A Spectral Analysis Approach for Experimental Designs

R. A. Bailey[1], Persi Diaconis[2], Daniel N. Rockmore[3], and Chris Rowley[4]

[1] School of Mathematics and Statistics, University of St Andrews, St Andrews, Fife, KY16 9SS, U.K. rab@mcs.st-and.ac.uk
and School of Mathematical Sciences, Queen Mary University of London (emerita), Mile End Road, London E1 4NS, U.K.
[2] Department of Statistics, Sequoia Hall, 390 Serra Mall, Stanford University, Stanford, CA 94305-4065, U.S.A. diaconis@math.stanford.edu
[3] Department of Mathematics and Neukom Institute for Computational Science, Dartmouth College, Hanover, NH 03755-3551, U.S.A.
rockmore@cs.dartmouth.edu
[4] Department of Mathematics and Statistics, Faculty of Mathematics, Computing and Technology, The Open University, Walton Hall, Milton Keynes, Buckinghamshire, MK7 6AA, UK

Summary. In this paper we show how the approach of spectral analysis generalizes the standard ANOVA-based techiques for studying data from designed experiments. Several examples are worked out in detail, including a thorough analysis of Calvin's famous ice cream data.

Key words: analysis of variance, block design, designed experiment, diallel experiment, irreducible subspace, orthogonal decomposition, permutation representation

1 Introduction

Designed experiments are used in many fields of enquiry. Government scientists may wish to compare the effects of different insecticides (including no insecticide) on colonies of bumble bees close to the fields where the insecticides are sprayed. Pharmaceutical companies experiment with new drugs, in various doses, to cure certain diseases, or alleviate symptoms. Psychologists may run a trial to see which of three teaching methods is most effective in helping autistic children to understand emotions in other people. In the manufacturing industry, there are frequent experiments to investigate how the process may be improved by changing the raw materials, changing their quantities, or altering parts of the process. (See [36] for examples and a classical statistical approach.)

For all of these, the items being compared, such as insecticides, drugs, teaching methods, or raw materials, are called *treatments*. There may be a single treatment factor, or treatments may consist of all combinations of two or more factors: for example, five varieties of cow-peas with three different methods of cultivation, giving fifteen treatments altogether. One of the treatments may also be 'untreated.'

In order to compare the treatments, they have to be applied to something or somebody: for example, a treatment may be applied to a field, a whole farm, an ill person for a certain amount of time, a child, a group of children, one part of the factory for a month, and so on. Measurements are made on these, either on each whole item, or on smaller units, such as each child in the class. We call these *observational units*.

We now formalize these ideas. In a designed experiment, a finite set Γ of treatments is applied to a finite set Ω of observational units. A measurement y_ω is made on each unit ω in Ω, thus giving a data vector y in the vector space $\mathbb{R}^\Omega$ of all real functions on Ω. Twin problems are how to design the experiment and how to analyze the data that it produces.

There is a long history of group theory being used to develop and design the combinatorial structures used in such experiments (see, for example, [1, 26]). As proposed by Diaconis in [23] and the many references therein, *spectral analysis* (that is, Fourier analysis for the symmetry group of choice) is a non-model-based approach to analyzing data that may be carried out in the presence a natural symmetry group. Spectral analysis seeks to approximate the data vector as a sum of its projections into orthogonal, symmetry-invariant and naturally interpretable subspaces of $\mathbb{R}^\Omega$, where orthogonality is with respect to the standard inner product. The paper [24] contains several spectral analysis examples as well as other examples of ways in which group theory enters statistical analysis.

In this paper we explore some new ideas relating to the spectral analysis of data from a designed experiment. We connect it to classical theory and show how, in a number of cases, this approach provides new information. Necessarily, this analysis depends on the representation theory of the associated symmetry group, and the attendant calculations depend on certain representation (Fourier) theoretic computations and algorithms. All of this is developed as we go along; the book [23] contains all the necessary representation-theoretic background.

In Section 2 we look at families of subspaces of $\mathbb{R}^\Gamma$ that may be used to model expectation, introducing several examples. In Section 3 we introduce a group of permutations of Γ, and compare the decomposition into irreducible subspaces with that obtained from the modeling approach.

Section 4 takes a similar approach to $\mathbb{R}^\Omega$, where structure on Ω might suggest models for the expectation of a random vector Y on Ω or for its variance-covariance matrix $\mathrm{Var}(Y)$. Again, these may be linked to a group of permutations of Ω.

The following sections combine these decompositions of $\mathbb{R}^\Gamma$ and $\mathbb{R}^\Omega$ when an allocation of treatments to observational units effectively makes $\mathbb{R}^\Gamma$ a subspace of $\mathbb{R}^\Omega$. The overall philosophy of analysis of variance is summarized in Section 5. In the most straightforward case, treated in Section 6, there is geometric orthogonality between all subspaces. Otherwise, as explained in Section 7, more complicated algebra is needed. In Section 8 we restrict attention to incomplete-block designs, which avoid some of the complications of the general case while still showing interesting behavior.

Section 9 introduces the subgroup of both previous groups that preserves structure on Γ and Ω as well as preserving the embedding of $\mathbb{R}^\Gamma$ in $\mathbb{R}^\Omega$. An example discussed in depth shows the utility of the approach from spectral analysis.

Finally, Sections 10–11 consider this subgroup in the contexts of orthogonal designs and incomplete-block designs. This subgroup is usually smaller than the previous ones, so its decomposition may have more subspaces. What meaning can we attach to them?

2 Treatments

In this section we ignore the observational units, and pretend temporarily that there is only one observation on each treatment. Different structure on the set of treatments can lead to different plausible models for the response. Under suitable conditions, a family of models leads to an orthogonal decomposition of the vector space of real functions on the treatments.

Let Γ be the finite set of treatments. A *linear model* is typically a subspace V of $\mathbb{R}^\Gamma$: if a measurement is made on each treatment then we expect the resulting vector of measurements to lie in V or close to V. In fact, linear models are frequently presented in notation that is shorthand for saying that a whole (finite) family $\mathcal{F}$ of subspaces is being considered. It is usual to assume that this family is closed under intersection and under vector-space summation.

Two subspaces V_1 and V_2 are defined in [54] to be *geometrically orthogonal* to each other if $V_1 \cap (V_1 \cap V_2)^\perp$ is orthogonal to $V_2 \cap (V_1 \cap V_2)^\perp$ (here $^\perp$ denotes orthogonal complement). If $\mathcal{F}$ is closed under $\cap$ and $+$ and every pair of subspaces in $\mathcal{F}$ is geometrically orthogonal then there is a collection of pairwise orthogonal subspaces $\{W_j : j \in J\}$ such that $\sum_{V\in\mathcal{F}} V = \bigoplus_{j\in J} W_j$ and every subspace in $\mathcal{F}$ is a direct sum of some of the spaces W_j. This is called *orthogonal treatment structure* in [5, 6]. Not all sums of W_j spaces need occur in $\mathcal{F}$.

Denote by V_0 the 1-dimensional subspace consisting of constant vectors. It is usually assumed that V_0 belongs to $\mathcal{F}$.

Example 1. Suppose that Γ consists of all combinations of the n levels of treatment factor C with the m levels of treatment factor D. For example, factor C might be three different non-zero quantities of aspirin and factor D might give two different times of day for taking the aspirin.

One obvious model subspace is V_C (of dimension n), which consists of all vectors which are constant on each level of C. This model is appropriate when the factor D has no effect. The m-dimensional subspace V_D is defined similarly. Then $V_C \cap V_D = V_0$ and V_C is geometrically orthogonal to V_D. The subspace $V_C + V_D$ is called the *additive model*. If this is appropriate then the difference between two given levels of C does not depend on the level of D. Finally, the whole nm-dimensional space $\mathbb{R}^\Gamma$ is the *full model*, allowing unrelated measurements on all treatments.

Figure 1 shows the Hasse diagram for this family of subspaces. The dimension of each is shown beside the corresponding dot.

Put $W_0 = V_0$, $W_C = V_C \cap V_0^\perp$, $W_D = V_D \cap V_0^\perp$, and $W_{CD} = (V_C + V_D)^\perp$. These spaces are called the *grand mean*, the *main effect* of C, the *main effect* of D, and the C-by-D *interaction*, respectively. Strictly speaking, it is the orthogonal projection of the vector of measurements onto each W-subspace that has this name. Now $\dim(W_0) = 1$, $\dim(W_C) = n - 1$, $\dim(W_D) = m - 1$ and $\dim(W_{CD}) = (n - 1)(m - 1)$. These subspaces are mutually orthogonal. Furthermore, $V_0 = W_0$, $V_C = W_0 \oplus W_C$, $V_D = W_0 \oplus W_D$ and $\mathbb{R}^\Gamma = W_0 \oplus W_C \oplus W_D \oplus W_{CD}$.

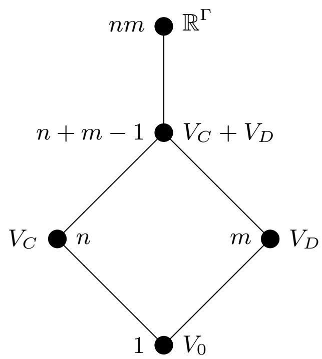

Fig. 1. Hasse diagram of subspaces in Example 1

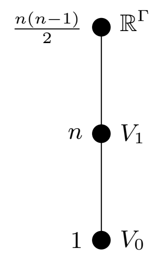

Fig. 2. Hasse diagram of subspaces in Example 2

Example 2. The half-diallel. Let Γ consist of all unordered pairs from a set of size n. This occurs in so-called *half-diallel* experiments in plant breeding, where $\{i, j\}$ denotes the cross between parental lines i and j (see [28]). Another example with $n = 5$ is an experiment to compare all fruit-juices made from equal quantities of two of orange, grapefruit, mango, pineapple and passionfruit, to find the effect on the drinker's blood-pressure.

The family of model subspaces consists of V_0, V_1 and $\mathbb{R}^\Gamma$, where V_1 consists of all functions f of the form $f(\{i, j\}) = \psi_i + \psi_j$. Figure 2 gives the Hasse diagram. Since each pair of subspaces is related by inclusion, geometric orthogonality is assured, and the whole space is decomposed as $W_0 \oplus W_1 \oplus W_2$, where $W_0 = V_0$, $W_1 = V_1 \cap V_0^\perp$ and $W_2 = V_1^\perp$. The last two are called *general combining ability* and *specific combining ability*, respectively. Now $\dim(W_0) = 1$, $\dim(W_1) = n - 1$ and $\dim(W_2) = n(n-3)/2$.

Example 3. Now let Γ consist of all ordered pairs of distinct elements from a set of size n. This occurs in breeding if the gender of the parents is relevant. It occurs in a modification of the fruit-juice example if the treatments consist of all instructions like 'drink orange juice at breakfast and mango juice at lunch'.

One model subspace is the space V_S of *symmetric* functions f, for which $f((i,j)) = f((j,i))$ for all i and j. Another is the space V_A of *antisymmetric* functions f, for which $f((i,j)) = -f((j,i))$ for all i and j. A further obvious space V_P consists of parental effects: f is in V_P if there are constants α_i and β_j such that $f((i,j)) = \alpha_i + \beta_j$ for all i and j, which is similar to the additive model in Example 1.

