

GENETIC ALGEBRAS ASSOCIATED WITH SEX LINKAGE

by P. HOLGATE

(Received 2nd August 1969)

1. Introduction

Most work in genetic algebras has been concerned with inheritance which is symmetric with respect to sex, in that the characters studied are determined by genes located at autosomal loci, and it is assumed that the segregation pattern is the same in males and females. When asymmetric situations are studied, the development of the theory is complicated by the higher dimensions of the algebras, and by a feature to which Etherington (3, p. 40) drew attention, namely the fact that the passage from the gametic to the zygotic algebra no longer quite corresponds to the process of duplication, as it does in the symmetric case. Etherington gave some results for the gametic and zygotic algebras of a single sex linked diallelic locus, and its properties were discussed further by Gonshor (4, p. 44). In a second paper (5, p. 334) Gonshor studied sex linkage in the case of multiple alleles, choosing a canonical basis which exhibited very clearly the multiplication table and ideal structure of the algebra. His treatment from the statement of the multiplication table in terms of the natural basis to its expression in terms of a canonical basis, is repeated in the displayed relations (4)-(8) below, for completeness and to establish the present notation.

In another direction, Reiersøl (10, pp. 33, 40) showed that a certain recurrent sequence $H(n)$ which could be but was not necessarily a sequence of non-associative powers, determined the distribution of a sex-linked character in successive generations, playing a rôle analogous to that of the plenary powers (3, p. 28) in the symmetric situation.

In the next two sections I study a single multiallelic locus, completely or partially linked to the sex determining locus, and obtain canonical forms for the gametic and zygotic algebras. In § 4 I then examine briefly the process of passing from the zygotic to the copular algebra.

2. Complete sex linkage. The gametic and zygotic algebras

Consider first the gametic algebra \mathcal{G} of a single locus with $r+1$ alleles which I have called an elementary algebra (9, p. 316). In terms of its natural basis $\{a_0, \dots, a_r\}$ whose elements correspond to the actual allelic forms, its multiplication table is

$$a_i a_j = \frac{1}{2}(a_i + a_j). \quad (1)$$

E.M.S.—H

$$\begin{aligned} c_0 &= a_0 \\ c_i &= a_0 - a_i, \quad i \neq 0 \end{aligned} \tag{2}$$

the multiplication table is

$$\left. \begin{aligned} c_0^2 &= c_0 \\ c_0 c_i &= \frac{1}{2} c_i \\ c_i c_j &= 0, \quad i, j \neq 0. \end{aligned} \right\} \tag{3}$$

Except on two occasions where it is explicitly mentioned, the convention will be followed that letters appearing as subscripts to a may take all integral values from 0 to n , while those appearing as subscripts to c are restricted to the range 1 to n .

The sex-linked zygotic algebra (\mathcal{L}) is defined to be the algebra whose natural basis consists of (i) the elements a_0, \dots, a_n , and (ii) the unordered pairs (a_i, a_j) , which may be written for brevity a_{ij} , and whose multiplication table is

$$\begin{aligned} a_i a_j &= a_{ij} a_{kl} = 0 \\ a_i a_{kl} &= \frac{1}{4} (a_k + a_l + a_{ik} + a_{il}) \end{aligned} \tag{4}$$

The vector space underlying \mathcal{L} is thus the direct sum of spaces equivalent to those underlying \mathcal{G} , and its commutative duplicate \mathcal{G}' . (A prime denotes duplication as defined in (2)). In terms of human genetics the basis of the first subspace represents possible male zygotic types, and that of the second, possible female types. The relations (4) express the fact that two individuals of the same sex cannot produce offspring.

