Genetic Algebras and Time Continuous Models

IVAR HEUCH*

Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48104

Received December 10, 1971

It is shown how linear genetic algebras, ordinarily applied in situations with discrete time, will also simplify certain systems of differential equations in time continuous models. These models describe the variation of genotype frequencies in infinite populations in different mating systems. The cases considered include matings between individuals randomly drawn from the population at each moment, a population which is continuously backcrossed to a second, constant population, and a population divided into two age groups, which take part in the matings with different intensities. For the first case the general theory is applied to an example with tetraploids having a mixture of chromatid and chromosome segregation.

1. INTRODUCTION

In order to study the behaviour of successive discrete generations in infinite populations under certain mating systems, Etherington (1939b, 1941a) introduced linear commutative nonassociative algebras in theoretical genetics. In such algebras a basis A_0, A_1 ,..., A_n is defined in which each element corresponds to one of the possible genotypes. Then a multiplication table of the form,

$$A_i A_j = \sum_{k=0}^{n} p_{ijk} A_k , \tag{1}$$

is given, where p_{ijk} is the relative frequency of the genotype A_k among the offspring in matings between individuals with genotypes A_i and A_j. (In some special cases it is more natural only to require that the p_{ijk} should be proportional to the frequencies.) Etherington (1939b) calls an algebra baric if it admits a non-trivial homomorphism w to the scalar field, and then w is called the weight function. It is seen that any algebra with a table (1) will be baric with $w(A_i) = 1$ for all i if $\sum_k p_{ijk} = 1$ for all i, j.

The distribution of genotypes in a population may now be specified by an element $\sum_i a_i A_i$ in the algebra defined, where a_i is the relative frequency of

* Present address: Institute of General Genetics, University of Oslo, Norway.

genotype A_i. This element will have weight 1 (but not all elements with weight 1 will give distributions). Contrary to the conventions, for instance in Raffin (1951), only such normalized elements will be interpreted as frequency distributions, and we assume that $\sum_k p_{ijk} = 1$ for all i, j. In situations where certain crosses will give no offspring at all, these restrictions may not be suitable, as for algebras describing sex-linked inheritance (see Gonshor (1960, 1965), Holgate (1970), Heuch (1972a)), but our treatment will not cover such cases, although the methods could be modified to apply here, too. Random mating between individuals in two infinite populations or parts of populations will now give offspring with frequencies found from the product of the two corresponding algebra elements. Various mating systems will be represented by different sequences of products.

Only a particular class of all possible linear algebras arising in the way described will be considered, the genetic algebras (Schafer, 1949; Gonshor, 1971). But this class includes many cases of interest, for example arbitrarily linked loci (Holgate, 1968a) and polyploidy with mutations (Gonshor, 1960, 1965, 1971; Holgate, 1966). An algebra is genetic (Gonshor, 1971) if it is possible to find a transformed basis C_0, C_1 ,..., C_n with multiplication table

$$C_i C_j = \sum_{k=0}^{n} \gamma_{ijk} C_k,\qquad\qquad(2)$$

where

$$\gamma_{000} = 1; \quad \gamma_{0jk} = 0 \quad \text{for} \quad k < j; \quad \gamma_{ijk} = 0 \quad \text{for} \quad k \leqslant \max(i, j), \quad(3)$$

when $i, j \geqslant 1$. The values $1, \gamma_{011}, ..., \gamma_{0nn}$ are the train roots of the algebra, and the basis C_0, C_1 ,..., C_n is said to be canonical. A genetic algebra is baric with $w(C_0) = 1$, $w(C_i) = 0$ for $i \geqslant 1$. In some cases the algebras considered have been shown to be special train algebras, as defined by Etherington (1939b, 1941b), but this implies in particular that they are genetic (Schafer, 1949, Theorem 2).

