SOME GENETIC PROBLEMS IN CONTROLLED POPULATIONS

EVERETT R. DEMPSTER
GENETICS DEPARTMENT, UNIVERSITY OF CALIFORNIA

The titles of this talk and of the succeeding one do not divide genetics into two fundamental categories. At the present stage of development, it may be more profitable to emphasize the common features rather than the differences between wild populations on the one hand, and domestic and laboratory populations on the other. The division indicated by the titles was adopted chiefly as a convenience for the purpose of preventing overlap in subject matter. Within what I take as my half of this subject, limitations of time prevent an over-all review and dictate a selection from the many interesting developments of a mathematical nature that have recently appeared. I am restricting myself therefore to a consideration of nonadditive genetic variance, a condition that must lie near the heart of the currently unpredictable behaviors of populations subjected to prolonged selection, or to inbreeding and subsequent crossing. It is on this particular field of nonadditive variance that a new method of attack has very recently been conceived and developed independently at three different research centers. The different but quite compatible developments of this new idea are themselves very interesting; in addition there may be value in a comparative study of the separate pathways by which individuals of different propensities and backgrounds can exploit the same basic idea. All this work is very recent and I am greatly indebted to the authors for permission to discuss their contributions prior to their appearance in print, although it will be possible to insert full references in the published version of this talk.

In studies of the variance of measurements of organisms, the measurement of each individual is considered to be the sum of a genotypic value, defined as the average measurement that replicates of individuals of the given genetic constitution would have if exposed to the whole array of environments, and of an environmental or nongenetic deviation from the genotypic value. In the present discussion only genotypic values will be considered. In order to subdivide the genotypic variance it is necessary to consider the general features of Mendelian heredity. The genotype of an individual consists of two alleles (related genes) at each of a large but fixed number of loci. One allele at each locus comes from the male parent, and one from the female parent. If the two alleles are the same at a particular locus, the individual is said to be homozygous at this locus; if the two alleles are different, the individual is heterozygous at this locus. The distributions of alleles at different loci, for the most part in the following discussion, are assumed to be independent, although deviations from this condition due to linkage or to linkage in combination with inbreeding are considered by two of the authors.

With this background we can consider the usual subdivision of genotypic values

and of the corresponding genotypic variance. The genotypic value is commonly taken as the sum of (1) the population mean, (2) an additive genetic deviation, (3) a dominance deviation, and (4) an epistatic deviation. The first plus the second of these is then the additive genetic value. The additive genetic deviation of an individual is defined as follows: each kind of gene in the population is assigned a value such that the sum of the values for the particular genes present in each individual comes as close as possible, in the least-squares sense, to the departure of the genotypic value of this individual from the population mean; this sum for an individual is its additive genetic deviation. If values are assigned to each possible *pair* of genes that can exist at each locus, instead of to the individual genes, the sum of values for all loci for each individual is the additive deviation plus the dominance deviation. The remainder that must be added to equal the genotypic value is the epistatic deviation, that is, nonadditive interaction *between* loci.

The methods to be discussed differ from previous work in the further subdivision of epistatic deviations. In the first place it has been found useful to divide these deviations into those that can be accounted for by considering the nonadditive effects of loci considered in pairs, those additional deviations that can be accounted for by considering loci in groups of three, and so on for higher numbers. This particular type of division is not new. Thus Fisher [1] derived the correlations between relatives due to nonadditive effects of pairs of loci, Wright [2] demonstrated that an important model of epistatic deviations, one in which a measured character depends on the squared deviation of a primary character from its optimum, is expressible in terms of interactions of loci in pairs, and Griffing [3] considered a model involving pairs of genes. Horner [4] considered variances and covariances in random bred and inbred populations in connection with a number of epistatic models, in some of which loci interact in groups of arbitrary number, but the proportion of variance due to interaction in groups of two, three, four, etc., were not considered. Other epistatic models, such as the threshold models considered by Lush, Lamoreux, and Hazel, Robertson and Lerner, and Dempster and Lerner [5], [6], [7], can be shown to involve nonadditive interactions of loci taken in groups larger than two, but analysis in this light was not attempted.

The new methods to be discussed differ from previous work in a further orthogonal subdivision of the epistatic deviations and variance. In the first example, this division is into interactions of heterozygous loci with heterozygous loci, of homozygous loci with homozygous loci, and of combinations of heterozygous and of homozygous loci. In the other examples, only superficially different, the subdivisions of nonlinear interactions between loci are additive deviations with additive deviations, dominance deviations with dominance deviations, and interactions of additive with dominance deviations.

We shall first consider the analysis, as developed by Anderson and Kempthorne of populations produced by the successive self-fertilizations of an F_1 population obtained by crossing two completely homozygous parents. The symbolism employed may be understood by reference to table I in which the numbers represent relative frequencies in the F_2 with respect to two loci. The letters a_0, a_1, and a_2 are used to represent the genotypes aa, Aa, and AA respectively, and similarly for genotypes at the B locus and other loci. Genotypic values are also denoted by the same symbols, so that a_0, for example, represents both the genotype aa and also its average

genotypic value in combination with all the combinations of genes at other loci in the population in the proportions in which they exist in the F_2. More complicated symbols are defined similarly; thus $a_0b_2c_1$ represents the genotype aaBBCc and also the average genotypic value of all individuals of the population possessing this combination of genes. The F_2 is thus used as a base population. With this definition it will be seen that the mean value of the F_2 population can be expressed by any one of the factors in the following expression or by the expansion of the product of two or more of them, the symbols in the expanded expression being replaced by genotypic values:

$$(1) \qquad v = [(a_0 + 2a_1 + a_2)/4][(b_0 + 2b_1 + b_2)/4][(c_0 + 2c_1 + c_2)/4] \cdots, \text{etc.}$$

TABLE I

DIHYBRID F_2
(Numbers are relative frequencies.)

Genotypes	aa a_0	Aa a_1	AA a_2
bb b_0	1	2	1
Bb b_1	2	4	2
BB b_2	1	2	1

(SOURCE: Anderson and Kempthorne [8])

The symbols A_0, A_1, A_2, and likewise B_0, B_1, B_2, etc., are now introduced to represent deviations of genotypes from the F_2 mean. Thus the value of A_0 can be expressed in the following way:

$$(2) \qquad A_0 = a_0 - (a_0 + 2a_1 + a_2)/4 = (3a_0 - 2a_1 - a_2)/4\,,$$

and if genotypic values are inserted this gives the deviation of the average of all genotypes containing aa at the A locus from the F_2 mean. Thus we can write:

$$(3) \qquad a_0 = A_0 + v = A_0 + (a_0 + 2a_1 + a_2)/4\,.$$

Note that $A_0 = a_0 - (a_0 + 2a_1 + a_2)/4$ so that sums like $A_0 + 2A_1 + A_2$ are equal to zero.