Now let the canonical basis for \mathcal{G} given by (2), and that given by the induced transformation of the basis of \mathcal{G}' , namely

$$\left. \begin{aligned} c_{00} &= a_{00} \\ c_{0i} &= a_{00} - a_{0i} \\ c_{ij} &= a_{00} - a_{0i} - a_{0j} + a_{ij} \end{aligned} \right\} \tag{6}$$

be taken as a basis for \mathcal{L} . In view of (4),

$$c_i c_j = c_{ij} c_{kl} = 0, \text{ all } i, j, k, l, \tag{7}$$

while from (4) and (5) the rest of the multiplication table is given by

$$\begin{array}{ccc} & c_{00} & c_{0j} & c_{jk} \\ \begin{array}{l} c_0 \\ c_i \end{array} & \left[\begin{array}{ccc} \frac{1}{2}(c_0 + c_{00}) & \frac{1}{4}(c_j + c_{0j}) & 0 \\ \frac{1}{2}c_{0i} & \frac{1}{4}c_{ij} & 0 \end{array} \right] & . \end{array} \tag{8}$$

It is of interest to note that as a consequence of the extremely simple structure of elementary algebras, the same multiplication table would be obtained if

instead of (2) the new basis were defined by any transformation of \mathcal{G} given by

$$c_0 = \sum x_{0j} a_j$$

$$c_i = \sum x_{ij} a_j$$

with $\sum x_{0j} = 1$, $\sum x_{ij} = 0$ ($i \neq 0$) and the transformation which this induces in \mathcal{G}' .

Before continuing with the study of \mathcal{L} , a concept will be introduced which characterises the equal division of offspring between the sexes.

Definition 1. The algebra \mathcal{S} with basis elements m and w , and commutative multiplication table

$$m^2 = w^2 = 0, mw = \frac{1}{2}(m + w)$$

will be called the sex differentiation algebra.

In terms of a new basis $p = m + w$, $q = m - w$ the multiplication table is

$$p^2 = p, q^2 = -p, pq = 0.$$

Thus the ideal \mathcal{S}^2 is isomorphic to the coefficient field.

Definition 2. An algebra will be said to be dibaric if it admits a homomorphism onto \mathcal{S} .

Proposition 1. *If an algebra \mathcal{A} is dibaric. then \mathcal{A}^2 is baric.*

Proof. Since S is a homomorphic image of A , S^2 is a homomorphic image of A^2 . But S^2 is isomorphic to the coefficient field.

Proposition 2. *The zygotic sex-linked algebra of a single $r + 1$ allelic locus, \mathcal{L} , is dibaric.*

Proof. The required correspondence may be defined in terms of the basis by taking

$$c_0 \rightarrow m; c_{00} \rightarrow w; c_i, c_{0j}, c_{kl} \rightarrow 0.$$

Reference to (8) shows that this is a homomorphism.

It follows that \mathcal{L}^2 is a baric algebra. The elements occurring in the cells of (8) provide a basis, but if a further transformation is applied by taking

$$e_0 = c_0 + c_{00},$$

$$u_i = c_i + 2c_{0i},$$

$$v_i = c_i - c_{0i},$$

the multiplication table takes the form

	e_0	u_k	v_k	c_{kl}	
e_0	e_0	$\frac{1}{2}u_k$	$-\frac{1}{4}v_k$	0	
u_i		c_{ik}	$\frac{1}{4}c_{ik}$	0	
v_i			$-\frac{1}{2}c_{ik}$	0	
c_{ij}				0	.

(9)

The sequence of principal powers of the nil ideal of \mathcal{L}^2 is

$$\begin{aligned} \mathcal{N}(\mathcal{L}^2) &= \{u_i, v_i, c_{ij}\} \\ \mathcal{N}^2(\mathcal{L}^2) &= \{c_{ij}\} \\ \mathcal{N}^3(\mathcal{L}^2) &= \{0\} \end{aligned}$$

all of which are ideals. Hence \mathcal{L}^2 is a special train algebra, and it can be seen from (9) that its dimension is $1 + 2r + \frac{1}{2}r(r+1)$, and its train roots are $1, \frac{1}{2}$ (r times), $-\frac{1}{4}$ (r times) and 0 ($\frac{1}{2}r(r+1)$ times). The principal train equation can also be seen from (9) to be

$$x(x-1)(x-\frac{1}{2})(x+\frac{1}{4}) = 0.$$

The plenary train equation may be obtained using the methods developed in (7). Let E be the squaring operator defined by $Ex = x^2$. Then for a general element of \mathcal{L} of unit weight