Several general theorems have been established for genetic algebras in situations with discrete time, in connection with nonoverlapping generations of the same population (Gonshor, 1960; Holgate, 1967), for overlapping generations (Heuch, 1972b), and for migration (Holgate, 1968b). The method of transformation to obtain a canonical basis is in the discrete case equivalent to the introduction of new variables to simplify certain systems of difference equations. We will show that the introduction of such a basis also makes it easier to solve differential equations in the time continuous case. Thus, the same substitution of variables can often be applied in both situations.

A connection between a linear algebra in genetics and the corresponding time continuous differential equations has been noted before by Andreoli (1960) for one autosomal locus in diploid organisms. Such connections between methods

used in discrete and continuous situations have also been pointed out when the algebraic treatment has not included linear algebras, as by Bennett (1954) for linked loci.

It would often be sufficient to consider a scalar field consisting of the real numbers. However, sometimes the transformation to canonical basis is not possible within this field (Heuch, 1972b), and the algebraic treatment of particular mating systems may also necessitate the use of complex numbers (as in Section 6). Therefore, all algebras are assumed to be defined over the field of complex numbers. In the final expressions for frequency distributions the imaginary parts may then be disregarded. Only the basis transformation technique will be used here, but it is also easy to extend the differential operator method applied in (Reiersøl, 1962) to time continuous situations.

2. DISTRIBUTION PATTERNS

Let B_0, B_1,..., B_n be a basis for the algebra $\mathscr{A}$, and consider functions G over the set R^+ of nonnegative real numbers with image in $\mathscr{A}$. They may be written as

$$G(t) = \sum_{i=0}^{n} b_i(t) B_i, \qquad t \geqslant 0,$$

where b_0, b_1,..., b_n are scalar-valued functions over R^+. The collection of all functions G will be designated as $F(\mathscr{A})$. If b_0, b_1,..., b_n all are differentiable a certain number of times, then G will also be said to be differentiable, and we define

$$\frac{d^m G(t)}{dt^m} = G^{(m)}(t) = \sum_{i=0}^{n} b_i^{(m)} B_i, \qquad m = 1, 2,... .$$

This definition is independent of the choice of basis.

Assume that $\mathscr{A}$ is baric. Let P be a function from $\mathscr{A}$ to $F(\mathscr{A})$ such that the element G in $F(\mathscr{A})$ is assigned to G_0 in $\mathscr{A}$. Then P will be called a distribution pattern if (i) $G(0) = G_0$ for all G_0 in $\mathscr{A}$ and (ii) $w(G(t)) = 1$ for all $t \geqslant 0$ whenever $w(G_0) = 1$. These patterns will correspond to particular ways of constructing sequences $G^{[n]}$ from elements $G^{[1]}$ in the discrete case, for instance the way of forming the plenary sequence $G^{[n]}$ (Etherington, 1939b) by the rule $G^{[n]} = (G^{[n-1]})^2$ from any element $G^{[1]}$ in $\mathscr{A}$. If $w(B_i) = 1$ for all i, we have in general

$$d(w(G(t)))/dt = \sum_{i=0}^{n} b_i'(t) = w(G'(t)).$$

Thus, if $G(t)$ is obtained from a distribution pattern, $w(G^{(m)}(t)) = 0$ for $m \geqslant 1$ if $w(G(0)) = 1$, when $G^{(m)}$ exists. In Section 6 we consider two P-functions at the same time, producing U and V in $F(\mathscr{A})$, dependent on each other. It is here natural to call these functions P distribution patterns if (i) $U(0) = U_0$, $V(0) = V_0$ and (ii) $w(U(t)) = w(V(t)) = 1$ for all t whenever both $w(U_0) = 1$ and $w(V_0) = 1$.

