noted above. To eliminate this possibility, it will be necessary to map both the original mutant stock and mutant clones derived from it via lambda transduction.

Summary.—A preliminary map of the galactose region of *E. coli* K-12, obtained by the study of crossing over in heterogenotes showing position effect is presented. The order, kinase cistron-operator region-transferase cistron is indicated. The relationship of the epimerase cistron to the above region is not known, since it has as yet not been possible to study epimerase mutants.

[1] Morse, M. L., *Genetics*, **40**, 586 (1955).

[2] Morse, M. L., E. M. Lederberg, and J. Lederberg, *Genetics*, **41**, 758 (1956).

[3] Morse, M. L., *Genetics*, **44**, 529 (1959).

[4] Lederberg, E., in *Microbial Genetics*, 10th Symposium, Society for General Microbiology (1961).

[5] Morse, M. L., *Genetics*, **45**, 1001 (1960).

[6] Morse, M. L., Abstracts Fifth Annual Meeting Biophysical Society (St. Louis, Feb. 1961).

[7] Morse, M. L., and M. L. Alire, *J. Bact.*, **76**, 270 (1958).

[8] Luria, S. E., and M. Delbrück, *Genetics*, **28**, 491 (1943).

[9] Lea, D. E., and C. A. Coulson, *J. Genetics*, **49**, 264 (1949).

[10] Armitage, P., *J. Hygiene*, **51**, 162 (1953).

[11] Kalckar, H. M., K. Kurahashi, and E. Jordan, these PROCEEDINGS, **45**, 1776 (1959).

[12] Yarmolinsky, M. B., and H. Wiesmeyer, these PROCEEDINGS, **46**, 1626 (1960).

[13] Soffer, R. L., *J. Bact.*, **82**, 471 (1961).

[14] Demerec, M., and P. E. Hartman, *Ann. Rev. Microbiol.*, **13**, 377 (1959).

[15] Jacob, F., and J. Monod, *J. Mol. Biol.*, **3**, 318 (1961).

[16] Pardee, A. B., F. Jacob, and J. Monod, *J. Mol. Biol.*, **1**, 165 (1959).

[17] Weigle, J., M. Meselson, and K. Paigen, in *Structure and Function of Genetic Elements*, Brookhaven Symposia in Biology, No. 12 (1959), p. 125.

[18] Morse, M. L., *Ann. N. Y. Acad. Sci.*, **68**, 324 (1957).

POPULATION STUDIES IN PREDOMINANTLY SELF-POLLINATED SPECIES, III. A MATRIX MODEL FOR MIXED SELFING AND RANDOM OUTCROSSING*

BY P. L. WORKMAN AND R. W. ALLARD

UNIVERSITY OF CALIFORNIA, DAVIS

Communicated by G. Ledyard Stebbins, June 22, 1962

The prevalence of self-pollination, and the importance of many of the species, both natural and agricultural, which have adopted self-pollination as their mode of reproduction indicate that this mating system is one of the most successful among the diverse types which occur in flowering plants. The success of predominant self-fertilization as a mating system has commonly been attributed to the genetic uniformity which it presumably encourages in populations. Under selfing, the population consists entirely or largely of homozygotes, and the effect of selection is postulated to favor the best adapted genotypes at the expense of the less adapted, leading to populations which consist of one or a few highly fit genotypes. The

favored individuals will produce offspring genetically like themselves, and such populations therefore are presumed to show high agreement with the optimum phenotype.

However, there is evidence indicating that heterozygotes are not uncommon in predominantly selfing populations[1] and that such populations are not as uniform genetically as commonly supposed.[2, 3] There is also evidence that optimum population structure may be attained when the population consists of many, not few, genotypes.[4] Additional investigations of the genetic structure of predominantly selfing species are therefore required in order to understand the evolutionary dynamics of the numerous species which have adopted this mating system.

