the two parts of the habitat corresponding to the two environments in (18) become connected at $t = 0$. If both subpopulations are at equilib-

rium at that time, then, regardless of the previous migration pattern in the two subpopulations, we have

$$p_n(0) = \begin{cases} 1, & 1 \le n \le j, \\ 0, & j+1 \le n \le i. \end{cases} \tag{20}$$

Since there appears to be no other natural initial condition for our problem, we shall always posit (20). With the restriction that $p_n(0) > 0$ only in a finite number of demes even if $i = \infty$, we expect our results to be a good qualitative guide for all initial conditions.

Having no rigorous proof, we first check numerically that the distribution of the extinction time T is, indeed, independent of i for sufficiently large i. We choose $s = 0.05$, $m = 0.4$, $N = 8$, and $j = 1$, and display the results of the simulation in Table 2. For various values of i, we present the median extinction time $\hat{T}$, the mean extinction time $\overline{T}$, the standard error $\sigma_{\overline{T}}$ of $\overline{T}$, and the coefficient of variation $\sigma_T/\overline{T}$ of the distribution of T. All times are in generations. Each line is based on 1000 extinctions; in addition, 15 fixations occurred with $i = 5$, 1 with $i = 9$, and none for higher values of i. Since extinction is certain with $i = \infty$, fixations indicate the presence of end effects. Nevertheless, Table 2 shows no statistically significant deviations of the extinction time distribution even for small i.

For each line of Table 2, indeed for all the simulations in this section, a positive skewness was found. This, and the fact that the mean always exceeds the median, are consequences of the expected long tail of the distribution of the extinction time. In Figure 12 we plot the probability density $h(T)$ corresponding to the last line of Table 2. We used unequal interval sizes for the underlying histogram in order to obtain roughly the same relative error in different parts of the distribution; the error bars are two standard deviations long. The median is approximately normally distributed with mean $\hat{T} = 44$ and standard deviation $[2h(\hat{T})\sqrt{1000}]^{-1} \approx 1.39$ (CRAMÉR 1946, p. 369). Since the median is much less influenced by extremely long extinction times than the mean, it is hardly surprising that it has a smaller standard error. The 25th and 75th percentile points are at $T = 25.5 = \hat{T} - 0.365\sigma_T$ and $T = 78.0 = \hat{T} + 0.671\sigma_T$, respectively; if the distribution of T were Gaussian, these points would be at $\hat{T} \pm 0.675\sigma_T$.

TABLE 2

Extinction times for various habitat lengths

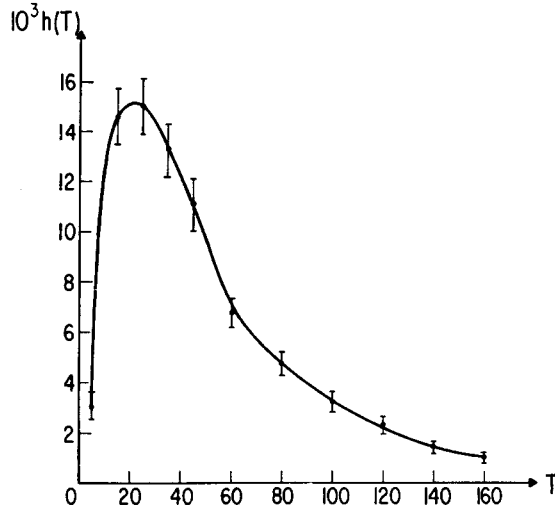| i	$\hat{T}$	$\overline{T}$	$\sigma_{\overline{T}}$	$\sigma_T/\overline{T}$
5	46	63.11	1.74	0.870
9	46	63.11	1.75	0.875
17	50	63.51	1.60	0.795
33	45	59.65	1.50	0.795
65	44	60.07	1.60	0.843

FIGURE 12.—The probability density of the extinction time for $s = 0.05$, $m = 0.4$, $N = 8$, $j = 1$ and $i = 65$.

We display the results of our simulations in Table 3. Since the simulations are extremely expensive, and the errors decline only as the reciprocal of the square root of the number of extinctions, 20 extinctions were observed for every parameter set. The parameters s, m, N, j and i specify the Markov chain. The The characteristic length for the deterministic cline in the diffusion limit is $l = \sqrt{m/(2s)}$ (SLATKIN 1973).