There are (at least) four interesting subspaces of V_P: $f \in V_P \cap V_S$ if $\alpha_i = \beta_i$ for all i; $f \in V_P \cap V_A$ if $\alpha_i = -\beta_i$ for all i; $f \in V_1$ if $\beta_j = 0$ for all j; and $f \in V_2$ if $\alpha_i = 0$ for all i. Now the four subspaces $V_P \cap V_S \cap V_0^\perp$, $V_P \cap V_A$, $V_1 \cap V_0^\perp$ and $V_2 \cap V_0^\perp$ all have dimension $n - 1$; the sum of any two is $V_P \cap V_0^\perp$; and no pair is geometrically orthogonal except for the first two.

On the other hand, $V_P^\perp$ is the orthogonal direct sum of W_S and W_A, where $W_S = V_S \cap V_P^\perp$, which has dimension $n(n-3)/2$ and is analogous to W_2 in Example 2, and $W_A = V_A \cap V_P^\perp$, which has dimension $(n-1)(n-2)/2$. See Figure 3.

Now the lack of a canonical orthogonal decomposition of $V_P \cap V_0^\perp$ can lead to difficulties in model choice.

Example 4. Suppose that we wish to compare six makes of strawberry ice cream. Sixty people take part in the experiment, so that each tastes four makes and rates one of these, giving it a score out of 100. Note that each make is tasted in the presence of all possible triples of other makes. Thus Γ consists of the 60 pairs $(i, \{j,k,l\})$ where i is rated in the presence of j, k and l. One model subspace V consists of functions f for which $f((i,\{j,k,l\})) = \gamma_i + \theta_{ij} + \theta_{ik} + \theta_{il}$, where $\theta_{ij}, \theta_{ik}, \theta_{il}$ each account for the effect of tasting i in the presence of j, k and l. Note that V has dimension 30 because the subspace V_1 with $\theta_{ij} = \alpha_i$ for all i and j is the same as the subspace with $\theta_{ij} = 0$ for all i and j. If we consider only subspaces of V, we appear to obtain the same family of models as in Example 3; however, their interpretation as subspaces of $\mathbb{R}^\Gamma$ is different.

Figure 4 represents the elements of Γ in Examples 3 and 4 when $n = 6$. Elements of Γ are identified by the labels of the rows and columns with "$\times$" marks those cells for pairings that do not occur as elements. The real numbers in the other cells thus represent a function in $\mathbb{R}^\Gamma$: in both cases it is the

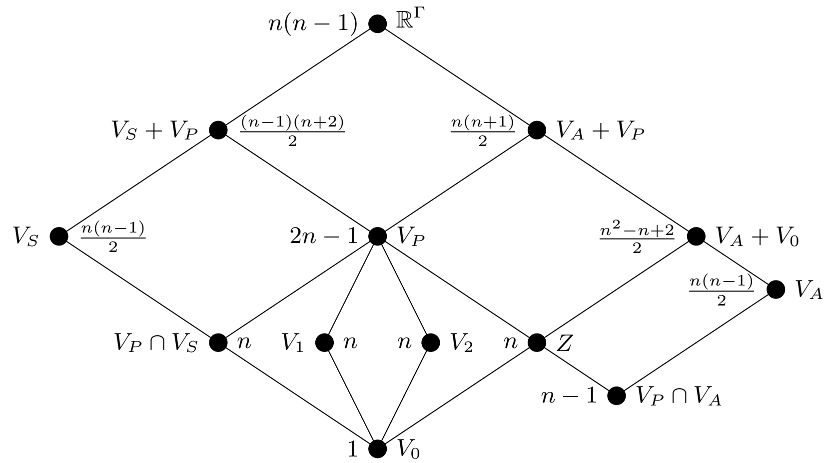

Fig. 3. Hasse diagram of subspaces in Example 3: Z is $(V_P \cap V_A) + V_0$

function in $V_P \cap V_A$ with $\alpha_1 = -\beta_1 = 1$ and $\alpha_i = \beta_i = 0$ if $i \neq 1$. In Example 3, this function is orthogonal to $V_S \cap V_P$; in Example 4 it is not. In Example 3, there seems to be no reason to prefer the orthogonal decomposition $(V_1 \cap V_0^\perp) \oplus (V_P \cap V_1^\perp)$ over $(V_2 \cap V_0^\perp) \oplus (V_P \cap V_2^\perp)$, whereas in Example 4 the explicit inclusion of γ_i in the formula seems to favour the former. In Example 4 there is a function h in $V_1 \cap V_0^\perp$ with $\gamma_1 = -10/3$ and $\gamma_i = 2/3$ if $i \neq 1$: when this is added to the function f shown in Figure 4(**b**), the result is constant on each column.

	1	2	3	4	5	6
1	×	1	1	1	1	1
2	−1	×	0	0	0	0
3	−1	0	×	0	0	0
4	−1	0	0	×	0	0
5	−1	0	0	0	×	0
6	−1	0	0	0	0	×

(**a**)

	1 2 3 4	1 2 3 5	1 2 3 6	1 2 4 5	1 2 4 6	1 2 5 6	1 3 4 5	1 3 4 6	1 3 5 6	1 4 5 6	2 3 4 5	2 3 4 6	2 3 5 6	2 4 5 6	3 4 5 6
1	3	3	3	3	3	3	3	3	3	3	×	×	×	×	×
2	−1	−1	−1	−1	−1	−1	×	×	×	×	0	0	0	0	×
3	−1	−1	−1	×	×	×	−1	−1	−1	×	0	0	0	×	0
4	−1	×	×	−1	−1	×	−1	−1	×	−1	0	0	×	0	0
5	×	−1	×	−1	×	−1	−1	×	−1	−1	0	×	0	0	0
6	×	×	−1	×	−1	−1	×	−1	−1	−1	×	0	0	0	0

(**b**)

Fig. 4. A function in $\mathbb{R}^\Gamma$: (**a**) Example 3 (**b**) Example 4.

3 Treatment Permutations

Often, the set Γ of treatments has some combinatorial symmetry which is preserved by a group G_1 of permutations of Γ. In this section we see how this may be used to derive an orthogonal decomposition of $\mathbb{R}^\Gamma$. We write the elements of G_1 on the right of their arguments, so that composition is done from left to right.

The permutation representation of G_1 associated with its action on Γ is an isomorphism ρ_1 from G_1 to group of permutation matrices in $\mathbb{R}^{\Gamma\times\Gamma}$, whose rows and columns are indexed by Γ: for g in G_1, the (α,β)-entry of $\rho_1(g)$ is 1 if $\alpha g = \beta$ and is 0 otherwise. These matrices act on $\mathbb{R}^\Gamma$, which can be decomposed as a direct sum of subspaces which are invariant under G_1 and irreducible under G_1. The centralizer algebra $\mathcal{C}(G_1)$ of G_1 is the set of matrices in $\mathbb{R}^{\Gamma\times\Gamma}$ which commute with $\rho_1(g)$ for all g in G_1.

Recall that, in general, a representation of degree N of a group G over a field F is just a homomorphism from G to the general linear group $\mathrm{GL}(N,F)$, where N is a non-negative integer. Two such representations are *equivalent* if and only if they differ by a change of basis. Thus, the trace function is an invariant of any equivalence class of representations. A representation is *irreducible* if and only if it is not equivalent to a direct sum of (non-trivial) representations. Up to equivalence, there are only a finite number of irreducible representations of a given group over a given field. Following community standards, we refer to both the collection of matrices as well as the underlying vector space with associated group action as "the representation."

Since every representation over $\mathbb{R}$ can also be considered to be a representation over $\mathbb{C}$, we sometimes need to write words like "real-irreducible" or "complex-irreducible" to make clear which field we are talking about. General theory is usually expressed over the complex numbers, giving decompositions of $\mathbb{C}^\Gamma$. However, only $\mathbb{R}^\Gamma$ occurs for actual data so we use real representations. See Section 11 for the modification from $\mathbb{C}^\Gamma$ to $\mathbb{R}^\Gamma$.

Let π_1 be the permutation character of G_1, so that $\pi_1(g) = \mathrm{trace}(\rho_1(g))$ for g in G_1. Let $\{\chi_i : i \in \mathcal{I}\}$ be the real-irreducible characters of G_1. Then there are non-negative integers m_i for i in $\mathcal{I}$ such that $\pi_1 = \sum_i m_i\chi_i$. If $m_i > 0$ then there is a corresponding *homogeneous* subspace U_i of $\mathbb{R}^\Gamma$: it has dimension $m_i \deg(\chi_i)$; it is G_1-invariant; it is orthogonal to U_j if $i \neq j$. If $m_i = 1$ then U_i is G_1-irreducible; otherwise, it can be decomposed as a direct sum of m_i irreducible subspaces, all admitting isomorphic actions of G_1, in infinitely many ways.

If $m_i \in \{0,1\}$ for all i in $\mathcal{I}$ then π_1 is said to be *real-multiplicity-free*. Then $\mathbb{R}^\Gamma$ has a unique decomposition as a direct sum of orthogonal G_1-irreducible subspaces. We may be able to use this decomposition to explain the data vector. Otherwise, there is a choice of such decompositions.

What light does this approach from representation theory throw on the previous examples? In Example 1, we may consider Γ to be an $n\times m$ rectangle. Then we may take G_1 to be $\mathrm{Sym}(n)\times\mathrm{Sym}(m)$ (where $\mathrm{Sym}(n)$ denotes the symmetric group on n letters) in its product action, so that $(\alpha, \beta)(g, h) = (\alpha g, \beta h)$. Its permutation character is real-multiplicity-free, and the corresponding irreducible subspaces are precisely the subspaces W_0, W_C, W_D and W_{CD} used by statisticians.

In Example 2, it is natural to take G_1 to be Sym(n) in its action on unordered pairs, so that $\{\alpha, \beta\}g = \{\alpha g, \beta g\}$. In the notation of [35], the vector space supporting the permutation representation is denoted $M^{n-2,2}$, which has a unique G_1-irreducible decomposition as $S^n \oplus S^{n-1,1} \oplus S^{n-2,2}$. In the notation of Example 2, S^n is the representation on W_0, $S^{n-1,1}$ is the representation on W_1 and $S^{n-2,2}$ is the representation on W_2.

In Example 3, we take G_1 to be Sym(n) in its action on ordered pairs, so that $(\alpha, \beta)g = (\alpha g, \beta g)$. The resulting permutation representation is denoted $M^{n-2,1,1}$. The decomposition (see [35]) of this representation is below.

	$M^{n-2,1,1}$	$=$	S^n	$\oplus$	$2S^{n-1,1}$	$\oplus$	$S^{n-2,2}$	$\oplus$	$S^{n-2,1,1}$
dim	$n(n-1)$		1		$2 \times (n-1)$		$\frac{n(n-3)}{2}$		$\frac{(n-1)(n-2)}{2}$

Here S^n, $S^{n-2,2}$ and $S^{n-2,1,1}$ are the representations on the subspaces W_0, W_S and W_A, respectively. The notation $2S^{n-1,1}$ denotes a representation which is the direct sum of two representations isomorphic to $S^{n-1,1}$. Note that it is such a sum in infinitely many ways, and there is no canonical decomposition of the corresponding homogenous subspace of dimension $2(n-1)$, which is precisely $V_P \cap V_0^\perp$. Thus the group theory reinforces the previous discussion.

In [57], Yates proposed decomposing $V_P \cap V_0^\perp$ as $(V_P \cap V_S \cap V_0^\perp) \oplus (V_P \cap V_A)$. Using the above representation theory in [32], James was able to show that this was, in some sense, an arbitrary choice. Fortini gave a different decomposition in [27].

