March, 1956 BU-63-M

AN ALTERNATIVE TO THE HAYMAN-MATHER MODEL FOR GENIC INTERACTION UNDER SELFING

D. S. Robson
Cornell University

Introduction. B. I. Hayman and K. Mather (1955) have constructed a model which completely describes phenotypes with respect to a quantitative character in a selfing series derived from the cross of two homozygous diploid parents. We shall here obtain an alternative construction which utilizes a device employed by Kempthorne (1954) for the random mating problem. The present result differs only slightly from the Hayman-Mather model and offers no particular advantage over their result.

Mather (1949) and Hayman and Mather (1955) use the expression "continuous variation" in place of "quantitative", however the intent appears to be the same. No attempt appears ever to have been made to describe precisely what is meant by a quantitative character in genetics and it would seem desirable to do so, first for the simple purpose of establishing a common understanding among geneticists, and second to aid in the establishment of logical consistency in the foundations of the science of genetics. To guarantee consistency, however, it would be necessary to examine such a definition in context with the other basic definitions of the science and these are not immediately available, at least to this writer. We shall, therefore, merely indicate some properities of a quantitative heritable character which seem to be implied by the present use of the term among geneticists.

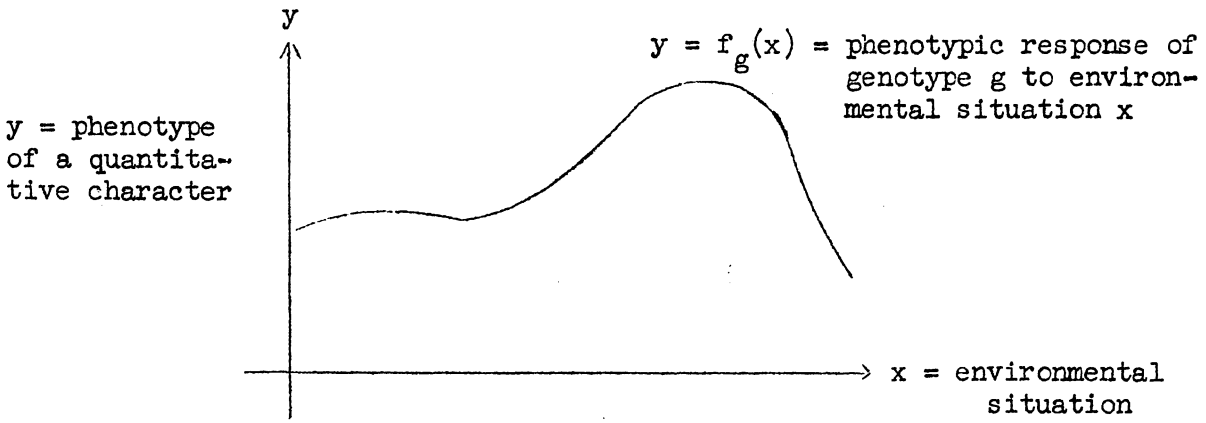
First, a quantitative character can be described numerically. This implies an ordering among the possible phenotypes and a scale of measurement which preserves this ordering. Furthermore, and particularly when the collection of possible phenotypes is uncountable, the ordering among phenotypes must be, in some sense, a natural ordering, for it is otherwise beyond description. A heritable character such as shape of an organ, for example, is difficult to quantify because there is no self-evident way of ordering the possible phenotypes. In any event, the order relationship is defined by the geneticist as a device to aid in interpreting the inheritance

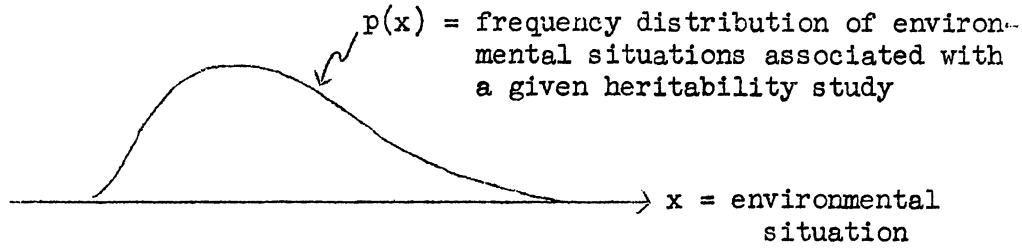
system; it is not an integral part of the system itself. Two different geneticists may use two quite different order relationships and measurement scales in describing the same inheritance system, and in this sense the ordering is artificial.