If i is sufficiently large to eliminate end effects, and the diffusion approximation applies, then the distribution of the extinction time, measured in natural units of $1/s$ generations, depends only on two, rather than four (i.e., s, m, N and j) parameters (NAGYLAKI 1978a). This important reduction of the number of parameters by two is possible because the units of length and time are arbitrary in the diffusion limit. The two dimensionless parameter combinations of which the scaled extinction time $\tau = sT$ is a function are most informatively chosen as $k = j/l$ and $\beta = 2\sqrt{2msN}$. The deterministic parameter k represents the size of the environmental pocket in natural units (NAGYLAKI 1975); as discussed before, β is an index of the amount of random drift. In the last four columns of Table 3 we present the median $\hat{\tau}$, the mean $\bar{\tau}$, the standard error $\sigma_{\bar{\tau}}$ of $\bar{\tau}$, and the coefficient of variation $\sigma_\tau/\bar{\tau}$.

To have no end effects and assure extinction, we require at least that $i \gg j$ and $i \gg l$. As Figure 1 indicates, however, for moderate or strong random drift ($\beta \leq 1$), the characteristic length of spatial variation appreciably exceeds l; then we enforce $i \gg l$ even more stringently. There are 560 extinctions in Table 3; no fixations were observed.

The diffusion approximation holds if $s \ll 1$, $m \ll 1$, $N \gg 1$, and $j \gg 1$ (NAGYLAKI 1978b). To test that the results depend only on k and β, we ran different combinations of s, m, N and j that yield the same k and β (see lines 2 and

TABLE 3

Extinction times for various parameters

s	m	N	j	i	l	k	β	$\hat{\tau}$	$\bar{\tau}$	$\sigma_{\bar{\tau}}$	$\sigma_{\tau}/\bar{\tau}$
0.00075	0.1	19	4	100	8.16	0.490	0.465	0.80	1.88	0.838	1.987
0.00075	0.1	38	4	60	8.16	0.490	0.931	1.61	2.68	0.698	1.165
0.0015	0.2	19	4	60	8.16	0.490	0.931	1.62	2.30	0.451	0.877
0.0015	0.2	38	4	36	8.16	0.490	1.862	2.09	2.21	0.316	0.640
0.0015	0.2	75	4	36	8.16	0.490	3.674	2.51	3.56	0.624	0.784
0.00075	0.4	150	8	72	16.33	0.490	7.348	3.98	5.36	0.814	0.680
0.003	0.4	75	4	36	8.16	0.490	7.348	4.03	4.41	0.546	0.553
0.0015	0.2	150	4	36	8.16	0.490	7.348	2.60	3.16	0.355	0.503
0.0015	0.2	300	4	36	8.16	0.490	14.697	4.92	5.36	0.416	0.347
0.0015	0.05	19	4	100	4.08	0.980	0.465	2.00	3.15	1.026	1.454
0.00075	0.025	38	4	100	4.08	0.980	0.465	2.04	4.69	1.570	1.498
0.0015	0.05	38	4	100	4.08	0.980	0.931	3.82	5.16	1.025	0.888
0.003	0.1	19	4	100	4.08	0.980	0.931	2.57	4.26	1.184	1.242
0.00075	0.1	38	8	100	8.16	0.980	0.931	4.17	5.98	1.112	0.831
0.003	0.1	38	4	60	4.08	0.980	1.862	5.27	6.38	1.249	0.875
0.003	0.1	75	4	36	4.08	0.980	3.674	5.00	6.88	1.097	0.713
0.003	0.1	150	4	36	4.08	0.980	7.348	9.21	12.42	2.153	0.775
0.006	0.2	150	4	36	4.08	0.980	14.697	21.39	25.93	3.508	0.605
0.003	0.025	19	4	100	2.04	1.960	0.465	5.50	8.78	1.922	0.979
0.006	0.05	19	4	100	2.04	1.960	0.931	6.79	8.95	2.507	1.252
0.006	0.05	38	4	60	2.04	1.960	1.862	22.00	27.03	4.756	0.787
0.006	0.05	75	4	36	2.04	1.960	3.674	29.88	53.25	12.17	1.022
0.012	0.1	75	4	36	2.04	1.960	7.348	727.5	840.2	160.0	0.851
0.006	0.0125	19	4	36	1.02	3.919	0.465	24.44	35.53	9.061	1.141
0.012	0.025	19	4	36	1.02	3.919	0.931	88.20	105.3	21.59	0.918
0.024	0.05	19	4	36	1.02	3.919	1.862	524.6	766.1	153.9	0.898
0.012	0.00625	19	4	36	0.51	7.838	0.465	234.9	270.0	44.44	0.736
0.024	0.0125	19	4	36	0.51	7.838	0.931	2231	2777	514.7	0.829