Genetic analyses of predominantly self-fertilizing populations would be aided if a theoretical model were available for computing the genotypic frequencies expected in any generation, given initial genotypic frequencies and the amount of selfing versus outcrossing. The models that are available either have employed algebraic recursions relating genotypic frequencies in successive generations[5, 6] or have approached the problem indirectly in terms of the inbreeding coefficient.[7] These methods have the disadvantage that they either do not provide formulas for direct computation of genotypic frequencies during passage to the limit, or equilibrium state, or are not readily extended to multilocus situations. We shall therefore develop a model, using the method of generation matrices,[8] which permits direct calculation of the genotypic array for any generation, given the initial genotypic frequencies and the proportion of selfing and random outcrossing. The model developed is generally applicable to the case of a single diallelic locus and can be extended, under certain restrictions, to diallelic-multilocus cases. In addition, as will be shown, the model can be extended to certain other mixed mating systems.

Preliminary Considerations.—To introduce concepts and notations that will be needed later we shall consider first the simplest cases, those of infinitely large diploid populations in which mating is either by complete self-fertilization, or alternatively, by complete random outcrossing. For a single locus with two alleles, the recurrences relating the frequencies of genotypes A_1A_1, A_1A_2, and A_2A_2 in generations (n) and $(n + 1)$, assuming full selfing and no selection, are

$$A_1A_1: \quad f_1^{(n+1)} = f_1^{(n)} + \tfrac{1}{4}f_2^{(n)},$$
$$A_1A_2: \quad f_2^{(n+1)} = \tfrac{1}{2}f_2^{(n)},$$
$$A_2A_2: \quad f_3^{(n+1)} = f_3^{(n)} + \tfrac{1}{4}f_2^{(n)},$$

or, in matrix form,

$$\begin{pmatrix} f_1^{(n+1)} \\ f_2^{(n+1)} \\ f_3^{(n+1)} \end{pmatrix} = \begin{pmatrix} 1 & \tfrac{1}{4} & 0 \\ 0 & \tfrac{1}{2} & 0 \\ 0 & \tfrac{1}{4} & 1 \end{pmatrix} \cdot \begin{pmatrix} f_1^{(n)} \\ f_2^{(n)} \\ f_3^{(n)} \end{pmatrix}, \tag{1}$$

where $f_i^{(n)}$ is the frequency of the ith genotype in the nth generation. Rewriting (1) in matrix notation we obtain

$$F^{(n+1)} = R \cdot F^{(n)}. \tag{2}$$

The general n-step relationship can be derived from (2) giving

$$F^{(n)} = R^n \cdot F^{(0)}. \tag{3}$$

The matrix R^n is found by the method of generation matrices[7] and the n-step transition formulas relating genotypic frequencies in the nth generation to those in the initial generation $(f_i^{(0)})$ can, for the case of full selfing, be shown to be

$$
\begin{aligned}
f_1^{(n)} &= f_1^{(0)} + \tfrac{1}{2}[1 - (\tfrac{1}{2})^n]f_2^{(0)}, \\
f_2^{(n)} &= (\tfrac{1}{2})^n f_2^{(0)}, \\
f_3^{(n)} &= f_3^{(0)} + \tfrac{1}{2}[1 - (\tfrac{1}{2})^n]f_2^{(0)}.
\end{aligned}
\right\}
\tag{4}
$$

In deriving the expectations under complete random mating, we denote by $D^{(n)}$ the column vector giving the genotypic array resulting from one generation of random mating following the nth generation. By the Hardy-Weinberg rule, we know that $D^{(l)} = D^{(m)}$ for all generations, l, m, so that $D^{(n)} = D$ for all n generations. In particular, if the initial population is denoted by $F^{(0)}$ and the frequencies of alleles A_1 and A_2 are given by p and q, respectively, then

$$
D = \begin{pmatrix} [f_1^{(0)} + \tfrac{1}{2}f_2^{(0)}]^2 \\ 2(f_1^{(0)} + \tfrac{1}{2}f_2^{(0)})(f_3^{(0)} + \tfrac{1}{2}f_2^{(0)}) \\ [f_3^{(0)} + \tfrac{1}{2}f_2^{(0)}]^2 \end{pmatrix} = \begin{pmatrix} p^2 \\ 2pq \\ q^2 \end{pmatrix}.
\tag{5}
$$