In Example 4, let G_1 be Sym(6) in its action on pairs (i, K) where K is a 4-subset of a 6-set and $i \in K$. Such a pair is equivalent to the partition with parts $\{i\}$, $K \setminus \{i\}$ and the complement of K, so the resulting permutation representation is $M^{3,2,1}$, whose decomposition (see [35]) is below.

	$M^{3,2,1}$	$=$	S^6	$\oplus$	$2S^{5,1}$	$\oplus$	$2S^{4,2}$	$\oplus$	$S^{4,1,1}$	$\oplus$	$S^{3,3}$	$\oplus$	$S^{3,2,1}$
dim	60		1		2×5		2×9		10		5		16.

(1)

Of course, S^6 is the representation on V_0. We have already seen that $V_P \cap V_0^\perp$ is the homogeneous subspace for $S^{5,1}$, that W_S affords one copy of $S^{4,2}$ and W_A affords $S^{4,1,1}$. The action of Sym(6) on 4-subsets (the column indices in Fig. 4(**b**)) is $M^{4,2}$, whose decomposition is $S^6 \oplus S^{5,1} \oplus S^{4,2}$. Thus the 15-dimensional subspace V_B of functions which are constant on each 4-subset includes a 5-dimensional subspace affording $S^{5,1}$, which must therefore be contained in $V_P \cap V_0^\perp$. Indeed, the end of Section 2 gives a non-zero function in $V_P \cap V_0^\perp$ which is constant on each 4-subset.

Therefore $V_B \cap V_P^\perp$ is a 9-dimensional subspace affording $S^{4,2}$. How is this related to W_S? Figure 5 displays the function in W_S defined by $\theta_{ij} = 1$ if

$\{i, j\} = \{1, 2\}$ or $\{3, 4\}$, $\theta_{ij} = -1$ if $\{i, j\} = \{1, 3\}$ or $\{2, 4\}$, and $\theta_{ij} = 0$ otherwise. This is in neither V_B nor $V_B^{\perp}$.

We return to this example in detail in Section 9.

	1 2 3 4	1 2 3 5	1 2 3 6	1 2 4 5	1 2 4 6	1 2 5 6	1 3 4 5	1 3 4 6	1 3 5 6	1 4 5 6	2 3 4 5	2 3 4 6	2 3 5 6	2 4 5 6	3 4 5 6
1	0	0	0	1	1	1	−1	−1	−1	0	×	×	×	×	×
2	0	1	1	0	0	1	×	×	×	×	−1	−1	0	−1	×
3	0	−1	−1	×	×	×	0	0	−1	×	1	1	0	×	1
4	0	×	×	−1	−1	×	1	1	×	0	0	0	×	−1	1
5	×	0	×	0	×	0	0	×	0	0	0	×	0	0	0
6	×	×	0	×	0	0	×	0	0	0	×	0	0	0	0

Fig. 5. A function in W_S which is neither constant on columns nor orthogonal to columns

4 Observational Units

Now we temporarily ignore the treatments, and think about the observational units to be used in the experiment. Again, there is a corresponding real vector space, and we seek meaningful orthogonal decompositions of this. Such a decomposition may be defined by inherent factors or by a group of symmetries.

In a designed experiment, there is a finite set Ω of observational units to which treatments are applied; later, some response is measured on each unit. Even before the treatments are applied, inherent features of Ω may suggest something about the pattern of the response.

Example 5. An experiment comparing different methods of soil preparation for a single cereal variety might use k fields on each of b farms, with a single method on each field. Then Ω consists of the bk fields. Let Y be the hypothetical random vector of responses. If some farms produce consistently better results than others, then, in the absence of treatment differences, the expected value of the response should just depend on the farm. On the other hand, differences between farms may change from season to season, but then it is plausible that fields within a farm are more alike than fields in different farms. In this case, the grouping of fields within farms affects the covariance (matrix) $\mathrm{Var}(Y)$.

More generally, if Ω consists of bk observational units grouped into b blocks of size k, let V_B be the subspace of $\mathbb{R}^{\Omega}$ consisting of vectors which are constant on each block. The first approach assumes that $E(Y) \in V_B$. In this case, blocks are said to have *fixed effects*. It is usual to assume that $\mathrm{Var}(Y) = \sigma^2 I$ in this case. The second approach assumes that $E(Y) \in V_0$ and that the covariance has the form $\mathrm{Var}(Y) = \sigma^2 I + k\sigma_B^2 Q_B$, where Q_B is the matrix of orthogonal projection onto V_B. Blocks then are said to have *random effects*. In this case, the eigenspaces of $\mathrm{Var}(Y)$ are V_B and $V_B^\perp$. Both approaches make it natural to consider the decomposition $V_0 \oplus W_B \oplus V_B^\perp$ of $\mathbb{R}^\Omega$, where $W_B = V_B \cap V_0^\perp$.

Let G_2 be a group of permutations of Ω that preserve structure on Ω, such as the partition into blocks, before treatments are allocated. Let π_2 be the permutation character of G_2. If π_2 is real-multiplicity-free then there is a unique decomposition of $\mathbb{R}^\Omega$ as a sum of G_2-irreducible subspaces, which may be pertinent for data analysis. Even without this uniqueness, a decomposition into G_2-invariant subspaces may give insight into the data.

When Ω consists of b blocks of size k, we may take G_2 to be the wreath product $\mathrm{Sym}(k)$ wr $\mathrm{Sym}(b)$ in its imprimitive action. This has a subgroup $\mathrm{Sym}(k)$ for each block, permuting just the units within it, and a subgroup $\mathrm{Sym}(b)$ permuting the set of whole blocks. The action is real-multiplicity-free, with irreducibles V_0, W_B and $V_B^\perp$.

Example 6. Another common structure for Ω is a rectangle with r rows and c columns. This may be an actual physical rectangle on the ground, or an abstract one where, for example, rows represent time-periods and columns represent people. As in Example 1, $\mathbb{R}^\Omega$ has a natural decomposition as $W_0 \oplus W_R \oplus W_C \oplus W_{RC}$. If rows and columns have fixed effects then $E(Y) \in W_0 \oplus W_R \oplus W_C$. If they have random effects then $\mathrm{Var}(Y) = \sigma^2 I + c\sigma_R^2 Q_R + r\sigma_C^2 Q_C$, whose eigenspaces are W_0, W_R, W_C and W_{RC}. These four subspaces are the irreducibles of $\mathrm{Sym}(r) \times \mathrm{Sym}(c)$ in its product action.

The key part of an experimental design is the function $\tau\colon \Omega \to \Gamma$, which allocates treatment $\tau(\omega)$ to observational unit ω. This allocation is normally *randomized* before treatments are applied: a permutation g is chosen at random from a suitable group G_2 of permutations of Ω, and τ is replaced by τ^g, where $\tau^g(\omega) = \tau(\omega g^{-1})$. It is argued in [1, 4] that this justifies the assumption that $\mathrm{Var}(Y) \in \mathcal{C}(G_2)$. Since $\mathrm{Var}(Y)$ is symmetric, its eigenspaces form a G_2-invariant decomposition of $\mathbb{R}^\Omega$. If π_2 is real-multiplicity-free, these subspaces are known even if the corresponding eigenvalues are not.

Hannan [29] and Speed [52, 53] considered random variables Y such that $\mathrm{Var}(Y) \in \mathcal{C}(G_2)$ without the complication of Γ and τ. See those papers for more examples.

The three most common inherent structures in designed experiments are the two that we have discussed and the unstructured one, in which G_2 consists of all permutations of the units in Ω. The operations of nesting (units within blocks) and crossing (rows and columns) can be iterated, to give simple orthogonal block structures: see [43]. Their automorphism groups are generalized wreath products of symmetric groups, for all of which the relevant action is real-multiplicity-free: see [8]. The remaining examples in this paper use only these three common structures.

5 Analysis of Variance

Analysis of variance is often the statistician's first step towards analyzing the data vector. See [50] for an extensive study. This section gives a quick summary of the method, in language more familiar to algebraists.

Classical analysis of variance (ANOVA) depends on a given orthogonal decomposition of $\mathbb{R}^\Omega$ into subspaces W_j for j in some set J. Denote by P_j the linear operator of orthogonal projection onto W_j. Then the data vector y is the sum $\sum_{j\in J} P_j y$, and $P_j y$ is orthogonal to $P_i y$ if $i \neq j$. The *sum of squares* SS_j for W_j is defined to be $\|P_j y\|^2$, which is just $y^\top P_j y$. Since $I = \sum_{j\in J} P_j$, these sums of squares add to give $y^\top y$, the total sum of squares.

Put $d_j = \dim(W_j)$. The *mean square* MS_j for W_j is defined to be SS_j/d_j. If the data are purely random, in the sense that the responses are mutually independent random variables with the same expectation and variance, then all mean squares except MS_0 have the same expectation, where $W_0 = V_0$. More precisely, if Y is a random vector on Ω and $\mathrm{Var}(Y) = \sigma^2 I$ then $E(\mathrm{MS}_j) = \|E(P_j Y)\|^2/d_j + \sigma^2$. In general, if W_j is contained in an eigenspace of $\mathrm{Var}(Y)$ with eigenvalue ξ_i then $E(\mathrm{MS}_j) = \|E(P_j Y)\|^2/d_j + \xi_i$.

Thus one approach to analyzing data is to calculate the mean squares and pick out those subspaces W_j whose mean squares are particularly large relative to the others: something interesting must be happening there. Part of the statistical theory of hypothesis testing is quantifying "particularly large." We do not go into details here.

The treatment allocation τ gives a subspace V_Γ of $\mathbb{R}^\Omega$ whose elements are constant on each treatment. Thus V_Γ is isomorphic to $\mathbb{R}^\Gamma$. To avoid complications, we assume that every treatment occurs on the same number of observational units. This ensures that any orthogonal decomposition of $\mathbb{R}^\Gamma$ into subspaces remains orthogonal when $\mathbb{R}^\Gamma$ is embedded into $\mathbb{R}^\Omega$ as V_Γ.

Here we assume that $\mathbb{R}^\Omega$ has a given G_2-invariant orthogonal decomposition and $\mathbb{R}^\Gamma$ has a given G_1-invariant orthogonal decomposition. When $\mathbb{R}^\Gamma$ is embedded in $\mathbb{R}^\Omega$ as V_Γ we need another orthogonal decomposition of $\mathbb{R}^\Omega$ which is related to the two previous ones and can be used for ANOVA. The next three sections show how to obtain such a decomposition in some cases, while indicating that it is difficult in general.

6 The Orthogonal Case

In this section, we start to combine the initial orthogonal decompositions of $\mathbb{R}^\Gamma$ and $\mathbb{R}^\Omega$. Of course, this depends on the way that the design map τ has embedded $\mathbb{R}^\Gamma$ in $\mathbb{R}^\Omega$ as V_Γ. The combination is relatively straightforward under a condition on τ known as orthogonality.

Given orthogonal decompositions of $\mathbb{R}^\Omega$ and $\mathbb{R}^\Gamma$, a design map τ is said to be *orthogonal* if every subspace in the given decomposition of V_Γ is geometrically orthogonal to every subspace in the given decomposition of $\mathbb{R}^\Omega$.

Then the non-zero intersections of these subspaces give a canonical orthogonal decomposition of $\mathbb{R}^\Omega$ that refines both of the previous two.

Example 7. When $G_2 = \mathrm{Sym}(\Omega)$, the initial decomposition of $\mathbb{R}^\Omega$ is $V_0 \oplus V_0^\perp$. Such a design is called *completely randomized*. We always assume that V_0 is in the decomposition of $\mathbb{R}^\Gamma$, so now the combined decomposition simply adjoins $V_\Gamma^\perp$ to the decomposition of V_Γ. The subspace $V_\Gamma^\perp$ is known as *residual* in ANOVA.