3, 6 to 8, etc.). For each (k,β) for which more than one set (s,m,N,j) was used, we divide the extinction times for each (s,m,N,j) into three groups, $\tau < \tau_1$, $\tau_1 < \tau < \tau_2$ and $\tau > \tau_2$, by choosing τ_1 and τ_2, depending only on k and β, so that the frequencies in the three classes are very roughly equal. This procedure yields a contingency table for each (k,β) (*e.g.*, a 2×3 table for lines 2 and 3), so a χ^2-test will tell us whether the samples of 20 extinction times for the same (k,β) and different (s,m,N,j) may come from the same distribution. No χ^2 was even close to significance. The total was $\chi^2_{12} = 8.35$, corresponding to a probability of 0.76 of a greater deviation from the diffusion approximation. Since we were forced to use some high values of m, and usually took $j = 4$ in order to limit computation costs, the agreement is surprisingly good.

We also checked the diffusion approximation in another way. In the deterministic diffusion limit, A_1 is preserved if and only if $k > \tan^{-1} 1 = 0.785$ (NAGYLAKI 1975). Choose $s = 0.003$, $j = 4$ and $i = 36$; then $m = 0.4$, 0.1, 0.025 and 0.00625 correspond to $k = 0.490$, 0.980, 1.960 and 3.919, respectively.

TABLE 4

Pooled extinction times for various parameters

$k \backslash \beta$	0.465	0.931	1.862	3.674	7.348	14.697
	1.88	2.49	2.21	3.56	4.31	5.36
0.490	0.838	0.336	0.316	0.624	0.361	0.416
	1.987	1.044	0.640	0.784	0.650	0.347
	3.92	5.13	6.38	6.88	12.42	25.93
0.980	0.762	0.635	1.249	1.097	2.153	3.508
	1.508	0.959	0.875	0.713	0.775	0.605
	8.78	8.95	27.03	53.25	840.2	—
1.960	1.922	2.507	4.756	12.17	160.0	—
	0.979	1.252	0.787	1.022	0.851	—
	35.53	105.3	766.1	—	—	—
3.919	9.061	21.59	153.9	—	—	—
	1.41	0.918	0.898	—	—	—
	270.0	2777	—	—	—	—
7.838	44.44	514.7	—	—	—	—
	0.736	0.829	—	—	—	—

Iterating the deterministic system (19), (20), we found convergence of the gene frequency to zero in the subcritical case, $k = 0.490$, and convergence to a cline in the three supercritical cases, $k > 0.785$. The maximum gene frequencies at equilibrium for $k = 0.980$, 1.960 and 3.919 are 0.317, 0.810 and 0.973. The diffusion values, read off roughly from Figure 3 of NAGYLAKI (1975), are 0.31, 0.82 and 0.98, in much closer agreement than one might expect.

Having convinced ourselves of the validity of the diffusion approximation, we may pool extinction times, τ, for the same (k,β). In Table 4 we display for each (k,β) the mean $\bar{\tau}$ (top line for each k), the standard error $\sigma_{\bar{\tau}}$ (middle line) and the coefficient of variation $\sigma_\tau / \bar{\tau}$ (bottom line). We omit the median $\hat{\tau}$ because its relation to the mean does not vary greatly. (All the distributions in Table 3 were skewed up.)

Let us dispose first of the coefficients of variation. These are all between 0.347 and 1.987, and have large standard errors. Table 4 suggests no obvious dependence on k, though the exact values are surely functions of k. We intuitively expect that the coefficients of variation should decrease as random drift becomes weaker. Therefore, we wish to examine Table 4 for decrease as β increases with k fixed. We can apply the exact Wilcoxon signed-ranks test (BROWNLEE 1965, pp. 258–260) in the following manner.

Each row of n coefficients of variation supplies $[n/2]$ independent adjacent pairs, and the pairs in different rows are independent. We obtain signed differences to which the test applies by subtracting the right from the left member of the pair. A low total rank sum for the negative differences indicates a pattern of decrease along rows. When n is odd, there are different ways of forming pairs.