We can now write certain products as in the following example:

$$(4) \qquad a_0b_2c_1 = [A_0+(a_0+2a_1+a_2)/4][B_2+(b_0+2b_1+b_2)/4][C_1+(c_0+2c_1+c_2)/4],$$

and if we expand the right-hand side without disturbing the symbols in parentheses, and appropriately substitute the meanings of certain symbols as described above we obtain:

$$(5) \qquad a_0b_2c_1 = v + A_0 + B_2 + C_1 + A_0B_2 + A_0C_1 + B_2C_1 + A_0B_2C_1\,.$$

In this expression A_0 is the deviation of the genotypic mean values of all individuals possessing aa from v, A_0B_2 is the deviation of genotypic mean values of all indi-

viduals possessing aaBB from $[\nu + A_0 + B_2]$ and $A_0B_2C_1$ is the deviation from $[\nu + A_0 + B_2 + C_1 + A_0B_2 + A_0C_1 + B_2C_1]$ of the mean genotypic value of all individuals possessing aa*BBCc*, and similarly for the other terms. Note that the expected values of terms like A_i, A_iB_j, $A_iB_jC_k$, etc., in the F_2 population are all zero. The zero value of expressions like the following shou ld also be noted:

$$[A_0 + 2A_1 + A_2][B_iC_jD_k \cdots, \text{etc.}] = 0, \tag{6}$$

where each subscript in the second factor is given a particular value, either 0, 1, or 2. In this model then terms like A_i are the additive deviations of individual loci, terms like A_iB_j are the epistatic deviations of two loci in a given genotype, terms like $A_iB_jC_k$ are epistatic deviations of three loci taken at a time, etc. The parameters are additive, the expectations of all of them in the F_2 population, with the exception of ν, are zero, the basic differential effects in the F_2 of genes at individual loci are given by terms like $(A_2 - A_0)/2$, the deviation of the heterozygotes at individual loci from the mean of the two homozygotes in the F_2 is given by expressions like $(2A_1 - A_0 - A_2)/2$. If it is assumed that interactions of certain classes do not exist the corresponding terms may be dropped and the model used without other alteration; thus if it is assumed that interactions involving three and more loci together are negligible then all terms like $A_iB_jC_k$ and $A_iB_jC_kD_l$, etc., can be simply omitted. It should be clearly understood that the present author is solely responsible for any faults in this brief paraphrase of the rigorous development of Anderson and Kempthorne [8].

The authors mentioned now apply this model to the population resulting from self-fertilizing the F_1 for m generations. Consider only the contribution to the mean of such a partially inbred population of interactions at the first p loci taken together. This may be expressed as:

$$\prod_{\substack{\text{first} \\ p \text{ loci}}} [qA_1 + (1 - q)(A_0 + A_2)/2] , \tag{7}$$

where q is the frequency of heterozygotes in this generation. Since from (6) we may substitute $-A_1$ in place of $(A_0 + A_2)/2$ in this expression it may be written:

$$\prod_{\substack{\text{first} \\ p \text{ loci}}} [qA_1 - (1 - q)A_1] = \prod_{\substack{\text{first} \\ p \text{ loci}}} (2q - 1)A_1 = (2q - 1)^pA_1B_1C_1 \cdots P_1 . \tag{8}$$

Thus the entire effect of nonadditive interactions on population means in the symmetrical situation produced by self-fertilization can be expressed in terms of heterozygous times heterozygous interactions. With self-fertilization the frequency of heterozygotes is halved at every generation, so we may write $1/2^m$ in place of q. The sum of all p factor interactions in the mth generation of selfing is then:

$$\sum_{\substack{\text{all sets} \\ \text{of } p \text{ loci}}} [(1/2^{m-1}) - 1]^p[A_1B_1C_1 \cdots] . \tag{9}$$

If we let β_p represent the summation of the second factor we can write for the mean of the population produced by m generations of selfing (which is the F_{m+1}) as follows:

$$\bar{F}_{m+1} = K_1 + \sum_{p=1}^{n} [(1/2^{m-1}) - 1]^p\beta_p . \tag{10}$$

In this expression K_1 is the mean of the F_2 population, and the β's are the sums of p-factor interactions of the various loci and their interactions with the fixed loci. (The latter phrase serves as a reminder that the interactions of factors are not independent of genetic background.)

If we can obtain the mean values of successive generations of a self-fertilized F_1, in the absence of disturbing factors such as selection and linkage, it should be possible to estimate the β's, and to determine in particular whether epistasis is of any great importance in the changes of mean with inbreeding. From expression (10) we may write expressions for generation means as follows:

$$\bar{F}_1 = K_1 + \beta_1 + \beta_2 + \cdots, \text{ etc .} \tag{11}$$

$$\bar{F}_2 = K_1 . \tag{12}$$

$$\bar{F}_3 = K_1 - (1/2)\beta_1 + (1/4)\beta_2 - (1/8)\beta_3 + \cdots . \tag{13}$$

$$\bar{F}_4 = K_1 - (3/4)\beta_1 + (9/16)\beta_2 - (27/64)\beta_3 + \cdots . \tag{14}$$

$$\bar{F}_5 = K_1 - (7/8)\beta_1 + (49/64)\beta_2 - (343/512)\beta_3 + \cdots . \tag{15}$$

Anderson and Kempthorne have made least-squares estimates of the first three parameters with respect to a number of characters in red peppers, using data on means of successive generations of selfing provided by Khambanonda [9]. The authors emphasize that the results are dependent on the absence of disturbances of means by selection or linkage, and that the tests of significance also depend on sampling distributions being normal. The estimates of β_2, the second order epistatic interactions, turn out to be significant at the 5 per cent level or better in all cases, even when the data are expressed in logarithms. The actual estimates of the β's depend on the assumption that higher order interactions can be neglected, but any error in this assumption could not of course invalidate the demonstration that epistatic interactions had significant effects on the means of the populations studied.

