ON THE EQUIVALENCE OF POLYKAYS OF THE SECOND DEGREE AND Σ'S[1]

BY EUGENE DAYHOFF[2]

Iowa State University

1. Introduction and summary. A generalization of Tukey's polykays [3], [4], and [5], was made by Hooke [1] in reference to sampling from a two-way array or population. These generalized polykays were christened "bipolykays." Smith [2] also developed these functions of degree two independently. Working with certain structural patterns in the analysis of variance, a set of functions denoted by Σ's were introduced by Wilk and Kempthorne [6] and formally defined and extended to include all "balanced" population structures by Zyskind [7]. Because of certain "nice" properties of symmetric means of degree two, the Σ expansions were found to be relatively simple and could be defined for a large class of structures. In contrast, the work on the extension of the polykays was limited only to sampling from a two-way population structure though polykays of higher degrees were also considered. Zyskind [8] recognized the equivalence of Hooke's bipolykays of degree two and a certain subset of the Σ's, and conjectured the equivalence of appropriately extended polykays with the whole set of Σ's for all balanced structures.

In this paper an extension of the bipolykays of degree two to "n-way-polykays" (henceforth referred to as generalized polykays) is made to encompass all balanced structures (as defined in [7]). (Since this paper was submitted, general definitions of polykays and symmetric means of all degrees have been formulated, properties of these developed, and the basic results applied to obtaining variances and co-variances of estimates of components of variation in certain two and three-factor balanced structures.) The equivalence of these generalized polykays and the Σ's is then shown.

2. Preliminaries. The symbol $\sum^{\neq}$ shall mean the sum over all subscripts that follow with the restriction that differently primed subscripts remain unequal. Symmetric means and polykays have been previously denoted by brackets and parentheses respectively [3], e.g. the symmetric mean $\langle ab\rangle = \sum^{\neq} x_i^a x_{i'}^b / N(N-1)$ where $i = 1, \cdots, N$, and the corresponding polykay is denoted (ab). In contrast to this "primary" notation Hooke [1] used a "secondary" notation, for example the secondary notation for $\langle ab \cdots d\rangle$ being $\langle p_1 p_2 \cdots p_a, q_1 q_2 \cdots q_b, \cdots, r_1 r_2 \cdots r_d\rangle$, where p_i, $i = 1, \cdots, a$, denotes the individual x_i and q_i, $i = 1, \cdots b$, the individual $x_{i'}$, etc. The entries $a, b, \cdots, d$ in the angle brackets

Received 3 September 1963.

[1] This Research was supported in whole or in part by the United States Air Force under Contract No. AF 33(616)-8269, monitored by the Wright Air Research Laboratory, Wright Air Development Center.

[2] Now at Texas A & M University.

are said to form a partition of the integer $m = a + b + \cdots + d$, where m is the degree of the symmetric mean. Thus in the secondary notation the comma separates the *parts* of the partition with $a, b, \cdots, d$, denoting the lengths of the partitions. Two partitions are said to be *equivalent*, or not *distinct* if they are identical, except possibly for the order of parts and the order of symbols within a part. For $m = 2$ the only partitions are 1 1 and 2 itself, or in the secondary notation p_1, q_1 and p_1p_2, respectively, which in this sequel will be denoted by p, q and pq.

Henceforth let α denote the partition pq and β the partition p, q. Let $\gamma, \delta, \cdots$ denote arbitrary partitions. A *subpartition* δ of a partition γ may be formed by inserting one or more commas between the letters of γ. Thus for $m = 2$, β is a subpartition of α. The following implicit definition of the polykays is given by Hooke:

Definition 2.1. $\langle\gamma\rangle = (\gamma) + \sum_i(\delta_i)$ where the summation is over all distinct subpartitions δ_i of γ. Two symmetric means, or polykays, are equivalent, or not distinct, if the partitions representing them can be made equivalent by renaming the symbols.

