SOME PROBLEMS OF STOCHASTIC PROCESSES IN GENETICS[1]

Motoo Kimura[2]

Department of Genetics, University of Wisconsin

Summary. In genetics, stochastic processes arise at all levels of organization ranging from subunits of the gene to natural populations. Types of stochastic processes involved are also diverse. In the present paper, the following five topics have been selected for mathematical discussion and new results are presented: (1) Random assortment of subunits of a gene. (2) Senescence in paramecium due to random assortment of chromosomes. (3) Process of natural selection in a finite population (interaction between selection and random genetic drift). (4) Chance of fixation of mutant genes. (5) Population structure and evolution. Finally it is pointed out that new mathematical techniques will be needed for a satisfactory treatment of Wright's theory of evolution.

"Elles n'auroient dû leur premiere origine qu'à quelques productions fortuites, dans lesquelles les parties élémentaires n'auroient pas retenu l'ordre qu'elles tenoient dans les animaux peres & meres: chaque degré d'erreur auroit fait une nouvelle espece:. . .

Des moyens différents des moyens ordinaires que la Nature emploie pour la production des animaux, loin d'être des objections contre ce systême, lui sont indifférents, ou lui seroient plutôt favorables."[3]

—Maupertuis (Oeuvres, 1756)

1. Introduction. These words, written two centuries ago, foreshadow the stochastic nature of genetic and evolutionary processes. Actually, stochastic processes are found in all levels of organization with which genetics is concerned, in the gene, the cell, the organism, and the population.

The types of stochastic processes involved are also diverse. Of special importance is the Markov process, which Kolmogorov [1] called stochastic definite; the exact treatment of regular systems of inbreeding is a typical example of a

finite Markov chain, though workers in this field have seldom used such a terminology [2] [3]. The fate of an individual mutant gene appearing in a population may best be studied by the theory of branching processes. The probability distribution of gene frequencies in natural populations is important in the mathematical theory of evolution developed by Fisher and Wright. It contains many difficult problems of continuous stochastic processes in which the Kolmogorov equation plays a fundamental role [4] [5] [6] [7].

In the present paper a few topics will be selected from various levels of organization ranging from subunits of the gene to natural populations.

2. Random assortment of subunits in chromosome division. The idea that each gene is composed of a number of subunits is a natural one, since analogous situations are familiar in physics and chemistry. If the subunits are of two or more different kinds, they will be sorted out in the process of chromosome division.

, In fact, such a model was proposed more than three decades ago to explain the high mutability of the so-called mutable genes. Unfortunately, precise experimental results in a few higher organisms apparently contradicted this model [8]. Later, however, as investigations of the finer structure of the chromosome have developed, a multiple-strand structure has been revealed and this encouraged the formulation of the same type of model again. Matsuura and Suto [9], in order to explain certain irregular segregations in maize, assumed that each chromosome contains 8 strands and that mutation may affect any one of the 8 gene replicates. A similar model was used by Auerbach [10] to explain the occurrence of mosaics in the offspring of *Drosophila* males treated with mustard gas. More recently, Friedrich-Freksa and Kaudewitz [11] carried out an interesting experiment with *Amoeba proteus* treated with radioactive P^{32}, in which they assumed that sorting-out of the radiation-damaged strands or subunits causes death to the organism in later generations.

Let us consider a model in which each chromosome consists of n subunits and suppose that a mutation has occurred in one of them. The subunits duplicate to produce $2n$ which separate at random into two groups of n subunits to form the daughter chromosomes. Thus the total number of subunits per chromosome is kept constant, but the number of the mutant subunits may change from generation to generation due to random segregation. We follow a single line of descent obtained by selecting randomly one of the pair of daughter chromosomes in each generation. We shall designate by $E_i (i = 0, 1, \cdots, n)$ the state in which a given chromosome contains exactly i mutant subunits. Let $a_i^{(t)}$ be the probability that the chromosome is in the state E_i at the tth generation (assuming that the mutation occurred at $t = 0$). In the present model, the transition probabilities $p_{j|i} \equiv \Pr\{E_j \leftarrow E_i\} (i, j = 0, 1, 2, \cdots, n)$ are given by

$$(2.1) \qquad p_{j|i} = \binom{2i}{j} \binom{2n - 2i}{n - j} \Big/ \binom{2n}{n}.$$

Thus the probability chain will be expressed by

(2.2) $$a^{(t)} = Pa^{(t-1)},$$

in which $a^{(t)}$ is a column vector whose ith element is $a_i^{(t)}$ and P is an $n \times n$ matrix whose element in the ith row, jth column is $p_{j|i}$. Obviously E_0 and E_n are absorbing barriers and the remaining states $E_1, \cdots, E_{n-1}$ are all transient. Eigenvalues of the matrix P, which satisfy $| P - \lambda I | = 0$, are

(2.3) $$\lambda_r = 2^r \binom{2n-r}{n-r} \bigg/ \binom{2n}{n} \qquad (r = 0, 1, \cdots, n).$$

This can be shown by following a procedure similar to that which Feller [5] gave in an appendix of his paper (p. 244), noting that a non-trivial set of y_j's which satisfy

$$\sum_{j=0}^{n} p_{j|i} y_j = \lambda_r y_i$$

can be written in the form

$$y_j = \sum_{\nu=0}^{r} c_\nu j^{(\nu)} \qquad (c_\nu \text{ is a constant})$$

and

$$\sum_{j=0}^{n} p_{j|i} j^{(\nu)} = (2i)^{(\nu)} \binom{2n-\nu}{n-\nu} \bigg/ \binom{2n}{n},$$

where $j^{(\nu)} = j(j-1) \cdots (j - \nu + 1)$ is the factorial of degree ν.

Though the general expressions for eigenvectors of P and its transpose P' corresponding to these eigenvalues do not seem easily obtainable, we can obtain numerical results for small values of n. Thus we can construct formulae giving probabilities of various states at a given generation.

I have worked out the cases of $n = 2, 3, 4, 5, 6, 8$, and 16, the details of which will be published elsewhere.