For a particular distribution pattern P, consider all G found when P operates on elements G_0 in $\mathscr{A}$ with $w(G_0) = 1$. P will be called a continuous train if all such G are differentiable a certain number of times and satisfy the same differential equation,

$$G^{(s)}(t) + \theta_1 G^{(s-1)}(t) + \cdots + \theta_{s-1} G'(t) + \theta_s G(t) = 0, \tag{4}$$

with constant coefficients $\theta_1, \theta_2, ..., \theta_s$. Taking the weight on each side of (4), we see that $\theta_s = 0$. This equation may be solved in the same way as differential equations of real functions. We assume that no such equation of lower order exists. The solutions of the characteristic equation for (4) will be called the train roots for P. When these are known, we may write a general expression for the solution of (4):

$$G(t) = \sum_{j=0}^{d} Q_j(t) \exp(y_j t), \tag{5}$$

where $y_0 = 0$, $y_1, ..., y_d$ are the different values of these roots, and $Q_0(t)$, $Q_1(t), ..., Q_d(t)$ are elements in $\mathscr{A}$ and polynomials in t. We could have used (5) in our definition of a continuous train, but (4) has been preferred to stress the correspondence with the definition of trains in the discrete case (Etherington, 1939b).

Now assume that $\mathscr{A}$ is genetic with $C_0, C_1, ..., C_n$ as canonical basis, and that P is a continuous train. When $w(G(0)) = 1$ we can then write

$$G(t) = C_0 + \sum_{k=1}^{n} c_k(t) C_k, \tag{6}$$

where

$$c_k(t) = \sum_{j=0}^{f} q_{kj}(t) \exp(r_{kj} t). \tag{7}$$

All $q_{kj}(t)$ are usual polynomials in t. The scalars $r_{k0}, r_{k1}, ..., r_{kf}$ will be called the roots for the basis element C_k with respect to P. Each $c_k(t)$ will satisfy a homogeneous differential equation with characteristic equation having $r_{k0}, r_{k1}, ..., r_{kf}$ as roots.

3. THE PLENARY PATTERN

Mating systems will be considered for a given infinite population where the frequency distribution at time $t = 0$ is specified by the algebra element G_0. The distributions at later points of time will give a function $G(t)$, and the distribution pattern itself connecting each $G_0 = G(0)$ with $G(t)$ will represent the mating system.

We first construct a pattern corresponding to random mating between individuals in the same population. Set $G(t) = \sum_i a_i(t) A_i$, where $a_i(t)$ is the frequency of genotype A_i at the time t. In the time interval $(t, \; t + h)$ a fraction $h + o(h)$ of the population will die and be replaced, except for a quantity $o(h)$, by offspring from random mating between individuals living at time t. With the multiplication table (1) a proportion $\sum_{i,j} a_i(t) a_j(t) p_{ijk}$ of this offspring will have genotype A_k. Thus, we get

$$a_k(t + h) = (1 - h) a_k(t) + h \sum_{i,j} a_i(t) a_j(t) p_{ijk} + o(h).$$

This equation and the corresponding one for the time interval $(t - h, t)$ then give in the usual way

$$a_k'(t) = -a_k(t) + \sum_{i,j} a_i(t) a_j(t) p_{ijk}. \tag{8}$$

After multiplication by A_k and summation over k, we finally obtain

$$G'(t) = -G(t) + G(t)^2. \tag{9}$$

The weight $w(t) = w(G(t))$ must satisfy the separable equation

$$w'(t) = -w(t) + w(t)^2,$$

which has $w(t) = 1$ as the only solution with $w(0) = 1$. Hence, in general, the function P assigning the particular solution of (9) with $G(0) = G_0$ to the element G_0 is really a distribution pattern. Corresponding to the discrete sequence of powers this pattern will be called the plenary pattern. The assumptions underlying a population model of this kind may not be very realistic. As pointed out by Moran (1962), such models imply that the lifetime distribution is exponential and the offspring distribution for each individual geometric.

Now assume that the algebra $\mathscr{A}$ used is genetic with canonical basis $C_0, C_1, ..., C_n$. The train roots $1, \gamma_{011}, ..., \gamma_{0nn}$ in $\mathscr{A}$ are written as $1, \lambda_1, ..., \lambda_n$. Then we have the following.