Thus, for example, the polykays of degree three are defined by the equations

$$\begin{aligned}\langle p, q, r\rangle &= (p, q, r),\\ \langle p, qr\rangle &= (p, qr) + (p, q, r),\\ \langle pqr\rangle &= (pqr) + (p, qr) + (q, pr) + (r, pq) + (p, q, r)\end{aligned}$$

which may be solved to yield

$$\begin{aligned}(p, q, r) &= \langle p, q, r\rangle,\\ (p, qr) &= \langle p, qr\rangle - \langle p, q, r\rangle,\\ (pqr) &= \langle pqr\rangle - \langle p, qr\rangle - \langle q, pr\rangle - \langle r, pq\rangle + 2\langle p, q, r\rangle,\end{aligned}$$

or in Tukey's primary notation,

$$\begin{aligned}k_{111} &= \langle 111\rangle,\\ k_{12} &= \langle 12\rangle - \langle 111\rangle,\\ k_3 &= \langle 3\rangle - 3\langle 12\rangle + 2\langle 111\rangle.\end{aligned}$$

In introducing bipolykays Hooke defined *generalized symmetric means* to be averages of monomial functions over a matrix, i.e., a polynomial of the form

$$(1/M) \sum_{p,q,\cdots,s,t}^{\neq} x_{pq}^{a_{pq}} \cdots x_{st}^{a_{st}}$$

where the symbol $\sum^{\neq}$ indicates the sum over all the subscripts with the restriction that subscripts represented by different letters must remain unequal throughout the summation, and M is the number of terms in the summation. The general term $x_{pq}^{a_{pq}} \cdots x_{st}^{a_{st}}$ contains m factors, of which a_{pq} are equal to x_{pq}, etc., the degree of the symmetric mean being $m = a_{pq} + \cdots + a_{st}$. To each factor a different symbol is assigned and the resulting set of symbols partitioned in two ways—once by rows and once by columns. Hence the secondary notation for the g.s.m. is an ordered pair $\langle\gamma_1/\gamma_2\rangle$ of partitions γ_1 and γ_2, each on the same set of symbols.

Following Hooke's formalization of the bipolykays we now define a non-commutative "dot-multiplication" for symmetric means from an n-way population structure as follows:

DEFINITION 2.2 $\langle\gamma_1\rangle\cdot\langle\gamma_2\rangle \cdot\cdots\cdot \langle\gamma_n\rangle = \langle\gamma_1/\gamma_2/ \cdots /\gamma_n\rangle$ if the γ_i consist of the same symbols where $\gamma_1, \cdots, \gamma_n$ are different partitions of the same set of letters, and is zero otherwise. Distributivity is assumed to provide dot-products for linear combinations of symmetric means.

3. Extension of bipolykays. The extension of bipolykays to generalized polykays is made simply in the following definition:

DEFINITION 3.1. $(\gamma_1/\gamma_2/ \cdots /\gamma_n) = (\gamma_1)\cdot(\gamma_2) \cdot\cdots\cdot (\gamma_n)$ where it is understood that the polykays are expressed in terms of symmetric means before dot multiplication.

EXAMPLE 3.1. Consider a two-way crossed population structure, i.e. $Y_{ij} = Y_{..} + (Y_{i.} - Y_{..}) + (Y_{.j} - Y_{..}) + (Y_{ij} - Y_{i.} - Y_{.j} + Y_{..})$ where $i = 1, \cdots, A$ and $j = 1, \cdots, B$, and a dot indicates a mean over the corresponding subscript has been taken. Then