Second, according to common usage, a quantitative heritable character is subject to environmental as well as genetic influence. This feature has been described by somewhat misleading expressions such as partial or incomplete heritability, thus reserving the concept of complete heritability to describe any character whose expression is uniquely determined by genotype. One way of defining the relationship between heredity and environment is to think of heredity in terms of the transmission from parents to offspring of a set of potential responses corresponding to the set of possible environmental situations, the particular set of potential responses in an individual being uniquely determined by the genes he receives from his parents. Every individual of a given genotype, then, would possess an identical set of potential responses to environment, and under identical environments two individuals of a given genotype would then exhibit the same response, or phenotype. The only reason for variation among individuals of the same genotype, then, is that these individuals are subjected to different environmental situations and hence realize different potentials from among a constant set of potentials. Viewed in this light, heritability is, in a sense, always complete.

We shall assume that in an experimental study of quantitative inheritance the collection of environmental situations is fixed and that the assignment of an individual to an environmental situation is made independent of the genotype of that individual. It is meaningful, then, to speak of the average phenotype of a given genotype since all genotypes are, on the average, exposed to the same range of environmental conditions. The average phenotype of a given genotype is called the genotypic value of that genotype; clearly, genetic value depends not only upon genotype but also upon the fixed set of environmental situations. Thus, if environment and phenotype were both one-dimensional real-valued variables, say x and y, respectively, then associated with each genotype g would be a set of potential

phenotypes defined by a function $f_g(x)$ of the environmental situation x and associated with a given heritability study would be a frequency distribution p(x) of environmental situations x:





The average phenotype associated with the genotype g is then given by the sum $\sum_x f_g(x)p(x)$ or, more generally, by the corresponding integral. Our assumption of independence is implicit in the notation that the function p does not depend upon g, and the statement that genotypic value depends upon the set of environmental situations appearing in the study is implied by the role of p(x) in the formula for genetic value, $\sum_x f_g(x)p(x)$.

The model. We shall use the device employed by Kempthorne of letting the same symbol denote both genotype and genotypic value. Since we are here concerned only with genotypes derived from two homozygous diploid parents the genotype of an individual may be regarded as completely specified by the allelic

pairs at the n loci at which the two parents differ. Thus, if $A^i_{j_i} A^i_{j_k}$ is the genotype at the i'th locus, then

$$(1) \qquad \prod_{i=1}^{n} A^i_{j_i} A^i_{j_k} = A^1_{j_1} A^1_{k_1} \cdots A^n_{j_n} A^n_{k_n}$$

denotes the complete genotype, or it may be interpreted as the genotypic value of that genotype. The subscripts j_i and k_i can assume only two values, say 0 and 1, since there are only two possible alleles at each locus.

For n = 1 the (mean) phenotypes may be completely described by the following model

$$(2) \qquad A^1_{j_1} A^1_{k_1} = \left[\frac{A^1_0 A^1_0 + A^1_1 A^1_1}{2}\right] + (1-j_1-k_1)\left[\frac{A^1_0 A^1_0 - A^1_1 A^1_1}{2}\right]$$

$$+ [1-(1-j_1-k_1)^2]\left[\frac{-A^1_0 A^1_0 + 2A^1_0 A^1_1 - A^1_1 A^1_1}{2}\right]$$

or, more briefly,

$$(3) \qquad A^1_{j_1} A^1_{k_1} = \mu_1 + \delta(j_1 k_1) a_1 + [1 - \delta^2(j_1 k_1)] d_1 \quad .$$

