Indicator function and complex coding for mixed fractional factorial designs [1]

Giovanni Pistone

Department of Mathematics - Politecnico di Torino

Maria-Piera Rogantin *

Department of Mathematics - Università di Genova

Abstract

In a general fractional factorial design, the n-levels of a factor are coded by the n-th roots of the unity. This device allows a full generalization to mixed-level designs of the theory of the polynomial indicator function which has already been introduced for two level designs by Fontana and the Authors (2000). the properties of orthogonal arrays and regular fractions are discussed.

Key words: Algebraic statistics, Complex coding, Mixed-level designs, Regular fraction, Orthogonal arrays

1 Introduction

Algebraic and geometric methods are widely used in the theory of the design of experiments. A variety of these methods exist: real linear algebra, $\mathbb{Z}_p$ arithmetic, Galois Fields $\mathrm{GF}(p^s)$ arithmetic, where p is a prime number as in Bose (1947). See, e.g., Raktoe et al. (1981) and the more recent books by Dey and Mukerjee (1999) and Wu and Hamada (2000).

Complex coding of levels has been used by many authors in various contexts, see e.g. Bailey (1982), Kobilinsky and Monod (1991), Edmondson (1994), Kobilinsky and Monod (1995), Collombier (1996) and Xu and Wu (2001).

* Corresponding author.
Email addresses: `giovanni.pistone@polito.it` (Giovanni Pistone), `rogantin@dima.unige.it` (Maria-Piera Rogantin).

[1] Partially supported by the Italian PRIN03 grant coordinated by G. Consonni

The use of a new background, called Commutative Algebra or Polynomial Ring Algebra, was first advocated by Pistone and Wynn (1996) and later discussed in detail in Pistone et al. (2001). Other relevant general references are Robbiano (1998), Robbiano and Rogantin (1998) and Galetto et al. (2003).

In the present paper, mixed-level (or asymmetric) designs with replicates are considered and the approach to the two-level designs discussed in Fontana et al. (1997) and Fontana et al. (2000) is generalized. In the latter, the fractional factorial design was encoded in its indicator function with respect to the full factorial design. In Tang and Deng (1999), entities related to coefficients of the polynomial indicator function were independently introduced into the construction of a generalized word length pattern. The coefficients themselves were called J-characteristics in Tang (2001), where it was shown that a two-level fractional design is uniquely determined by its J-characteristics. The representation of a fraction by its indicator polynomial function was generalized to designs with replicates in Ye (2003) and extended to non two-level factors using orthogonal polynomials with an integer coding of levels in Cheng and Ye (2004).

Sections 2 and 3 are a self-contained introduction of the indicator function representation of a factorial design using complex coding. The main results are in Sections 4 to 6. The properties of the indicator polynomial are discussed in Section 4. If the factor levels are coded with the n-th roots of the unity, the coefficients of the indicator polynomial are related to many interesting properties of the fraction in a simple way: orthogonality among the factors and interactions, projectivity, aberration and regularity. Combinatorial orthogonality vs. geometrical orthogonality is discussed in Section 5. A type of generalized regular fraction is defined and discussed in Section 6. The usual definition, where the number of levels is prime for all factors is extended to asymmetric design with any number of levels. With such a definition, all the monomial terms of any order are either orthogonal or totally aliased. However, our framework does not include the GF(p^s) case. Some examples are shown in Section 7.

A first partial draft of the present paper was presented in the GROSTAT V 2003 Workshop. Some of the results of Proposition 5 have been obtained independently by Ye (2004).

2 Coding of factor levels

Let m be the number of factors of a design. We denote the factors by A_j, $j = 1, \ldots, m$, and the number of levels of the factor A_j by n_j. We consider only *qualitative factors*.

We denote the full factorial design by $\mathcal{D}$, $\mathcal{D} = A_1 \times \cdots \times A_m$, and the space of all real responses defined on $\mathcal{D}$ by $\mathcal{R}(\mathcal{D})$.

In some cases, it is of interest to *code* qualitative factors with numbers, especially when the levels are ordered. Classical examples of numerical coding with rational numbers $a_{ij} \in \mathbb{Q}$ are: (1) $a_{ij} = i$, or (2) $a_{ij} = i - 1$, or (3) $a_{ij} = (2i - n_j - 1)/2$ for odd n_j and $a_{ij} = 2i - n_j - 1$ for even n_j, see (Raktoe et al., 1981, Tab. 4.1). The second case, where the coding takes value in the additive group $\mathbb{Z}_{n_j}$, i.e. integers mod n_j, is of special importance. We can define the important notion of *regular fraction* in such a coding. The third coding is the result of the orthogonalization of the linear term in the second coding with respect to the constant term. The coding $-1, +1$ for two-level factors has a further property in that the values $-1, +1$ form a multiplicative group. This property was widely used in Fontana et al. (2000), Ye (2003), Tang and Deng (1999) and Tang (2001).