The Model of Mixed Selfing and Random Outcrossing.—Consider a diploid population, infinitely large, in which there is a constant probability, t, of random outcrossing, and a constant probability, s, of selfing ($s + t = 1$). We shall temporarily restrict consideration to one locus with two alleles, A_1 and A_2.

Starting with a population in which the distribution of the genotypic frequencies is specified by $F^{(0)}$ and assuming no selection, the recurrence relationships under the mixed system of partial random outcrossing and partial selfing are

$$
\begin{aligned}
f_1^{(n+1)} &= t[f_1^{(n)} + \tfrac{1}{2}f_2^{(n)}]^2 + s[f_1^{(n)} + \tfrac{1}{4}f_2^{(n)}], \\
f_2^{(n+1)} &= 2t[f_1^{(n)} + \tfrac{1}{2}f_2^{(n)}][f_3^{(n)} + \tfrac{1}{2}f_2^{(n)}] + s[\tfrac{1}{2}f_2^{(n)}], \\
f_3^{(n+1)} &= t[f_3^{(n)} + \tfrac{1}{2}f_2^{(n)}]^2 + s[f_3^{(n)} + \tfrac{1}{4}f_2^{(n)}],
\end{aligned}
\right\}
\tag{6}
$$

or in matrix form,

$$
\begin{pmatrix} f_1^{(n+1)} \\ f_2^{(n+1)} \\ f_3^{(n+1)} \end{pmatrix} = t\begin{pmatrix} [f_1^{(n)} + \tfrac{1}{2}f_2^{(n)}]^2 \\ 2[f_1^{(n)} + \tfrac{1}{2}f_2^{(n)}][f_3^{(n)} + \tfrac{1}{2}f_2^{(n)}] \\ [f_3^{(n)} + \tfrac{1}{2}f_2^{(n)}]^2 \end{pmatrix} + s\begin{pmatrix} 1 & \tfrac{1}{4} & 0 \\ 0 & \tfrac{1}{2} & 0 \\ 0 & \tfrac{1}{4} & 1 \end{pmatrix}\begin{pmatrix} f_1^{(n)} \\ f_2^{(n)} \\ f_3^{(n)} \end{pmatrix}.
\tag{7}
$$

From the discussion of the case of full selfing and by definition of $D^{(n)}$, we see that (7) can be written in matrix notation as

$$
F^{(n+1)} = tD + sRF^{(n)},
\tag{8}
$$

where D is a matrix with constant terms which can be obtained from the initial distribution of genotypes given by $F^{(0)}$. From (8), it is seen that

$$
\begin{aligned}
F^{(1)} &= tD + sRF^{(0)}, \\
F^{(2)} &= tD + sRF^{(1)}, \\
&= tD + stRD + s^2R^2F^{(0)}, \text{ and so on.}
\end{aligned}
$$

Repeated substitutions lead to the general result, which can be proved by induction, that

$$
F^{(n)} = tD + \{t[I - sR]^{-1}[sR - (sR)^n]\}D + (sR)^nF^{(0)},
\tag{9}
$$

where I is the appropriate identity matrix. This equation can be written more simply as

$$F^{(n)} = tD + BD + s^n R^n F^{(0)}, \qquad (10)$$

where B is the matrix enclosed by $\{\ \}$ in (9). The general n-step formulas for predicting the genotypic frequencies expected in any generation can be obtained by substitution in (9) of the matrices R and R^n given by (2) and (3). The formulas are