$$
\begin{aligned}
(pq/pq) &= (pq)\cdot(pq) = (\langle pq\rangle - \langle p, q\rangle)\cdot(\langle pq\rangle - \langle p, q\rangle) \\
&= \langle pq\rangle\cdot\langle pq\rangle - \langle p, q\rangle\cdot\langle pq\rangle - \langle pq\rangle\cdot\langle p, q\rangle + \langle p, q\rangle\cdot\langle p, q\rangle \\
&= \langle pq/pq\rangle - \langle p, q/pq\rangle - \langle pq/p, q\rangle + \langle p, q/p, q\rangle \qquad (3.1)\\
&= \textstyle\sum Y_{ij}^2/AB - \sum^{\neq} Y_{ij}Y_{i'j}/AB(A-1) \\
&\quad - \textstyle\sum^{\neq} Y_{ij}Y_{ij'}/AB(B-1) \\
&\quad + \textstyle\sum^{\neq} Y_{ij}Y_{i'j'}/AB(A-1)(B-1).
\end{aligned}
$$

Note that the g.s.m.'s above occur with a plus or minus sign according to whether the letters in an odd or even number of the polykays contained in $(\alpha)\cdot(\alpha)$ have been subpartitioned. The generalization of this result to products of many polykays is obvious.

Hence for a n-way structure the following definition of polykays of degree two in terms of g.s.m.'s may be given:

DEFINITION 3.2. $(\theta_1/ \cdots /\theta_n) = \langle\theta_1/ \cdots /\theta_n\rangle + \Sigma(-1)^{\pi} \langle\theta_1'/ \cdots /\theta_n'\rangle$ where θ_i' is a subpartition of θ_i. If, for instance, θ_i is α, θ_i' can be α or β and if θ_i is β, θ_i' must be β. The sum is over all possible subpartitions of the θ_i and π is the number of changes α to β.

A general structure, however, consists of both nested and crossed factors. If in Example 3.1, the subscript j is considered nested within i, rather than crossed with it, sums of the form $\sum^{\neq} Y_{ij}Y_{i'j}$ and $\sum^{\neq} Y_{ij}Y_{i'j'}$ will be taken to have the same meaning since the sum over the subscript j is independent from one i-level to another, and the corresponding g.s.m. may be denoted

$$\sum^{\neq} Y_{ij} Y_{i'j'} / AB^2 (A - 1).$$

In view of this, g.s.m.'s of the form $\langle p, q/pq \rangle$ and $\langle p, q/p, q \rangle$ will be considered the same and in (3.1) may be canceled when they appear with opposite signs. Hence for a situation with the second factor nested within the first factor Equation (3.1) becomes

$$(pq/pq) = \langle pq/pq \rangle - \langle pq/p, q \rangle = \sum Y_i^2/AB - \sum^{\neq} Y_{ij} Y_{ij'} / AB(B - 1),$$

the polykay (pq/pq), of course, having a different meaning now.

Before formalizing a definition of generalized polykays to include nesting in structures we shall review a few definitions given by Zyskind [7].

DEFINITION 3.3. An *admissible mean* corresponding to a given population structure is a mean in which whenever a nested index appears then all the indices which nest it appear also.

Thus in Example 3.1, $Y_{..}$, $Y_{i.}$, $Y_{.j}$, and Y_{ij} are admissible means, while in the nested situation, $Y_{..}$, $Y_{i.}$, and $Y_{i(j)}$ are admissible means.

DEFINITION 3.4. Indices of the *rightmost bracket* are those indices of an admissible mean which nest no other subscripts of the mean. For convenience these indices are distinguished from the others by enclosure in parentheses.

Thus in Example 3.1 the indices i, j of the mean Y_{ij} are included in the rightmost bracket, (since i and j are both in the rightmost bracket the parentheses are omitted) while in the case of nesting only j is in the rightmost bracket and the admissible mean is denoted by $Y_{i(j)}$. To denote a general admissible mean we shall use the symbol $Y_{L(R)}$ where R is the set of subscripts in the rightmost bracket and L is the set of remaining subscripts.

We now propose a general definition of generalized symmetric means and polykays of degree two, for any balanced structure.