There may be some value in considering further the meanings of the β's. K_1 is of course the F_2 mean. The dominance parameter, β_1, decreases by increments that are halved in successive generations of selfing as would be expected. However the *increase*, with selfing, of β_2, the sum of heterozygote by heterozygote interactions, may appear puzzling at first, since combinations of pairs of heterozygous loci must *decrease* in frequency. The explanation is in the definition of dominance, which is based on the F_2 generation and therefore in a sense includes a good deal of what might, in physiological terms, be considered heterozygote by heterozygote interaction. The use of the F_2 as a base is not obligatory and some further light may be shed on meanings of the parameters by alternative procedures. The general method of Anderson and Kempthorne is easily applied to the F_1 or any other generation as a base, but the use of F_∞, a population which consists of all possible completely homozygous lines, as a base seems to be of particular interest. In this case the physiological meanings of the parameters are rather easily formulated; the dominance parameter is then the sum of the effects of making each factor heterozygous in turn while all other factors are homozygous; β_2 is the sum of the additional effects due to making all possible combinations of two loci heterozygous, with the remaining loci homozygous; and similarly for β_3 and higher order parameters. Such definitions

have a direct relationship to certain theories of heterosis. For example Shultz and Briles [10] present some evidence, in the case of blood-group genes in poultry, that heterozygosity at two loci may boost certain superficially unrelated characters more than the sum of the boosts due to heterozygosity of each locus alone. Such a situation, if of general occurrence, would correspond to a positive value of β_2 and perhaps also of higher order parameters, as they are defined in terms of the F_∞ population.

Anderson and Kempthorne are not responsible for the following manipulation to which their general model has been subjected. The following expressions apply with F_∞ used as a base:

$$\nu = \frac{(a_0 + a_2)(b_0 + b_2)(c_0 + c_2) \cdot \cdot \cdot (p_0 + p_2)}{2^p}, \tag{16}$$

$$A_2 = \frac{(a_2 - a_0)}{2}, \tag{17}$$

$$A_1 = a_1 - \frac{(a_0 + a_2)}{2} = \frac{-a_0 + 2a_1 - a_2}{2}. \tag{18}$$

Then the interaction deviation of, say, AA with Bb is

$$A_2B_1 = \frac{(a_2 - a_0)(-b_0 + 2b_1 - b_2)}{4} \tag{19}$$

and, as an example, we can write

$$a_2b_1 = \nu + A_2 + B_1 + A_2B_1 = \frac{(a_2 + a_0)(b_2 + b_0)}{4} \tag{20}$$

$$+ \frac{a_2 - a_0}{4}(b_2 + b_0) + \frac{-b_0 + 2b_1 - b_2}{4}(a_2 + a_0) + \frac{(a_2 - a_0)(-b_0 + 2b_1 - b_2)}{4}.$$

Now consider the mean effect of the p-factor interactions for the case where there is q proportion of Aa and $(1 - q)/2$ proportion each of AA and aa, and similarly at other loci. The p-factor interactions for the first p loci can be obtained, as before, by expanding

$$[(1-q)A_0/2 + qA_1 + (1-q)A_2/2][(1-q)B_0/2 + qB_1 + (1-q)B_2/2] \cdot \cdot \cdot, \text{etc.} \tag{21}$$

In this case however expressions like the following equal zero:

$$(1 - q)A_0XYZ + (1 - q)A_2XYZ = 0, \tag{22}$$

where as before X, Y, and Z represent symbols like B_1, C_0, D_2, etc. Making substitutions accordingly, expression (21) can be written

$$[qA_1][qB_1][qC_1] \cdot \cdot \cdot [qP_1] = q^pA_1B_1C_1 \cdot \cdot \cdot P_1 = (1/2^m)^pA_1B_1C_1 \cdot \cdot \cdot P_1. \tag{23}$$

Summing over all sets of p loci, and letting

$$\beta_p = \sum_{\substack{\text{all sets} \\ \text{of } p \text{ loci}}} (1/2^m)^p\beta_p \tag{24}$$

we have the following very simple result:

$$\bar{F}_{m+1} = K_0 + (1/2)^m\beta_1 + (1/2)^{2m}\beta_2 + (1/2)^{3m}\beta_3 + \cdots, \text{ etc.} \tag{25}$$

Note that K_0 is the mean of the completely inbred populations. The expressions for the first few generations, excluding four-factor and higher order interactions then become

$$\bar{F}_1 = K_0 + \beta_1 + \beta_2 + \beta_3 , \tag{26}$$

$$\bar{F}_2 = K_0 + (1/2)\beta_1 + (1/4)\beta_2 + (1/8)\beta_3 , \tag{27}$$

$$\bar{F}_3 = K_0 + (1/4)\beta_1 + (1/16)\beta_2 + (1/64)\beta_3 , \tag{28}$$

$$\bar{F}_4 = K_0 + (1/8)\beta_1 + (1/64)\beta_2 + (1/512)\beta_3 . \tag{29}$$

TABLE II

TWO-FACTOR MODELS. THE NUMBERS ARE DEVIATIONS FROM AN ARBITRARY BASE OF THE GENOTYPES INDICATED AT THE MARGINS

Model	*a*			*a'*			*a''*			*b* (Duplicate Factors)			*c* (Complementary Factors)		
Genotypes	*AA*	*Aa*	*aa*	*AA*	*Aa*	*aa*	*AA*	*Aa*	*aa*	*AA*	*Aa*	*aa*	*AA*	*Aa*	*aa*
BB	0	0	0	2	1	0	1	1	0	0	0	0	1	1	0
Bb	0	1	0	1	1	−1	0	1	−1	0	0	0	1	1	0
bb	0	0	0	0	−1	−2	0	0	−1	0	0	1	0	0	0

With these definitions of dominance and of epistatic deviations, the increments of all factors in successive generations form simple geometric series, and the higher order interactions rapidly become of relatively less importance as inbreeding continues. It becomes obvious that the relative values of the parameters that measure dominance, two-factor interactions, three-factor interactions, etc., can be quite different where different populations are used as bases, and have exact meanings only when defined with respect to a specified generation. An exception to this statement must be made for the case in which there are no interactions of higher order than the one-factor or dominance interactions, in which case the particular generation chosen as a base is immaterial.