$$
\left.
\begin{aligned}
f_1{}^{(n)} &= (1 - s^n)p^2 + 2pq\left[\frac{s}{4-2s} - \tfrac{1}{2}(s^n) - \left(\frac{s}{2}\right)^n\left(\frac{t}{2-s}\right)\right] \\
&\qquad + s^n \cdot f_1{}^{(0)} + \tfrac{1}{2}(s^n)[1 - (\tfrac{1}{2})^n]f_2{}^{(0)}, \\
f_2{}^{(n)} &= 2pq\left[\frac{2t}{2-s}\right]\left[1 - \left(\frac{s}{2}\right)^n + \left(\frac{s}{2}\right)^n\right]f_2{}^{(0)}, \\
f_3{}^{(n)} &= (1 - s^n)q^2 + 2pq\left[\frac{s}{4-2s} - \tfrac{1}{2}(s^n) - \left(\frac{s}{2}\right)^n\left(\frac{t}{2-s}\right)\right] \\
&\qquad + s^n \cdot f_3{}^{(0)} + \tfrac{1}{2}(s)^n[1 - (\tfrac{1}{2})^n]f_2{}^{(0)}.
\end{aligned}
\right\} \qquad (11)
$$

The limiting distribution for any set of assumptions about $F^{(0)}$ can be obtained from (11) by considering the limits of the $f_i{}^{(n)}$ as $n \to \infty$. These limiting values are

$$
\left.
\begin{aligned}
f_1 &= p^2 + 2pq[s/(4-2s)], \\
f_2 &= 2pq[2t/(2-s)], \\
f_3 &= q^2 + 2pq[s/(4-2s)].
\end{aligned}
\right\} \qquad (12)
$$

Since in many experiments the initial population is composed solely of heterozygotes, i.e., $f_2{}^{(0)} = 1$, we are especially interested in this case. The proportions in the nth generation can be obtained from (11) as

$$f_1, f_3 = [1/(4-2s)]\left[1 - \left(\frac{s}{2}\right)^n\right] \text{ with limiting value } 1/(4-2s),$$

$$f_2 = [t/(2-s)] + [1/(2-s)]\left(\frac{s}{2}\right)^n \text{ with limiting value } t/(2-s). \qquad (13)$$

One of the main issues in predominantly self-pollinated populations is the amount of heterozygosity that can be maintained indefinitely, given various assumptions about the initial composition of the population and the amount of outcrossing. Table 1 gives limiting frequencies of heterozygotes expected under various assump-

TABLE 1

EXPECTED PROPORTIONS OF HETEROZYGOTES IN THE LIMIT $(n \to \infty)$ UNDER VARIOUS ASSUMPTIONS ABOUT t AND p (INITIAL FREQUENCY) OF A_1

t	p					
	0.05	0.10	0.20	0.30	0.40	0.50
0.1	0.0173	0.0327	0.0582	0.0764	0.0783	0.0909
0.2	0.0317	0.0600	0.1067	0.1400	0.1600	0.1667
0.3	0.0439	0.0831	0.1477	0.1939	0.2216	0.2308
0.4	0.0543	0.1029	0.1829	0.2610	0.2743	0.2857
0.5	0.0633	0.1200	0.2133	0.2800	0.3200	0.3333
0.6	0.0713	0.1350	0.2400	0.3150	0.3600	0.3750
0.7	0.0783	0.1482	0.2636	0.3459	0.3953	0.4118
0.8	0.0844	0.1600	0.2844	0.3733	0.4266	0.4444
0.9	0.0900	0.1705	0.3022	0.3969	0.4548	0.4737
1.0	0.0950	0.1800	0.3200	0.4200	0.4800	0.5000

tions about t and p (initial frequency of A_1). By the symmetry of the model, the tabular value of p is that of q when $p > 0.5$.