DEFINITION 3.5. Consider a general structure with n factors. Let S denote the set of subscripts $i, j, k, \cdots$, etc., and let the range of the subscripts be N_1, N_2, N_3, $\cdots$, N_n. Then

$$\langle \theta_1/\theta_2/ \cdots /\theta_n \rangle = \sum^{\neq} x_s x_{s'} / N_1^{\theta_1} N_2^{\theta_2} \cdots N_n^{\theta_n}$$

where

θ_i can be α or β if the ith factor is not nested by any other factor and $N_i^\alpha = N_i$, $N_i^\beta = N_i(N_i - 1)$;

θ_i can be α or β if all factors which nest the ith factor have α and then $N_i^\alpha = N_i$, $N_i^\beta = N_i(N_i - 1)$;

$\theta_i = \beta$ if one or more of the factors which nest the ith factor have β and then $N_i^\beta = N_i^2$; and where the sum is over all subscripts of S with the restriction that those corresponding to β are unequal.

In this definition we use the rule that a nested factor has a different subscript for every one of its levels [e.g. if j is nested in i then the values j takes for different values of i are different].

As an example consider again the case of factor i nesting a factor j. Then

$$\langle pq/pq\rangle = \sum x_{ij}^2/AB,$$

$$\langle pq/p,\, q\rangle = \sum\nolimits^{\neq} x_{ij}x_{ij'}/AB(B-1),$$

$$\langle p,\, q/pq\rangle = \text{is not defined},$$

$$\langle p,\, q/p,\, q\rangle = \sum\nolimits^{\neq} x_{ij}x_{i'j'}/AB^2(A-1).$$

It is obvious that there is a one-to-one correspondence of generalized symmetric means of degree two and admissible means. In the above case the admissible means are $Y_{i(j)}$, $Y_{i\cdot}$ none, $Y_{\cdot\cdot}$ respectively.

DEFINITION 3.6. Given a population structure P and an admissible mean $Y_{L(R)}$, denote by S the set of all subscripts of P where $S = L + R + Q$, Q being the set of remaining subscripts. Then

$$(\theta_1/\theta_2/\cdots/\theta_n) = \langle\theta_1/\theta_2/\cdots/\theta_n\rangle + \sum(-1)^{\pi}\langle\theta_1'/\theta_2'/\cdots/\theta_n'\rangle$$

$$\begin{aligned} \text{where}\quad \theta_i &= \alpha \text{ if } i\,\varepsilon\, L \quad \text{and} & \theta_i' &= \theta_i \text{ if } i\,\varepsilon\, L\\ &= \alpha \text{ if } i\,\varepsilon\, R & &= \alpha \text{ or } \beta \text{ if } i\,\varepsilon\, R\\ &= B \text{ if } i\,\varepsilon\, Q & &= \theta_i \text{ if } i\,\varepsilon\, Q \end{aligned}$$

and π is the number of α's which are changed to β, and the sum is over all possible subpartitions of θ_i where $i\,\varepsilon\, R$.

4. The Σ functions. The following two definitions from Zyskind [7] are stated before introducing the Σ's.

DEFINITION 4.1. A *component* corresponding to the admissible mean $Y_{L(R)}$ is a linear combination of admissible means obtained by selecting all those means which are yielded by $Y_{L(R)}$ when some, all, or none of its rightmost bracket subscripts are omitted in all possible ways. Whenever an odd number of indices is omitted the coefficient of the mean is minus 1 and whenever an even number is omitted the coefficient of the mean is plus 1.

Then it can be easily seen that a typical response or observation can be expressed as the sum of the components. This expression is commonly called the population identity.

A consequence of this definition of components is that the component corresponding to $Y_{L(R)}$ vanishes when summed over any of the subscripts in R.

EXAMPLE 4.1. In a two-way crossed structure the basic identity was seen to be $Y_{ij} = Y_{\cdot\cdot} + (Y_{i\cdot} - Y_{\cdot\cdot}) + (Y_{\cdot j} - Y_{\cdot\cdot}) + (Y_{ij} - Y_{i\cdot} - Y_{\cdot j} + Y_{\cdot\cdot})$ where there are four components, each corresponding to an admissible mean. This identity could be written $Y_{ij} = \mu + A_i + B_j + (AB)_{ij}$.