THEOREM 1. (i) *The plenary pattern in $\mathscr{A}$ is a continuous train.*

  (ii) C_0 *has the root 0.*

(iii)   *For $k \geqslant 1$, the roots of C_k will be $-(1 - 2\lambda_k)$ and all sums of pairs of roots for C_i and C_j with $0 \leqslant i, j < k$ and $\gamma_{ijk} \neq 0$.*

(iv)   *The set of roots for $\mathcal{A}$ consists of all roots for the basis elements.*

Proof. We consider a fixed initial value $G(0) = G_0$ with weight 1 and express $G(t)$ as in (6). It will be proved by induction with respect to k that all $c_k(t)$ may be written in the form (7), with r_{kj} determined by (ii) and (iii) previously. Since $c_0(t) = 1$ this is evidently true for $k = 0$. Now assume that it is correct for $k \leqslant g - 1$. Inserting (6) in (9) and considering the coefficients of C_g, we then get

$$c_g'(t) = -(1 - 2\lambda_g)\, c_g(t) + \sum_{i=0}^{g-1} \sum_{j=0}^{g-1} \gamma_{ijg} c_i(t)\, c_j(t), \tag{10}$$

where we have applied (2) and (3). It is known that $c_i(t)$ and $c_j(t)$ may be expressed as in (7) and, thus, (10) may be regarded as a usual inhomogeneous differential equation for $c_g(t)$, with the last term on the right side equal to

$$\sum_{i,j<g} \sum_{m,s} k_{ijms}(t) \exp((r_{im} + r_{js})t),$$

where $k_{ijms}(t)$ is a polynomial in t. Hence, (10) may be solved in the ordinary way, giving an expression (7) for $c_g(t)$, with quantities r_{gj} obtained in accordance with (iii). The unknown constant in the solution is found using the initial value $c_g(0)$, and the induction is finished. All expressions (7) may now be put together to an expression (5) for $G(t)$, and so the plenary pattern is a train.

In special cases not all roots given are necessary. The proof follows the same lines as the one given for Theorem 5.1 in (Heuch, 1972b) for the discrete case. Rules for the multiplicities of the roots in our Theorem 1 may also be found, corresponding to those given by the discrete theorem for the sequence of plenary powers. The terms $\exp(-(1 - 2\lambda_k)t)$ obviously correspond to terms $(2\lambda_k)^n$ for this sequence. There is, however, one difference between the two situations. Knowledge of the roots for the complete algebra with multiplicities in addition to $G(0)$, will not be sufficient for constructing the final expression (5) in the continuous case, since this contains too many unknown algebra elements included in the polynomials $Q_j(t)$. In the discrete situation it is usually possible from a finite number of multiplications to find values $G^{[n]}$ for a sufficient number of different n to determine the constants, but no similar easy method exists in the continuous case. However, the inductive method applied in the proof for Theorem 1 may be used to find all $c_k(t)$ recursively, since one unknown constant is found at each step during this process.

Suppose that $\mathrm{Re}(\lambda_k) < 1/2$ for $k \geqslant 1$. Then it is seen from (iii) that $Q_0(t)$ in (5) corresponding to $y_0 = 0$ will never contain any terms in t. All other y_j will have negative real parts, and so $G(t)$ will converge to a limit $\lim G(t) = Q_0(t)$ when

$t \to \infty$. Furthermore, $G'(t) \to 0$, and from (9) it follows that $(\lim G(t)) = (\lim G(t))^2$. According to Gonshor (1960, Theorem 2.1) only one such nonzero idempotent exists in the algebra in this case. This will give a stable genotype distribution.

The method of constructing $G(t)$ recursively may easily be extended to the case where (9) is replaced by $G'(t) = -G(t) + (G(t))^s$, where $(G(t))^s$ is any nonassociative power of $G(t)$, given by the shape s (Etherington, 1939a). This corresponds to the generalization from plenary powers to an arbitrary sequence of powers, treated in (Heuch, 1972b).