Example 8. For a so-called *complete-block design*, there are b blocks of size k, where $k = |\Gamma|$, and every treatment occurs once in each block. If $G_1 = \mathrm{Sym}(k)$ and $G_2 = \mathrm{Sym}(k) \text{ wr } \mathrm{Sym}(b)$ then the initial decompositions are $V_0 + W_T$ (for $\mathbb{R}^\Gamma$) and $V_0 \oplus W_B \oplus V_B^\perp$ (for $\mathbb{R}^\Omega$). The combined decomposition is $V_0 \oplus W_B \oplus W_T \oplus (V_B + V_T)^\perp$, whose subspaces are usually called *grand mean*, *blocks*, *treatments* and *residual* respectively. If G_1 gives a finer decomposition of $\mathbb{R}^\Gamma$ then this gives a finer decomposition of W_T without affecting W_B or $(V_B + V_T)^\perp$.

Example 9. A third popular orthogonal design is the Latin square. Here Ω is an $n \times n$ rectangle, where $n = |\Gamma|$. The initial decomposition of $\mathbb{R}^\Omega$ is $V_0 \oplus W_R \oplus W_C \oplus (V_R + V_C)^\perp$. When treatments are applied in a Latin square, V_Γ is orthogonal to $W_R \oplus W_C$, so $(V_R + V_C)^\perp$ is decomposed as $W_T \oplus (V_R + V_C + V_T)^\perp$, with the second part being called *residual*. Again, any finer initial decomposition of $\mathbb{R}^\Gamma$ simply gives a decomposition of W_T.

Example 10. The simplest case in which more than one subspace in the initial decomposition of $\mathbb{R}^\Omega$ is split up is the so-called *split-plot design*. The treatments are as in Example 1. For Ω, there are rn blocks, each containing m observational units, so that $\mathbb{R}^\Omega = V_0 \oplus W_B \oplus V_B^\perp$. Each level of C is applied to r whole blocks, and each level of D is applied to one observational unit per block. Thus $V_C < V_B$, while W_D and W_{CD} are both orthogonal to V_B. The combined decomposition is

$$\begin{array}{ccccccccccc} V_0 & \oplus & W_C & \oplus & (W_B \cap W_C^\perp) & \oplus & W_D & \oplus & W_{CD} & \oplus & (V_B^\perp \cap V_\Gamma^\perp) \\ 1 & & n-1 & & n(r-1) & & m-1 & & (n-1)(m-1) & & n(m-1)(r-1) \end{array},$$

where the dimension is shown underneath each subspace. The third subspace is called *block residual*. Under the assumption that treatments affect expectation and blocks have random effects, MS_C is compared with the mean square for block residual while MS_D and MS_{CD} are both compared with the mean square for $V_B^\perp \cap V_\Gamma^\perp$.

In general, if W is a subspace in the initial decomposition of $\mathbb{R}^\Omega$ and $W \cap V_\Gamma^\perp$ is non-zero then $W \cap V_\Gamma^\perp$ is called a *residual subspace*.

7 The General Non-Orthogonal Case

This section gives a quick overview of some difficulties that can occur when the design is not orthogonal. We show that the worst of these are avoided if the structure on Ω is a partition into blocks of equal size. The development follows [30, 44].

When there is not geometric orthogonality between the original decompositions of $\mathbb{R}^\Omega$ and V_Γ, it is normal to start with the decomposition of $\mathbb{R}^\Omega$ and then try to refine it. Let U be one of the subspaces in the original decomposition of $\mathbb{R}^\Omega$. Given an orthogonal decomposition $\bigoplus_{j\in J} W_j$ of V_Γ, one obvious step is to project each W_j onto U. There are two difficulties. The first is that, even if W_i is orthogonal to W_j, their projections onto U may no longer be orthogonal to each other.

The second difficulty needs more explanation. Suppose that $P_j y = v$. Denote by Q the (matrix of) orthogonal projection onto U. If ϕ is the angle between v and Qv then the sum of squares for $Q(W_j)$ is $(\cos^2\phi)\left\|v\right\|^2$ plus other non-negative pieces. Even if there are no other contributions to this sum of squares, we need to know $\cos^2\phi$ in order to make a judgement about the size of v. It is therefore helpful if all vectors in W_j make the same angle with U.

Fortunately, both difficulties are solved if each space W_j is an eigenspace of PQP, where $P = \sum_j P_j$. If this eigenspace decomposition of V_Γ and the original decomposition of V_Γ have a common refinement, then it is used. If not, there are disagreements about how to proceed: see [47].

Now consider two different subspaces U_1 and U_2 in the original decomposition of $\mathbb{R}^\Omega$, with projectors Q_1 and Q_2. If $V_0 \oplus U_1 \oplus U_2 = \mathbb{R}^\Omega$ and $W_j \perp V_0$ then $P_j(Q_1+Q_2)P_j = P_j$, and so W_j is an eigenspace of $P_jQ_1P_j$ if and only if it is an eigenspace of $P_jQ_2P_j$. However, if $V_0 \oplus U_1 \oplus U_2$ is not the whole of $\mathbb{R}^\Omega$ then PQ_1P may not commute with PQ_2P, in which case these matrices do not have common eigenspaces.

Given an original decomposition of $\mathbb{R}^\Omega$ into subspaces $U_1, \ldots, U_s$ with projectors $Q_1, \ldots, Q_s$, Houtman and Speed [30] defined the design to have *general balance* if the matrices $PQ_1P, \ldots, PQ_sP$ commute with each other, where P is the projector onto V_Γ. This implies that, if the structure on Ω is defined simply by a partition into blocks of equal size, then all designs are generally balanced.

For the rest of this paper, we restrict attention to block designs or orthogonal designs, so that general balance is assured. Even with this restriction, there are still plenty of complications.

8 Incomplete-Block Designs

Apart from split-plot designs like those in Example 10, block designs in which the blocks are too small to hold all the treatments are not orthogonal. In this section we give the algebraic approach to studying these, starting with the seminal work of James [31], which was extended by Mann [39] to linear models which may not arise from experimental designs.

When there are b blocks of size k, the initial decomposition of $\mathbb{R}^\Omega$ is $V_0 \oplus W_B \oplus V_B^\perp$, with corresponding projectors Q_0, Q_B and $I - Q_0 - Q_B$. Here $Q_0 = (bk)^{-1}J$, where J is the all-1 matrix, and $Q_0 + Q_B = k^{-1}J_B$, where the (ω_1, ω_2)-entry of J_B is 1 if ω_1 and ω_2 are in the same block and is 0 otherwise. Thus the linear operator $Q_0 + Q_B$ simply replaces each entry of y by the average value on each block, so it is similar to a Radon transformf [10, 37].

Suppose that there are t treatments each occurring r times, so that $tr = bk$. In an *incomplete-block design*, $k < t$ and no treatment occurs more than once in any block. Such a design is *balanced* if there is a constant λ such that each pair of treatments occur together in exactly λ blocks.

Put $W_T = V_\Gamma \cap V_0^\perp$. The orthogonal projector P_T onto W_T is $r^{-1}J_T - Q_0$, where the (ω_1, ω_2)-entry of J_T is 1 if $\tau(\omega_1) = \tau(\omega_2)$ and is 0 otherwise. Thus $Q_0 + P_T$ is another averaging operator. Let $W_{T|B}$ be the orthogonal projection of W_T onto $V_B^\perp$, which is $(W_T + V_B) \cap V_B^\perp$. The classical analysis of data from an incomplete-block design uses the decomposition $V_0 \oplus W_B \oplus W_{T|B} \oplus (V_B + V_T)^\perp$. The second and third subspaces are called *blocks ignoring treatments* and *treatments eliminating blocks* respectively.

James seems to have been one of the first to have studied ANOVA for balanced incomplete-block designs from the point of view of the algebraic properties of the idempotents which yield the sum of squares decomposition. In [31] he defined the "relationship algebra of an experimental design" as the complex algebra $\mathcal{A}$ generated by the matrices I, J, J_B and J_T. He showed that $P_T Q_B P_T = (1 - e)P_T$, where $e = t(k-1)/(t-1)k$, which is called the *efficiency factor* of the design. As James and Wilkinson later showed in [34], every vector in W_T has angle ϕ with W_B, where $\cos^2 \phi = 1 - e$.

James proposed refining the classical ANOVA by decomposing W_B into $Q_B(W_T)$ and its orthogonal complement. The latter is zero if and only if $b = t$; the former has dimension $t - 1$ (this gives an easy proof of Fisher's Inequality, which states that $b \geq t$ for a balanced incomplete-block design). He showed that the algebra $\mathcal{A}$ has dimension six if $b = t$ and dimension seven otherwise. Since the generating matrices are symmetric, the algebra $\mathcal{A}$ is semisimple and thus may be decomposed as a direct sum of matrix algebras, in this case two or three one-dimensional algebras and one 2×2 matrix algebra (which is a four-dimensional algebra). The one-dimensional algebras correspond to the subspaces V_0, $(V_B + V_\Gamma)^\perp$ and, if it is non-zero, $W_B \cap V_\Gamma^\perp$. The projector onto their orthogonal complement can be written as a sum of two idempotents in the algebra in infinitely many ways.

James closed [31] by remarking that "For certain designs, the relationship algebra is the commutator algebra of the representation of a group expressing the symmetry of the experimental design." We take up the relevance of this remark in the Section 11.

This work was extended to arbitrary incomplete-block designs in [34]. If $P_T Q_B P_T$ has rank less than $t-1$ then $V_\Gamma \cap V_B^\perp$ is non-zero. The vectors in this subspace are said to have *canonical efficiency factor* 1. Let the distinct non-zero eigenvalues of $P_T Q_B P_T$ be $\mu_1, \ldots, \mu_s$, where μ_i has multiplicity d_i. Then there are corresponding d_i-dimensional subspaces U_i of W_B and V_i of W_T such that every vector in V_i makes angle ϕ_i with U_i, and vice versa, where $\cos^2 \phi_i = \mu_i$, and $1 - \mu_i$ is the canonical efficiency factor for vectors in V_i. If $i \neq j$ then $U_i \perp V_j$, $U_i \perp U_j$ and $V_i \perp V_j$.

Now $\mathbb{R}^\Omega$ has an orthogonal decomposition as

$$V_0 \oplus (V_B \cap V_\Gamma^\perp) \oplus (V_\Gamma \cap V_B^\perp) \oplus (V_B + V_\Gamma)^\perp \oplus (U_1 + V_1) \oplus \cdots \oplus (U_s + V_s), \quad (2)$$

where the second or third subspace may be zero. The algebra $\mathcal{A}$ is the sum of scalar algebras on the first four subspaces, plus one 2×2 matrix algebra on each of $U_1 + V_1, \ldots, U_s + V_s$.

When $s = 1$ but the design is not balanced then there are just two canonical efficiency factors, one of which is 1. Such designs are called *partial geometric designs* in [11, 12], *C-designs* in [49], and $1\frac{1}{2}$*-designs* in [45]. They appear to be rather useful (see [16]). One way of obtaining them is to exchange the roles of blocks and treatments (thus forming the *dual* design) in a balanced design with $b > t$. Other examples include transversal designs and the lattice designs of Yates [56]. Some classes of such designs have been shown to be optimal (see [7, 22]).