EXAMPLE 4.2. In a two-factor nested structure the population identity is $Y_{i(j)} = Y_{\cdot\cdot} + (Y_{i\cdot} - Y_{\cdot\cdot}) + (Y_{i(j)} - Y_{i\cdot})$ or $Y_{i(j)} = \mu + A_i + A(B)_{ij}$.

DEFINITION 4.2. The *component of variation* corresponding to the admissible mean $Y_{L(R)}$ is defined as

$$\sigma_{L(R)}^2 = \sum [L(R)]^2/L \prod_i (R_i - 1)$$

where $L(R)$ represents the component in the population identity corresponding to $Y_{L(R)}$, and the summation is over all subscripts of the leading mean of $L(R)$ and L_i and R_i denote the population range of the respective subscripts with $\prod_i L_i = L$ and $\prod_i R_i = R$.

For example, consider the two-factor nested structure whose identity is given in Example 4.2. Corresponding to the means $Y_{i(j)}$, $Y_{i\cdot}$, and $Y_{\cdot\cdot}$ are the components of variation $\sigma^2_{A(B)} = \sum_{ij} (Y_{i(j)} - Y_{i\cdot})^2/A(B - 1)$, $\sigma^2_A = \sum (Y_i - Y_{\cdot\cdot})^2/A - 1$, $\sigma^2 = Y^2_{\cdot\cdot}$ respectively.

Using the above notation, Zyskind's definition of the Σ's is as follows:

DEFINITION 4.3. Consider a particular component $L(R)$ and all σ^2's of the following form:

(i) the set of subscripts of σ^2 includes the set of subscripts corresponding to the leading term of the component as a subset,

(ii) the excess subscripts lie exclusively in the rightmost bracket of σ^2.

The linear combination of all such σ's, where the coefficient of a particular σ^2 with k excess subscripts is

$$(-1)^k \, 1/\text{Product of population ranges of the excess indices},$$

is defined as the Σ corresponding to the component $L(R)$ and is denoted by $\Sigma_{L(R)}$.

Again taking the two-factor crossed structure as an example we have according to this definition

$$\Sigma_\phi = Y^2_{\cdot\cdot}, \qquad \sum_A = \sigma^2_A - (1/B)\sigma^2_{AB},$$
$$\Sigma_{AB} = \sigma^2_{AB}, \qquad \sum_B = \sigma^2_B - (1/A)\sigma^2_{AB}.$$

5. The equivalence of the polykays and the Σ's. Before proceeding to the equivalence of the Σ's and generalized polykays a few lemmas will first be demonstrated.

LEMMA 5.1. *The sum of the coefficients of the* g.s.m.'s *in any generalized polykay except* $(\beta/\cdots/\beta)$ *add to zero. The coefficient of the* g.s.m. *in* $(\beta/\cdots/\beta)$ *is unity.*

PROOF. By definition, $(\beta/\cdots/\beta) = \langle\beta/\cdots/\beta\rangle$. Consider now the generalized polykay corresponding to $Y_{L(R)}$. Let k denote the number of $R_i \in R$. Then the number of possible subpartitions of the α corresponding to the subscripts R_i will be $2^k = \sum_{i=0}^k \binom{k}{k-i}$ and by Definition 3.6 the sum of the coefficients of the g.s.m.'s is

$$\sum_{i=0}^{k} (-1)^{k-i} \binom{k}{k-i} = (1-1)^k = 0.$$

LEMMA 5.2. *A* g.s.m. *of population values is equal to the sum of corresponding* g.s.m.'s *of all components, a corresponding* g.s.m. *being obtained by deleting all factors in the name of the population* g.s.m. *which do not occur in the name of the component and using the same partition of the remaining factors. For example in* *the two-way crossed structure*

$$\frac{\sum^{\neq} Y_{ij} Y_{ij'}}{AB(B-1)} = \mu^2 + \frac{\sum A_i^2}{A} + \frac{\sum^{\neq} B_j B_{j'}}{B(B-1)} + \frac{\sum^{\neq} (AB)_{ij}(AB)_{ij'}}{AB(B-1)}.$$