Thus, for example,

$$A^1_0 A^1_0 = \left[\frac{A^1_0 A^1_0 + A^1_1 A^1_1}{2}\right] + (1)\left[\frac{A^1_0 A^1_0 - A^1_1 A^1_1}{2}\right] + [1-1]\left[\frac{-A^1_0 A^1_0 + 2A^1_0 A^1_1 - A^1_1 A^1_1}{2}\right]$$

or

$$A^1_0 A^1_0 = \mu_1 + a_1 \, .$$

The quantity $\mu_1 = \frac{1}{2}[A^1_0 A^1_0 + A^1_1 A^1_1]$ is here the midparent while

$$a_1 = \frac{A^1_0 A^1_0 - A^1_1 A^1_1}{2} = A^1_0 A^1_0 - \frac{A^1_0 A^1_0 + A^1_1 A^1_1}{2} = A^1_0 A^1_0 - \mu_1$$

is the deviation of the genotypic value $A^1_0 A^1_0$ from the midparent and is called the additive effect of A^1_0 and

$$d_1 = \frac{-A^1_0 A^1_0 + 2A^1_0 A^1_1 - A^1_1 A^1_1}{2} = A^1_0 A^1_1 - \frac{A^1_0 A^1_0 + A^1_1 A^1_1}{2} = A^1_0 A^1_1 - \mu_1$$

is the deviation of the heterozygote from the midparent and is called the dominance effect of A_o^1.

In the F_m generation of a selfing series the genotypes $A_o^1A_o^1$, $A_o^1A_1^1$, $A_1^1A_1^1$ appear with relative frequencies $\frac{1}{2} - \frac{1}{2^m}$, $\frac{1}{2^{m-1}}$, $\frac{1}{2} - \frac{1}{2^m}$, respectively; or, letting $\lambda_m = \frac{1}{2^{m-1}}$, these frequencies may be written $\frac{1}{2}(1-\lambda_m)$, λ_m, $\frac{1}{2}(1-\lambda_m)$. Since $\delta(j_1,k_1)$ possesses the same symmetric properities as genotype; i.e., $\delta(j_1,k_1) = \delta(k_1,j_1)$ then the frequency distribution of δ in F_m is likewise

$$\lambda_m = P_m\left\{\delta(j_1,k_1) = 0\right\} = \tfrac{1}{2}\left[1 - P_m\left\{\delta(j_1,k_1) = -1\right\}\right].$$

$$= \tfrac{1}{2}\left[1 - P_m\left\{\delta(j_1,k_1) = +1\right\}\right].$$

The mean phenotype in F_m, written $E_m\left\{A_{j_1}^1 A_{k_1}^1\right\}$, is therefore

$$E_m\left\{A_{j_1}^1 A_{k_1}^1\right\} = \mu_1 + a_1 E_m\left\{\delta(j_1k_1)\right\} + d_1 E_m\left\{1 - \delta^2(j_1k_1)\right\}.$$

where

$$E_m\left\{\delta(j_1k_1)\right\} = (+1)\left[\tfrac{1}{2}(1-\lambda_m)\right] + (0)\left[\lambda_m\right] + (-1)\left[\tfrac{1}{2}(1-\lambda_m)\right] = 0$$

and

$$E_m\left\{\delta^2(j_1k_1)\right\} = (+1)^2\left[\tfrac{1}{2}(1-\lambda_m)\right] + (0)^2\left[\lambda_m\right] + (-1)^2\left[\tfrac{1}{2}(1-\lambda_m)\right] = 1-\lambda_m$$

so

$$E_m\left\{A_{j_1}^1 A_{k_1}^1\right\} = \mu_1 + \lambda_m d_1$$

The genetic variance, or the variance among genotypic values, in F_m is therefore (writing simply δ_1, with the argument understood, instead of $\delta(j_1,k_1)$)