9 Ice Cream Data

So far, we have assumed that the measurement y_ω on ω is influenced by ω itself and its inherent relation to the rest of Ω, as well as the treatment $\tau(\omega)$ and its relation to the rest of Γ. This paradigm does not cover experiments where y_ω might also be influenced by the treatments on observational units which are, in some sense, near to ω. For example, tall varieties of sunflower will shade their shorter neighbors, or the taste of one ice cream may affect the score given by the taster to another ice cream.

In such circumstances, it seems appropriate to consider the group G of *strong* symmetries of the design. This is the subgroup of the group G_2 of permutations of Ω which preserve the partition of Ω defined by the inverse images under τ (in particular, they stabilize V_Γ) and whose induced permutations on Γ are in G_1.

Example 4 is, in fact, silly. It wastes resources, because 240 items are tasted but only 60 are rated. In the actual experiment reported by Calvin in [17], only 15 people took part, each tasting four items and scoring each one. The six treatments were six quantities of vanilla added to the basic ice cream. There was one taster for each subset of size four, and they were asked to give each tasted item an integer score between 0 and 5 inclusive. It is not clear whether the tasters were told that their four treatments were all different (e.g., four tasters gave the same score to two of their items).

In this case, $t = |\Gamma| = 6$ and G_1 is Sym(6) in the natural action. Also, $|\Omega| = 60$. Calvin considered the tasters as blocks, so that G_2 is Sym(4) wr Sym(15) in its imprimitive action. The design was a balanced incomplete-block design with fifteen blocks of size four. Thus G is Sym(6) in the action described at the end of Section 3.

Let $B(\omega)$ be the block containing ω. Calvin proposed the model that $E(Y) = f + d$, where $d(\omega) = \delta_{B(\omega)}$ and $f(\omega) = \gamma_i + \theta_{ij} + \theta_{ik} + \theta_{il}$ if $\tau(\omega) = i$ and the other treatments in $B(\omega)$ are j, k and l. Furthermore, $\theta_{ij} = -\theta_{ji}$ for all i and j. Thus $d \in V_B$; the γ-parameters give a vector in V_1; and the θ-parameters give a vector in V_A, in the notation used in Example 4.

As we saw in Sections 2 and 3, there is a 5-dimensional subspace $\tilde{V}_B$ of V_B such that the sum of any two of $\tilde{V}_B$, $V_1 \cap V_0^\perp$ and $V_P \cap V_A$ is the same 10-dimensional subspace, the homogeneous subspace for $S^{5,1}$. Thus the δ-, γ- and θ-parameters are not all identifiable. Calvin got around this problem by restricting the θ-parameters to give a vector in the 10-dimensional subspace W_A, which is the homogeneous subspace for $S^{4,1,1}$.

Calvin gave the ANOVA in Table 1. As is common for statisticians, he omitted the line for V_0, he wrote "d.f." (degrees of freedom) for "dimension," and he called the residual line "Error." We have added the column for subspaces to clarify what he meant by "Source of variation."

Table 1. Expanded version of the ANOVA table given by Calvin in [17]

Subspace	Source of variation	d.f.	S.S.	M.S.
V_0	Grand mean	1	390.15	390.15
W_B	Blocks (unadjusted)	14	18.10	1.29
$(V_1 + V_B) \cap V_B^\perp$	Treatments (adjusted)	5	71.17	14.23
W_A	Correlations (adjusted)	10	27.17	2.72
$(V_A + V_B)^\perp$	Error	30	40.41	1.35
V	Total	60	547.00	—

In (1) we gave the decomposition of V into homogeneous subspaces. The approach outlined in [48] (via the discrete Radon transform [10]) gives the sum of squares (S.S.) for each of these as follows.

$$\begin{array}{c|ccccccc} & M^{3,2,1} = & S^6 & \oplus\ 2S^{5,1} & \oplus\ 2S^{4,2} & \oplus\ S^{4,1,1} & \oplus\ S^{3,3} & \oplus\ S^{3,2,1} \\ \hline \text{S.S.} & 547 & = 390.15 & + 77.667 & + 23.40 & + 27.167 & + 8.083 & + 20.533 \end{array} \quad (3)$$

Let U_1 and U_2 be the reducible homogeneous subspaces of dimensions 10 and 18 respectively. We now explore different ways of decomposing these two subspaces into orthogonal irreducibles.

Each block can be labelled by the pair of treatments which are not present in it. Thus the fifteen blocks have structure similar to that in Example 2.

The method given by Yates in [57] decomposes the sum of squares for blocks into 6.50 for $\tilde{V}_B$ and 11.60 for $W_B \cap \tilde{V}_B^\perp$. These two subspaces are $U_1 \cap V_B$ and $U_2 \cap V_B$ respectively. We have already seen that $\tilde{V}_B^\perp \cap U_1 = V_P \cap V_A = (V_1 + V_B) \cap V_B^\perp$, whose sum of squares is given in Table 1 as 71.17 (to two decimal places). Then $6.50 + 71.17 = 77.67$, which gives the sum of squares for U_1, as confirmed in (3). The sum of squares for U_2 is 23.40, so the sum of squares for $U_2 \cap V_B^\perp$ is $23.40 - 11.60 = 11.80$.

A statistician expects there to be differences between the blocks. It is therefore standard to begin with the decomposition $V_0 \oplus W_B \oplus V_B^\perp$ and then refine that. Refining it into irreducibles gives the ANOVA in Table 2.

Table 2. ANOVA table obtained by refining the original decomposition (defined by blocks) into group-irreducibles

Original	Group refinement	d.f.	S.S.	M.S.
V_0	V_0	1	390.15	390.15
W_B	$U_1 \cap V_B$	5	6.50	1.30
	$U_2 \cap V_B$	9	11.60	1.29
$V_B^\perp$	$U_1 \cap V_B^\perp = V_P \cap V_A$	5	71.17	14.23
	W_A	10	27.17	2.72
	$U_2 \cap V_B^\perp$	9	11.80	1.31
	$S^{3,3}$	5	8.08	1.62
	$S^{3,2,1}$	16	20.53	1.28
V	Total	60	547.00	—

It is not unusual for the mean square for V_0 to be much larger than the rest. The interesting question for data analysis is: which other mean squares are significantly larger than the rest?

One notable feature of Table 2 is that the four smallest mean squares are all approximately equal (about 1.3). In particular, the two subspaces of W_B are among these, so it appears that there are no differences between blocks.

In fact, given the way that the experiment was carried out, this is not surprising. Each taster had to give four integer scores in the range $[0, 5]$, and most of them thought that they should not give the same score twice. It was therefore almost impossible for one taster to give consistently higher scores than another.

If blocks are not important, what other natural subspaces of U_1 and U_2 should we look at? To Calvin, the next most obvious subspace of U_1 was W_T, where $W_T = V_1 \cap V_0^\perp$ and $f \in V_1$ if there are constants $\gamma_1, \ldots, \gamma_6$ such that

$$f((i, \{j, k, l\})) = \gamma_i. \tag{4}$$

The constant γ_i is estimated by the mean of the responses for treatment i, and this gives the sum of squares for W_T as 63.35.

However, Calvin also proposed that treatments should affect each other asymmetrically: the taster has a fixed, short scale, so if one treatment's score goes up then another comes down. Another submodel of his model is

$$f((i,\{j,k,l\})) = \mu + 3\alpha_i - \alpha_j - \alpha_k - \alpha_l, \tag{5}$$

where μ is an overall constant. This corresponds to the subspace $V_0 \oplus (V_P \cap V_A)$. Since $V_P \cap V_A = U_1 \cap V_B^{\perp}$, we have already seen that the sum of squares for this is 71.17. Thus model (5) fits the data better than model (4).

In Table 3, the tasters' scores (the data) are shown at the top of each box. Below that are the fitted values for model (5), which can be easily calculated from Calvin's results. They fit the data rather well. The third row gives the fitted values for the more general asymmetric model

$$f((i,\{j,k,l\})) = \mu + \theta_{ij} + \theta_{ik} + \theta_{il} \qquad \text{where } \theta_{rs} = -\theta_{sr} \text{ for all } r \text{ and } s\text{:} \tag{6}$$

this corresponds to the subspace $V_0 \oplus (V_P \cap V_A) \oplus W_A = V_0 \oplus V_A$.

Table 3. Ice cream analysis: data (top row); fitted values in model (5) (second row); fitted values in model (6) (third row); fitted values in model (7) (bottom row)

	1 2 3 4	1 2 3 5	1 2 3 6	1 2 4 5	1 2 4 6	1 2 5 6	1 3 4 5	1 3 4 6	1 3 5 6	1 4 5 6	2 3 4 5	2 3 4 6	2 3 5 6	2 4 5 6	3 4 5 6
1	2 1.02 1.49 2.42	0 0.99 0.70 2.33	0 0.86 −0.05 2.17	1 0.75 0.78 3.25	0 0.62 0.01 3.08	0 0.60 −0.76 3.00	2 0.65 2.09 2.50	0 0.52 1.32 2.33	0 0.49 0.55 2.25	2 0.25 0.61 3.17	×	×	×	×	×
2	4 2.46 3.82 2.75	4 2.44 3.05 3.33	3 2.31 3.28 2.83	3 2.20 2.11 3.25	2 2.06 2.34 2.75	1 2.04 1.57 3.33	×	×	×	×	0 1.73 1.34 3.17	1 1.60 1.57 2.67	2 1.58 0.80 3.25	1 1.33 −0.14 3.17	×
3	0 2.88 1.90 0.92	3 2.85 2.47 2.50	4 2.72 1.86 2.17	×	×	×	3 2.51 2.84 1.33	0 2.38 2.24 1.00	2 2.36 2.80 2.58	×	2 2.15 2.42 2.00	1 2.02 1.82 1.67	3 1.99 2.38 3.25	×	4 1.65 2.76 2.08
4	3 3.85 2.99 1.75	×	×	4 3.58 3.28 2.58	4 3.45 3.17 3.08	×	1 3.48 2.92 1.50	3 3.35 2.82 2.00	×	4 3.08 3.11 2.83	3 3.12 3.51 1.17	2 2.99 3.40 1.67	×	4 2.72 3.70 2.50	3 2.62 3.34 1.42
5	×	5 3.91 3.97 3.83	×	5 3.67 4.03 3.08	×	4 3.52 4.65 3.00	4 3.56 2.34 3.00	×	3 3.41 2.97 2.92	1 3.17 3.03 2.17	1 3.20 2.92 3.33	×	5 3.05 3.55 3.25	5 2.81 3.61 2.50	1 2.70 1.92 2.42
6	×	×	5 4.31 5.11 2.83	×	5 4.06 4.67 2.92	5 4.04 4.73 2.33	×	4 3.96 3.82 3.00	3 3.94 3.88 2.42	4 3.70 3.45 2.50	×	4 3.60 3.40 3.00	4 3.58 3.47 2.42	2 3.33 3.03 2.50	2 3.23 2.17 2.58

The asymmetric effect is sometimes known as *competition*x. It occurs when neighbouring treatments compete for finite resources, be they food or tasters' good opinions. In some situations, a symmetric effect is more natural: that is, $\theta_{ij} = \theta_{ji}$ for all i and j. In many wildlife habitats, there is synergy between organisms filling different niches, to each others' mutual benefit, so that θ_{ij} is positive. On the other hand, antagonism gives a symmetric effect with negative θ_{ij}.