An Alternative Derivation.—The genetic model can also be represented as a simple branching process in which the population is treated as if it were broken into lines arising either by selfing, with probability s, or by random outcrossing, with probability t. This process, which corresponds to a coin-tossing game, is shown graphically in Figure 1, in which the original population is represented by $F^{(0)}$, S denotes a line derived by selfing, and T denotes a line derived by random outcrossing. Figure 1 is thus a "family tree" giving all possible outcomes in any
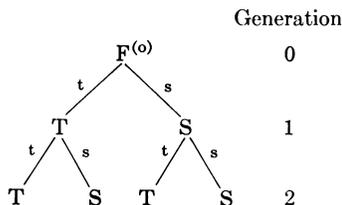
FIG. 1.

generation, with the probability of each outcome being the product of the respective probabilities along the branches leading to that outcome. The genotypic arrays of the 2^n outcomes in the nth generation are determined as follows. Any line ending in T has a genotypic array given by D, where D is as defined in (5). In any branch in which there is a T, the line at that point is brought back to the Hardy-Weinberg equilibrium with its genotypic array given by D. Hence, with one exception, we need only consider k generations of selfing from the equilibrium state for any line ending in S in the nth generation, where k represents the number of S's following the last T. This can be shown by (3) to be equal to $R^k D$, where R is the appropriate generation matrix. The exceptional line is the one consisting solely of S's, which by (3) has a genotypic array given by $R^n F^{(0)}$. We now have a means of specifying for each outcome in the nth generation both its genotypic composition and its probability of occurrence. Summing over all outcomes in the nth generation, we obtain a specification of the total population in that generation. As an example, in Figure 1, $F^{(2)} = t^2 D + stRD + stD + s^2 R^2 F^{(0)} = tD + stRD + s^2 R^2 F^{(0)}$. Results for $F^{(3)}, F^{(4)}, \ldots$ lead to formula (9), which was derived solely from the matrix recursion (8).

It is evident from Figure 1 and the above discussion that the model of a branching process can be applied to all systems of mating involving the combination of random outcrossing with any mating system which can be treated by the generation matrix method (e.g., random outcrossing and sib-mating). Substitution into formula (9) of the appropriate generation matrices, say Q and Q^n for R and R^n, where Q is the generation matrix for some other system of mating, yields the appropriate n-step transition formulas.

Extension of the Model to More than One Locus.—Matrix formula (9) developed for the one-locus case can also be applied to cases involving two or more unlinked loci, provided that certain restrictions are placed on $F^{(0)}$. The proof for the one locus case required only that $D^{(l)} = D^{(m)}$ for all l, m. That this is the case can be seen from an examination of (8), which states that the genotypic array in any

generation is the sum of the selfing component and the outcrossing component of the previous generation. If the genotypic array of the outcrossing component is constant for all generations, i.e., if $D^{(l)} = D^{(n)}$ for all l, m, then (8) and hence the general matrix solution (9) will be independent of the number of loci involved and hence valid for multilocus cases.

Suppose there are two loci with alleles A_1 and A_2 at one locus and alleles B_1 and B_2 at the other. Let the proportions of gametes A_1B_1, A_2B_2, A_1B_2, and A_2B_1 be denoted by G_1, G_2, G_3, and G_4, respectively. Then, if $G_1G_2 = G_3G_4$ for some generation, l, it can be shown that $D^{(l)} = D^{(n)} = D$ for all $n > l$. Thus, if we restrict $F^{(0)}$ to a population in which $G_1G_2 = G_3G_4$, the model will be valid for all subsequent generations. In the following derivations, it is assumed that $F^{(0)}$ is so restricted.