$$\begin{aligned} E(e_0 + \sum_i \alpha_i u_i + \sum_i \beta_i v_i + \sum_i \gamma_{ii} c_{ii} + 2 \sum_{i < j} \sum_{i < j} \gamma_{ij} c_{ij}) \\ = e_0 + \sum_i \alpha_i u_i - \frac{1}{2} \sum_i \beta_i v_i + \sum_i (\alpha_i^2 - \frac{1}{2} \beta_i^2 + \frac{1}{2} \alpha_i \beta_i) c_{ii} \\ + 2 \sum_{i < j} \sum_{i < j} (\alpha_i \alpha_j - \frac{1}{2} \beta_i \beta_j + \frac{1}{4} \alpha_i \beta_j + \frac{1}{4} \beta_i \alpha_j) c_{ij}. \end{aligned}$$

The set of coordinates $\alpha_i, \beta_j, \gamma_{ij}$ can be supplemented with a set of quadratic functions of them, namely the set $\alpha_i \alpha_j, \alpha_i \beta_j, \beta_i \alpha_j, \beta_i \beta_j$ for $i = 1, \dots, n; j = 1, \dots, n$ and the operation of E represented by a matrix multiplication given essentially after expansion of the vector replacing suffixed components by all possible values, by

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{4} & -\frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} \alpha_i \\ \beta_i \\ \alpha_i \alpha_j \\ \alpha_i \beta_j \\ \beta_i \alpha_j \\ \beta_i \beta_j \\ \gamma_{ij} \end{bmatrix} = \begin{bmatrix} \alpha_i \\ -\frac{1}{2} \beta_i \\ \alpha_i \alpha_j \\ -\frac{1}{2} \alpha_i \beta_j \\ -\frac{1}{2} \beta_i \alpha_j \\ \frac{1}{4} \beta_i \beta_j \\ \alpha_i \alpha_j + \frac{1}{4} \alpha_i \beta_j + \frac{1}{4} \beta_i \alpha_j - \frac{1}{2} \beta_i \beta_j \end{bmatrix}.$$

The minimal polynomial of the matrix on the left, even when expanded, is $E(E-1)(E+\frac{1}{2})(E-\frac{1}{4})$, and hence the plenary train equation of \mathcal{L} is

$$x^{[2]}[x-1][x+\frac{1}{2}][x-\frac{1}{4}] = 0.$$

3. Partial sex linkage

The phenomenon discussed here exists in some mammals, but although still a subject of controversy, its existence in man appears doubtful (11, p. 251).

In this situation $(r + 1)^2$ types of male zygotes have to be considered, labelled $a_{i(j)}$ where the first suffix indicates the allele linked to the X gene, and the bracketed suffix that linked to the Y gene. If crossing over occurs between the locus being studied and the sex-determining locus in a proportion θ of the cases, then $a_{i(j)}$ will behave like a_i of the previous section in a proportion $1 - \theta$ of cases, and like a_j in the remaining θ of them. Thus the multiplication table for the zygotic algebra of partial sex linkage, which will be denoted by \mathcal{L}_θ , is

$$a_{i(j)}a_{kl} = \frac{1}{4}(1 - \theta)(a_{k(j)} + a_{l(j)} + a_{ik} + a_{il}) + \frac{1}{4}\theta(a_{k(i)} + a_{l(i)} + a_{jk} + a_{jl}). \tag{10}$$

It may be noted that the direct summand of the vector space which corresponds to the male zygotic types is obtained by non-commutative duplication of that underlying G . The induced transformation of its basis, which replaces the transformation (6) is

$$\begin{aligned} c_{0(0)} &= a_{0(0)} \\ c_{0(i)} &= a_{0(0)} - a_{0(i)} \\ c_{i(0)} &= a_{0(0)} - a_{i(0)} \\ c_{i(j)} &= a_{0(0)} - a_{0(j)} - a_{i(0)} + a_{i(j)}. \end{aligned} \tag{11}$$