To illustrate the relationship of the parameters to the generation used as a base, the model is applied to the hypothetical two-locus populations shown in table II. The numbers shown are deviations of genotypic values from an arbitrary base. Note that the change per generation of selfing, and hence the dominance and higher order parameters, are the same for model *a* (in which only the double heterozygote has a differential effect) and models *a'* and *a''* where additive effects and opposing dominance effects respectively are introduced for all other genotypes. Model *b* is the duplicate factor case, and model *c* the complementary factor case. Table III shows the values of the parameters, as dependent on the generation used as a base, and

also the decrease in mean between certain generations attributable to dominance and to two-factor interactions in the different cases. It may be noted in passing that, when the F_2 is the base population, the change from F_1 to F_∞ due to all even-number factor interactions is always zero. The rather large differences in the changes attributed to dominance as compared to epistasis, depending on the base population used, illustrate the limitations inherent in the use of fixed parameters in a rapidly changing population; the method is perfectly valid, however, for estimating parameters as they exist in a particular generation (assuming data can be obtained corresponding to the assumptions on which the model is based) and it is therefore of

TABLE III

Quantity	Base Population	Model *a*	Model *b* (Duplicate)	Model *c* (Complementary)
β_1 (dominance)	F_2	1/2	−1/8	3/8
	F_∞	0	−1/2	1/2
β_2 (Two-factor epistasis)	F_2	1/4	1/16	1/16
	F_∞	1	1/4	1/4
Decrease from F_1 to F_2 attributed to dominance and epistasis respectively	F_2	1/2 : 1/4	−1/8 : 1/16	3/8 : 1/16
	F_∞	0 : 3/4	−1/4 : 3/16	1/4 : 3/16
Decrease from F_2 to F_∞ attributed to dominance and epistasis respectively	F_2	1/2 : −1/4	−1/8 : −1/16	3/8 : −1/16
	F_∞	0 : 1/4	−1/4 : 1/16	1/4 : 1/16

some importance to understand the meanings of the parameters as applied to the particular base population adopted.

It should be mentioned that Anderson and Kempthorne adapt their model to utilize data from parental lines, backcrosses, repeated direct and reciprocal backcrosses, selfed backcrosses, and the offspring of backcrosses by F_1's. In particular they explore the scaling tests of Mather [11] and show that it will detect interactions of the second order if higher order interactions are absent, but could by chance fail in some cases if higher order interactions are present. These interesting developments cannot be discussed further at this time.

Hayman and Mather [12] have also studied the partition of epistatic variance to inbred and backcrossed populations. Their method of attack is quite different from that discussed above and, as it provides a logical transition to the application of the same ideas to crossbreeding populations with unrestricted allelic frequencies, will be briefly reviewed at this point. Table IV shows their symbolism as applied to a dihybrid F_2. The capital letter gene symbols, A and B, refer to alleles that have a plus effect on the character measured and are not necessarily either dominant or recessive. The symbols d_a and $-d_a$ are the additive values of genes A and a, $2d_a$ being the difference in mean genotypic values of AA and aa individuals in the F_2; d_b is similarly defined. The symbol h_a, the dominance deviation of Aa, is the excess

of the mean genotypic value of Aa F_2 individuals over the mean of AA and aa individuals. Because of the assignment of h_a and h_b solely to Aa and Bb individuals respectively, the mean of the genotypic values listed in the table is $(h_a + h_b)/2$ rather than zero. This arrangement seems to have no particular disadvantage in the computation of variances where gene frequencies are equal. The further symbols, $i_{ab/}$, $j_{a/b}$, $j_{b/a}$, and $l_{/ab}$, represent deviations due to the following kinds of interactions respectively: additive with additive, additive at locus A with dominance at locus B, the reverse, and dominance with dominance. All these interactions, with the exception of h_a and h_b, add to zero in the F_2.

TABLE IV

Genotypes	$\underset{d_a}{AA}$	$\underset{h_a}{Aa}$	$\underset{-d_a}{aa}$
	$d_a + d_b$	$h_a + d_b$	$-d_a + d_b$
BB	$+ i_{ab/}$		$- i_{ab/}$
d_b	$-(1/2)j_{a/b} - (1/2)j_{b/a}$	$+(1/2)j_{b/a}$	$(1/2)j_{a/b} - (1/2)j_{b/a}$
	$(1/4)l_{/ab}$	$-(1/4)l_{/ab}$	$(1/4)l_{/ab}$
	$d_a + h_b$	$h_a + h_b$	$-d_a + h_b$
Bb	$(1/2)j_{a/b}$		$-(1/2)j_{a/b}$
h_b	$-(1/4)l_{/ab}$	$(1/4)l_{/ab}$	$-(1/4)l_{/ab}$
	$d_a - d_b$	$h_a - d_b$	$-d_a - d_b$
bb	$- i_{ab/}$		$i_{ab/}$
$-d_b$	$-(1/2)j_{a/b} + (1/2)j_{b/a}$	$(1/2)j_{b/a}$	$(1/2)j_{a/b} + (1/2)j_{b/a}$
	$+(1/4)l_{/ab}$	$-(1/4)l_{/ab}$	$+(1/4)l_{/ab}$

It will simplify our comparison of this table with the work of Kempthorne [13] and of Cockerham [14] to be discussed later, if we note that the definitions of the different kinds of interactions by the three groups of investigators would be identical if, in table IV, the dominance interactions were distributed equally between the heterozygotes and homozygotes. Thus instead of adding h_a to Aa genotypes, if $(1/2)\,h_a$ is added to these genotypes and $(1/2)\,h_a$ subtracted from AA and aa genotypes, and similarly with h_b, the F_2 mean of genotypic deviations becomes zero. Such a change would somewhat clarify the terminology in that the coefficient of any compound parameter for any genotype becomes simply the product of the coefficients of the corresponding elementary ones. Thus the coefficient of $j_{ab/}$, the additive by additive interaction, for genotype $AABb$ is the product of unity, the coefficient of d_a, and zero, the coefficient of d_b; the coefficient of $l_{/ab}$, the dominance by dominance interaction, for the same genotype would be the product of the coefficients for h_a and h_b or $(-1/2)(1/2) = -1/4$ after the change suggested above had been made. With this superficial change the definitions correspond also to those used by Anderson

and Kempthorne in the work just discussed; in the latter work, since means only were being considered, the additive by additive interactions cancelled and the other interactions were expressible in terms of dominance by dominance interactions because of the symmetry of gene combinations in the populations derived by successive self-fertilizations.

There are thus eight parameters in the table which completely define the deviations of the nine genotypes from the mean. With three segregating loci affecting the character there are 27 genotypes and 26 parameters as follows: three additive, three additive by additive, one additive by additive by additive, three dominant, three dominant by dominant, one dominant by dominant by dominant, six additive by dominant, three additive by additive by dominant, and three additive by dominant by dominant. It is easy to assign the proper parameters, and to determine their coefficients, for interactions of higher orders than two by the same scheme mentioned above for the two-factor case.