The recurrence equations for the two locus model are

$$
\begin{aligned}
A_1A_1B_1B_1 \quad & f_1^{(n+1)} = t(G_1)^2 + s[f_1^{(n)} + \tfrac{1}{4}f_3^{(n)} + \tfrac{1}{4}f_7^{(n)} + \tfrac{1}{16}f_9^{(n)}], \\
A_1A_1B_2B_2 \quad & f_2^{(n+1)} = t(G_3)^2 + s[f_2^{(n)} + \tfrac{1}{4}f_3^{(n)} + \tfrac{1}{4}f_8^{(n)} + \tfrac{1}{16}f_9^{(n)}], \\
A_1A_1B_1B_2 \quad & f_3^{(n+1)} = 2t(G_1G_3) + s[\tfrac{1}{2}f_3^{(n)} + \tfrac{1}{8}f_9^{(n)}], \\
A_2A_2B_1B_1 \quad & f_4^{(n+1)} = t(G_4)^2 + s[f_4^{(n)} + \tfrac{1}{4}f_6^{(n)} + \tfrac{1}{4}f_7^{(n)} + \tfrac{1}{16}f_9^{(n)}] \\
A_2A_2B_2B_2 \quad & f_5^{(n+1)} = t(G_2)^2 + s[f_5^{(n)} + \tfrac{1}{4}f_6^{(n)} + \tfrac{1}{4}f_8^{(n)} + \tfrac{1}{16}f_9^{(n)}], \\
A_2A_2B_1B_2 \quad & f_6^{(n+1)} = 2t(G_2G_4) + s[\tfrac{1}{2}f_6^{(n)} + \tfrac{1}{8}f_9^{(n)}], \\
A_1A_2B_1B_1 \quad & f_7^{(n+1)} = 2t(G_1G_4) + s[\tfrac{1}{2}f_7^{(n)} + \tfrac{1}{8}f_9^{(n)}], \\
A_1A_2B_2B_2 \quad & f_8^{(n+1)} = 2t(G_2G_3) + s[\tfrac{1}{2}f_8^{(n)} + \tfrac{1}{8}f_9^{(n)}], \\
A_1A_2B_1B_2 \quad & f_9^{(n+1)} = 2t(G_1G_2 + G_3G_4) + s[\tfrac{1}{4}f_9^{(n)}],
\end{aligned}
\tag{14}
$$

where the gametic frequencies are given by

$$
\begin{aligned}
G_1 &= (A_1B_1) = f_1^{(n)} + \tfrac{1}{2}f_3^{(n)} + \tfrac{1}{2}f_7^{(n)} + \tfrac{1}{4}f_9^{(n)}, \\
G_2 &= (A_2B_2) = f_5^{(n)} + \tfrac{1}{2}f_6^{(n)} + \tfrac{1}{2}f_8^{(n)} + \tfrac{1}{4}f_9^{(n)}, \\
G_3 &= (A_1B_2) = f_2^{(n)} + \tfrac{1}{2}f_3^{(n)} + \tfrac{1}{2}f_8^{(n)} + \tfrac{1}{4}f_9^{(n)}, \\
G_4 &= (A_2B_1) = f_4^{(n)} + \tfrac{1}{2}f_6^{(n)} + \tfrac{1}{2}f_7^{(n)} + \tfrac{1}{4}f_9^{(n)}.
\end{aligned}
$$

Both the R matrix and the D matrix for the two-locus case are taken from the above recursions and R^n is computed using generation matrix methods. As in the one-locus case, standard matrix methods were used to derive the general n-step formulas, which follow:

$$
\begin{aligned}
f_1^{(n)} &= t(G_1)^2 + U(G_1)^2 + 2V(G_1G_3 + G_1G_4) + 2W(G_1G_2 + G_3G_4) \\
&\qquad\qquad + s^n[f_1^{(0)} + Mf_3^{(0)} + Mf_7^{(0)} + Nf_9^{(0)}], \\
f_2^{(n)} &= t(G_3)^2 + U(G_3)^2 + 2V(G_1G_3 + G_2G_3) + 2W(G_1G_2 + G_3G_4) \\
&\qquad\qquad + s^n[f_2^{(0)} + Mf_3^{(0)} + Mf_8^{(0)} + Nf_9^{(0)}], \\
f_3^{(n)} &= t(2G_1G_3) + 2X(G_1G_3) + 2Y(G_1G_2 + G_3G_4) + s^n[(\tfrac{1}{2})^n f_3^{(0)} + Lf_9^{(0)}], \\
f_4^{(n)} &= t(G_4)^2 + U(G_4)^2 + 2V(G_2G_4 + G_1G_4) + 2W(G_1G_2 + G_3G_4) \\
&\qquad\qquad + s^n[f_4^{(0)} + Mf_6^{(0)} + Mf_7^{(0)} + Nf_9^{(0)}], \\
f_5^{(n)} &= t(G_2)^2 + U(G_2)^2 + 2V(G_2G_4 + G_2G_3) + 2W(G_1G_2 + G_3G_4) \\
&\qquad\qquad + s^n[f_5^{(0)} + Mf_6^{(0)} + Mf_8^{(0)} + Nf_9^{(0)}], \\
f_6^{(n)} &= t(2G_2G_4) + X(2G_2G_4) + 2Y(G_1G_2 + G_3G_4) + s^n[(\tfrac{1}{2})^n f_6^{(0)} + Lf_9^{(0)}], \\
f_7^{(n)} &= t(2G_1G_4) + X(2G_1G_4) + 2Y(G_1G_2 + G_3G_4) + s^n[(\tfrac{1}{2})^n f_7^{(0)} + Lf_9^{(0)}], \\
f_8^{(n)} &= t(2G_2G_3) + X(2G_2G_3) + 2Y(G_1G_2 + G_3G_4) + s^n[(\tfrac{1}{2})^n f_8^{(0)} + Lf_9^{(0)}], \\
f_9^{(n)} &= 2t(G_1G_2 + G_3G_4) + 2Z(G_1G_2 + G_3G_4) + (s/4)^n f_9^{(0)},
\end{aligned}
$$

where

$$L = \frac{1}{2}[(\frac{1}{2})^n - (\frac{1}{4})^n],$$
$$M = \frac{1}{2}[1 - (\frac{1}{2})^n],$$
$$N = \frac{1}{4} - (\frac{1}{2})^{n+1} + (\frac{1}{4})^{n+1},$$
$$U = s - s^n,$$

$$V = \frac{s - (2 - s)s^n}{2(2 - s)} + \left(\frac{s}{2}\right)^n \left(\frac{1 - s}{2 - s}\right),$$

$$W = \frac{s(2 + s)}{4(4 - s)(2 - s)} - \frac{s^n}{4} + \left(\frac{s}{4}\right)^n \left(\frac{s - 1}{4 - s}\right) + \left(\frac{s}{2}\right)^n \left(\frac{1 - s}{2 - s}\right),$$

$$X = \frac{2t}{2 - s}\left[\left(\frac{s}{2}\right) - \left(\frac{s}{2}\right)^n\right],$$

$$Y = \frac{st}{(4 - s)(2 - s)} + \left(\frac{s}{4}\right)^n \left(\frac{2t}{4 - s}\right) - \left(\frac{s}{2}\right)^n \frac{t}{2 - s},$$

$$Z = \frac{4t}{4 - s}\left[\left(\frac{s}{4}\right) - \left(\frac{s}{4}\right)^n\right].$$

In many experiments, the initial population is made up solely of double heterozygotes. Hence, the case $f_9{}^{(0)} = 1$ and $G_i = \frac{1}{4}$ for $i = 1, \ldots 4$ holds special interest. Due to the symmetry of the resulting system, the equations are reduced to three: one for the four double homozygotes, one for the four single heterozygotes, and one for the double heterozygote. The n-step formulas from which proportions in the nth generation can be computed are

Double homozygotes $\dfrac{2 + s}{4(4 - s)(2 - s)} - \left(\dfrac{s}{2}\right)^n\left[\dfrac{1}{2(2 - s)}\right] + \left(\dfrac{s}{4}\right)^n\left[\dfrac{3}{4(4 - s)}\right]$

Single heterozygotes $\dfrac{t}{(2 - s)(4 - s)} + \left(\dfrac{s}{2}\right)^n\left[\dfrac{1}{2(2 - s)}\right] - \left(\dfrac{s}{4}\right)^n\left[\dfrac{3}{2(4 - s)}\right]$

Double heterozygote $\dfrac{t}{(4 - s)} + \left(\dfrac{s}{4}\right)^n\left[\dfrac{3}{(4 - s)}\right]$

with limiting values $f_i{}^{(n)}$, $n \to \infty$,

Double homozygotes $(2 + s)/[4(4 - s)(2 - s)]$
Single heterozygotes $t/(2 - s)(4 - s)$
Double heterozygote $t/(4 - s)$.