After some calculation, the multiplication table is found to be

	c_{00}	c_{0k}	c_{kl}
$c_{0(0)}$	$\frac{1}{2}c_{0(0)} + \frac{1}{2}c_{00}$	$\frac{1}{4}c_{k(0)} + \frac{1}{2}c_{0k}$	0
$c_{i(0)}$	$\frac{1}{2}\theta c_{0(i)} + \frac{1}{2}(1 - \theta)c_{0i}$	$\frac{1}{4}\theta c_{k(i)} + \frac{1}{4}(1 - \theta)c_{ki}$	0
$c_{0(i)}$	$\frac{1}{2}(1 - \theta)c_{0(i)} + \frac{1}{2}\theta c_{0i}$	$\frac{1}{4}(1 - \theta)c_{k(i)} + \frac{1}{4}\theta c_{ki}$	0
$c_{i(j)}$	0	0	0

(12)

together with

$$c_{ik}c_{kl} = c_{i(j)}c_{k(l)} = 0, \text{ all } i, j, k, l.$$

The mapping $c_{0(0)} \rightarrow m; c_{00} \rightarrow w; c_{0(i)}, c_{i(0)}, c_{i(j)}, c_{0i}, c_{i(j)} \rightarrow 0$ exhibits \mathcal{L}_θ as a dibaric algebra. The elements $c_{0(0)} + c_{00}, c_{i(0)}, c_{0(i)}, c_{ij}, c_{0i}, c_{i(j)}$ may be chosen as a basis for \mathcal{L}_θ^2 . If $\theta \neq \frac{1}{2}$, the principal powers of its nil ideal are

$$\begin{aligned} \mathcal{N}(\mathcal{L}_\theta^2) &= \{c_{i(0)}, c_{0(i)}, c_{ij}, c_{0i}, c_{i(j)}\} \\ \mathcal{N}^2(\mathcal{L}_\theta^2) &= \{c_{ij}, c_{i(j)}\} \\ \mathcal{N}^3(\mathcal{L}_\theta^2) &= \{0\}. \end{aligned}$$

If $\theta = \frac{1}{2}$, \mathcal{N}^2 is smaller, in fact $\mathcal{N}^2(\mathcal{L}_\theta^2) = \{c_{ij} + c_{i(j)}\}$. In either case the principal powers of \mathcal{N} are ideals, and so \mathcal{L}_θ^2 is a special train algebra. It can also be seen from (12) that $e_0 = c_{0(0)} + c_{00}$ is an idempotent. A further transformation producing a table analogous to (9) cannot be carried through in an

equally simple way for general values of θ . The new entries in the multiplication table involving e_0 can be written as

$$e_0 \begin{bmatrix} e_0 \\ c_{i(0)} \\ c_{0(i)} \\ c_{0i} \\ c_{i(j)} \\ c_{ij} \end{bmatrix} = \begin{bmatrix} e_0 \\ \frac{1}{2}\theta c_{0(i)} - \frac{1}{2}(1-\theta)c_{0i} \\ \frac{1}{2}(1-\theta)c_{0(i)} + \frac{1}{2}\theta c_{0i} \\ \frac{1}{4}c_{i(0)} + \frac{1}{4}c_{0i} \\ 0 \\ 0 \end{bmatrix} \tag{13}$$

Thus the operation of multiplication by e_0 corresponds to multiplication by a matrix which is the direct sum of (i) the unit matrix of order 1, (ii) r matrices of order 3, each of which has the form

$$\begin{bmatrix} 0 & \frac{1}{2}\theta & \frac{1}{2}(1-\theta) \\ 0 & \frac{1}{2}(1-\theta) & \frac{1}{2}\theta \\ \frac{1}{4} & 0 & \frac{1}{4} \end{bmatrix}, \tag{14}$$

and (iii) a zero matrix of order $r^2 + \frac{1}{2}r(r+1)$.