We are particularly interested in the probability $(d^{(t)})$ with which a mutant chromosome (i.e., a chromosome containing only mutant subunits) first appears by the sorting out process in a given generation $t(t = 1, 2, 3, \cdots)$. This is obtained as $d_n^{(t)} = a_n^{(t)} - a_n^{(t-1)}$ for the case of n subunits. For $n = 2, 4, 8$ and 16, we have

$$d_2^{(t)} = \frac{1}{6} \left(\frac{2}{3}\right)^{t-1},$$

$$d_4^{(t)} = \frac{1}{3080} \left\{ 195 \left(\frac{6}{7}\right)^{t-1} - 330 \left(\frac{4}{7}\right)^{t-1} + 135 \left(\frac{16}{70}\right)^{t-1} \right\},$$

$$d_8^{(t)} = \frac{1}{12870} \left\{ (242.0) \left(\frac{14}{15}\right)^{t-1} - (709.1) \left(\frac{12}{15}\right)^{t-1} \right.$$

$$\left. + (904.5) \left(\frac{120}{195}\right)^{t-1} - (686.9) \left(\frac{80}{195}\right)^{t-1} + \cdots \right\},$$

TABLE 1

n	t_{max}	d_{max} (%)
2	1	16.667
4	5	2.287
8	15	0.481
16	35	0.111

$$d_{16}^{(t)} = 0.0051626 \left(\frac{30}{31}\right)^{t-1} - 0.0193172 \left(\frac{28}{31}\right)^{t-1} + 0.0358431 \left(\frac{728}{899}\right)^{t-1}$$

$$- 0.0457009 \left(\frac{624}{899}\right)^{t-1} + 0.0451643 \left(\frac{13728}{24273}\right)^{t-1} + \cdots .$$

If the mutant chromosome causes death to the Ameba as in the model of Friedrich-Freksa et al. [11], $d^{(t)}$ gives the probability of a death at the tth generation.

It may be expected on intuitive grounds that the larger the number of sub-units, the later will be the appearance of the mutant chromosome. One way of expressing this tendency is to calculate the value of t which maximizes $d_n^{(t)}$. Table 1 shows these values and corresponding values of d.

Next we shall consider the situation in which n is very large. The proportion of mutant subunits $x = i/n$ ($0 \leq x \leq 1$) may be treated as a continuous variable with good approximation. Let $\phi(x, t)$ be the probability density of x at time t measured in generations. If δx is the amount of change in x per generation,

$$E(\delta x) = 0, \qquad E(\delta x)^2 = x(1 - x)/(2n - 1) \qquad \text{and, for } k \geq 3,$$

$E(\delta x)^k$ is $o(1/n)$. Therefore we can use the following differential equation to obtain $\phi(x, t)$ (see [16]):

$$(2.4) \qquad \frac{\partial \phi(x, t)}{\partial t} = \frac{1}{2(2n - 1)} \frac{\partial^2}{\partial x^2} \{x(1 - x)\phi(x, t)\} \qquad (0 < x < 1),$$

with the initial condition

$$(2.5) \qquad \phi(x, 0) = \delta \left(x - \frac{1}{n}\right),$$

where δ is the Dirac function. The singular equation (2.4) is equivalent to the one describing the process of random genetic drift in natural populations if we put $2n - 1 = 2N$, N being the effective population size. The complete solution of this equation has been worked out (see Section 4). The points $x = 0$ and $x = 1$ act as absorbing barriers and the rate of fixation at $x = 1$ is given by $\phi(1, t)/4N$ which reduces to

$$d_n^{(t)} = \frac{1}{4Nn} \sum_{i=1}^{\infty} (-1)^{i-1} i(i + 1)(2i + 1) e^{-i(i+1) t/4N}.$$

The value of t giving $d_n^{(t)}$ maximum is obtained by solving

$$1 - 15e^{-2r} + 84e^{-5r} - 300e^{-9r} + 825e^{-14r} - \cdots = 0,$$

where $r = t/2N$. The required root of this transcendental equation can be obtained numerically; $r = 1.2940 \cdots$.

From this we can derive the following remarkable asymptotic relations $(n \to \infty)$:

$$(2.6) \qquad \begin{cases} t_{\max} \sim 2.59n, \\[2mm] d_{\max} \sim 1.08/(4n^2). \end{cases}$$

Namely, $t_{\max}$ will be proportional to the number of subunits and $d_{\max}$ will be inversely proportional to the square of that number.

Finally it is desirable to consider the cases in which the initial number of mutant subunits is more than one, that is, some number, say np, where p is the initial proportion of the subunits $(0 < p < 1)$. The limiting probability of absorption at n, starting at i, is the ith component of an eigenvector corresponding to $\lambda = 1$, and from the third formula after equation (2.3) it follows that $p = i/n$ is the required probability. The total frequency of mutant chromosomes in the tth generation will be expressed in the form:

$$(2.7) \qquad f_n^{(t)}(p) = \sum_{\tau=0}^{t} d_n^{(\tau)} = p + \sum_{i=2}^{n} C_i(p, n)\lambda_i^t,$$

where the $C_i(p, n)$'s are functions of p and n but not of t. If the mutant chromosome changes the fitness of its possesser, this type of formula must be applied with caution. Generally $d_n^{(t)}$ should be used as a basis of comparing expectation with observed results.

For very large n, (2.7) should approach

$$(2.8) \qquad p + \sum_{i=1}^{\infty} (2i + 1)pq(-1)^i F(1 - i, i + 2, 2, p)e^{-[i(i+1)t/4N]},$$

where $q = 1 - p$ and F designates the hypergeometric function (see (5.3) in Section 5).

The experimental data of Friedrich-Freksa [11] appear to agree with the model for $n = 16$.

3. Senescence in Paramecium. It has been known to biologists for a long time that if cultures of the protozoon, *Paramecium*, are kept under exclusive asexual reproduction, they lose vigor and eventually die. This phenomenon is known as senescence or aging of paramecium and in fact is one of the old problems in biology. Recently Dr. T. M. Sonneborn has made extensive studies of this phenomenon and discussed a hypothesis that the aging is due to an accumulation of chromosome aberrations in the macronucleus (cf. e.g. Sonneborn and Schneller [12]). Following the suggestion of Dr. J. Lederberg, I have tried to work out the logical consequences of the stochastic model involved.