In the present paper, an approach is taked to parallel our theory for two-level factors with coding $-1, +1$. The n levels of a factor are coded by the complex solutions of the equation $\zeta^n = 1$:

$$\omega_k = \exp\left(\mathrm{i}\,\frac{2\pi}{n}k\right) \quad , \qquad k = 0, \ldots, n-1 \ . \tag{1}$$

We denote such a factor with n levels by Ω_n, $\Omega_n = \{\omega_0, \ldots, \omega_{n-1}\}$. With such a coding, a complex orthonormal basis of the responses on the full factorial design is formed by all the monomials.

For a basic reference to the algebra of the complex field $\mathbb{C}$ and of the n-th complex roots of the unity references can be made to Lang (1965); some useful points are collected in Section 8 below.

As $\alpha = \beta \mod n$ implies $\omega_k^\alpha = \omega_k^\beta$, it is useful to introduce the residue class ring $\mathbb{Z}_n$ and the notation $[k]_n$ for the residue of $k \mod n$. For integer α, we obtain $(\omega_k)^\alpha = \omega_{[\alpha k]_n}$. The mapping

$$\mathbb{Z}_n \longleftrightarrow \Omega_n \subset \mathbb{C} \qquad \text{with} \qquad k \longleftrightarrow \omega_k \tag{2}$$

is a group isomorphism of the additive group of $\mathbb{Z}_n$ on the multiplicative group $\Omega_n \subset \mathbb{C}$. In other words,

$$\omega_h \omega_k = \omega_{[h+k]_n} \ .$$

We drop the sub-n notation when there is no ambiguity.

We denote by:

- $\#\mathcal{D}$: the number of points of the full factorial design, $\#\mathcal{D} = \prod_{j=1}^m n_j$.
- L: the full factorial design with integer coding $\{0, \ldots, n_j - 1\}$, $j = 1, \ldots, m$,

and $\mathcal{D}$ the full factorial design with complex coding:

$$L = \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_m} \qquad \text{and} \qquad \mathcal{D} = \mathcal{D}_1 \times \cdots \mathcal{D}_j \cdots \times \mathcal{D}_m \ \text{ with } \mathcal{D}_j = \Omega_{n_j}$$

According to map (2), L is both the integer coded design and the exponent set of the complex coded design;

- α, β, …: the elements of L:

$$L = \{\alpha = (\alpha_1, \ldots, \alpha_m) : \alpha_j = 0, \ldots, n_j - 1, j = 1, \ldots, m\} \ ;$$

  that is, α is both a treatment combination in the integer coding and a multi-exponent of an interaction term;
- $[\alpha - \beta]$: the m-tuple $\left([\alpha_1 - \beta_1]_{n_1}, \ldots, [\alpha_j - \beta_j]_{n_j}, \ldots, [\alpha_m - \beta_m]_{n_m}\right)$; the computation of the j-th element is in the ring $\mathbb{Z}_{n_j}$.

3 Responses on the design

The responses on the design and the linear models are discussed in this section. According to the generalization of the algebraic approach by Fontana et al. (2000), the design $\mathcal{D}$ is identified as the zero-set of the system of polynomial equations

$$\zeta_j^{n_j} - 1 = 0 \quad , \qquad j = 1, \ldots, m \ .$$

A *complex* response f on the design $\mathcal{D}$ is a $\mathbb{C}$-valued function defined on $\mathcal{D}$. This response can be considered as the restriction to $\mathcal{D}$ of a complex polynomial.

We denote by:

- X_i; the i-th component function, which maps a point to its i-th component:

$$X_i : \quad \mathcal{D} \ni (\zeta_1, \ldots, \zeta_m) \ \longmapsto \ \zeta_i \ .$$

  The function X_i is called *simple term* or, by abuse of terminology, *factor*.
- X^α, with $\alpha \in L$: the *interaction term* $X_1^{\alpha_1} \cdots X_m^{\alpha_m}$, i.e. the function

$$X^\alpha : \quad \mathcal{D} \ni (\zeta_1, \ldots, \zeta_m) \ \mapsto \ \zeta_1^{\alpha_1} \cdots \zeta_m^{\alpha_m} \quad , \qquad \alpha \in L \ .$$

  The function X^α is a special response that we call *monomial response* or *interaction term*, in analogy with current terminology.

In the following, we shall use the word *term* to indicate either a simple term or an interaction term.