PROOF. This result is obvious as regards products of like components when the population identity is substituted for the population values appearing in the generalized symmetric mean. All products consisting of unlike components will vanish for there will be at least one subscript not common to the rightmost brackets of the two components of a product and the sum over this subscript will be zero by definition of the components.

LEMMA 5.3. *Each numerator of a* g.s.m. *of any component can be expressed as a sum of squares of that component, if it does not vanish. The sum of squares will be preceded by either a plus or minus sign. For example* $\sum^{\neq} A(B)_{i(j)}A(B)_{i'(k)} = 0$ *but,* $\sum^{\neq} A_i A_{i'} = -\sum A_i^2$.

PROOF. If the name of a g.s.m. of a component contains β partitions corresponding to factors in the non-rightmost bracket of the component that component g.s.m. vanishes because the factors of the non-rightmost bracket will nest at least one factor contained in the rightmost bracket and the sum over subscripts of a nested factor is independent from one level of the nesting factor to another—which is the case when there is a β partition corresponding to a nesting subscript— and by definition this sum is zero.

Thus the name of a g.s.m. of a component contains α or β partitions for the factors of the rightmost bracket of the set of factors involved in the component and contains only the α partitions for all other factors involved in the component. So with regard to the α partitioned factors contained in both the rightmost bracket and the non-rightmost bracket we have a sum of squares over levels. With regard to a factor in the rightmost bracket for which the name of the g.s.m. involves β we have a sum over all pairs of unequal levels. But

$$\sum^{\neq} x_j x_{j'} = -\sum x_j^2 \quad \text{if} \quad \sum x_j = 0.$$

So with regard to the numerator of a component g.s.m. each β in the name of the g.s.m. for a rightmost bracket factor can be replaced by α with multiplication by (-1). The numerator then becomes a sum of squares of the components if the particular type with a coefficient of $(-1)^r$ where r is the number of such β's corresponding to factors in the rightmost bracket.

Consider again the population structure P and the admissible mean $Y_{L(R)}$. As before let S denote the set of all subscripts of P where $S = L + R + Q$, R being the set of rightmost bracket subscripts and Q the set of subscripts not contained in the name of the admissible mean. Then we state the following theorem:

THEOREM 5.1. $\Sigma_{L(R)}$ = *the generalized polykay, as defined in Definition* 3.6, *whose name contains* α *for the factors of* L, α *for the factors of* R, *and* β *for the factors of* Q.

PROOF. Consider a given component g.s.m. whose corresponding component of variation is denoted by $\sigma_{S'}^2$, where $S' = L' + R'$, R' being the set of rightmost bracket factors in the name of the given component. When the generalized polykay above is expressed in terms of the component g.s.m.'s (by Lemma 5.2) it will be shown that g.s.m.'s of all components vanish except those whose subscripts S' are of the form $S' = L + R + Q^*$ where Q^* is a subset of Q and lies exclusively in R', the set of rightmost bracket subscripts of S'. These component g.s.m.'s will then be expressed as components of variation of the form $\sigma^2_{L(R+Q^*)}$ with the coefficient given in the definition of $\Sigma_{L(R)}$.

CASE 1. Consider first the case where the set of subscripts S', is null. Then the component g.s.m. $Y^2 = \sigma^2_{S'} = \sigma^2_{\phi}$ occurs in each g.s.m. of the polykay and by Lemma 5.1 vanishes for all polykays except $(\beta/\beta/\cdots/\beta)$ in which case it occurs with a coefficient of $(+1)$.

Now consider g.s.m.'s of components whose set of subscripts S' is non-empty. Then we distinguish the following three cases of component g.s.m.'s according to the relationship of the subscripts S' to the subscripts R of the generalized polykay.