The macronucleus is considered to be polyploid, consisting of, say, m chromosome sets each with n chromosomes. As in a polyploid nucleus of higher plants, the various chromosomes are mixed at random inside the nucleus[4]. If we designate the chromosomes of a set by successive letters A, B, C, $\cdots$, N and designate sets by subscripts, then the normal constituent of the nucleus will be expressed in the form:

$$m \begin{cases} \overbrace{\begin{matrix} A_1 & B_1 & C_1 \cdots N_1 \\ A_2 & B_2 & C_2 \cdots N_2 \\ \vdots & \vdots & \vdots \qquad \vdots \\ A_m & B_m & C_m \cdots N_m \end{matrix}}^{n} \end{cases}$$

We designate the total number of chromosomes by $M(=mn)$.[5]

On this model we assume that at the division of the macronucleus, each chromosome duplicates itself followed by the random distribution of chromosomes into two groups of equal number to form the daughter macronuclei. The death by aging is assumed to occur whenever chromosomes of any one type are lost entirely from the nucleus by chance. Various states of the nucleus will be expressed as n dimensional vectors. Here we have a hierarchical structure of absorbing barriers, and a direct attack on the problem may seem extremely difficult. However, because of the symmetry of the model, we can find an easier approach. Suppose we start from an individual with normal macronucleus, and each generation takes one of the daughters to continue the lineage. Our purpose is to calculate the probability that all the n chromosome types coexist in the individual at the tth generation. Since the process of loss of any type of chromosome is irreversible, we can treat the problem as if all possible chromosome constituents are viable and then remove unsuitable parts afterwards.

Let us fix our attention on the tth generation. We designate by P_1 the probability that all the chromosomes except those of one specific type have been lost by that time, by P_2 the probability that all but 2 specific types have been lost and that these two coexist. Generally P_i will be defined in a similar way. Since we can classify n chromosome types into two alternative groups like A vs. non A, A or B vs. neither A nor B, etc.,

$$\phi_1 = P_1,$$

$$\phi_2 = \binom{2}{1} P_1 + P_2,$$

$$\phi_3 = \binom{3}{1} P_1 + \binom{3}{2} P_2 + \binom{3}{3} P_3, \text{ etc.,}$$

---

[4] This model is essentially different from the one considered by Kimball and Householder [13] to explain the delayed lethal effect of radiation.

[5] n in this section has a different meaning from that of the previous section. Generally the same symbol in different sections may not have the same meaning.

up to

$$\phi_{n-1} = \binom{n-1}{1} P_1 + \binom{n-1}{2} P_2 + \cdots + \binom{n-1}{n-1} P_{n-1},$$

where ϕ_i represents $f_M^{(t)}(i/n)$ in (2.7) $(i = 1, \cdots, n-1)$. For example, ϕ_2 is the probability that all the chromosomes except A or B or both have been lost and this is a sum of the probabilities that all but A have been lost (P_1), that all but B have been lost (P_1), and that A and B coexist but all others have been lost (P_2). It is convenient to consider the above relations as a linear transformation of P_i's into ϕ_i's with an $(n-1) \times (n-1)$ matrix whose element in the ith row and jth column is $\binom{i}{j}$. The inverse transformation can be shown to have a matrix whose element in the ith row and jth column is $(-1)^{i+j} \binom{i}{j}$. Let $\Omega_t (\equiv P_n)$ be the probability that all the chromosome types coexist in the macronucleus at the tth generation. Since $\sum_{i=1}^{n} \binom{n}{i} P_i = 1$,

$$\Omega_t = 1 - \sum_{i=1}^{n-1} \binom{n}{i} P_i = 1 - \sum_{j=1}^{n-1} (-1)^{n-1+j} \binom{n}{j} \phi_j.$$

Substituting $\phi_j = f_M^{(t)}(j/n)$ from (2.7) and noting that

$$\sum_{j=1}^{n-1} (-1)^{n-1+j} \binom{n}{j} (j/n) = 1,$$

we have

$$\Omega_t = \sum_{j=1}^{n-1} \sum_{i=2}^{M} (-1)^{n+j} \binom{n}{j} C_i \left(\frac{j}{n}, M\right) \lambda_i^t,$$

where

$$\lambda_i = 2^i \binom{2M-i}{M-i} \Big/ \binom{2M}{M}.$$

According to Sonneborn, the usual strains of *Paramecium* have a chromosome number of $n \doteqdot 41$, but also there are strains with $n \doteqdot 35$ and 50. There is a good reason to believe that the macronucleus is at least 100 ploid $(m = 100)$. Thus the total number of chromosomes M in the macronucleus would be of the order of 5000. This fact will enable us to use the asymptotic formula for $f_M^{(t)}(j/n)$ given in (2.8).

$$\Omega_t = \sum_{j=1}^{n-1} (-1)^{n+j} \binom{n}{j} \sum_{i=1}^{\infty} (2i + 1)(j/n)(1 - j/n)(-1)^i$$
$$\times F(1 - i, i + 2, 2, j/n) e^{-[i(i+1)t/4M]}$$

or if we put

$$\alpha_i \equiv \sum_{j=1}^{n-1} (-1)^j \binom{n}{j} (j/n)(1 - j/n) F(1 - i, i + 2, 2, j/n)$$

(3.1)
$$\Omega_t = \sum_{i=1}^{\infty} (2i + 1)(-1)^{n+i}\alpha_i\, e^{-[i(i+1)t/4M]}.$$

It can be shown that $\alpha_i = 0$ for $i < n - 1$. For $i \geqq n - 1$

$$\alpha_{n-1} = -\frac{(2n - 2)!}{(n - 1)!n^n}, \qquad \alpha_n = 0,$$

and in general, writing $i = (n - 1) + \nu(\nu \geqq 0)$,

$$\alpha_{n-1+\nu} = \frac{-(n - 1)!}{(n + \nu)!n^{n-1}} \sum_{\mu=0}^{\nu} (-1)^{\mu} \binom{\nu + n - 2}{\mu + n - 2} \frac{[\mu + \nu + 2(n - 1)]!}{(\mu + n - 1)!n^{\mu}} \times \mathfrak{S}_{\mu+n-1}^{n-1},$$

where $\mathfrak{S}_{\mu+n-1}^{n-1}$ is Stirling's number of the second kind defined by

$$j^n = \sum_{\nu=1}^{n} \mathfrak{S}_n^{\nu}\, j^{(\nu)}$$

(see [14]). Examination of the absolute values of $\alpha_{n-1+\nu}$ at $\nu = 1, 2, \cdots$ (small values of ν) suggests that they are at most of the order of $1/n$ relative to that of α_{n-1}. This enables us to write down the following asymptotic formula for small Ω_t:

(3.2)
$$\Omega_t \sim \frac{(2n - 1)!}{(n - 1)!} \left(\frac{1}{n}\right)^n e^{-(n-1)t/4m} \qquad\qquad (t \to \infty);$$

or applying Stirling's formula for $n!$,

(3.2′)
$$\Omega_t \sim \frac{1}{\sqrt{2}} \exp\left\{(2 \log 2 - 1)n - \left(\frac{n - 1}{4m}\right)t\right\}.$$

Formula (3.2) can also be derived by a different method, by using the multivariate Kolmogorov forward equation. To reach a given small probability of survival (Ω), the approximate number of generations required will be given by

$$\hat{t} = \frac{4m}{n - 1} (0.39n - 0.35 - 2.3 \log_{10} \Omega).$$

In the case of $m = 100, n = 41$, we have $\hat{t} = 156.4 - 23 \log_{10} \Omega$ and the generations giving 99%, 99.9% and 99.99% deaths are respectively about 202, 225, and 248 generations. This agrees reasonably well with the finding of Sonneborn that under exclusively asexual reproduction, many of the lines die before 200 fissions and almost all die before 324.