The symmetric model is

$$f((i,\{j,k,l\})) = \mu + \theta_{ij} + \theta_{ik} + \theta_{il} \qquad \text{where } \theta_{rs} = \theta_{sr} \text{ for all } r \text{ and } s. \quad (7)$$

The corresponding subspace V_S is similar to the whole space in Example 2, and decomposes as $V_0 \oplus (V_P \cap V_S) \oplus W_S$. A slight modification of Yates's method gives the corresponding sums of squares as 390.15, 7.64 and 16.54, respectively. The fitted vector in V_S is shown in the last row of Table 3. It is clearly not as good a fit to the data as either of the two rows above.

We now have three natural ways of decomposing U_1 as a pair of orthogonal irreducible subspaces. Table 4 shows the corresponding sums of squares and mean squares. Starting with blocks or with the asymmetric treatment model (5) gives (**a**); starting with direct effects of treatments, which is model (4), gives (**b**); and starting with the symmetric treatment model (7) gives (**c**). Of these, the only one where the larger mean square corresponds to a meaningful subspace and the other mean square is about 1.3 is the first. Thus consideration of U_1 suggests that we should include the subspace $V_P \cap V_A$ in the explanation for the data but that the rest of U_1 is just random noise.

Of course, there is a fourth decomposition, into the G-irreducible subspace containing the projection of y onto U_1 and its orthogonal complement, with mean squares 15.53 and 0 respectively. This is the most extreme decomposition of U_1 into orthogonal irreducibles, but we cannot consider it seriously for data analysis. In the first place, this decomposition is not known before the data are obtained. In the second place, the zero mean square is just *too* small: when four others are around 1.3 then anything much smaller is suspicious.

Table 4. Three natural ways of decomposing the subspace U_1

subspace	S.S.	M.S.
$U_1 \cap V_B$	6.50	1.30
$V_P \cap V_A$	71.17	14.23
U_1	77.67	

(**a**) blocks or asymmetric model

subspace	S.S.	M.S.
W_T	63.35	12.67
$U_1 \cap W_T^\perp$	14.32	2.86
U_1	77.67	

(**b**) direct treatment effects

subspace	S.S.	M.S.
$V_P \cap V_S$	7.64	1.53
$U_1 \cap V_S^\perp$	70.03	14.01
U_1	77.67	

(**c**) symmetric model

For decomposing U_2, we have the two possibilities shown in Table 5 (ignoring the extra one defined by the data). Starting with blocks gives (**a**), while starting with the symmetric treatment model gives (**b**). In the first, both mean squares are about 1.3, which is consistent with random noise. In the second, the larger mean square is 1.84. This is less than twice 1.3, so is unlikely to indicate anything meaningful. Moreover, it corresponds to the subspace W_S. Figure 3 shows that any natural treatment subspace containing W_S must contain the whole of V_S, in particular $V_P \cap V_S$, whose contribution to the data we have already decided is just random noise. These considerations suggest that no part of U_2 is anything more than random noise.

Table 5. Two natural ways of decomposing the subspace U_2

subspace	S.S.	M.S.
$U_2 \cap V_B$	11.60	1.29
$U_2 \cap V_B^\perp$	11.80	1.31
U_2	23.40	

(**a**) blocks

subspace	S.S.	M.S.
W_S	16.54	1.84
$U_2 \cap W_S^\perp$	6.86	0.76
U_2	23.40	

(**b**) symmetric model

There are three remaining subspaces in Table 2. Of these, $S^{3,2,1}$ has the smallest mean square, while the mean square for $S^{3,3}$ is 1.62, less than that for W_S, which we have already decided to ignore. That leaves just W_A, with a mean square of 2.72. As Figure 3 shows, including W_A in the treatment subspace that already includes V_0 and $V_P \cap V_A$ gives the rather natural subspace $V_0 + V_A$. This corresponds to model (6).

The conclusion from the spectral analysis is that model (6) explains the data well. That is, the different quantities of vanilla compete with each other for the tasters' scores, but there is no evidence of any direct effect of quantities or any differences between tasters. These conclusions differ from those in [17], because Calvin assumed that the most important effects would be the differences between tasters and the differences between the direct effects of the quantities of vanilla.

Note that the computational aspects of this are an instance of computing projections of a data vector onto the isotypic components of a representation of the symmetric group (see e.g., [25]). The general computational problem of isotypic projection for arbitrary groups is considered in [40] as well as [41].

10 Strong symmetries of orthogonal designs

Even without the complication of the effects of neighboring treatments, we can define the group G of strong symmetries of the design. Do its irreducible subspaces help us to analyze the data? In this section we revisit Examples 7–10 and consider their strong symmetries.

The simplest orthogonal case is the completely randomized design in Example 7, where Ω is unstructured and each treatment is applied to r observational units, for some integer r. Then $G = \mathrm{Sym}(r) \,\mathrm{wr}\, G_1$, and [19] shows that every decomposition of $\mathbb{R}^\Omega$ into orthogonal G-irreducible subspaces has the form $\left(\bigoplus_{j\in J} W_j\right) \oplus V_\Gamma^\perp$, where $\bigoplus_{j\in J} W_j$ is an orthogonal decomposition of V_Γ into G_1-irreducible subspaces. Here $V_\Gamma^\perp$ is the subspace which is classically called "Error" or "residual" in the ANOVA, and so the approach using strong symmetries gives nothing new.

For a complete-block design in b blocks of size k, as in Example 8, we have $|\Gamma| = k$, $G_2 = \mathrm{Sym}(k) \,\mathrm{wr}\, \mathrm{Sym}(b)$, and $G = G_1 \times \mathrm{Sym}(b)$ in its product action. Let U_0 and U_1 be the irreducibles of $\mathrm{Sym}(b)$ in its natural action, of dimensions 1 and $b-1$ respectively. If $\bigoplus_{j\in J} W_j$ is an orthogonal G_1-irreducible decomposition of $\mathbb{R}^\Gamma$ then the subspaces in an orthogonal G-irreducible decomposition of $\mathbb{R}^\Omega$ are $U_i \otimes W_j$ for i in $\{0,1\}$ and j in J. Here $U_0 \otimes W_0 = V_0$, and $U_1 \otimes W_0 = W_B$, which is the subspace for differences between blocks. The subspaces $U_0 \otimes W_j$, for j in $J \setminus \{0\}$, give the decomposition of $V_\Gamma \cap V_0^\perp$ specified by G_1. If $|J| = 2$ then the only remaining subspace is $U_1 \otimes W_1$, which is $(V_B + V_\Gamma)^\perp$, the unique residual subspace. This case was discussed, from the point of view of strong symmetries, in [29, 38]. However, if $|J| \geq 3$ then $(V_B + V_\Gamma)^\perp$ is not G-irreducible.

For example, suppose that $k = mn$ and that Γ is as in Example 1. The approach of Section 6 gives the ANOVA in Table 6(**a**), while consideration of strong symmetries gives the decomposition in Table 6(**b**). Which should be used?

There is disagreement among statisticians about how to answer this question. The approach described by Nelder in [43, 44] is that in Sections 6–7: start with the decomposition of $\mathbb{R}^\Omega$ determined by G_2 and refine it using the decomposition of V_Γ. If there are simply fifteen treatments then $(V_B + V_\Gamma)^\perp$ is used as the residual subspace: why should this be decomposed if the fifteen treatments are all combinations of five varieties of cow-peas with three methods of cultivation? This gives the decomposition in Table 6(**a**). A popular alternative approach is to start with a list of *factors* (that is, partitions of Ω with named parts) and close it under infima, where the infimum of two partitions is their coarsest common refinement. This gives the decomposition in Table 6(**b**). These two approaches are contrasted in [13].

It is not uncommon for the initial structure on Ω to be defined by a family $\mathcal{P}_2$ of partitions of Ω (such as the partitions into blocks, rows or columns) and the structure on Γ to be defined by a family $\mathcal{P}_1$ of partitions of Γ (such as those defined by factors C and D in Example 1). Each partition defines the subspace of vectors which are constant on each of its parts. Two partitions are said to be *orthogonal* to each other if their corresponding subspaces are geometrically orthogonal. The design function τ enables us to consider partitions of Γ to be partitions of Ω which refine the partition defined by the inverse images of τ. If $\mathcal{P}_1 \cup \mathcal{P}_2$ contains the two trivial partitions of Ω, is closed under suprema, and has the property that each pair of partitions is orthogonal, then it defines an orthogonal decomposition $\mathcal{W}$ of $\mathbb{R}^\Omega$ [54]. Now G is the group of

Table 6. Two different decompositions for a factorial design in complete blocks

Subspace	Source of variation	d.f.
$U_0 \otimes W_0 = V_0$	Grand mean	1
$U_1 \otimes W_0$	Blocks	$b-1$
$U_0 \otimes W_C$	Main effect of C	$n-1$
$U_0 \otimes W_D$	Main effect of D	$m-1$
$U_0 \otimes W_{CD}$	C-by-D interaction	$(n-1)(m-1)$
$(V_B + V_\Gamma)^\perp$	Residual	$(b-1)(mn-1)$
V	Total	bmn

(**a**) Method of Section 6

Subspace	Source of variation	d.f.
$U_0 \otimes W_0 = V_0$	Grand mean	1
$U_1 \otimes W_0$	Blocks	$b-1$
$U_0 \otimes W_C$	Main effect of C	$n-1$
$U_1 \otimes W_C$	Residual for main effect of C	$(b-1)(n-1)$
$U_0 \otimes W_D$	Main effect of D	$m-1$
$U_1 \otimes W_D$	Residual for main effect of D	$(b-1)(m-1)$
$U_0 \otimes W_{CD}$	C-by-D interaction	$(n-1)(m-1)$
$U_1 \otimes W_{CD}$	Residual for C-by-D-interaction	$(b-1)(n-1)(m-1)$
V	Total	bmn

(**b**) Irreducible subspaces of group of strong symmetries

permutations of Ω which preserve every partition in $\mathcal{P}_1 \cup \mathcal{P}_2$. In order for the subspaces in $\mathcal{W}$ to be G-irreducible, it is necessary that each partition have all its parts of the same size (otherwise, G cannot be transitive on Ω) and that $\mathcal{P}_1 \cup \mathcal{P}_2$ be closed under infima. It is arguable that the problem with the preceding example of a factorial design in complete blocks is the lack of closure under infima.

If, in addition to satisfying the preceding properties, the lattice of partitions in $\mathcal{P}_1 \cup \mathcal{P}_2$ is distributive, then G is a generalized wreath product and its irreducible subspaces are precisely those in $\mathcal{W}$ [8]. In this case, the strong symmetries give the same decomposition as that in Section 6. The split-plot design in Example 10 is a case in point.

To show that lack of closure under infima is not the whole explanation, we conclude this section by considering the Latin-square design in Example 9, and suppose that $G_1 = \mathrm{Sym}(n)$. The partitions of Ω into rows, columns, and letters, together with the two trivial partitions, satisfy all the aforementioned conditions, except that the lattice is not distributive. Now G is the subgroup of $\mathrm{Sym}(n) \times \mathrm{Sym}(n)$ which preserves the partition into letters. If the Latin square is not the Cayley table of a group then G may not even be transitive on Ω: indeed, it may be trivial. Even when it is such a Cayley table, the results in [2] show that there may be surprisingly many G-irreducibles in a decomposition of $\mathbb{R}^\Omega$. However, neither of the common approaches to ANOVA described above uses any finer decomposition than the one in Section 6.

Thus considerations of symmetries, partitions, combinatorial conditions, or models, may lead to different analyses. The Latin square seems to be a relatively straightforward design, yet subtle differences in assumptions have led to arguments over the correct data analysis ever since Neyman [46].