The proper values of (14) are $\lambda = \frac{1}{2}$ and

$$\lambda = \frac{1}{8} - \frac{1}{4}\theta \pm \left(\frac{1}{16}\theta^2 - \frac{5}{16}\theta + \frac{9}{64} \right)^{\frac{1}{2}}. \tag{15}$$

On putting $\phi = \frac{1}{4}\theta - \frac{1}{8}$ this can be written as

$$\lambda = -\phi \pm \{\phi(1-\phi)\}^{\frac{1}{2}}.$$

The case $\theta = 0$ corresponds to complete linkage. It is then possible to “identify” all the natural basis elements $a_{i(j)}$ for $j = 0, \dots, r$ in the table (10), and (15) gives $\lambda = \frac{1}{2}, -\frac{1}{4}$ as in the previous section. The case $\theta = \frac{1}{2}$ corresponds to no linkage. Here $a_{i(j)}$ and $a_{j(i)}$ may be identified in (10), and the resulting algebra is then the direct product $Z \times S$ of the autosomal zygotic algebra and the sex differentiation algebra. The values given by (15) are $\lambda = 0$ (twice). If $0 \leq \theta \leq \frac{1}{2}$ the values of λ given by (15) are real, while if $\frac{1}{2} < \theta \leq 1$ they are complex, but in all cases $|\lambda| < \frac{1}{2}$ and consequently (4, p. 42) the sequence of plenary powers of an element of unit weight tends to a limiting idempotent.

If a circle denotes the product of two elements of the vector space of \mathcal{L}_θ evaluated according to the rules of \mathcal{L}_0 , and a dot that evaluated according to those of $\mathcal{L}_{\frac{1}{4}}$, then (10) can be written as

$$a_{i(j)}a_{kl} = (2\theta - 1)a_{i(j)} \circ a_{kl} + 2(1-\theta)a_{i(j)} \cdot a_{kl}.$$

This exhibits \mathcal{L}_θ as a mixture of two special train algebras, a concept which has appeared in studies of polyploidy (6, p. 6) and linkage between autosomal loci (9). However, although \mathcal{L}_0 and $\mathcal{L}_{\frac{1}{4}}$ admit a common weight function and the decompositions induced by the sequences $\mathcal{N}, \mathcal{N}^2, \mathcal{N}^3$ are compatible,

it is impossible to find a basis such that the matrices corresponding to multiplication by e_0 are simultaneously diagonalised, and hence the train roots of \mathcal{L}_0 are not given as mixtures of those of \mathcal{L}_0 and \mathcal{L}_+ .

4. The copular algebra

Returning to the case of complete linkage, consider the duplicate of \mathcal{L} , abbreviating (a_i, a_j) , (a_i, a_{jk}) , (a_{ij}, a_{kl}) to $a_{i,j}$, $a_{i,jk}$ and $a_{ij,kl}$. The multiplication table may be obtained from (5). The elements a_{ij} and $a_{ij,kl}$ are annihilators, while

$$a_{i,jk}a_{l,st} = \frac{1}{16}(a_{j,ls} + a_{j,lt} + a_{k,ls} + a_{k,lt} + a_{s,ij} + a_{s,ik} + a_{t,ij} + a_{t,ik} + a_{ij,ls} + a_{ij,lt} + a_{ik,ls} + a_{ik,lt} + a_{j,s} + a_{j,t} + a_{k,s} + a_{k,t}). \tag{16}$$

In an asymmetric genetic situation, the square of the zygotic algebra contains all those elements representing populations which are half male and half female, and is thus the natural object to study. In \mathcal{L}' , the duplicate of \mathcal{L} , such populations are represented by the sub vector space spanned by $\{a_{i,jk}\}$. The complement of this, spanned by $\{a_{i,j}, a_{ij,kl}\}$ is an ideal of annihilators in \mathcal{L}' , which will be denoted by \mathcal{I} . Now in (16) in which multiplication in \mathcal{L}' is specified in terms of a basis consisting of elements of unit weight, the first 8 elements in the brackets on the right represent elements in the difference algebra $\mathcal{L}' - \mathcal{I}$, and the last eight belong to \mathcal{I} . Then to preserve the baric property let a multiplication denoted by $*$ be defined on $\mathcal{L}' - \mathcal{I}$ by

$$a_{i,jk} * a_{l,st} = 2a_{i,jk}a_{l,st} \pmod{\mathcal{I}}. \tag{17}$$

The multiplication table of the new algebra in terms of the transformed basis $c_{i,jk}$ can be obtained by duplicating (8) and applying (17), and is