4. An Example with Tetraploidy

Consider one locus with two alleles A_1 and A_2 in autotetraploids. The population will be described by the distribution of diploid gametes that fused to form the actual tetraploid individuals. The multiplication rule will be defined as if mating consisted of copulation of diploid individuals and immediate meiosis. In an infinite population with random mating this will not give an entirely wrong picture. The population at time t will, thus, be given by

$$G(t) = a_{11}(t)\, A_1A_1 + a_{12}(t)\, A_1A_2 + a_{22}(t)\, A_2A_2\,, \tag{11}$$

where A_1A_1, A_1A_2, A_2A_2 is the (original) basis in the appropriate algebra. We will allow any mixture of chromosome and chromatid segregation, giving weight q to the chromatid component and $1 - q$ to the chromosome one. As the parameter describing this intermediate situation Fisher and Mather (1943) used the probability α for a gamete to contain genes from the same chromosome in the zygote producing it. Then $q = 7\alpha$.

The multiplication table corresponding to (1) will be

$$(A_iA_j)(A_kA_m) = (7 - q)/42 \cdot (A_iA_j + A_iA_k + A_iA_m + A_jA_k$$
$$+ A_jA_m + A_kA_m) + q/28 \cdot (A_iA_i + A_jA_j + A_kA_k + A_mA_m),$$

$i, j, k, m = 1, 2$. Holgate (1966) has shown that introduction of the basis C_0, C_1, C_2 given by

$$C_0 = A_1A_1\,, \qquad C_1 = A_1A_1 - A_1A_2\,, \qquad C_2 = A_1A_1 - 2A_1A_2 + A_2A_2\,,$$

produces a multiplication table

$$C_0{}^2 = C_0\,, \qquad C_0C_1 = 1/2C_1 - q/28C_2\,,$$
$$C_0C_2 = C_1{}^2 = (1/6 - q/42)\,C_2\,, \qquad C_1C_2 = C_2{}^2 = 0.$$

Thus, we have a genetic algebra with train roots 1, 1/2, and $1/6 - q/42$.

According to Theorem 1 the plenary pattern will be a continuous train, with C_0 and C_1 both having the root 0, and C_2 the roots 0 and $-(2/3 + q/21)$. Writing $G(t) = C_0 + c_1(t) C_1 + c_2(t) C_2$ for the normalized element, Eq. (10) is found to be $c_1'(t) = 0$ for $g = 1$, giving $c_1(t) = c_1(0)$, and

$$c_2'(t) = -(2/3 + q/21) c_2(t) - q/14 \cdot c_1(0) + (1/6 - q/42) c_1(0)^2$$

for $g = 2$, giving

$$c_2(t) = c_2(0) \exp(-(2/3 + q/21)t)$$
$$+ ((7 - q) c_1(0)^2 - 3qc_1(0))/(28 + 2q) \cdot [1 - \exp(-(2/3 + q/21)t].$$

Here $c_1(0) = -a_{12}(0) - 2a_{22}(0)$ and $c_2(0) = a_{22}(0)$. Hence, the frequency distribution at time t is given by (11) with $a_{11}(t) = 1 + c_1(0) + c_2(t)$, $a_{12}(t) = -c_1(0) - 2c_2(t)$, $a_{22}(t) = c_2(t)$.

It is seen that $G(t)$ converges when $t \to \infty$. Introducing $\alpha = q/7$ and noting that $c_1(0) = -2p(A_2)$, where $p(A_2)$ is the (initial) frequency of the gene A_2, we find that the limit is identical with the one given for discrete generations in (Geiringer, 1949, Eq. (23)).