11 Strong symmetries of incomplete-block designs

In this section we return to the incomplete-block designs of Section 8, and use the notation introduced there. Thus Ω consists of b blocks of size k, and Γ consists of t treatments, where $t > k$. We assume no structure on Γ. The group G of strong symmetries consists of all permutations of Ω which preserve the partition into blocks and the partition into treatments.

James argued in [32] that ANOVA should use a decomposition of $\mathbb{R}^\Omega$ into orthogonal G-irreducible subspaces. Here we compare this approach with that of Section 8.

Let ρ be the permutation representation of this action of G, with permutation character π. If $g \in G$ then $\rho(g)$ fixes the subspaces V_0, W_B and W_T. Therefore $\rho(g)$ commutes with Q_B and P_T as well as with I and J, and so $\mathcal{A} \subseteq \mathcal{C}(G)$. Hence each of the summands in (2) is G-invariant, while U_i is G-isomorphic to V_i for $i = 1, \ldots, s$.

For simplicity, write $V_B \cap V_\Gamma^\perp$ as W_{B-T}, $V_\Gamma \cap V_B^\perp$ as W_{T-B} and $(V_B + V_\Gamma)^\perp$ as W. Assume that $k \geq 2$ and $r \geq 2$, so that V_0 and W are both non-zero. Let δ be the number of subspaces among W_{B-T} and W_{T-B} that are non-zero, so that $\delta \in \{0, 1, 2\}$.

A block design is said to be *resolvable* if there is a partition of Ω into replicates, coarser than the partition into blocks, such that each treatment occurs once in each replicate. For a resolvable design, define W_R analogously to W_B. Then $\dim(W_R) = r - 1$ and $W_R \leq W_{B-T}$: hence the latter cannot be zero and so $\delta \geq 1$.

Returning to the general case, recall that the *rank* p of G is defined to be the number of orbits of G in its induced action on $\Omega \times \Omega$ (see [55]). If G is transitive on Ω then p is equal to the number of orbits on Ω of the stabilizer in G of any element of Ω. Less obviously, p is also equal to the sum of the squares of the multiplicities of complex-irreducible characters in π.

As in Section 3, there are non-negative integers m_i such that $\pi = \sum_{i \in \mathcal{I}} m_i \chi_i$, where $\{\chi_i : i \in \mathcal{I}\}$ is the set of real-irreducible characters of G. The relation to complex-irreducibles is explained in [4, 9, 18, 42, 51], as follows. The set $\mathcal{I}$ is the disjoint union of $\mathcal{I}_1$, $\mathcal{I}_2$ and $\mathcal{I}_3$. If $\chi \in \mathcal{I}_1$ then χ is also complex-irreducible, of real type; if $\chi \in \mathcal{I}_2$ then $\chi = 2\eta$, where η is a complex-irreducible of quaternionic type; if $\chi \in \mathcal{I}_3$ then $\chi = \zeta + \bar{\zeta}$, where ζ is complex-irreducible of complex type. Therefore

$$p = \sum_{i \in \mathcal{I}_1} m_i^2 + 4 \sum_{i \in \mathcal{I}_2} m_i^2 + 2 \sum_{i \in \mathcal{I}_3} m_i^2. \tag{8}$$

For an incomplete-block design, comparison of (8) with (2) shows that

$$p \geq 2 + \delta + 4s,$$

with equality if and only if each of V_0, W_{B-T}, W_{T-B}, W, U_1, …, U_s is G-irreducible, admitting a complex-irreducible character of real type, and there are no G-isomorphisms among this list of subspaces. In particular, if $p \leq 9$ then $s = 1$ and U_1 is G-irreducible.

Furthermore, if $p = 6$ then $s = 1$, $\delta = 0$, the design is a balanced incomplete-block design with $b = t$, and W is G-irreducible. If $p = 7$ then $s = 1$. In this case either $\delta = 1$, $t \neq b$, the design or its dual is a balanced incomplete-block design, and W is G-irreducible; or $\delta = 0$, the design is a balanced incomplete-block design with $b = t$, and W is the sum of two G-irreducible subspaces. If $p = 8$ then $s = 1$, then either $\delta = 2$, the design and its dual are both partial geometric designs, and W is G-irreducible; or $\delta = 1$, $t \neq b$, the design or its dual is a balanced incomplete-block design and one of W, W_{B-T}, W_{T-B} is the sum of two G-irreducible subspaces; or $\delta = 0$, the design is a balanced incomplete-block design with $b = t$, and W is the sum of three G-irreducible subspaces.

We now specialize these results to several well-known families of incomplete-block designs.

Example 11. Let q be a prime power. Then there is a Desarguesian projective plane Π of order q. Its points and lines can be used as the treatments and blocks in an incomplete-block design with $t = b = q^2 + q + 1$, $r = k = q + 1$ and $\delta = 0$. The group of strong symmetries is $\mathrm{P\Gamma L}(3, q)$, which is transitive on sets of four points in general position. This can be used to show that G has rank 6: the details are in [4, 14, 15, 20, 21, 32]. Hence a G-irreducible decomposition of $\mathbb{R}^\Omega$, with dimensions, is

	V_0 $\oplus$	W_B $\oplus$	W_T $\oplus$	W
dim	1	q^2+q	q^2+q	q^3

where only the middle two subspaces are non-orthogonal to each other. In this case, the G-decomposition is the same as that used in classical ANOVA.

Example 12. Consider the affine plane Δ obtained from the projective plane Π in Example 11 by deleting one line and all points on it. This gives a resolvable balanced incomplete-block design with $t = q^2$, $b = q(q + 1)$, $r = q + 1$, $k = q$, and $\delta = 1$. This design is also known as a *balanced square lattice design.* Its group G of strong symmetries is the stabilizer in $\mathrm{P\Gamma L}(3, q)$ of the omitted line. This is transitive on the units in Ω, which may be identified with the flags (x, λ) where x is a point of Δ incident with the line λ of Δ. The stabilizer in G of (x, λ) has the following orbits on Ω:

$$\{(x, \lambda)\}$$
$$\{(x, \mu) : x \in \mu \neq \lambda\}$$

$$\begin{aligned}
&\{(z,\lambda) : x \neq z \in \lambda\}\\
&\{(z,\mu) : z \in \mu, z \in \lambda, x \notin \mu\}\\
&\{(z,\mu) : z \in \mu, z \notin \lambda, x \in \mu\}\\
&\{(z,\mu) : z \in \mu, \mu \| \lambda\}\\
&\{(z,\mu) : z \in \mu, z \notin \lambda, \mu \not\| \lambda, x \notin \mu\}.
\end{aligned}$$

Thus the rank of G is 7, and so a G-irreducible decomposition of $\mathbb{R}^\Omega$, with dimensions, is

	V_0	$\oplus$	W_R	$\oplus$	$(W_B \cap W_R^\perp)$	$\oplus$	W_T	$\oplus$	W
dim	1		q		q^2-1		q^2-1		$(q-1)^2(q+1)$

All pairs of subspaces are orthogonal, apart from the two with dimension q^2-1. Moreover, $W_R = W_{B-T}$. This decomposition was obtained by Burton and Chakravarti in [15].

For a resolvable block design, the classical ANOVA normally splits W_B into W_R and $W_B \cap W_R^\perp$. These subspaces are called *replicates* and *blocks within replicates* respectively. So this is another example where the G-decomposition is the same as that used in classical ANOVA.

Example 13. The simple square lattice design introduced by Yates in [56] is resolvable with $r = 2$. The treatments are identified with an abstract $n \times n$ array, so that $t = n^2$, where $n > 1$. In the first replicate, the rows of the array are blocks; in the second replicate, the columns of the array are blocks. Thus $k = n$ and $b = 2n$.

Now $\dim(W_R) = 1$ and so $W_B \cap W_{B-T}^\perp$ has dimension at most $b-2$, which is $2(n-1)$. Therefore $\dim(W_{T-B}) = t - 1 - \dim(W_B \cap W_{B-T}^\perp) \geq n^2 - 1 - 2(n-1) = (n-1)^2 > 0$. Hence W_{B-T} and W_{T-B} are both non-zero and so $\delta = 2$. Both the design and its dual are partial geometric designs.

Now the group G of strong symmetries is $\mathrm{Sym}(n) \,\mathrm{wr}\, \mathrm{Sym}(2)$ in its product action. It is generated by all permutations of the set of rows ($\mathrm{Sym}(n)$), all permutations of the set of columns ($\mathrm{Sym}(n)$), and the interchange of rows and columns ($\mathrm{Sym}(2)$). It is transitive on flags. If x is the treatment in row λ and column μ then the stabilizer in G of the flag (x, λ) has the following orbits on Ω:

$$\begin{aligned}
&\{(x,\lambda)\}\\
&\{(x,\mu)\}\\
&\{(z,\lambda) : x \neq z \in \lambda\}\\
&\{(z,\mu) : x \neq z \in \mu\}\\
&\{(z,\nu) : z \in \lambda, z \in \nu \neq \mu, \nu \text{ is a column}\}\\
&\{(z,\nu) : z \in \mu, z \in \nu \neq \lambda, \nu \text{ is a row}\}\\
&\{(z,\nu) : z \notin \lambda, z \in \nu \neq \mu, \nu \text{ is a column}\}\\
&\{(z,\nu) : z \notin \mu, z \in \nu \neq \lambda, \nu \text{ is a row}\}.
\end{aligned}$$

Hence the rank of G is 8, and the subspaces V_0, W_{B-T}, W_{T-B}, U_1, V_1 and W are all G-irreducible. Since W_R is G-invariant and $W_R \leq W_{B-T}$, we must have $W_R = W_{B-T}$. Hence the decomposition, with dimensions, is

	V_0	$\oplus$	W_R	$\oplus$	U_1	$\oplus$	V_1	$\oplus$	W_{T-B}	$\oplus$	W
dim	1		1		$2(n-1)$		$2(n-1)$		$(n-1)^2$		$(n-1)^2$.

Only the pair U_1 and V_1 are non-orthogonal. In spite of the equality of their dimensions, the subspaces W_{T-B} and W are not G-isomorphic, because otherwise the rank would be greater.

Example 14. Projective spaces of higher dimension are also considered in [4, 15]. Here we consider dimension 3. Let q be a prime power and let Θ be the projective space of dimension 3 over the field with q elements. Take the treatments and blocks to be the points and planes of Θ, so that $t = b = q^3 + q^2 + q + 1$ and $r = k = q^2 + q + 1$. The design and its dual are both balanced, and so $\delta = 0$. Now $G = \mathrm{P\Gamma L}(4, q)$, which is transitive on ordered sets of five points in general position, so the stabilizer in G of a flag (x, Ψ) has the following orbits on Ω:

$$\begin{aligned}
&\{(x, \Psi)\}\\
&\{(x, \Phi) : x \in \Phi \neq \Psi\}\\
&\{(z, \Psi) : x \neq z \in \Psi\}\\
&\{(z, \Phi) : x \neq z \in \Phi \neq \Psi, z \in \Psi, x \in \Phi\}\\
&\{(z, \Phi) : z \in \Phi, z \in \Psi, x \notin \Phi\}\\
&\{(z, \Phi) : z \in \Phi, z \notin \Psi, x \in \Phi\}\\
&\{(z, \Phi) : z \in \Phi, z \notin \Psi, x \notin \Phi\}.
\end{aligned}$$

Thus G has rank 7, and so U_1 and V_1 are G-irreducible while W is the sum of two G-irreducibles.