Functional dependence should be clearest toward the right, where β varies more rapidly. Hence, with n odd, it is reasonable to ignore the first element in each row; *e.g.*, for $k = 1.960$ we have the differences $1.252 - 0.787 = 0.465$ and $1.022 - 0.851 = 0.171$. This procedure yields a rank sum of 5 for Table 4; the probability of a rank sum less than or equal to 5 is $10/2^{10} = 0.010$. If we disregard our *a priori* argument, however, and neglect the last element in rows with odd n, we lose significance: the rank sum is now 16, and the probability level rises to $141/2^{10} = 0.138$. This occurs entirely due to fluctuations for $k = 1.960$. Decrease can be established independently for $k = 0.490$ and $k = 0.980$. From Appendix Table 2 of KENDALL (1962), we find that the probabilities of Spearman rank correlations as low or lower than those calculated for these rows are 0.029 and 0.0083, respectively. We conclude that the coefficient of variation decreases with increasing β.

A glance at the mean extinction times and their errors in Table 4 immediately reveals that the means increase with both k and β. In the subcritical case ($k = 0.490$), the increase with β is extremely slow. Even without appeal to normality and the estimated errors, however, the increase is statistically significant: the probability of obtaining a Spearman rank correlation as high or higher by chance is 0.0083. The slow increase is expected on intuitive grounds: very roughly, the process is analogous to the elimination of a deleterious allele by selection and random drift. Approximating a result of NEI (1971), we can easily see that in this simpler case $\tau \sim \ln N$ as $N \to \infty$ with s fixed, leading us to conjecture that $\bar{\tau}$ grows like $\ln \beta$ as $\beta \to \infty$ with k fixed at a subcritical value.

The increase with β is faster in the supercritical case, becoming more rapid for higher values of k. The dependence on the deterministic parameter k is even stronger than on the relative random drift parameter β. For higher values of β, very long mean extinction times are reached for lower values of k.

We conclude that, if a cline exists deterministically, mean extinction times will be very long in almost all cases. This is particularly so because, as discussed in Section I, in most natural populations β will greatly exceed unity unless selection is extremely weak. Thus, spatial diversity is extremely effective in maintaining polymorphism: random drift generally causes gene frequency fluctuation, rather than loss of genetic variability. The model will usually cease to apply long before extinction.

Although the cost of extending our simulations to higher k and β is prohibitive, we can support our assertions by analogy. Conceptually, our problem resembles that of an island of N individuals on which A_1 is favored, the fitness pattern being identical to the one in the environmental pocket and all immigrants being A_2A_2. This island model has the same qualitative behavior as the full cline problem in the deterministic case (NAGYLAKI 1975). Our island model, in turn, is equivalent to selection and irreversible mutation in an isolated panmictic population.

LI and NEI (1977) have evaluated the expected extinction times in the latter situation by numerical integration. In their notation, $S = 4Ns$ (we retain our convention, $s > 0$, for consistency) and $M = 4Nm$. The dependence of k and β

TABLE 5

Mean extinction times of Li *and* Nei *for the island model*

$k \backslash \beta$	1.41	2.24	3.16	4.47	7.07	10.0	14.1	22.4	31.6
1.41	5.6	—	—	—	—	—	86	—	—
2.24	—	50	—	—	—	—	—	3.1×10^{10}	—
3.16	—	—	2800	—	—	—	—	—	1.2×10^{29}
4.47	—	—	—	2.6×10^7	—	—	—	—	—
7.07	—	—	—	—	1.1×10^{20}	—	—	—	—
10.0	—	—	—	—	—	2.5×10^{41}	—	—	—
14.1	3.4×10^9	—	—	—	—	—	—	—	—
22.4	—	3.4×10^{22}	—	—	—	—	—	—	—
31.6	—	—	1.6×10^{44}	—	—	—	—	—	—

on S and M is obvious: $k = \sqrt{S/M}$ and $\beta = \sqrt{SM}$. Since the island and cline models are analogous, but not identical, these choices may be multiplied by numerical factors of order unity. With our convention, the critical value of k is one (HALDANE 1930; NAGYLAKI 1975). In their Table 3, Li and Nei (1977) present the mean extinction times in units of $4N$ generations; to obtain $\bar{\tau}$ for their model, we must evidently multiply their results by S. We display their values for large k and β in Table 5. [Their results for small k and β are similar to ours. The third entry in their first row should be 0.61, not 0.41 (W.-H. Li, private communication).]

Table 5 agrees with all the observations based on Table 4. The rapid increase of $\bar{\tau}$ to astronomical values as either k or β significantly exceed one (for $k > 1$) is manifest.