5. THE PRINCIPAL PATTERN

We will consider a population where the individuals dying in the time interval $(t, t + h)$ are replaced by offspring of matings between individuals in this population at time t and individuals in a constant population with the same genotype distribution as we originally had at time $t = 0$. If we use the same notation as in Section 3, we find

$$a_k'(t) = -a_k(t) + \sum_{i,j} a_i(t) a_j(0) p_{ijk} \tag{12}$$

corresponding to (8) and

$$G'(t) = -G(t) + G(t) G(0) \tag{13}$$

corresponding to (9). Equation (13) implies that $d(w(G(t)))/dt = 0$ if $w(G(0)) = 1$, and, thus, with $G(0) = G_0$ (13) defines a distribution pattern. This will be called the principal pattern, since it is analogous with the sequence of principal powers (Etherington, 1939b) in the discrete case. For this pattern we now have the following.

THEOREM 2. *The principal pattern is a continuous train in a genetic algebra $\mathscr{A}$. The roots for $\mathscr{A}$ are all values 0, $-(1 - \lambda_1),..., -(1 - \lambda_n)$.*

This is easily shown in the same way as Theorem 1, with

$$c_g'(t) = -(1 - \lambda_g) c_g(t) + \sum_{i=0}^{g-1} \sum_{j=0}^{g-1} \gamma_{ijg} c_i(t) c_j(0) + \lambda_g c_g(0),$$

instead of (10). The roots for a particular basis element C_k will be $-(1 - \lambda_k)$ and all roots for C_i with $i < k$ where $\gamma_{ijk} \neq 0$ for at least one $j < k$. If $\mathrm{Re}(\lambda_k) < 1$ for all $k \geqslant 1$, then when $t \to \infty$ $G(t)$ will converge to a limit satisfying $\lim G(t) = (\lim G(t)) G(0)$.

It is easy for the principal pattern to give an example of an algebra where the pattern is no train. Consider the algebra with basis A_0, A_1 and multiplication table $A_0^2 = A_1^2 = A_0$, $A_0 A_1 = A_1$. This would correspond to a situation with two genotypes A_0 and A_1 where a mating between individuals of the same kind would produce A_0 individuals only, while the result of matings between different individuals would always be A_1. Even if this case is not likely to be found in the real world, it will serve as a theoretical example. This algebra is not a train algebra as defined in (Etherington, 1939b), since its rank equation (and characteristic equation) is found to be

$$G^3 - 2a_0 G^2 + (a_0^2 - a_1^2)G = 0,$$

for the general element $G = a_0 A_0 + a_1 A_1$, and the coefficients in this equation are not given as functions of the weight $w(G) = a_0 + a_1$. Then, according to (Schafer, 1949, Theorem 1), it cannot be a genetic algebra either.

In this case Eqs. (12) are

$$a_0'(t) = -a_0(t) + a_0(t) a_0(0) + a_1(t) a_1(0),$$
$$a_1'(t) = -a_1(t) + a_0(t) a_1(0) + a_1(t) a_0(0).$$

When $w(G(0)) = 1$, the solutions are found to be given by

$$\begin{aligned} G(t) &= a_0(t) A_0 + a_1(t) A_1 \\ &= (A_0 + A_1)/2 + (G(0) - (A_0 + A_1)/2) \exp(-2a_1(0)t). \end{aligned}$$

Thus, $G(t)$ satisfies the equation $G''(t) + 2a_1(0) G'(t) = 0$, but the principal pattern is still no train, since the coefficient $2a_1(0)$ depends on which element $G(0)$ with weight 1 was considered. It is seen that $G(t)$ will converge when $t \to \infty$, but an essential difference from what happens in genetic algebras is that the rate of convergence will not be the same for all initial values representing distributions. (The polynomials in (5) may depend on the initial value even in genetic algebras, but the exponential terms will not.)

For other patterns, as the plenary one, the system of equations analogous to (12) may be far more difficult to solve in the nongenetic case.