$$V_m = E_m\left\{A_{j_1}^1 A_{k_1}^1 - \mu_1 - \lambda_m d_1\right\}^2$$

$$= E_m\left\{\delta_1 a_1 + \left[1 - \delta_1^2 - \lambda_m\right]d_1\right\}^2$$

$$= E_m\left\{ \delta_1^2 a_1^2 + [(1-\lambda_m) - \delta_1^2]^2 d_1^2 + 2\delta_1(1-\delta_1^2-\lambda_m)a_1 d_1 \right\}$$

$$= (1-\lambda_m)a_1^2 + [(1-\lambda_m)^2 - 2(1-\lambda_m)^2 + (1-\lambda_m)]d_1^2$$

$$(4) \qquad V_m = (1-\lambda_m)[a_1^2 + \lambda d_1^2]$$

because $E_m\left\{ \delta^3(j_1 k_1) \right\} = E_m\left\{ \delta(j_1 k_1) \right\} = 0$.

A model for the general case of n loci may now be constructed simply by substituting in (1) the expression for $A^i_{j_i} A^i_{k_i}$ defined by (2), or, with the appropriate interpretation, by (3). Admittedly, the symbols in (2) and (3) were intended to refer to genotypic values rather than genotypes; as a device in constructing the model, however, equation (2) may be regarded as an identity function on the space of genotypes and substituted into the genotypic expression (1). And to further simplify the notation we shall use (3) as an identity on the space of genotypes; thus the following identity obtains on the space of genotypes

$$(5) \qquad \prod_{i=1}^{n} A^i_{j_i} A^i_{k_i} = \prod_{i=1}^{n}\left\{ \mu_i + \delta_i a_i + (1-\delta_i^2)d_i \right\}$$

$$= \sum_{h=0}^{n} \sum_{t=0}^{n-h} \sum_{i_\upsilon \neq j_\alpha} \prod_{\upsilon=1}^{h} \delta_{i_\upsilon} \prod_{\alpha=1}^{t} (1-\delta^2_{j_\alpha}) \prod_{\upsilon=1}^{h} a_{i_\upsilon} \prod_{\alpha=1}^{t} d_{j_\alpha} \prod_{i \neq i_\upsilon, j_\alpha} \mu_i$$

$$i_1 < \cdots < i_h \qquad \upsilon = 1, \ldots, h$$

$$j_1 < \cdots < j_t \qquad \alpha = 1, \ldots, t$$

In its expanded form the right side of (5) is a function of the 3^n genotypes of the form $\prod_{1}^{n} A^i_{j_i} A^i_{k_i}$ and upon substituting genotypic values for genotypes we obtain a model which completely describes (mean) phenotypes for the case of n loci. Letting $\mu = \prod_{1}^{n} \mu_i$ and

$$\prod_{\upsilon=1}^{h} a_{i_\upsilon} \prod_{\alpha=1}^{t} d_{j_\alpha} \prod_{i\neq i_\upsilon, j_\alpha} \mu_i = (a^h d^t)_{i_1\cdots i_h j_1\cdots j_t}$$

$$\upsilon=1,\ldots,h$$

$$\alpha=1,\ldots,t$$

with the understanding that $\prod_{\upsilon=1}^{o} a_{i_\upsilon} = \prod_{\alpha=1}^{o} d_{j_\alpha} = \prod_{i=1}^{o} \mu_i = 1$ we have as an expression of genotypic value

$$(6)\quad \prod_{i=1}^{n} A^i_{j_i} A^i_{k_i} = \mu + \sum_{h=0}^{n} \sum_{t=0}^{n-h} \sum_{i_\upsilon\neq j_\alpha} \Big[\prod_{\upsilon=1}^{h} \delta_{i_\upsilon} \prod_{\alpha=1}^{m} (1-\delta^2_{j_\alpha})\Big](a^h d^t)_{i_1\cdots i_h j_1\cdots j_t}.$$