The space Θ contains $(q^2+1)(q^2+q+1)$ lines, each incident with $(q+1)^2$ flags. Let V_L be the subspace of $\mathbb{R}^\Omega$ spanned by the characteristic vectors of the lines. Then V_L is G-invariant. Analysis of the permutation characters of G on lines, flags and points shows that V_L is the sum of three G-irreducible subspaces: one is V_0; one is G-isomorphic to both W_B and W_T; the third is orthogonal to $V_B + V_T$ and has dimension $q^4 + q^2$. Hence the G-irreducible subspaces of W are $V_L \cap W$ and $W \cap V_L^\perp$. The decomposition, with dimensions, is

	V_0	$\oplus$	W_{B-T}	$\oplus$	W_{T-B}	$\oplus$	$(V_L \cap W)$	$\oplus$	$(W \cap V_L^\perp)$
dim	1		q^3+q^2+q		q^3+q^2+q		q^4+q^2		$q^5+q^4+q^3$.

In this case, the group of strong symmetries decomposes the classical residual subspace into two parts.

Example 15. A *triple square lattice* is made from a simple one by adding an extra replicate. A Latin square is superimposed on the $n \times n$ array. A block in the third replicate contains all treatments in a given letter of the square.

Now G is the group of all permutations of the square array which preserve the set of three partitions into rows, columns and letters, in its action on the $3n^2$ flags. Depending on the Latin square, G may not be transitive, and finding a meaningful G-irreducible decomposition may be as difficult as for the case of a Latin-square design discussed at the end of Section 10.

These examples show that the decomposition defined by the group of strong symmetries may be the same as the classical one, may give further decomposition of the residual subspace, or may prove intractable. James [33] and Bailey [3] have both suggested that using the group of *weak* symmetries of Γ may give a meaningful decomposition of V_Γ. This group consists of those permutations of Γ whose permutation matrices commute with $X^\top Q_B X$, where X is the $\Omega \times \Gamma$ incidence matrix whose (ω, i)-entry is equal to 1 if $\tau(\omega) = i$ and is equal to 0 otherwise. However, this approach does not get the extra residual subspace in Example 14, nor does it make Example 15 tractable.

References

1. Bailey, R. A.: A unified approach to design of experiments. J. Roy. Statist. Soc., Series A, **144**, 214–223 (1981)
2. Bailey, R. A.: Latin squares with highly transitive automorphism groups. J. Austral. Math. Soc., Series A, **33**, 18–22 (1982)
3. Bailey, R. A.: Automorphism groups of block structures with and without treatments. In: Ray-Chaudhuri, D. (ed) Coding Theory and Design Theory, Part II, Design Theory. Springer, New York (1990), 24–41.
4. Bailey, R. A.: Strata for randomized experiments (with discussion). J. Roy. Statist. Soc., Series B, **53**, 27–78 (1991)
5. Bailey, R. A.: Design of Comparative Experiments. Cambridge University Press, Cambridge (2008)
6. Bailey, R. A.: Structures defined by factors. In: Bingham, D., Dean, A. M., Morris, M., Stufken, J. (eds) Handbook on Design of Experiments. Chapman and Hall, Boca Raton (2014)
7. Bailey, R. A., Monod, H., Morgan, J. P.: Construction and optimality of affine-resolvable designs. Biometrika, **82**, 187–200 (1995)
8. Bailey, R. A., Praeger, C. E., Rowley, C. A., Speed, T. P.: Generalized wreath products of permutation groups. Proc. London. Math. Soc., **47**, 69–82 (1983)
9. Bailey, R. A., Rowley, C. A.: General balance and treatment permutations. Linear Alg. Appl., **127**, 183–225 (1990)
10. Bolker, E.: The finite Radon transform. Contemporary Math., **63**, 27–50 (1987)
11. Bose, R. C., Bridges, W. G., Shrikhande, M. S.: A characterization of partial geometric designs. Discrete Math., **16**, 1–7 (1976)
12. Bose, R. C., Shrikande, S. S., Singhi, N. M.: Edge regular multigraphs and partial geometric designs with an application to the embedding of quasi-residual designs. In Teorie Combinatoire (Rome, 1973), Atti dei Convegni Lincei, **17**, Vol. I, Accademia Nazionale dei Lincei, Roma (1976), 49–81.

13. Brien, C. J., Harch, B. D., Correll, R. L., Bailey, R. A.: Multiphase experiments with at least one later laboratory phase. I. Orthogonal designs. J. Agric. Biol. Env. Statist., **16**, 422–450 (2011)
14. Burton, C. T.: Automorphism groups of balanced incomplete block designs and their use in statistical model construction and analysis. PhD Thesis, University of North Carolina at Chapel Hill (1980)
15. Burton C. T., Chakravarti, I. M.: On the commutant algebras corresponding to the permutation representations of the full collineation groups of PG(k, s) and EG(k, s), $s = p^r$, $k \geq 2$. J. Math. Anal. Appl., **89**, 489–514 (1982)
16. Caliński, T.: On some desirable patterns in block designs. Biometrics, **27**, 275–292 (1971)
17. Calvin, L. D.: Doubly balanced incomplete block designs for experiments in which the treatment effects are correlated. Biometrics, **10**, 61–88 (1954)
18. Cameron, P. J.: Bounding the rank of certain permutation groups. Math. Z., **124**, 243–352 (1972)
19. Ceccherini–Silberstein, T., Scarabotti, F., Tolli, F.: Representation Theory and Harmonic Analysis of Wreath Products of Finite Groups. London Mathematical Society Lecture Notes Series **410**. Cambridge University Press, Cambridge (2014)
20. Chakravarti, I. M., Burton, C. T.: On the algebras of symmetries (groups of collineations) of designs from finite Desarguesian planes with applications in statistics. J. Math. Anal. Appl., **89**, 515–529 (1982)
21. Chakravarti, I. M., Burton, C. T.: Symmetries (groups of automorphisms) of Desarguesian finite projective and affine planes and their role in statistical model construction. In: Kallianpur, G., Krishnaiah, P. R., Ghosh, J. K. (eds) Statistics and Probability: Essays in Honor of C. R. Rao. North-Holland, Amsterdam (1982), 169–178.
22. Cheng, C. S., Bailey, R. A.: Optimality of some two-associate-class partially balanced incomplete-block designs. Ann. Statist., **19**, 1667–1671 (1991)
23. Diaconis, P.: Group Representations in Probability and Statistics. IMS Lecture Notes—Monograph Series **11**. Institute of Mathematical Statistics, Hayward, CA (1988)
24. Diaconis, P.: Applications of group representations to statistical problems. In Proceedings of the ICM, Kyoto, Japan, 1990. Springer-Verlag, Tokyo (1991), 1037–1048.
25. Diaconis, P., Rockmore D.: Efficient computation of isotypic projections for the symmetric group. In: Finkelstein, L., Kantor W. (eds) DIMACS Series in Discrete Math. Theoret. Comp. Sci. Volume **11**, 87–104 (1993)
26. Fisher, R. A.: The theory of confounding in factorial experiments in relation to the theory of groups. Ann. Eugen., **11**, 341–353 (1942)
27. Fortini, P.: Representations of groups and the analysis of variance. PhD Thesis, Harvard University (1977)
28. Griffing, B.: Concept of general and specific combining ability in relation to diallel crossing system. Austral. J. Biol. Sci., **9**, 463–493 (1956)
29. Hannan, E. J.: Group representations and applied probability. J. Appl. Prob., **2**, 1–68 (1965)
30. Houtman, A. M., Speed T. P.: Balance in designed experiments with orthogonal block structure. Ann. Statist., **11**, 1069–1085 (1983)
31. James, A. T.: The relationship algebra of an experimental design. Ann. Math. Statist., **28**, 993–1002 (1957)

32. James, A. T.: Analysis of variance determined by symmetry and combinatorial properties of zonal polynomials. In: Kallianpur, G., Krishnaiah, P. R., Ghosh, J. K. (eds) Statistics and Probability: Essays in Honor of C. R. Rao. North-Holland, Amsterdam (1982), 329–341.
33. James A. T.: Contribution to the discussion of 'Strata for randomized experiments' by R. A. Bailey. J. Roy. Statist. Soc., Series B, **53**, 71 (1991)
34. James, A. T., Wilkinson, G. N.: Factorization of the residual operator and canonical decomposition of nonorthogonal factors in the analysis of variance. Biometrika, **58**, 279–294 (1971)
35. James, G., Kerber, A.: The Representation Theory of the Symmetric Group. Encyclopedia of Mathematics and its Applications **16**. Addison–Wesley, Reading, Mass. (1981)
36. John, P. M. W.: Statistical Design and Analysis of Experiments. Macmillan, New York (1971)
37. Kung, J.: The Radon transform of a combinatorial geometry. J. Comb. Theory, Series A, **26**, 97–102 (1979)
38. Ledermann, W.: Representation theory and statistics. In Séminaire Dubreil-Pisot (Algèbre et Théorie des Nombres), **15**, 15.01–15.08. University of Paris (1967)
39. Mann, H. B.: The algebra of a linear hypothesis. Ann. Math. Statist., **31**, 1–15 (1960)
40. Maslen, D. K., Orrison, M. E., Rockmore, D. N.: Computing isotypic projections with the Lanczos iteration. SIAM J. Matrix Anal. Appl., **25**, 784–803 (2003)
41. Maslen, D. K., Rockmore, D. N.: Separation of variables and the computation of Fourier transforms on finite groups, I. J. Amer. Math. Soc., **10**, No. 1, 169–214 (1997)
42. McLaren, A. D.: On group representations and invariant stochastic processes. Proc. Camb. Phil. Soc., **59**, 431–450 (1963)
43. Nelder, J. A.: The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. Proc. Roy. Soc., Series A, **283**, 147–162 (1965)
44. Nelder, J. A.: The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance. Proc. Roy. Soc., Series A, **283**, 163–178 (1965)
45. Neumaier, A.: $t\frac{1}{2}$-Designs. J. Comb. Theory, Series A, **28**, 226–248 (1980)
46. Neyman, J.: Statistical problems in agricultural experimentation. J. Roy. Statist. Soc., Suppl., **2**, 107–154 (1935)
47. Pearce, S. C.: The Agricultural Field Experiment. Wiley, Chichester (1983)
48. Rockmore, D. N.: Some applications of generalized FFTs. In Finkelstein, L., Kantor, W. (eds.), DIMACS Series in Discrete Math. Theoret. Comput. Sci, Volume **28**, 329–370 (1997)
49. Saha, G. M.: On Caliński's patterns in block designs. Sankhyā, Series B, **38**, 383–392 (1976)
50. Scheffé, H.: The Analysis of Variance. J. Wiley and Sons, Inc., New York (1959)
51. Serre, J.-P.: Linear Representations of Finite Groups. Springer, New York (1977)
52. Speed, T. P.: ANOVA models with random effects: an approach via symmetry. In: Gani, J., Priestly, M. B. (eds) Essays in Time Series and Allied Processes: Papers in Honour of E. J. Hannan. Applied Probability Trust, Sheffield (1986), 355–368.

53. Speed, T. P.: What is an analysis of variance? (with discussion). Ann. Statist., **15**, 885–910 (1987)
54. Tjur, T.: Analysis of variance models in orthogonal designs (with discussion). Internat. Statist. Rev., **52**, 33–81 (1984)
55. Wielandt, H.: Finite Permutation Groups. Academic Press, New York (1964)
56. Yates, F.: A new method of arranging variety trials involving a large number of varieties. J. Agric. Sci., **26**, 424–455 (1936)
57. Yates, F.: Analysis of data from all possible reciprocal crosses between a set of parental lines. Heredity, **1**, 287–301 (1947)