At this point it might be asked whether there is much utility in replacing individual genotypic values by an equal number of parameters. The authors believe that the higher order interactions are likely to be negligible and that, if this is the case, the scheme permits the combination of different types of epistasis, that may be acting in a population, into a single model in which each component can be considered somewhat independently of the others. We have already seen how this was done by Kempthorne and Anderson and further examples will be considered later. In addition some consideration will be given to the question of higher order parameters.

Hayman and Mather next show certain simple relations among the parameters that result in classical types of epistatic interactions. As examples we may note that the 9:3:3:1 ratio occurs when $d_a = h_a$, $d_b = h_b$ and $i_{ab/} = j_{a/b} = j_{b/a} = l_{/ab}$ and that when, in addition, $d = (3/2)i_{ab/}$ the 9:3:4 ratio is produced. These authors also discuss methods for estimating the parameters in the two-locus case where the genes are "associated" and where they are "dispersed," that is, where the two plus genes come from the same parent in the original cross and where they come from different parents. In this connection, they also give the linear expressions involving population means, which in scaling tests have the value of zero in the absence of epistasis, in terms of the above parameters. A number of expressions for variances and covariances of backcross and selfed populations are derived and the effects of linkage on these various relationships discussed. Some of these relations are given in table V. Time is not available for a consideration of these very interesting exploratory investigations in further detail.

The application of the same basic idea to populations having any gene frequencies and any degree of inbreeding, but restricted to two alleles at any locus, is illustrated in table VI, extracted from a table by Cockerham [14]. The two subscripts to the Y's (genotypic values) and the f's (genotypic frequencies) indicate the alleles present at the A locus and the B locus in that order. Thus, subscript 2 represents homozygous AA or BB, subscript 1 the heterozygote, and 0 the aa or bb homozygous. Y_{10} is thus the genotypic value of $Aabb$ and f_{10} its frequency. Dots indicate averages, so that $Y_{1.}$ is the mean of $AaBB$, $AaBb$, and $Aabb$ genotypic values in the population. The rows headed by W_1, W_2, etc., are orthogonal scales and it will be noted that multiplying the genotypic values by the corresponding f's and W_1's yields the

product of $2uw$ and an expression representing the additive effect at the A–a locus, namely, the mean of all the AA's and half the Aa's minus the mean of all the aa's

TABLE V

SCALING TESTS

Test	Associated	Dispersed
$\bar{P}_1 + \bar{F}_1 - 2\bar{B}_1$	$(1/2)(i - j - j + l)$	$(1/2)(-i - j + j + l)$
$\bar{P}_2 + \bar{F}_1 - 2\bar{B}_2$	$(1/2)(i + j + j + l)$	$(1/2)(-i + j - j + l)$
$\bar{P}_1 + \bar{P}_2 + 2\bar{F}_1 - 4\bar{F}_2$	$2i + l$	$-2i + 1$
$\bar{P}_1 + \bar{P}_2 + 2\bar{F}_2 - 4\bar{F}_3$	$2i + (1/4)l$	$-2i + (1/4)l$

EXAMPLES OF VARIANCES

$$V_{F_2} = (1/2)d_a^2 + (1/2)d_b^2 + (1/4)h_a^2 + (1/4)h_b^2 + (1/4)i_{ab/}^2 + (1/8)j_{a/b}^2 + (1/8)j_{b/a}^2 + (1/16)l_{/ab}^2$$

$$V_{F_3} = (1/2)[d_{\bar{a}} - (1/4)j_{a/b}]^2 + (1/2)[d_b - (1/4)j_{b/a}]^2 + (1/16)[h_{\bar{a}} - (1/4)l_{/ab}]^2 + (1/16)[h - (1/4)l_{/ab}]^2 + (1/4)i_{ab/}^2 + (1/32)j_{a/b}^2 + (1/32)j_{b/a}^2 + (1/256)l_{/ab}^2$$

(SOURCE: Hayman and Mather [12])

TABLE VI

ORTHOGONAL SCALES AND PARTITION OF VARIANCE

Scale	Genetic Type								
Y	Y_{22}	Y_{21}	Y_{20}	Y_{12}	Y_{11}	Y_{10}	Y_{02}	Y_{01}	Y_{00}
f	f_{22}	f_{21}	f_{20}	f_{12}	f_{11}	f_{10}	f_{02}	f_{01}	f_{00}
W_1	$2v$*	$2v$	$2v$	$v-u$	$v-u$	$v-u$	$-2u$	$-2u$	$-2u$
W_2	$1/f_{2.}$	$1/f_{2.}$	$1/f_{2.}$	$-2/f_{1.}$	$-2/f_{1.}$	$-2/f_{1.}$	$1/f_{0.}$	$1/f_{0.}$	$1/f_{0.}$
W_3	$2y$	$y-x$	$-2x$	$2y$	$y-x$	$-2x$	$2y$	$y-x$	$-2x$
W_4	$1/f_{.2}$	$-2/f_{.1}$	$1/f_{.0}$	$1/f_{.2}$	$-2/f_{.1}$	$1/f_{.0}$	$1/f_{.2}$	$-2/f_{.1}$	$1/f_{.0}$

$W_5 = W_1W_3$, $W_6 = W_1W_4$, $W_7 = W_3W_2$, $W_8 = W_2W_4$.

*$u = f_{2.} + f_{1.}/2.$, $v = 1 - u$, $x = f_{.2} + f_{.1}/2$, $y = 1 - x$.

$\sigma_t^2 = (\sum f_{ij}Y_{ij}W_{tij})^2 / \sum f_{ij}W_{tij}^2 = (CovYW_t)^2/\sigma_{W_t}^2 = \rho_{YW_t}^2\sigma_Y^2$.

(SOURCE: Cockerham [14])

and half the aa's. The W_3 scale does the same for the B locus, while W_2 and W_4 are related to the dominance effects at these two loci. Thus the summation of the products of the Y's and f's with W_2 yields a value proportional to the mean of the Aa's minus half the means of AA and of aa. These scales are orthogonal in that the summation of the products of the W's of any one scale and the corresponding f's is zero (thus ensuring that deviations around the means are dealt with) and also in that the summation of the products of two scales and the corresponding f's is zero.