$$h+t\neq 0 \quad i_1<\cdots<i_h$$

$$j_1<\cdots<j_t$$

If segregation is independent at the several loci; i.e., if linkage is absent, then the δ_i may be regarded as independent and identically distributed chance variables, so that

$$E_m\Big\{\prod_{\upsilon=1}^{h}\delta_{i_\upsilon} \prod_{\alpha=1}^{t}(1-\delta^2_{j_\alpha}) \Big| i_\upsilon\neq j_\upsilon,\ i_1<\cdots<i_h, j_1<\cdots<j_t\Big\} = \prod_{\upsilon=1}^{h} E_k\{\delta_1\} \prod_{\alpha=1}^{t} E_k\{1-\delta_1^2\}$$

$$= \begin{cases} 0 \text{ for } h>0 \\ \lambda_m^t \text{ for } h=0. \end{cases}$$

The phenotypic mean in F_m is therefore

$$E_m\Big\{\prod_{i=1}^{n} A^i_{j_i} A^i_{k_i}\Big\} = \mu + \sum_{t=1}^{n} \lambda_m^t \sum_{j_1<\cdots<j_t} (d^t)_{j_1\cdots j_t}.$$

We note than $E_m\Big\{\prod_{i=1}^{n} A^i_{j_i} A^i_{k_i}\Big\}$ may also be computed by regarding

$$E_m(A^i_{j_i} A^i_{k_i}) = \frac{A^i_o A^i_o + A^i_1 A^i_1}{2} + \lambda_m \frac{-A^i_o A^i_o + 2A^i_o A^i_1 - A^i_1 A^i_1}{2} = \mu_i + \lambda_m d_i$$

as a function on the space of genotypes. Then

$$(7) \qquad E_m\left\{\prod_{i=1}^{n} A^i_{j_i} A^i_{k_i}\right\} = \prod_{i=1}^{n} E_m\left\{A^i_{j_i} A^i_{k_i}\right\} = \prod_{i=1}^{n}(\mu_i + \lambda_m d_i)$$

$$= \mu + \sum_{t=1}^{n} \lambda_m^t \sum_{j_1<\cdots<j_t} (d^t)_{j_1\cdots j_t}.$$

The genetic variance in F_m is, under the assumption of no linkage,

$$V_m = E_m\left\{ \sum_{\substack{h=0 \\ h+t\neq 0}}^{n} \sum_{t=0}^{n-h} \sum_{\substack{i_\upsilon \neq j_\upsilon \\ i_1<\cdots<i_h \\ j_1<\cdots<j_t}} \left[\prod_{\upsilon=1}^{h} \delta_{i_\upsilon} \prod_{\alpha=1}^{t}(1-\delta^2_{j_\alpha})\right](a^h d^t)_{i_1\cdots i_h j_1\cdots j_t} \right.$$

$$\left. - \sum_{t=1}^{n} \lambda_m^t \sum_{j_1<\cdots<j_t} (d^t)_{j_1\cdots j_t} \right\}^2$$

$$(9)^* = \sum_{\upsilon=1}^{n}(1-\lambda_m)^\upsilon \sum_{i_1<\cdots<i_\upsilon} \sum_{k=0}^{\upsilon} \lambda_m^k \left\{ \sum_{\alpha=0}^{n-\upsilon} \lambda^\alpha \sum_{\substack{j_1<\cdots<j_\alpha \\ j_1,\ldots,j_\alpha \neq i_1,\ldots,i_\upsilon}} (a^{\upsilon-k} d^{k+\alpha})_{i_1\cdots i_\upsilon j_1\cdots j_\alpha} \right\}^2$$