A slightly modified model was suggested to the author by Dr. J. Lederberg: After chromosomes have reduplicated themselves in the macronucleus, they are distributed into two daughter nuclei in such a way that each chromosome has an independent and equal chance of going to either daughter. This differs from the previous model in that the total number of chromosomes per cell does not remain constant. This leads to the following asymptotic formula for the probability of survival:

$$\Omega_t \sim (1 - e^{-4m/(t+4)})^n \qquad\qquad (t \to \infty).$$

For $n = 41$, $m = 100$, the number of generations for 99%, 99.9% and 99.99% deaths are respectively about 175, 210, 246, rather similar to the previous model.

Also these models allow predictions for time of death of a lineage derived from repeated regeneration from a small fraction of the macronucleus and for segregation of recessive factors, thus permitting two additional independent tests of the models by comparison with data.

4. Process of natural selection in finite population (interaction between selection and random genetic drift). From the standpoint of population genetics, the most elementary step in evolution is the change in gene frequency, especially the one due to natural selection. It may not be difficult to imagine that the process of change is not entirely deterministic, since there exist various factors which introduce an element of indeterminacy into the process, among which random sampling of gametes due to finite population size is of special interest. Let A and A' be a pair of alleles whose frequencies are respectively x and $1 - x$ in the population. In natural populations, the number of individuals is usually large and there may be overlapping of generations, so that gene frequency and time parameter (t) may be treated as continuous variables with advantage. We shall designate by $\phi(x, p; t)$ the probability density that the gene frequency lies between x and $x + dx$ at the tth generation given that the initial gene frequency is p at $t = 0$.

The simplest situation is obviously that of pure random genetic drift in which no mutation, selection, or migration is involved. The gene frequency changes randomly from generation to generation due to random sampling of gametes in reproduction. In this case if N is the number of reproducing individuals in the random mating population, ϕ satisfies the following partial differential equation [15], [16].

$$(4.1) \qquad \frac{\partial \phi}{\partial t} = \frac{1}{4N} \frac{\partial^2}{\partial x^2} \{x(1 - x)\phi\} \qquad (0 < x < 1),$$

with the initial condition

$$\phi(x, p; 0) = \delta(x - p),$$

where δ represents Dirac's delta function. Equation (4.1) is a special case of the Kolmogorov forward (or Fokker-Planck) equation, and its pertinent solution is given by

$$(4.2) \qquad \phi(x, p; t) = \sum_{i=1}^{\infty} \frac{(2i + 1)(1 - r^2)}{i(i + 1)} T_{i-1}^1 (r) T_{i-1}^1 (z) e^{-i(i+1)t/4N},$$

where $r = 1 - 2p$, $z = 1 - 2x$ and $T_{i-1}^1(r)$ is the Gegenbauer Polynomial defined by

$$T_{i-1}^1 (r) = \frac{i(i + 1)}{2} F\left(i + 2, 1 - i, 2, \frac{1 - r}{2}\right).$$

Boundaries $x = 0$ and $x = 1$ act as absorbing barriers and as the number of generations increases, the probability distribution of the classes in which A and A' coexist ("heterallelic," or unfixed, classes) approaches a definite form and decays at the constant rate of $1/2N$. The process ultimately leads to complete fixation or loss of one of the alleles.

When linear pressures (mutation, migration) are involved, the problem becomes a little more complicated. A thoroughgoing analysis of the solutions of the differential equations in this case has been made by Goldberg [17]. The present author also obtained the pertinent solution[6] by studying the law of change in the moments of the distribution [6]. Malécot [4] [18] studied interesting problems of migration and decrease of correlation with distance in the case of no selection

For the evolution of the genetic system, however, natural selection which acts on mutant genes will be of utmost importance. The simplest situation here is genic selection in which no dominance exists. Suppose gene A has selective advantage s over A', measured in Malthusian parameters [19], that is to say, the rate of geometric growth. The partial differential equation now becomes

$$(4.3) \qquad \frac{\partial \phi}{\partial t} = \frac{1}{4N} \frac{\partial^2}{\partial x^2} \{x(1 - x)\phi\} - s\frac{\partial}{\partial x} \{x(1 - x)\phi\} \qquad (0 < x < 1),$$

with the same initial condition as before. Recently this equation was used by Wright and Kerr [20] in connection with their selection experiment in very small populations. The state of steady decay of the heterallelic classes was successfully analysed by Wright. The complete solution of the above equation, which reduces to that of pure random drift for $s = 0$, is given in terms of oblate spheroidal functions studied by Stratton and others [21]:

$$\phi(x, p; t) = \sum_{k=0}^{\infty} C_k e^{-\lambda_k t + 2cx} V_{1k}^{(1)}(z),$$

where $c = Ns$ and $z = 1 - 2x$. The spheroidal function $V_{1k}^{(1)}(z)$ is expressed as a series of Gegenbauer polynomials:

$$V_{1k}^{(1)}(z) = \sideset{}{'}\sum_{n=0,1} f_n^k T'_n^1(z),$$

where f_n^k's are constants, and primed summation is over even values of n if k is even, odd values of n if k is odd. For details of the solution see [7]. The boundaries $x = 0$ and $x = 1$ act as absorbing barriers as in the preceding cases and the gene A will ultimately be fixed in the population or completely lost from it. The probability of fixation is larger, the larger the value of s.

---

[6] Strictly speaking, the existing solution which treats boundaries as reflecting barriers is not entirely satisfactory, because for small populations boundaries should act as elastic barriers

At the state of steady decay, the probability distribution ϕ decreases in value at a constant rate λ_0 :

$$\lim_{t \to \infty} \frac{1}{\phi} \frac{d\phi}{dt} = -\lambda_0 .$$

For small values of c, we can expand λ_0 into a power series in c.

$$(4.4) \qquad 2N\lambda_0 = 1 + \frac{2}{5}c^2 - \frac{2}{5^3 \cdot 7}c^4 - \frac{2^2}{3 \cdot 5^5 \cdot 7}c^6 - \frac{2 \cdot 31}{5^6 \cdot 7^3 \cdot 11}c^8 - \cdots .$$

This suggests that genic selection increases the rate of decay as compared with the case of no selection ($c = 0$), at least when c is small. Values of λ_0 for larger values of c will be found in the above reference [7] in which values of $2N\lambda_0$ up to $Ns = 8$ have been studied.