The additive variance due to the A locus is now obtained by summing the triple products of corresponding Y's, f's, and W_1's, squaring the sum and dividing the square by the sum of the frequencies multiplied by the squares of the corresponding W_1's. This division standardizes the scales, the terms of which need only be in correct proportions. The W_2, W_3, and W_4 scales, when treated similarly, give, respectively, the dominance variance due to the A locus, and the additive and dominance variances due to the B locus. By multiplying the terms of W_1 by those of W_3, and similarly for other scales involving the A and B loci as indicated near the bottom of table VI, four new orthogonal scales are obtained which can be used to give the additive by additive variance, the additive by dominance, the dominance by additive, and the dominance by dominance. The sum of the eight components equals the total variance, and these components correspond to those obtained by Hayman and Mather, and by Kempthorne in work to be discussed presently. Where three or more loci are to be considered it is only necessary to add new W scales for the additive and dominance effects of the new locus and then to obtain a whole new set of scales by multiplication of the appropriate terms of the previously existing scales. It is thus a straightforward process to obtain scales permitting the calculation of all interaction components of dominance and additive variance taking two loci at a time, three at a time, and so on up to the number of loci considered. The total variance is the sum of these components which again correspond to those obtained by Hayman and Mather, and by Kempthorne in work to be discussed presently.

Table VII gives examples of single locus and two-locus variance components. In these expressions F is Wright's inbreeding coefficient. As is well known, complete inbreeding ($F = 1$) ordinarily leads to an increase in additive variance, always to the disappearance of dominance variance, and as the new methods demonstrate, to the disappearance of any variance with dominance in its nomenclature. However these expressions show that dominance variance may actually increase in the early stages of inbreeding, as has previously been noted by Robertson [15]. The variance components given here apply strictly to the particular population on which the computations are made, and not only components and frequencies, but also orthogonal scales themselves, change with inbreeding.

These scales are used not only to obtain variance partitions, but also (and chiefly) to obtain correlations between relatives. This is done by setting up joint tables for the two sets of relatives, and obtaining covariances of, in theory, all terms in one with all terms in the other. Although the processes are too involved to detail at this time, a number of simplifications in the relationships make the task usually much less formidable than might seem to be the case. In the first place, it is only necessary to obtain covariances between the scales, taking account of joint frequencies, and of the variance components, rather than between the genotypic deviations. Secondly, one-factor deviations in one relative are correlated only with one-factor

deviations in the other and of the same locus, and similarly for two-factor and higher order deviations. In fact only the same type of n-factor deviations (as for example two-factor dominance by additive) are correlated in a noninbred population at equilibrium. The latter condition does not hold for inbred relatives, and linkage partially invalidates the method for some conditions of inbreeding, even when the population is assumed originally to be in equilibrium. We will reproduce here only the very simple results that were obtained for random mated populations in equilibrium, namely that the correlation between relatives for a given order of dominance-by-additive deviations is equal to $(p)^A(q)^D$, where A is the number of loci

TABLE VII

EXAMPLES OF PARTITIONS OF VARIANCE IN INBRED POPULATIONS FOR TWO LOCI

Orthogonal Scale	Partitions of the Variance
W_1 (add.)	$\frac{2uw}{1+F}[(u+Fv)(Y_{2.}-Y_{1.})+(v+Fu)(Y_{1.}-Y_{0.})]^2$
W_2 (dom.)	$\frac{uw(u+Fv)(v+Fu)(1-F)}{1+F}(Y_{2.}-2Y_{1.}+Y_{0.})^2$
W_5 (add. × add.)	$\frac{4uwxy}{(1+F)^2}[(u+Fv)(x+Fy)e^*_{22}+(u+Fv)(y+Fx)e_{21}+(v+Fu)(x+Fy)e_{12}+(v+Fu)(y+Fx)e_{11}]^2$
W_8 (dom. × dom.)	$\frac{uwxy(1-F)^2}{(1+F)^2}(x+Fy)(y+Fx)(u+Fv)(v+Fu)(e_{22}-e_{21}-e_{12}+e_{11})^2$

$^*e_{22} = Y_{22} - Y_{21} - Y_{12} + Y_{11}$, $\quad e_{12} = Y_{12} - Y_{11} - Y_{02} + Y_{01}$

$e_{21} = Y_{21} - Y_{20} - Y_{11} + Y_{10}$, $\quad e_{11} = Y_{11} - Y_{10} - Y_{01} + Y_{00}$

(SOURCE: Cockerham [14])

entering into the deviation with additive nomenclature, D is the number with dominance nomenclature, and p and q are the correlations for single factor additive and dominance deviations respectively. Wright's coefficient of relationship is the value of p, and q is well known for simple cases such as for parent-offspring correlations ($q = 0$) and for full sibs ($q = 1/4$). Thus the correlation between full sibs of dominance-by-dominance deviations would be $(1/4)^2 = 1/16$, between additive-by-dominance deviations $(1/4)(1/2) = 1/8$, and between additive-by-additive deviations $(1/2)^2 = 1/4$. The higher order correlations thus diminish rather rapidly with decrease in genetic relationship. The value of q can be obtained by the orthogonal scales for inbred populations, but in such cases there may be n-factor correlations between dominance deviations of one relative and additive deviations of the other, which can greatly complicate the results.

Although analogy suggests that the results obtained by the orthogonal scales may also apply to cases of multiple alleles, Cockerham states that this can be only a conjecture at the present time. The results of Kempthorne [13] to be discussed demonstrate, however, that his conjecture is valid at least for the case of noninbred

populations. The latter attacks the problem by means of the general factorial model. The total genotypic array of a random bred population without linkage is given by the expansion of the following expression:

$$\prod_{i=1}^{n}\left\{\sum p_j^i A_j^i\right\}^2 . \tag{30}$$

The superscripts are not exponents but refer to loci, the subscripts refer to particular alleles at a given locus, A refers to the gene, p is the relative frequency of the particular allele at the given locus, and n is the number of loci. The expansion thus gives every genotype, and the coefficient of each is the relative frequency with which it appears in the population. The gene symbols are also used to indicate genotypic values as in the work of Anderson and Kempthorne discussed above.