Discussion. The Hayman-Mather model may be constructed in an entirely similar manner starting with the basic identity

$$A^1_{j_1} A^1_{k_1} = \left[\frac{A^1_o A^1_o + 2A^1_o A^1_1 + A^1_1 A^1_1}{4}\right] + (1-j_1-k_1)\left[\frac{A^1_o A^1_o - A^1_1 A^1_1}{2}\right]$$

$$+ \frac{1-2(1-j_1-k_1)^2}{2}\left[\frac{-A^1_o A^1_o + 2A^1_o A^1_1 - A^1_1 A^1_1}{2}\right]$$

instead of (2), or

$$A^1_{j_1} A^1_{k_1} = \mu^1_1 + \delta(j_1 k_1) a_1 + \frac{1-2\delta^2(j_1 k_1)}{2} d_1$$

*See appendix for derivation and correction.

instead of (3). The difference in the end result may be illustrated by considering the two models for the case n = 2. For the model of the present paper

$$A^1_{j_1} A^1_{k_1} A^2_{j_2} A^2_{k_2} = \mu + \delta_1(a)_1 + \delta_2(a)_2 + (1-\delta_1^2)(d)_1 + (1-\delta_2^2)(d)_2$$
$$+ \delta_1\delta_2(a^2)_{12} + \delta_1(1-\delta_2^2)(ad)_{12} + \delta_2(1-\delta_1^2)(ad)_{21}$$
$$+ (1-\delta_1^2)(1-\delta_2^2)(d^2)_{12}.$$

Thus,

	$A^1_0A^1_0$	$A^1_0A^1_1$	$A^1_1A^1_1$
	μ	μ	μ
$A^2_0A^2_0$	$+(a)_1+(a)_2$ $+(a^2)_{12}$	$+(d)_1+(a)_2$ $+(ad)_{21}$	$-(a)_1+(a)_2$ $-(a^2)_{12}$
	μ	μ	μ
$A^2_0A^2_1$	$+(a)_1+(d)_2$ $+(ad)_{12}$	$+(d)_1+(d)_2$ $+(d^2)_{12}$	$-(a)_1+(d)_2$ $-(ad)_{12}$
	μ	μ	μ
$A^2_1A^2_1$	$+(a)_1-(a)_2$ $-(a^2)_{12}$	$+(d)_1-(a)_2$ $-(ad)_{21}$	$-(a)_1-(a)_2$ $+(a^2)_{12}$

For the Hayman-Mather model

$$A^1_{j_1} A^1_{k_1} A^2_{j_2} A^2_{k_2} = \mu' + \delta_1(a)'_1 + \delta_2(a)'_2 + (\tfrac{1}{2} - \delta_1^2)(d)'_1 + (\tfrac{1}{2} - \delta_2^2)(d)'_2$$
$$+ \delta_1\delta_2(a^2)'_2 + \delta_1(\tfrac{1}{2} - \delta_2^2)(ad)'_{12} + \delta_2(\tfrac{1}{2} - \delta_1^2)(ad)'_{21}$$
$$+ (\tfrac{1}{2} - \delta_1^2)(\tfrac{1}{2} - \delta_2^2)(d^2)'_{12}.$$