CASE 2. Assume R is non-empty and take g.s.m.'s of components whose subscripts S_2' do not contain any of the subscripts R, i.e. $S_2' \subseteq L + Q$, where either L or Q may be empty. By Definition 3.6 the partitions corresponding to the factors in L and Q are α and β respectively in each g.s.m. of the generalized polykay and by Lemma 5.1 the g.s.m.'s of components vanish.

CASE 3. Consider again R to be non-empty but component g.s.m.'s whose subscripts S_3' contain some, but not all, of the subscripts of R, i.e. $S_3' \subseteq L + Q + R^*$ where $R^* \subset R$. In the polykay any g.s.m. has a corresponding g.s.m. differing with regard to one subscript in $R - R^*$ and therefore in sign but these two g.s.m.'s give the same component g.s.m. and therefore cancel each other.

CASE 4. Now take the components whose subscripts S_4' contain all the subscripts R, where R may be null as in the case of the polykay $(\beta/\beta/\cdots/\beta)$. Since the subscripts of R are contained in S_4' so must be the set L, and hence $S_4' = L + R + Q^*$, where $Q^* \subseteq Q$. This case may be subdivided into the following two cases depending on whether Q^* lies exclusively or not in the set R_4', the set of rightmost subscripts of S_4'.

CASE 4a. If the set Q^* is non-empty and does not lie exclusively in R_4', then a subset of Q^* must be contained in L_4', the non-rightmost bracket part of S_4', but according to Lemma 5.3 the component g.s.m.'s whose subscripts are of this form vanish.

CASE 4b. Consider now component g.s.m.'s whose subset of subscripts Q^* is either null or lies exclusively in R_4'. Then each component g.s.m. may be expressed as a sum of squares of the corresponding component by Lemma 5.3, preceded by a plus or minus sign. In fact, the sum of squares will appear with like sign in each g.s.m. of the polykay, positive if the number of subscripts in Q^* is even and negative if the number is odd. For suppose Q^* has an odd number of subscripts. The original polykay has α partition for the set L, α partitions for the set R and β partitions for the set Q^* and when the polykay is expressed in terms of g.s.m.'s we get the leading term appearing with a plus sign and with the same number of α's and β's as in the name of the polykay. But since Q^* contains an odd number of subscripts the sign of the sum of squares of the component in question will be $(+1)(-1) = -1$ in the leading g.s.m. by Lemma 5.3. Now a second g.s.m. of the polykay will have one α of R changed to β and thus preceded by a negative sign because of the additional β. But the component g.s.m. now has an additional β partition so that the sum of squares will be preceded by a $(-1)(-1)(-1) = -1$ again. Another g.s.m. will have two of the α's changed to β and hence preceded by a plus one so that the sum of squares will be preceded by a $(+1)(-1)(-1)(-1) = -1$ and so on. A similar argument shows the sum of squares of a component in this case will occur with a plus sign if the number of subscripts in Q^* is even.

Thus we see that the sums of squares of components in Case 4b do not vanish in the polykay whose name contains α partitions for L and R and β partitions for Q. We must now find the proper coefficient for these sums of squares.

Let a component of the type in question be denoted by $\Gamma(\psi)$ and the corresponding sum of squares by $\sum [\Gamma(\psi)]^2$ where Γ contains ρ subscripts and ψ contains k subscripts. Let Γ_i denote the population range of the corresponding subscripts with $\Gamma = \prod_{i=1}^{\rho} \Gamma_i$. Similarly let $\psi = \prod_{i=1}^{k} \psi_i$. Further let $N = (\prod_i \Gamma_i)(\prod_j \psi_j) = \Gamma\psi$.