DISCUSSION

The diffusion approximation investigated numerically in Section I and analytically in Section II neglects all third central moments of the gene frequency distribution (NAGYLAKI 1978a,b). Nevertheless, all the results are entirely reasonable. The problem of establishing the absence of pathologies mathematically (NAGYLAKI 1978a,b) and examining the accuracy of the diffusion approximation remains open.

In Section III, we employed the diffusion approximation only to reduce the number of parameters from four to two, facilitating thorough numerical analysis and straightforward biological interpretation. Various tests indicate that the diffusion approximation is more than adequate for this purpose.

The simulations of Section III show that in the single-locus diallelic model, provided the allele favored in the environmental pocket would be maintained in the population deterministically, the extinction time in a natural population due to random drift would almost invariably be extremely long. Other evolutionary forces, such as demographic or environmental changes, would supervene

long before extinction. Hence, at the level of the classical model, random drift is reflected in stochastic fluctuation—usually small—of gene frequencies. When spatial diversity acts to preserve genetic variability, it does so too effectively to permit its loss sufficiently fast to matter.

With many alleles at a single locus, low gene frequencies will be common, and one expects reduction of the number of alleles to occur more rapidly than complete loss of polymorphism. But we have no results even for the deterministic theory with multiple alleles.

The studies presented here suggest that much greater genetic complexity is required for the operation of WRIGHT's (1977) shifting balance theory of evolution. As WRIGHT has insisted, control by many loci with strong epistasis is certainly a prerequisite; pleiotropy may also be necessary.

LITERATURE CITED

BROWNLEE, K. A., 1965 *Statistical Theory and Methodology*, second edition. John Wiley & Sons, New York.

CRAMÉR, H., 1946 *Mathematical Methods of Statistics*. Princeton University Press, Princeton, N.J.

FELSENSTEIN, J., 1975 Genetic drift in clines which are maintained by migration and natural selection. Genetics **81**: 191–207.

FIEDLER, M., 1969 Numerical solution of algebraic and transcendental equations. pp. 1168–1182. In: *Survey of Applicable Mathematics*. Edited by K. REKTORYS. The M.I.T. Press, Cambridge, Mass.

FISHER, R. A., 1950 Gene frequencies in a cline determined by selection and diffusion. Biometrics **6**: 353–361.

FLEMING, W. H., 1975 A selection-migration model in population genetics. J. Math. Biol. **2**: 219–233.

HALDANE, J. B. S., 1930 A mathematical theory of natural and artificial selection. Part VI. Isolation. Proc. Camb. Phil. Soc. **26**: 220–230. ——, 1948 The theory of a cline. J. Genet. **48**: 277–284.

KENDALL, M. G., 1962 *Rank Correlation Methods*, third edition. Hafner, New York.

LI, W.-H., and M. NEI, 1977 Persistence of common alleles in two related populations or species. Genetics **86**: 901–914.

LUSKIN, M. and T. NAGYLAKI, 1979 Numerical analysis of weak random drift in a cline. Genetics **92**: 297–303.

MAY, R. M., J. A. ENDLER and R. E. McMURTRIE, 1975 Gene frequency clines in the presence of selection opposed by gene flow. Amer. Nat. **109**: 659–676.

NAGYLAKI, T., 1975 Conditions for the existence of clines. Genetics **80**: 595–605. ——, 1978a Random genetic drift in a cline. Proc. Natl. Acad. Sci. U.S. **75**: 423–426. ——, 1978b A diffusion model for geographically structured populations. J. Math. Biol. **6**: 375–382. ——, 1978c The geographical structure of populations. pp. 588–624. In: *Studies in Mathematics*, Vol. 16: *Studies in Mathematical Biology*. Edited by S. A. LEVIN. The Mathematical Association of America, Washington. ——, 1978d Clines with asymmetric migration. Genetics **88**: 813–827.

NEI, M., 1971 Extinction time of deleterious mutant genes in large populations. Theoret. Pop. Biol. **2**: 419–425.

SLATKIN, M., 1973 Gene flow and selection in a cline. Genetics **75**: 733–756.

SLATKIN, M. and T. MARUYAMA, 1975 Genetic drift in a cline. Genetics **81**: 209–222.

VILHELM, V., 1969 Arithmetic and algebra. pp. 39–106. In: *Survey of Applicable Mathematics.* Edited by K. REKTORYS. The M.I.T. Press, Cambridge, Mass.

WRIGHT, S., 1977 *Evolution and the Genetics of Populations,* Vol. III. The University of Chicago Press, Chicago. ——, 1978 *Ibid.* Vol. IV.

Corresponding editor: W. J. EWENS