These methods can be applied to cases involving any number of genes but, in view of rapidly increasing complexity as gene numbers increase, further extension appears hardly worth while unless need arises.

Summary.—A genetic model was constructed which permits direct calculation of expected genotypic frequencies in any generation for populations mating under mixed selfing and random outcrossing. The model is generally applicable to the case of a single diallelic locus and, under certain restrictions about the initial composition of the population, can be extended to diallelic multilocus cases. General n-step formulas are given for the one- and two-locus cases. Applications of the results to other mixed mating systems are discussed.

[1] Jain, S. K., and R. W. Allard, these PROCEEDINGS, **46**, 1371–1377 (1960).

[2] Morley, F. H. W., in *Genetics and Twentieth Century Darwinism*, Cold Spring Harbor Symposia on Quantitative Biology, vol. 24 (1959), pp. 47–56.

[3] Allard, R. W., and S. K. Jain, *Evolution*, **16**, 90–101 (1962).

[4] Allard, R. W., *Crop Sci.*, **1**, 127–133 (1961).

[5] Hayman, B. I., *Heredity*, **7**, 185–192 (1953).

[6] Ali, M., and H. H. Hadley, *Agron. Jour.*, **47**, 589–590 (1955).

[7] Kempthorne, O., *An Introduction to Genetic Statistics* (New York: John Wiley and Sons, 1957).

[8] Fisher, R. A., *The Theory of Inbreeding* (Edinburgh: Oliver and Boyd, 1949).

ON A NEW APPROACH TO THE COMPUTATIONAL SOLUTION OF PARTIAL DIFFERENTIAL EQUATIONS

BY RICHARD BELLMAN, ROBERT KALABA, AND BELLA KOTKIN

THE RAND CORPORATION, SANTA MONICA, CALIFORNIA

1. *Introduction.*—At the present time, there exist several approaches to the numerical solution of partial differential equations, of which the most versatile and frequently employed is that of approximation by means of difference equations. The technique has many advantages: conceptual simplicity, wide applicability, and ready suitability for digital computation. It also has disadvantages: predilection toward instability and frequent requirement for large storage capabilities and excessive computing time, particularly in the treatment of multidimensional equations.

In this note, we wish to indicate a modified approach which may be superior in some situations. It is based upon two themes. The first is that of using a more efficient way of recreating a function than by storing its values at grid points, and the second is the idea that an approximating algorithm should as clearly as possible exhibit the properties of the actual solution. Thus, for example, if the solution is nonnegative, this fact should be evident from the relations used to obtain it computationally. It is to be expected that algorithms with these replicating properties will be more stable than those without these properties. Whether or not algorithms of the desired type always exist is an interesting and unsolved problem at the present time.

2. *The Equation $u_t = uu_x$.*—To illustrate these ideas in a simple setting, let us consider the equation

$$u_t = uu_x, \qquad u(x,0) = g(x), \tag{2.1}$$

which has the great merit of possessing an explicit analytic solution,

$$u = g(x + ut), \tag{2.2}$$

and which displays a "shock" phenomenon. Both of these characteristics are extremely useful for testing computational techniques. Let us assume, for convenience, that $g(x)$ is an odd function, periodic of period 2.

In place of the usual type of difference approximation, we use the relation

$$u(x, t + \Delta) = u(x + u(x,t)\Delta, t), \tag{2.3}$$