The value of a particular genotype, A_jA_k, where A_j and A_k are particular alleles at a single locus, may be expressed by the expansion of the following identity:

$$\begin{aligned} A_jA_k = {} & \Big[\sum_m P_mA_m\Big]\Big[\sum_n P_nA_n\Big] + \Big[A_j - \sum_m P_mA_m\Big]\Big[\sum_n P_nA_n\Big] \\ & + \Big[\sum_m P_mA_m\Big]\Big[A_k - \sum_n P_nA_n\Big] + \Big[A_j - \sum_m P_mA_m\Big]\Big[A_k - \sum_n P_nA_n\Big] . \end{aligned} \tag{31}$$

The subscripts m and n refer to the first and second alleles at the locus distinguished as having been contributed by, say, the sire and dam respectively, and the value of every term in the expansion is to be interpreted as averaged over all genotypes at other loci. The four terms on the right are the population mean, the additive effect of the jth allele, the additive effect of the kth allele, and the dominance effect of the two together, and this may be written

$$A_jA_k = \nu + a_j + a_k + d_{jk} . \tag{32}$$

If every symbol on the right of expression (31) is given a subscript, indicating that it refers to the ith locus, and if the entire expression is then multiplied by a similar one for the i'th locus and then by a similar one for the i''th locus and so on for all n loci, the expansion will indicate a genotypic value for a particular genotype expressed at every locus. Typical terms in the expansion of such an expression are as follows:

$$\nu = \prod_{i=1}^{n}\Big[\sum_m p_m^i A_m^i\Big]\Big[\sum_n p_n^i A_n^i\Big] \quad \text{(mean of population)} \tag{33}$$

$$a_{j_i}^i = \Big[A_j^i - \sum_m p_m^i A_m^i\Big]\Big[\sum_n p_n A_n\Big]\prod_{i'\neq i}\Big[\sum_m p_m^{i'} A_m^{i'}\Big]\Big[\sum_n p_n^{i'} A_n^{i'}\Big] \quad \begin{array}{l}\text{(additive effect}\\ \text{of the } j\text{th allele}\\ \text{at the } i\text{th locus)}\end{array} \tag{34}$$

$$\begin{aligned} (a^i a^{i'})_{j_i j_{i'}} = {} & \Big[A_{j_i}^i - \sum_m p_m^i A_m^i\Big]\Big[\sum_n p_n^i A_n^i\Big]\Big[A_{j_{i'}}^{i'} - \sum_m p_m^{i'} A_m^{i'}\Big]\Big[\sum_n p_n^{i'} A_n^{i'}\Big] \cdot \\ & \prod_{i'\neq i\neq i'}\Big[\sum_m p_m^{i'} A_m^{i'}\Big]\Big[\sum_n p_n^{i''} A_n^{i''}\Big] \end{aligned} \quad \begin{array}{l}\text{(additive by additive effect of the}\\ j_i\text{th allele at the } i\text{th locus, and the}\\ j_{i'}\text{th allele at the } i'\text{th locus.)}\end{array} \tag{35}$$

The general form taken by the expressions for dominance effects, additive × dominance effects, etc., can be inferred from these examples. In this manner Kempthorne shows that the genotype indicated by $\sum_{i=1}^{n} A^i_{j_i} A^i_{k_i}$ can be expressed by a series of terms like the following:

$$(36) \qquad \nu + \sum_{i=1}^{m} (a^i_{j_i} + a^i_{k_i}) \quad \text{the mean plus all additive effects of individual genes}$$

$$+ \sum_{i=1}^{n} d^{ii'}_{j_i k_i} \quad \text{the sum of all dominant effects at individual loci}$$

$$+ \sum_{i=1}^{n} \{(a^i a^{i'})_{j_i j_{i'}} + (a^i a^{i'})_{j_i k_{i'}} + (a^i a^{i'})_{k_i j_{i'}} + (a^i a^{i'})_{k_i k_{i'}}\}$$

the sum of all additive by additive interactions of genes at two loci.

Similar expressions may be written for additive-by-dominance interactions, dominance-by-dominance, additive-by-additive-by-additive at all sets of three loci, etc. Kempthorne demonstrates that these are all additive parameters and that all of them (except ν) have expected values of zero and are uncorrelated. The definitions of single locus additive effects and dominance effects are equivalent to those used in the past and they correspond, for the cases of two alleles and no inbreeding, to the definitions of Cockerham, and for the case of equal frequencies of the alleles to the definitions of Hayman and Mather with the exception noted in the discussion of the latter paper.

To compute correlations between individuals related in a given manner, the two groups with common ancestors are designated populations I and II. With the aid of an ingenious symbolism, which cannot be reproduced here because of lack of space, expressions indicating the genotypic array of the two populations are expanded and multiplied together in order to obtain the covariance. Only terms of the same order and kind can contribute to covariance, and only if genes are received from a common ancestor. In the case of dominance deviations, or of additive-by-additive deviations, two alleles have to trace to a common ancestor for the expected covariance to be other than zero, and three alleles in the case of dominance-by-additive or of additive-by-additive-by-additive deviations, and similarly for deviations of higher order. Inasmuch as the present model is restricted to noninbred populations, a sire and dam cannot be related and two relatives will be related either through their two sires or through their two dams, or both, or in crisscross fashion—not from the sire of one, for example, to both the sire and dam of the other. Kempthorne adapts the procedure that Malécot [16] has applied to one locus to the case of many loci. If Wright's path coefficient between gametes contributed to the two related individuals by one set of parents (two sires, or sire of one and dam of the other) is given the designation ϕ, and the coefficient through the other set of parents is given the designation ϕ' (using Malécot's terminology), then it is shown that the p and the q of Cockerham's results (given above) are equal to $(\phi + \phi')/2$ and $\phi\phi'$ respectively. The result now demonstrated is equivalent to that given by Cockerham, namely that the correlation of nth order deviations in-

volving additive nomenclature A times, and dominance nomenclature D times, is $p^A q^D$, where p and q are defined as mentioned above.

I regret that it is impossible to more than suggest the beautiful and rigorous development of Kempthorne in a brief presentation. The ideas that form the basis of the final result are not, however, difficult to grasp. Consider two individuals related separately through their two sires and through their two dams. The probability that the genes at a given locus contributed by the sires to the two individuals are descendants of the same gene in a common ancestor is Wright's path coefficient computed solely through paths passing through the sires. The equivalent probability with respect to the genes contributed at the same locus by the dams is Wright's path coefficient computed solely through the dams. The probability that both alleles at this locus in both individuals trace back to common genes in the common ancestors is thus the product of the two paths. Only in this event will the expected covariance of the dominance deviations be other than zero. Where there are n-order deviations n genes have to trace back to common ancestors; if less than n trace back in this way, there is no covariance contribution from the n-order deviations, although there may well be contributions with respect to lower order deviations.

Kempthorne and Cockerham by remarkably different methods have applied the same basic idea to Mendelian populations and arrived at entirely compatible results. Kempthorne restricts himself to noninbred populations, whereas Cockerham's results, though applicable to all kinds of inbreeding, are based on populations with only two alleles at each locus.