6. Patterns for Overlapping Generations

The plenary pattern represents a situation with overlapping generations, but it imposes heavy restrictions on the mating system, since the parents of each new individual are drawn at random from the population at the time of the birth. The following time continuous model corresponds more closely to some of the cases treated in (Heuch, 1972b, Section 5) with discrete sequences. Let the total population be divided into two classes K_U and K_V. In the small time interval $(t, t + h)$ a fraction h of the class K_U is transferred to class K_V, and is replaced by offspring produced by matings of which a proportion b_{11} are between individuals in class K_U, a proportion $2b_{01}$ between one K_U and one K_V individual, and a proportion b_{00} between two K_V individuals ($b_{00} + 2b_{01} + b_{11} = 1$). During the same time interval a fraction h of group K_V dies and is replaced by individuals from K_U. A disadvantage with this model is the introduction of the two age classes, when the process of aging really ought to be continuous.

Separate genotype frequency distributions will be used for the two classes, corresponding to algebra elements $U(t)$ and $V(t)$ at time t, each with weight 1. In the same manner as we found (9), we get

$$U'(t) = -U(t) + b_{11}U(t)^2 + 2b_{01}U(t)\,V(t) + b_{00}V(t)^2, \qquad (14)$$

$$V'(t) = -V(t) + U(t). \qquad (15)$$

It follows from (15) that $V''(t)$ exists and that $U(t) = V'(t) + V(t)$, $U'(t) = V''(t) + V'(t)$. Inserting this in (14) we find

$$V''(t) + 2V'(t) + V(t) = b_{11}V'(t)^2 + 2bV'(t)\,V(t) + V(t)^2, \qquad (16)$$

with $b = b_{11} + b_{01}$.

Now consider any normalized $U(t)$ and $V(t)$, not necessarily representing distributions. Expressing $V(t)$ as a linear combination of the elements in a canonical basis $C_0, C_1, ..., C_n$ in a genetic algebra,

$$V(t) = C_0 + \sum_{k=1}^{n} v_k(t)\, C_k,$$

we then find that such a solution of (16) may be constructed recursively in the same way as for (9) or (13), and that each $v_k(t)$ may be written as $c_k(t)$ in (7). Suppose that $v_1(t), ..., v_{g-1}(t)$ have been found. Then, equating the coefficients of C_g in (16), we get

$$v_g''(t) + 2(1 - b\lambda_g)\,v_g'(t) + (1 - 2\lambda_g)\,v_g(t)$$
$$= \sum_{i=0}^{g-1}\sum_{j=0}^{g-1} \gamma_{ijg}[b_{11}v_i'(t)\,v_j'(t) + 2bv_i(t)\,v_j'(t) + v_i(t)\,v_j(t)].$$

The right side is here known as a linear combination $\sum_m k_m(t) \exp(r_m t)$, and the second order differential equation may be solved by the standard procedure. The characteristic roots for the homogeneous equation found by omitting the right side are

$$b\lambda_g - 1 \pm (2(1 - b)\lambda_g + b^2\lambda_g{}^2)^{1/2}. \tag{17}$$

Thus, $v_g(t)$ may be written as

$$v_g(t) = \sum_{j=1}^{f} q_{gj}(t) \exp(r_{gj} t),$$

with the quantities r_{gj} found from (17) and as sums of r-values for pairs $v_i(t)$, $v_j(t)$ belonging to C_i, C_j with i, $j < g$ and $\gamma_{ijg} \neq 0$. If the initial values $U(0)$ and $V(0)$ are known, then we also know $V'(0) = -V(0) + U(0)$, and the two unknown constants in the solution for $v_g(t)$ may be determined.

Eventually all $v_k(t)$ are found in this fashion. Since it really is possible for all normalized initial elements $U(0) = U_0$, $V(0) = V_0$ to construct a $V(t)$ with weight 1 satisfying (16), it may be concluded that there is a distribution pattern giving such a $V(t)$ in all genetic algebras, and that this pattern forms a train, where the roots are found in an obvious way. As $U(t) = V'(t) + V(t)$, there must be a train giving $U(t)$, too, and an explicit expression is obtained from the one for $V(t)$.