	$c_{0,00}$	$c_{0,0k}$	$c_{k,00}$	$c_{k,0l}$
$c_{0,00}$	$c_{0,00}$	$\frac{1}{4}c_{0,0k} + \frac{1}{4}c_{k,00}$	$\frac{1}{2}c_{0,0k}$	$\frac{1}{4}c_{0,kl}$
$c_{0,0i}$		$\frac{1}{8}c_{k,0i} + \frac{1}{8}c_{i,0k}$	$\frac{1}{4}c_{i,0k}$	$\frac{1}{8}c_{i,kl}$
$c_{i,00}$			0	0
$c_{i,0j}$				0

while $c_{0,ij}$ and $c_{i,jk}$ are annihilators. This is clearly a baric algebra with a weight function defined by $c_{0,00} \rightarrow 1$; $c_{0,0i}, c_{i,00}, c_{i,0j}, c_{0,ij}, c_{i,jk} \rightarrow 0$. The powers of the nil ideal are

$$\begin{aligned} \mathcal{N} &= \{c_{0,0i}, c_{i,00}, c_{i,0j}, c_{0,ij}, c_{i,jk}\} \\ \mathcal{N}^2 &= \{c_{i,0j}, c_{i,jk}\} \\ \mathcal{N}^3 &= \{0\}. \end{aligned}$$

However, since $c_{0,00} * c_{i,0j} = \frac{1}{4}c_{0,ij}$, it is not a special train algebra.

Corrigenda

In (6) p. 2, l. 14 read $D_i D_j$ for $D_a D_b$; p. 3, l. 6 read $j < m$ for $j \leq m$; p. 4 in the second displayed equation read E^{k-1} for E^k ; and in the next line 2^k for k .

In (8) there is a confusion between the symmetric and the asymmetric ways of writing forms in doubly suffixed quantities. If on p. 292 one works with both d_{ij} and d_{ji} , remembering that they represent the same element, the definition of the typical element should be $x = \sum_{i=0}^n \sum_{j=0}^n \alpha_{ij} d_{ij}$ with $\alpha_{ij} = \alpha_{ji}$, and the factor $\frac{1}{2}$ on the right of the equation in the proof of Proposition 3 should be deleted. Alternatively, one may work only with d_{ij} for $j \geq i$, leading to slightly less elegant formulae. I am indebted to Dr E. Wallace of Leeds University for drawing my attention to this point.

REFERENCES

- (1) I. M. H. ETHERINGTON, Genetic algebras, *Proc. Roy. Soc. Edinburgh* **59** (1939), 242-258.
- (2) I. M. H. ETHERINGTON, Duplication of linear algebras, *Proc. Edinburgh Math. Soc.* (2) **6** (1941), 222-230.
- (3) I. M. H. ETHERINGTON, Non-associative algebra and the symbolism of genetics, *Proc. Roy. Soc. Edinburgh Sect. B* **61** (1941), 24-42.
- (4) H. GONSHOR, Special train algebras arising in genetics, *Proc. Edinburgh Math. Soc.* (2) **12** (1960), 41-53.
- (5) H. GONSHOR, Special train algebras arising in genetics II, *Proc. Edinburgh Math. Soc.* (2) **14** (1965), 333-358.
- (6) P. HOLGATE, Genetic algebras associated with polyploidy, *Proc. Edinburgh Math. Soc.* (2) **15** (1966), 1-9.
- (7) P. HOLGATE, Sequences of powers in genetic algebras, *J. London Math. Soc.* **42** (1967), 489-496.
- (8) P. HOLGATE, Jordan algebras arising in population genetics, *Proc. Edinburgh Math. Soc.* (2) **15** (1967), 291-294.
- (9) P. HOLGATE, The genetic algebra of k linked loci, *Proc. London Math. Soc.* (3) **18**, (1968), 315-327.
- (10) O. REIERSØL, Genetic algebras studied recursively and by means of differential operators, *Math. Scand.* **10** (1962), 25-44.
- (11) C. STERN, *Principles of human genetics* (Freeman, San Francisco, 2nd Ed., 1960).

BIRKBECK COLLEGE
LONDON, W.C.1