or

	$A_o^1A_o^1$	$A_o^1A_1^1$	$A_1^1A_1^1$
$A_o^2A_o^2$	μ' $+(a)'_1+(a)'_2$ $-\frac{1}{2}(d)'_1-\frac{1}{2}(d)'_2$ $+(a^2)'_{12}$ $-\frac{1}{2}(ad)'_{12}-\frac{1}{2}(ad)'_{21}$ $+\frac{1}{4}(d^2)'_{12}$	μ' $+(a)'_2$ $+\frac{1}{2}(d)'_1-\frac{1}{2}(d)'_2$ $+\frac{1}{2}(ad)'_{21}$ $-\frac{1}{4}(d^2)_{12}$	μ' $-(a)'_1+(a)'_2$ $-\frac{1}{2}(d)'_1-\frac{1}{2}(d)'_2$ $-(a^2)'_{12}$ $+\frac{1}{2}(ad)_{12}-\frac{1}{2}(ad)_{21}$ $+\frac{1}{4}(d^2)'_{12}$
$A_o^2A_1^2$	μ' $+(a)'_1$ $-\frac{1}{2}(d)'_1+\frac{1}{2}(d)'_2$ $+\frac{1}{2}(ad)'_{12}$ $-\frac{1}{4}(d^2)'_{12}$	μ' $+\frac{1}{2}(d)'_1+\frac{1}{2}(d)'_2$ $+\frac{1}{4}(d^2)'_{12}$	μ' $+(a)'_1$ $-\frac{1}{2}(d)'_1+\frac{1}{2}(d)'_2$ $-\frac{1}{2}(ad)'_{12}$ $-\frac{1}{4}(d^2)'_{12}$
$A_1^2A_1^2$	μ' $+(a)'_1-(a)'_2$ $-\frac{1}{2}(d)'_1-\frac{1}{2}(d)'_2$ $-(a^2)'_{12}$ $-\frac{1}{2}(ad)'_{12}+\frac{1}{2}(ad)'_{21}$ $+\frac{1}{4}(d^2)_{12}$	μ' $-(a)'_2$ $+\frac{1}{2}(d)'_1-\frac{1}{2}(d)'_2$ $-\frac{1}{2}(ad)'_{21}$ $-\frac{1}{4}(d^2)'_{12}$	μ' $-(a)'_1-(a)'_2$ $-\frac{1}{2}(d)'_1-\frac{1}{2}(d)'_2$ $+(a^2)'_{12}$ $+\frac{1}{2}(ad)'_{12}-\frac{1}{2}(ad)'_{21}$ $+\frac{1}{4}(d^2)_{12}$

This latter model lacks, perhaps, a certain intuitive appeal of the former. It would be difficult to visualize, for example, why the additive x dominance and dominance x dominance interaction effects should contribute to the genotypic value of a homozygote.

The practical utility of either model appears to be virtually nil since genetic variance parameters of the model are not estimable when the number n of loci is unknown. It is possible, of course, to make certain further simplifying assumptions and thereby render certain parameters estimable, but this defeats the purpose of the model. An obvious simplification is the assumption of no dominance x (——) interaction effects, in which case the F_m mean becomes, from (8),

$$E_m \left\{ \prod_{i=1}^{n} A^i_{j_i} A^i_{k_i} \right\} = \mu + \lambda_m \sum_{j=1}^{n} (d)_j$$

and the variance in F_m becomes, from (9),

$$V_m = (1-\lambda_m)\Sigma(a)_i^2 + \lambda_m(1-\lambda_m)\Sigma(d)_i^2 + \sum_{\upsilon=1}^{n} (1-\lambda)^{\upsilon} \sum_{i_1 < \cdots < i_\upsilon} (a^\upsilon)^2_{i_1 \cdots i_\upsilon}.$$

Cutting off the above series at an arbitrary value of n would then permit the estimation of the additive type interaction variances -- provided, of course, that no linkage is present. A partial check on the realism of the assumption of no dominance x (———) interaction effects could be obtained from the means of the selfed generations. Clearly, by formula (8), the sums $\Sigma(d^\upsilon)$ are estimable up to an arbitrary value of υ.

References

Bateman, A. J. and Mather, K. (1955) The description of genic interactions in continuous variation. Biometrics, 11:69-82.

Kempthorne, O. (1954) The correlation between relatives in a random mating population. Proc. Roy. Soc. of London, B, 143:103-113.

Mather, K. (1949) Biometrical Genetics. Methuen, London.