Even with only real train roots λ_g, (17) may give complex r-values if some train root is contained in the interval $(-2(1 - b)/b, 0)$. This will give rise to oscillations in the frequency distributions. With the algebra for linked loci and $b_{00} = b_{11} = 0$, $b_{01} = \frac{1}{2}$, $U(t)$ and $V(t)$ may also be used to construct the distributions for sex-linked genes in the X-chromosome in a mating situation corresponding to the one for the plenary pattern. $U(t)$ gives the distribution for the homogametic sex, $V(t)$ for the heterogametic. This corresponds to the sequence $H(n) = H(n - 1)H(n - 2)$ applied by Reiersøl (1962) for this purpose with discrete generations.

REFERENCES

ANDREOLI, G. 1960. Algebre non associative e sistemi differenziati di Riccati in un problema di Genetica, *Ann. Mat. Pura Appl.* **49,** 97–116.

BENNETT, J. H. 1954. On the theory of random mating, *Ann. Eugen.* **18,** 311–317.

ETHERINGTON, I. M. H. 1939a. On non-associative combinations, *Proc. Roy. Soc. Edinburgh* **59,** 153–162.

ETHERINGTON, I. M. H. 1939b. Genetic algebras, *Proc. Roy. Soc. Edinburgh* **59,** 242–258.

ETHERINGTON, I. M. H. 1941a. Non-associative algebra and the symbolism of genetics, *Proc. Roy. Soc. Edinburgh B* **61,** 24–42.

ETHERINGTON, I. M. H. 1941b. Special train algebras, *Quart. J. Math. Oxford* **12,** 1–8.

FISHER, R. A. AND MATHER, K. 1943. The inheritance of style length in Lythrum salicaria, *Ann. Eugen.* 12, 1–23.

GEIRINGER, H. 1949. Chromatid segregation of tetraploids and hexaploids, *Genetics* 34, 665–684.

GONSHOR, H. 1960. Special train algebras arising in genetics, *Proc. Edinburgh Math. Soc.* 12, 41–53.

GONSHOR, H. 1965. Special train algebras arising in genetics II, *Proc. Edinburgh Math. Soc.* 14, 333–338.

GONSHOR, H. 1971. Contributions to genetic algebras, *Proc. Edinburgh Math. Soc.* 17, 289–298.

HEUCH, I. 1972a. k loci linked to a sex-factor in haploid individuals, *Biomet. Z.* 13, 57–68.

HEUCH, I. 1972b. Sequences in genetic algebras for overlapping generations, *Proc. Edinburgh Math. Soc.* 18, 19–29.

HOLGATE, P. 1966. Genetic algebras associated with polyploidy, *Proc. Edinburgh Math. Soc.* 15, 1–9.

HOLGATE, P. 1967. Sequences of powers in genetic algebras, *J. London Math. Soc.* 42, 489–496.

HOLGATE, P. 1968a. The genetic algebra of k linked loci, *Proc. London Math. Soc.* 18, 315–327.

HOLGATE, P. 1968b. Interaction between migration and breeding studied by means of genetic algebras, *J. Appl. Probability* 5, 1–8.

HOLGATE, P. 1970. Genetic algebras associated with sex linkage, *Proc. Edinburgh Math. Soc.* 17, 113–120.

MORAN, P. A. P. 1962. "The Statistical Processes of Evolutionary Theory," Oxford Univ. Press (Clarendon), London.

RAFFIN, R. 1951. Axiomatisation des algèbres génétiques, *Acad. Roy. Belg. Bull. Cl. Sci.* 37, 359–366.

REIERSØL, O. 1962. Genetic algebras studied recursively and by means of differential operators, *Math. Scand.* 10, 25–44.

SCHAFER, R. D. 1949. Structure of genetic algebras, *Amer. J. Math.* 71, 121–135.