Suppose ψ_i, $i = 1, \cdots, \nu$ corresponds to subscripts which are contained in R and ψ_i, $i = \nu + 1, \cdots, k$ correspond to subscripts in Q^*. Then the partitions corresponding to the latter are of the form β while those of the former are sub-partitioned according to Definition 3.6. By the argument given in Case 4b the sign of the components is $(-1)^{k-\nu}$, so the coefficient of $\sum_\alpha [\Gamma(\psi)]_\alpha^2$ will be:

$$(-1)^{k-\nu}\left[\frac{1}{N\prod_{\nu+1}^{k}(\psi_j - 1)} + \frac{1}{N(\psi_1 - 1)\prod_{\nu+1}^{k}(\psi_j - 1)}\right.$$

$$+ \cdots + \frac{1}{N(\psi_\nu - 1)\prod_{\nu+1}^{k}(\psi_j - 1)} + \frac{1}{N(\psi_1 - 1)(\psi_2 - 1)\prod_{\nu+1}^{k}(\psi_j - 1)}$$

$$\left. + \cdots + \frac{1}{N(\psi_1 - 1)\cdots(\psi_\nu - 1)\prod_{\nu+1}^{k}(\psi_j - 1)}\right]$$

$$= (-1)^{k-\nu}\left\{\frac{1}{N\prod_{1}^{k}(\psi_j - 1)}\left[1 + \sum_{i=1}^{\nu}(\psi_i - 1)\right.\right.$$

$$\left.\left. + \sum_{i<j=1}^{\nu}(\psi_i - 1)(\psi_j - 1) + \cdots + (\psi_1 - 1)\cdots(\psi_\nu - 1)\right]\right\}$$

$$= (-1)^{k-\nu}\left\{\frac{\prod_{i=1}^{\nu}[(\psi_i - 1) + 1]}{N\prod_{1}^{k}(\psi_j - 1)}\right\} = (-1)^{k-\nu}\frac{\prod_{i=1}^{\nu}\psi_i}{N\prod_{1}^{k}(\psi_j - 1)}$$

$$= (-1)^{k-\nu} \frac{1}{\left(\prod_{\nu+1}^{k} \psi_i\right)\left(\prod_{1}^{p} \Gamma_i\right)\left(\prod_{1}^{k} (\psi_j - 1)\right)}.$$

Thus

$$\frac{(-1)^{k-\nu} \sum [\Gamma(\psi)]^2}{\left(\prod_{\nu+1}^{k} \psi_i\right)\left(\prod_{1}^{p} \Gamma_i\right) \prod_{1}^{k} (\psi_j - 1)} = \frac{(-1)^{k-\nu}}{\prod_{\nu+1}^{k} \psi_i} \sigma^2_{\Gamma(\psi)}$$

which is the typical term in the expansion of $\Sigma_{\Gamma(\Psi)}$ according to Definition 4.3, q.e.d.

REFERENCES

[1] HOOKE, ROBERT (1956). Symmetric functions of a two-way array. *Ann. Math. Statist.* **27** 55–79.

[2] SMITH, H. FAIRFIELD (1955). Variance components, finite population, and experimental inference. Mimeo Series No. 135, Institute of Statistics, Univ. of North Carolina.

[3] TUKEY, J. W. (1950). Some sampling simplified. *J. Amer. Statist. Assoc.* **45** 501–519.

[4] TUKEY, J. W. (1951). Finite sampling simplified. Mimeo. Report No. 45, Statistical Research Group, Princeton Univ.

[5] TUKEY, J. W. (1956). Keeping moment-like sampling computations simple. *Ann. Math. Statist.* **27** 37–54.

[6] WILK, M. B. and KEMPTHORNE, OSCAR (1957). Non-additivities in a latin square design. *Amer. Statist. Assoc.* **52**.

[7] ZYSKIND, GEORGE (1962). On structure, relation, Σ, and expectation of mean squares. *Sankhyā* **24** (Series A, Part 2) 115–148.

[8] ZYSKIND, GEORGE (1958). Error structures in experimental designs. Unpublished Ph.D. Thesis, Iowa State Univ.