APPENDIX

DERIVATION OF THE GENETIC VARIANCE V_m IN GENERATION F_m

Since

$$V_m = E\left\{\prod_{i=1}^{n} A^i_{j_i} A^i_{k_i}\right\}^2 - \left[E\left\{\prod_{i=1}^{n} A^i_{j_i} A^i_{k_i}\right\}\right]^2$$

becomes, with the appropriate interpretation of the symbols

$$V_m = E\left\{\prod_{i=1}^{n}(\mu_i + \delta_i a_i + (1-\delta_i^2)d_i)\right\}^2 - \left[E\left\{\prod_{i=1}^{n}(\mu_i + \delta_i a_i + (1-\delta_i^2)d_i)\right\}\right]^2$$

then, as in (7), we may write

$$V_m = \prod_{i=1}^{n} E\left\{\mu_i + \delta_i a_i + (1-\delta_i^2)d_i\right\}^2 - \prod_{i=1}^{n}\left[E\left\{\mu_i + \delta_i a_i + (1-\delta_i^2)d_i\right\}\right]^2$$

$$= \prod_{i=1}^{n}\left[(1-\lambda_m)(a_i^2 + \lambda_m d_i^2) + (\mu_i + \lambda_m d_i)^2\right] - \prod_{i=1}^{n}(\mu_i + \lambda_m d_i)^2$$

$$= \sum_{\upsilon=1}^{n}(1-\lambda_m)^{\upsilon} \sum_{I_\upsilon} \prod_{i \text{ in } I_\upsilon}(a_i^2 + \lambda_m d_i^2) \prod_{j \text{ in } \bar{I}_\upsilon}(\mu_j + \lambda_m d_j)^2$$

where $I_\upsilon = (i_1, \ldots, i_\upsilon)$ is at set of υ integers such that $1 \leq i_1 < \ldots < i_\upsilon \leq n$ and $\bar{I}_\upsilon$ is the complement of I_υ with respect to the set $(1, \ldots, n)$; i.e., $\bar{I}_\upsilon$ is the set of $n - \upsilon$ integers contained in $(1, \ldots, n)$ but not in I_υ, or we may write $\bar{I}_\upsilon = (1, \ldots, n) - I_\upsilon$. Thus,

$$V_m = \sum_{\upsilon=1}^{n}(1-\lambda_m)^{\upsilon} \sum_{I_\upsilon} \sum_{k=0}^{\upsilon} \lambda_m^k \sum_{I_k \text{ in } I_\upsilon}\left[\prod_{i \text{ in } I_\upsilon - I_k} a_i \prod_{j \text{ in } I_k} d_j\right]^2$$

$$\left[\sum_{\alpha=0}^{n-\upsilon} \lambda_m^\alpha \sum_{I_\alpha \text{ in } \bar{I}_\upsilon} \prod_{i \text{ in } \bar{I}_\upsilon - I_\alpha} \mu_i \prod_{j \text{ in } I_\alpha} d_j\right]^2$$

$$= \sum_{\upsilon=1}^{n}(1-\lambda_m)^{\upsilon} \sum_{I_\upsilon} \sum_{k=0}^{\upsilon} \lambda_m^k \sum_{I_k \text{ in } I_\upsilon}\left[\sum_{\alpha=0}^{n-\upsilon} \lambda_m^\alpha \sum_{I_\alpha \text{ in } \bar{I}_\upsilon} \prod_{i \text{ in } I_\upsilon - I_k} a_i \prod_{j \text{ in } I_k + I_\alpha} d_j \prod_{i \text{ in } I_\upsilon - I_\alpha} \mu_i\right]^2$$

or, applying the definition on page 7 we get

$$V_m = \sum_{\upsilon=1}^{n}(1-\lambda_m)^{\upsilon} \sum_{I_\upsilon} \sum_{k=0}^{\upsilon} \lambda_m^k \sum_{I_k \text{ in } I_\upsilon}\left[\sum_{\alpha=0}^{n-\upsilon} \lambda_m^\alpha \sum_{I_\alpha \text{ in } \bar{I}_\upsilon} (a^{\upsilon-k} d^{k+\alpha})_{I_\upsilon - I_k, I_k + I_\alpha}\right]^2$$

for the genetic variance in F_m.