Very often, however, there is some dominance between alleles, and usually "complete" dominance. The main purpose of this section is to develop a method to analyse this situation.

Let us suppose that A is dominant over A' and the dominant genotypes AA and AA' have selective advantage s, measured in Malthusian parameters, over the homozygous recessive ($A'A'$). The differential equation for the probability distribution ϕ is

$$(4.5) \qquad \frac{\partial \phi}{\partial t} = \frac{1}{4N} \frac{\partial^2}{\partial x^2} \{x(1-x)\phi\} - \frac{\partial}{\partial x} \{sx(1-x)^2\phi\} \qquad (0 < x < 1),$$

with the initial condition

$$\phi(x, p, 0) = \delta(x - p).$$

If we apply the transformation

$$\begin{cases} \phi = e^{-\lambda t} e^{2cx(1-x/2)} w, \\ x = (1-z)/2, \end{cases}$$

to (4.5), we obtain the following ordinary differential equation:

$$(4.6) \quad (1-z^2)w'' - 4zw' + \left[\Lambda - 2 - \frac{c}{2}(z^2-1) + \frac{c^2}{4}(z^2-1)(1+z)^2\right]w = 0,$$

in which $\Lambda = 4N\lambda$ and $c = Ns$. We note that for the case of no selection ($c = 0$) the pertinent solution is the Gegenbauer polynomial. So we try to expand the solution into a series of Gegenbauer polynomials, which are known to form a complete orthogonal system in the interval $[-1, 1]$. Let

$$w = \sum_{n=0}^{\infty} d_n T_n^1(z),$$

in which the d_n's are constants. If we substitute this into (4.6) and use repeatedly the recurrence relation,

$$zT_n^1(z) = \frac{n+2}{2n+3} T_{n-1}^1(z) + \frac{n+1}{2n+3} T_{n+1}^1(z) \qquad (\text{set } T_{-1}^1(z) \equiv 0)$$

we obtain a 9-term recursion formula for the d_n's. Now we expand Λ and d_n's into power series of c:

$$\Lambda = k_0 + k_1c + k_2c^2 + k_3c^3 + \cdots ,$$
$$d_1 = (\alpha_1^1 c + \alpha_2^1 c^2 + \alpha_3^1 c^3 + \cdots) d_0 ,$$
$$d_2 = (\alpha_1^2 c + \alpha_2^2 c^2 + \alpha_3^2 c^3 + \cdots) d_0 ,$$
$$d_3 = (\alpha_2^3 c^2 + \alpha_3^3 c^3 + \alpha_4^3 c^4 + \cdots) d_0 ,$$
$$d_4 = (\alpha_2^4 c^2 + \alpha_3^4 c^3 + \cdots) d_0 ,$$
$$d_5 = (\alpha_3^5 c^3 + \cdots) d_0 , \text{ etc.},$$

and substitute these into the recursion formula. By picking out coefficients of equal powers of c, we can determine the k's and α's, by means of which the eigen-value λ (or Λ) and the eigenfunction w are expressed. The most important information is the smallest eigenvalue (λ_0) which gives the "rate of decay," and the corresponding eigenfunction. To get λ_0, we set $k_0 = 2$, since for $c = 0$, $\Lambda(= 4N\lambda)$ should be 2, as shown in the previous treatment of pure random drift in which the final rate of decay is $1/2N$.

Though the calculation involved is quite tedious, we can obtain the desired coefficients step by step. For the smallest eigenvalue λ_0 we have:

$$(4.7) \quad 2N\lambda_0 = 1 - \tfrac{1}{5}c + \frac{199}{2 \cdot 5^3 \cdot 7} c^2 + \frac{17}{2 \cdot 5^5 \cdot 7} c^3 - \frac{23 \cdot 41 \cdot 29599}{2^3 \cdot 3^3 \cdot 5^6 \cdot 7^3 \cdot 11} c^4 \cdots .$$

The coefficients of the eigenfunction are:

$$\alpha_1^1 = 0, \quad \alpha_2^1 = -\frac{1}{2 \cdot 3 \cdot 7}, \quad \alpha_3^1 = \frac{11}{3^4 \cdot 5^2 \cdot 7},$$

$$\alpha_1^2 = -\frac{1}{2 \cdot 3 \cdot 5^2}, \quad \alpha_2^2 = \frac{13}{3^2 \cdot 5^4}, \quad \alpha_3^2 \doteqdot -7.31 \times 10^{-5}, \cdots ,$$

$$\alpha_2^3 = \frac{1}{2 \cdot 3^2 \cdot 5 \cdot 7}, \cdots ,$$

$$\alpha_2^4 \doteqdot 2.49 \times 10^{-4}, \text{ etc.}$$

The same method may be applied to get similar expansions for other eigenvalues and eigenfunctions.

The shape of the distribution curve at the state of steady decay is given by

$$(4.8) \qquad \phi(x) = e^{2cx(1-x/2)}w_0 .$$

It will be convenient to adjust d_0 so that $\int_0^1 \phi(x) \, dx = 1$ (fixed classes excluded). The rate of fixation and loss of the gene A per generation at this state is given by $\phi(0)/4N$ and $\phi(1)/4N$ and therefore

$$(4.9) \qquad 4N\lambda_0 = \phi(1) + \phi(0).$$

This can be derived from (4.5) noting that $\phi(x)$ is finite at the boundaries ($x = 0$ and $x = 1$), as shown for the case of no dominance in [7].