The methods described are likely to be most useful if genotypic variance is found in practice to be due chiefly to main effects and low order interactions. The suggestion that this is probably true is made both by Anderson and Kempthorne and by Hayman and Mather in the articles discussed above. The ultimate resolution of this problem must of course be empirical, and the methods themselves provide the basis of the necessary tests. As an illustration of the usefulness of the new methods, as well as of possible techniques of test, we may consider the demonstration of Henderson [17], by the use of Cockerham's results for inbred populations, that higher order interactions will contribute little to specific combining ability unless the inbreeding coefficient of the parents is close to unity. Conversely, if the specific combining ability were found to change considerably in the last stages of inbreeding, we could conclude that higher order interactions did exist to an important degree.

It may be worthwhile as a preliminary gesture to see what arguments there may be in mathematical theory or general genetic knowledge for supposing variance might or might not be largely composed of the low order varieties. Consider first a large panmictic dihybrid population with two equally frequent alleles at each locus. Suppose we assign each genotype a value at random from a normal population with unit variance, so that no two genotypes are alike. The sampling variance of the genotypic means of such populations is 9/64 so that the expected variance of the population itself is 53/64. This population, in which the values are assigned without any reference to the genes has, nevertheless, the variance components indicated in table VIII, computed by means of the appropriate variance expressions given by Cockerham [14]. The correctness of these components has been tested by computing the correlation between half-sibs in such a population by the variance component methods that have just been described above and independently in terms of

TABLE VIII

PARTITION OF VARIANCE IN DIHYBRID F_2 WHERE GENOTYPES ARE ASSIGNED VALUES AT RANDOM

Type of Variance	Proportion
A locus additive	6/55
A locus dominance	9/55
B locus additive	6/55
B locus dominance	9/55
Add. × add	4/55
Add. × dom	6/55
Dom. × add	6/55
Dom. × dom	9/55

the frequency of identical genotypes in related and unrelated individuals. Both methods give a correlation of 13/220. With assurance of the correctness of the method employed, then, the proportion of variance due to single factor additive effects was computed for similar "random" populations with higher numbers of loci, and these results are shown in table IX. The proportion for five loci is rather

TABLE IX

PROPORTION OF TOTAL SINGLE FACTOR ADDITIVE VARIANCE IN VARIOUS MODELS

Number of Loci	Type of Model		
	"Random"	Duplicate	Complementary
1	.4	.667	.667
2	.218	.571	.267
3	.111	.095	.486
4	.0527	.031	.411
5	.0255	.0098	.346

small and it is obvious that large amounts of high order interactions can not be shown to be improbable unless whatever pertinent information may be available on the mode of gene action is taken into account.

Of the simple well-known models of epistasis, probably the most extreme is that of duplicate factors, in which only the homozygous recessive at n loci produces a differential result. Actual instances that correspond closely to such models have been observed, up to the case of three loci, in plants of polyploid origin. The simple complementary model, in which a minimum of one dominant gene at each of n loci is sufficient to produce a result seems, from considerations of the relationships of genes to chain reactions in biochemical syntheses in lower organisms, to be more realistic. The proportions of total single factor additive variance for both these models for various number of loci are also shown in table IX. If the alleles do not produce all-or-none effects, the amount of additive variance can be greatly increased.

The conclusions to be drawn from this table and similar calculations must be exceedingly tentative and provisional. On the one hand, the possibility that higher

order interactions may be of general importance is certainly not excluded by these considerations, but neither do they preclude the view, and perhaps they encourage it, that most genetic variation, in the case of continuously variable characters in crossbred populations, may be treated as single locus effects plus low order multilocus interactions. With the new mathematical tools described in the preceding pages for studying interactions between loci, it is to be hoped that progress in the elucidation of these questions will be more rapid in the future than it has been in the past.

REFERENCES

[1] R. A. FISHER, "The correlation between relatives on the supposition of Mendelian inheritance," *Trans. Roy. Soc. Edin.*, Vol. 52 (1918–19), pp. 399–433.

[2] SEWALL WRIGHT, "The analysis of variance and the correlations between relatives with respect to deviations from an optimum," *Jour. of Genetics*, Vol. 30 (1935), pp. 243–256.

[3] BRUCE GRIFFING, "Analysis of quantitative gene action by constant parent regression and related techniques," *Genetics*, Vol. 35 (1950), pp. 303–321.

[4] T. W. HORNER, "Non-allelic gene interaction and the interpretation of quantitative genetic data." Unpublished Ph.D. thesis, North Carolina State College, 1952.

[5] J. L. LUSH, W. F. LAMOREUX, and L. N. HAZEL, "The heritability of resistance to death in the fowl," *Poultry Science*, Vol. 27 (1948), pp. 375–388.

[6] A. ROBERTSON and I. M. LERNER, "The heritability of all-or-none traits: viability of poultry," *Genetics*, Vol. 34 (1949), pp. 395–411.

[7] EVERETT R. DEMPSTER and I. MICHAEL LERNER, "Heritability of threshold characters," *Genetics*, Vol. 35 (1950), pp. 212–236.

[8] V. L. ANDERSON and OSCAR KEMPTHORNE, "A model for the study of quantitative inheritance," *Genetics*, Vol. 39 (1954), pp. 883–898.

[9] I. KHAMBANONDA, "Polygenic inheritance of fruit size in red pepper (*capsicum frutescen L.*)," Unpublished Ph.D. thesis, Iowa State College, 1948.

[10] FRED T. SCHULTZ and W. E. BRILES, "The adaptive value of blood group genes in chickens," *Genetics*, Vol. 38 (1953), pp. 34–50.

[11] K. MATHER, *Biometrical Genetics*, London, Methuen and Co., Ltd., 1949.

[12] B. I. HAYMAN and K. MATHER, "The description of genic interactions in continuous variation," *Biometrics*, Vol. 11 (1955), pp. 69–82.

[13] OSCAR KEMPTHORNE, "The correlation between relatives in a random mating population," *Proc. Roy. Soc. Lond.*, Ser. B., Vol. 143 (1954–55), pp. 103–113.

[14] C. CLARK COCKERHAM, "An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present," *Genetics*, Vol. 39 (1954), pp. 859–882.

[15] ALAN ROBERTSON, "The effect of inbreeding on the variation due to recessive genes," *Genetics*, Vol. 37 (1952), pp. 189–207.

[16] G. MALÉCOT, *Les Mathématiques de l'Hérédité*, Paris, Masson et Cie, 1948.

[17] C. R. HENDERSON, "The effect of inbreeding on genetic components of general and specific combining ability" (abstract), *Jour. of Animal Science*, Vol. 13 (1954), p. 959.