TABLE 2

x	$2Ns = 1$	$2Ns = -1$
0.0	0.688	1.389
0.1	0.764	1.251
0.2	0.838	1.142
0.3	0.910	1.056
0.4	0.977	0.990
0.5	1.037	0.940
0.6	1.088	0.903
0.7	1.128	0.879
0.8	1.155	0.865
0.9	1.168	0.860
1.0	1.166	0.866

A numerical example will be given here. For weak selection favoring domi-
nants; $2Ns = 2c = 1$, we get

$$2N\lambda_0 = 0.928$$

and

$$w_0 = d_0[T_0^1(z) - 0.0058T_1^1(z) - 0.0028T_2^1(z)$$

$$+ 0.0004T_3^1(z) + 0.00006T_4^1(z) + \cdots],$$

in which $T_0(z) = 1$, $T_1(z) = 3z$, $T_2(z) = \frac{3}{2}(5z^2 - 1)$, $T_3(z) = \frac{5}{2}(7z^3 - 3z)$, $\cdots$.
Values of $\phi(x)$ at 0, 0.1, 0.2, $\cdots$ and 1 are listed in Table 2. They are adjusted
by Simpson's rule so that the area under the curve is unity. $\phi(1) + \phi(0)$ comes
out 1.855, while $4N\lambda_0$ is 1.856. The agreement is satisfactory for this level of
approximation. As a second example, we assume weak selection against the
dominants: $2Ns = 2c = -1$. $2N\lambda_0$ is 1.128 and values of $\phi(x)$ are given in Table
2. In this case $\phi(1) + \phi(0)$ comes out 2.254 while $4N\lambda_0$ is 2.256. Again the
agreement is satisfactory.

The above treatment leading to the power series expansion of eigenvalues and
of coefficients of eigenfunctions is rather heuristic. For the more rigorous treat-
ment of the problem, further investigation of these series will be required.

As to uniqueness of the solutions of the type of singular partial differential
equations considered in this section, an investigation could presumably be based
on Section 23 of Feller's paper [22].

The most remarkable fact suggested by the above analysis seems to be that
as compared with the case of pure random drift, selection toward dominants
$(s > 0)$ decreases the final rate of decay, while selection against dominants
$(s < 0)$ increases it. At least for weak selection the above results follow from
(4.7), since the most influential term $-\frac{1}{5}c$ is negative if $c(= Ns)$ is positive and
positive if e is negative.

For this continuous treatment to be applicable, the population number N

TABLE 3

Classes	Frequencies %
$7A + 1A'$	14.80
$6A + 2A'$	16.48
$5A + 3A'$	16.32
$4A + 4A'$	15.51
$3A^{\cdot} + 5A'$	14.28
$2A + 6A'$	12.64
$1A + 7A'$	9.97
Total......................................	100.00

should be fairly large so that $1/N$ is negligible as compared with 1. If the population is extremely small, we must treat the problem by the methods of finite Markov chains. The transition probability that the number of A genes in the population becomes j in the next generation, given that it is i in the present generation will be given by

$$p_{j|i} = \binom{2N}{j} x'^{j}(1 - x')^{2N-j} \qquad (i, j = 0, 1, \cdots, 2N),$$

where $x' = x + \delta x$, in which $x = i/(2N)$, and δx is the change of gene frequency by selection per generation and is $sx(1 - x)^2$ if s is small. The rate of decay of the unfixed classes and their limiting distribution may be obtained by iteration. For example, if $N = 4$, $2Ns = 1$, the limiting form of the distribution (fixed classes excluded) becomes as follows (Table 3), with rate of decay (λ_0) 11.875 %, giving $2N\lambda_0 = 0.9500$. If there is no selection $(s = 0)$, it turns out that the rate of decay becomes $1/2N = 0.125$ or 12.5 %. Note that with selection for dominants, the rate of decay is smaller.

5. Chance of fixation of mutant genes. In any large natural population, gene mutations may be occurring in each generation. Most of the mutant genes are likely to be deleterious but a few of them may turn out to be advantageous. Such advantageous mutant genes have a tendency to increase their frequencies in later generations thus having a positive chance of establishing themselves even in a very large population. Because of its importance in evolutionary genetics, the probability of fixation of mutant genes has been studied by Fisher [23], Haldane [24] and Wright [25] [26]. However, due to mathematical difficulties involved, so far only a few cases have been successfully worked out.

In this section I will try to present the solution under quite general conditions and will show that the previous results are obtained as special cases.

We will designate the selective advantage of the mutant homozygote (AA) by s and that of the heterozygote (AA') by sh. Let $u(p, t)$ be the conditional probability that the mutant gene reaches fixation by the tth generation, given that its initial frequency is p. Under the assumption of a continuous model and

random mating it is possible to show that $u(p, t)$ satisfies the following partial differential equation:

$$(5.1) \qquad \frac{\partial u}{\partial t} = \frac{p(1 - p)}{4N} \frac{\partial^2 u}{\partial p^2} + sp(1 - p)\{h + (1 - 2h)p\} \frac{\partial u}{\partial p}.$$

Here we have inevitably the following boundary conditions:

$$(5.2) \qquad\qquad u(0, t) = 0, \qquad u(1, t) = 1.$$

For the special case of neutral genes ($s = 0$), the pertinent solution is

$$(5.3) \quad u(p, t) = p + \sum_{i=1}^{\infty} (2i + 1)pq(-1)^i F(1 - i, i + 2, 2, p)e^{-[i(i+1)t/4N]},$$

which agrees exactly with the results obtained by the study of moments [16]. Usually the process of evolution extends over an enormous period of time and hence the probability of ultimate fixation will be of special importance. We will designate such probability by $u(p)$ which is defined by

$$u(p) = \lim_{t \to \infty} u(p, t).$$

For the neutral mutant gene, $u(p) = p$. If ν is the initial number of mutant genes, $u(p) = \nu/2N$ for this case and hence the probability of fixation per mutant gene is $1/2N$.

For the general case the probability may be obtained by setting $\partial u/\partial t = 0$ in (5.1). This leads to

$$(5.4) \qquad u(p) = \int_0^p e^{-2cDx(1-x)-2cx} \, dx \Big/ \int_0^1 e^{-2cDx(1-x)-2cx} \, dx,$$

where $c = Ns$ and $D = 2h - 1$.

The rate of approach to the ultimate state of complete fixation or loss may be given by the smallest eigenvalue λ_0 of equation (5.1). For a small value of c, we can expand λ_0 into a power series in c as follows:

$$(5.5) \qquad 2N\lambda_0 = 1 + K_1 c + K_2 c^2 + K_3 c^3 + K_4 c^4 + \cdots,$$

where

$$K_1 = -\tfrac{1}{5}D, \qquad K_2 = \frac{1}{2 \cdot 5} + \frac{2^2 \cdot 3}{5^3 \cdot 7} D^2, \qquad K_3 = \frac{1}{2 \cdot 5^3 \cdot 7} D - \frac{2^2}{5^5 \cdot 7} D^3,$$

$$K_4 = -\frac{1}{2^3 \cdot 5^3 \cdot 7} - \frac{7^3}{2 \cdot 3^3 \cdot 5^5} D^2 - \frac{2^2 \cdot 3^5}{5^6 \cdot 7^3 \cdot 11} D^4, \text{ etc.}$$

It should be noted that for the case of no dominance $D = 0$ and the above series (5.5) agrees with (4.4) provided that $2s$ is used instead of s to express the selective advantage of the homozygous mutants. For the case of complete dominance, $D = 1$ or -1 according as the mutant gene is either dominant or recessive. In the former case of $D = 1$, (5.5) agrees with (4.7). Returning to formula (5.4), we will consider a few cases of special importance in evolution. To obtain the

chance of fixation of individual mutant gene denoted by u we may put $p = 1/2N$. For the case of no dominance $(D = 0)$, we have

$$u = (1 - e^{-s})/(1 - e^{-2c}),$$

or denoting the selective advantage of the homozygote by $2s$,

$$u = (1 - e^{-2s})/(1 - e^{-4Ns}).$$

Thus for a slightly advantageous mutant gene we may write

(5.6) $$u = 2s/(1 - e^{-4Ns})$$

with good approximation. The result agrees with Fisher [23] who used the method of branching processes and also with Wright [25] who used the method of integral equations. For a large N this chance is very close to $2s$ as given by Haldane [24]. For a slightly disadvantageous mutant gene $(s < 0)$, we have

(5.7) $$u = 2s'/(e^{4Ns'} - 1),$$

where $s' = -s$. The chance is not negligible if Ns' is small. The result agrees with that obtained by Wright [25].

For the completely dominant gene $(D = 1)$ with small selective advantage $s(s > 0)$ we may use the formula $u = 2s$ unless Ns is small.

The case of a completely recessive mutant gene $(D = -1)$ with small selective advantages s $(s > 0)$ in the homozygous state is of special interest. Haldane [24] estimated the chance of fixation as of the order of $\sqrt{s/N}$ using the method of branching processes and Wright [26] estimated it as of the order of $\sqrt{s/2N}$ by his method of integral equations. Our formula (5.4) gives

(5.8) $$u = \sqrt{2s/(\pi N)}$$

as the best simple approximation for a large N. Since

$$\sqrt{2s/\pi N} = \sqrt{2/\pi}\sqrt{s/N} = \sqrt{4/\pi}\sqrt{s/2N},$$

it may readily be seen that our result lies between those of Haldane and Wright. Furthermore it is interesting to note that Wright [26] obtained numerically the formula $1.1(s/2N)^{\frac{1}{2}}$ as the average chance of fixation for values of s ranging from $4/2N$ to $64/2N$. The factor 1.1 is indeed very close to $\sqrt{4/\pi}$ which is $1.128 \cdots$.

Finally our general formula (5.4) allows us to calculate the chance of fixation of a nearly recessive mutant gene with selective advantage s $(s > 0)$ in the homozygous state. Namely for $0 < h \ll 1$, we may have

(5.9) $$u = e^{-2Nsh^2/(1-2h)} \sqrt{\frac{2s(1 - 2h)}{\pi N}} \Big/ \{1 - 2\Phi(\sqrt{4Nsh^2/(1 - 2h)})\}$$

as a good approximation, unless $2Ns$ is small. Here $\Phi(x)$ stands for the error function

$$\Phi(x) = (1/\sqrt{2\pi}) \int_0^x e^{-x^2/2} \, dx.$$

As an example consider a case with $N = 10^3$ and $s = 10^{-1}$. If the mutant gene is completely recessive $(h = 0)$, $u \approx 0.8 \times 10^{-2}$. With slight phenotypic effect of $h = 0.01$ in the heterozygote $u \approx 0.9 \times 10^{-2}$, while with

$$h = 0.1, \qquad u \approx 2.3 \times 10^{-2}.$$

6. Population structure and evolution. So far we have considered the process of change in gene frequency in an isolated population in which mating is random and the number of individuals remains constant through generations. This may be an over-simplification for the study of evolutionary processes in general, since most species in nature may have a much more complicated breeding structure. Unfortunately this immediately brings us baffling problems, for the solution of which new techniques will be required.

First let us suppose that a species is subdivided into numerous isolated colonies, each of which may receive, from time to time, migrants taken as random samples from the whole population. Mating is assumed to be random within each colony. Following Wright [27] we will call this the "island model." The model may be realistic to describe a species inhabiting an archipelago such as the Galapagos Islands studied by Darwin. The number of individuals may fluctuate from generation to generation not only due to fluctuation in environmental conditions but also due to change in the genetic make up of each colony which in turn is influenced by the population number. If the number of reproducing individuals per colony is small, say less than 100, and if isolation is so severe that less than one migrant is expected per thousand generations, the chance of disadvantageous mutant genes reaching fixation may be considerable, as suggested by (5.7), and accumulation of such genes will lead to extinction of colonies. We would like to know then what is the chance that an isolated colony becomes extinct before a migrant comes in to start a new colony. What is the joint distribution of the population number and the gene frequency among colonies at the steady state? These questions may have to be answered before we reach conclusions on the optimum structure of populations for the evolution of a species.

Next we will consider the continuum model of a population. The model is realistic for representing a species inhabiting a wide range with more or less uniform density. Here the whole population can not be a random mating unit since a tendency toward "isolation by distance" may arise due to limitation in the locomotive ability of the organism [27]. In the course of time advantageous mutant genes may arise with exceedingly low rate in various spots in the continuum and these will spread into the population. If the local fluctuation of gene frequencies is negligible, the process of spread will be very similar to diffusion of physical particles in a medium, except here that differential rate of multiplication is involved among particles.

Let $x(\mu, \nu)$ be the relative frequency of a mutant gene denoted by A at a point (μ, ν) in the continuum with rectangular coordinate system. The process of spread of the advantageous dominant gene may be described by the equation

(6.1)
$$\frac{\partial x}{\partial t} = m\nabla^2 x + sx(1 - x)^2,$$

where m represents locomotive ability of an individual and corresponds to a diffusion constant in physical systems, ∇^2 denotes the two-dimensional Laplace operator $(\partial^2/\partial\mu^2 + \partial^2/\partial\nu^2)$, and s is the selective advantage of the dominant gene A to its allele A'. The simplest situation is that s is constant throughout the continuum. The mutant gene will spread in the form of concentric circles from the point of origin which we may take as $(0, 0)$. Introducing the polar coordinates (r, θ) and assuming that $\partial^2 x/\partial^2\theta = 0$, (6.1) becomes

(6.2)
$$\frac{\partial x}{\partial \tau} = \frac{\partial^2 x}{\partial r^2} + \frac{1}{r}\frac{\partial x}{\partial r} + \alpha x(1 - x)^2,$$

where $\tau = mt$ and $\alpha = s/m$. At an early stage when the frequency of A is still low, the distribution may be approximated by

$$x(r, \tau) = x_0 e^{\alpha\tau - r^2/4\tau} r/2\tau,$$

where x_0 is the initial frequency of A at the origin. Beyond this stage, however, we face a difficult problem of solving a non-linear diffusion equation.

The problem of steady state distribution is worthwhile to investigate if the mutant gene is advantageous within a closed region but disadvantageous outside, as in the case of melanic genes in many lepidopteran species which in recent years have increased their relative frequencies in a spectacular fashion in many industrial areas but remain in low frequencies in rural districts—a phenomenon known as "industrial melanism" [28].

Real mathematical difficulties arise, however, when we take random fluctuation of local gene frequencies into consideration. The fluctuation may be due to random sampling of gametes in reproduction or due to random fluctuation of selection intensities brought about by chance fluctuation of local environments. Notable contributions have been made by Wright [27] [29] [30] and Malécot [4] [18] for the case of no selection, but more important cases involving selection are yet to be worked out.

Such studies should be indispensable for our understanding of the process of speciation and also of the mechanism of evolution in general.

In his theories of evolution Wright [31] put forward an important concept of "balance," especially of balance between directional factors such as selection, mutation, and migration and undirectional or stochastic factors such as random sampling of gametes and random fluctuation of environmental conditions. It appears that new methods of stochastic processes will be needed for a satisfactory treatment of Wright's theory of evolution.

ing some of the eigenvectors in the Section 2. Finally the author wishes to express his indebtness to Professor Sewall Wright for guidance and help in many ways.

REFERENCES

[1] A. Kolmogorov, "Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung," *Math. Ann.*, Vol. 104 (1931), pp. 415–458.

[2] M. S. Bartlett and J. B. S. Haldane, "The theory of inbreeding in autotetraploids," *J. Genetics*, Vol. 24 (1934), pp. 175–180.

[3] R. A. Fisher, *The Theory of Inbreeding*, Oliver and Boyd, Ltd., 1949.

[4] G. Malécot, *Le Mathematiques de L'Hérédité*, Masson et Cie, Paris, 1948.

[5] W. Feller, "Diffusion processes in genetics," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1951, pp. 227–246.

[6] J. F. Crow and M. Kimura, "Some genetic problems in natural populations," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 4 (1956), pp. 1–22.

[7] M. Kimura, "Stochastic processes and distribution of gene frequencies under natural selection," *Cold Spring Harbor Symposium*, Vol. 20 (1955), pp. 33–55.

[8] M. Demerec, "Behavior of two mutable genes of Delphinium ajacis," *J. Genetics*, Vol. 24 (1931), pp. 179–193.

[9] H. Matsuura and T. Suto, "Genic analysis in maize V. mode of calico character" (in Japanese), *Japanese J. Genetics*, Vol. 23 (1948), p. 31.

[10] Ch. Auerbach, "Problems in chemical mutagenesis," *Cold Spring Harbor Symposium*, Vol. 16 (1951), pp. 199–213.

[11] Von H. Friedrich-Freksa und F. Kaudewitz, "Letale Spätfolgen nach Einbau von ^{32}P in Amoeba proteus und ihre Deutung durch genetische Untereinheiten," *Z. Naturforsch.*, Vol. 86 (1953), pp. 343–355.

[12] T. M. Sonneborn and M. V. Schneller, "Genetic consequences of aging in variety 4 of Paramecium aurelia," *Records Genetics Soc. America*, Vol. 24 (1955), p. 596.

[13] A. W. Kimball and A. S. Householder, "A stochastic model for the selection of macronuclear units in paramecium growth," *Biometics*, Vol. 10 (1954), pp. 361–374.

[14] Ch. Jordan, *Calculus of Finite Differences*, New York, 1950.

[15] S. Wright, "The differential equation of the distribution of gene frequencies," *Proc. Nat. Acad. Sci.*, Vol. 31 (1945), pp. 382–389.

[16] M. Kimura, "Solution of a process of random genetic drift with a continuous model," *Proc. Nat. Acad. Sci.*, Vol. 41 (1955), pp. 144–150.

[17] S. Goldberg, "On a Singular Diffusion Equation," Ph.D. Thesis (Cornell University), 1950, unpublished.

[18] G. Malécot, "Un traitement stochastique des problèmes linéaires [mutation, linkage, migration] en Génétique de Population," *Sciences Mathématiques et Astronomie XIV, Annales de l'Université de Lyon* (1951), pp. 79–117.

[19] R. A. Fisher, *The Genetical Theory of Natural Selection*, Clarendon Press, 1930.

[20] S. Wright and W. E. Kerr, "Experimental studies of the distribution of gene frequencies in very small populations of Drosophila melanogaster, II. Bar," *Evolution*, Vol. 8 (1954), pp. 225–240.

[21] J. A. Stratton, P. M. Morse, L. J. Chu, and R. A. Hutner, *Elliptic Cylinder and Spheroidal Wave Functions*, John Wiley & Sons, New York, 1941.

[22] W. Feller, "The parabolic differential equations and the associated semi-group of transformations," *Ann. Mathematics*, Vol. 55 (1952), pp. 468–519.

[23] R. A. FISHER, "The distribution of gene ratios for rare mutations," *Proc. Roy. Soc. Edinburgh*, Vol. 50, Pt. II (1930), pp. 205–220.

[24] J. B. S. HALDANE, "A mathematical theory of natural and artificial selection. Part V: Selection and mutation," *Proc. Cambridge Philos. Soc.*, Vol. 23 (1927), pp. 838–844.

[25] S. WRIGHT, "Evolution in Mendelian populations," *Genetics*, Vol. 16 (1931), pp. 97–159.

[26] S. WRIGHT, "Statistical genetics and evolution," *Bull. Amer. Math. Soc.*, Vol. 48 (1942), pp. 223–246.

[27] S. WRIGHT, "Isolation by distance," *Genetics*, Vol. 28 (1943), pp. 114–138.

[28] H. B. D. KETTLEWELL, "Selection experiments on industrial melanism in the *Lepidoptera*," *Heredity*, Vol. 9 (1955) pp. 323–342.

[29] S. WRIGHT, "Isolation by distance under diverse systems of mating," *Genetics*, Vol. 31 (1946), pp. 39–59.

[30] S. WRIGHT, "The genetical structure of populations," *Ann. Eugenics*, Vol. 15 (1951), pp. 323–354.

[31] S. WRIGHT, "Population structure as a factor in evolution," in *Moderne Biologie*, Berlin, 1950, pp. 275–287.