We say term X^α has *order* (or order of interaction) k if k factors are involved, i.e. if the m-tuple α has k non-null entries.

If f is a response defined on $\mathcal{D}$ then its *mean value* on $\mathcal{D}$, denoted by $E_{\mathcal{D}}(f)$, is:

$$E_{\mathcal{D}}(f) = \frac{1}{\#\mathcal{D}} \sum_{\zeta \in \mathcal{D}} f(\zeta) \ .$$

We say that a response f is *centered* if $E_{\mathcal{D}}(f) = 0$. Two responses f and g are *orthogonal on* $\mathcal{D}$ if $E_{\mathcal{D}}(f\ \overline{g}) = 0$.

it should be noticed that the set of all the responses is a complex Hilbert space with the Hermitian product $f \cdot g = E_{\mathcal{D}}(f\ \overline{g})$.

Two basic properties connect the algebra to the Hilbert structure, namely

(1) $X^\alpha \overline{X^\beta} = X^{[\alpha-\beta]}$;
(2) $E_{\mathcal{D}}(X^0) = 1$, and $E_{\mathcal{D}}(X^\alpha) = 0$ for $\alpha \neq 0$, see Section 8 Item (3).

The set of functions $\{X^\alpha\ ,\ \alpha \in L\}$ is an orthonormal basis of the *complex responses* on design $\mathcal{D}$. From properties (1) and (2) above it follows that:

$$E_{\mathcal{D}}(X^\alpha \overline{X^\beta}) = E_{\mathcal{D}}(X^{[\alpha-\beta]}) = \begin{cases} 1 & \text{if } \alpha = \beta \\ 0 & \text{if } \alpha \neq \beta \end{cases}$$

Moreover, $\#L = \#\mathcal{D}$.

Each response f can therefore be represented as a unique $\mathbb{C}$-linear combination of constant, simple and interaction terms:

$$f = \sum_{\alpha \in L} \theta_\alpha\ X^\alpha, \quad \theta_\alpha \in \mathbb{C} \tag{3}$$

where the coefficients are uniquely defined by: $\theta_\alpha = E_{\mathcal{D}}\left(f \overline{X^\alpha}\right)$. In fact,

$$\sum_{\zeta \in \mathcal{D}} f(\zeta) \overline{X^\alpha}(\zeta) = \sum_{\zeta \in \mathcal{D}} \sum_{\beta \in L} \theta_\beta X^\beta \overline{X^\alpha}(\zeta) = \sum_{\beta \in L} \theta_\beta \sum_{\zeta \in \mathcal{D}} X^\beta(\zeta) \overline{X^\alpha}(\zeta) = \#\mathcal{D}\ \theta_\alpha \ .$$

We can observe that a function is centered on $\mathcal{D}$ if, and only if, $\theta_0 = 0$.

As $\overline{\theta_\alpha} = E_{\mathcal{D}}\left(\overline{f} X^\alpha\right)$, the conjugate of response f has the representation:

$$\overline{f(\zeta)} = \sum_{\alpha \in L} \overline{\theta_\alpha}\ \overline{X^\alpha}(\zeta) = \sum_{\alpha \in L} \overline{\theta_{[-\alpha]}} X^\alpha(\zeta) \ .$$

A response f is *real valued* if, and only if, $\overline{\theta_\alpha} = \theta_{[-\alpha]}$ for all $\alpha \in L$.

We suggest the use of the roots of the unity because of the mathematical convenience we are going to show. In most of the applications, we are interested in *real valued* responses, e.g. measurements, on the design points. Both the real vector space $\mathcal{R}(\mathcal{D})$ and the complex vector space $\mathcal{C}(\mathcal{D})$ of the responses on the design $\mathcal{D}$ have a real basis, see (Kobilinsky, 1990, Prop. 3.1) and Pistone and Rogantin (2005), where a special real basis that is common to both spaces is computed. The existence of a real basis implies the existence of real linear models even though the levels are complex.

4 Fractions

A fraction $\mathcal{F}$ is a subset of the design, $\mathcal{F} \subseteq \mathcal{D}$. We can algebraically describe a fraction in two ways, namely using generating equations or the indicator polynomial function.

4.1 Generating equations

All fractions can be obtained by adding further polynomial equations, called *generating equations*, to the design equations $X_j^{n_j} - 1 = 0$, for $j = 1, \ldots, m$, in order to restrict the number of solutions.

For example, let us consider a classical 3^{4-2}_{III} regular fraction, see (Wu and Hamada, 2000, Table 5A.1), coded with complex numbers according to the map in Equation (2). This fraction is defined by $X_j^3 - 1 = 0$ for $j = 1, \ldots, 4$, together with the generating equations $X_1 X_2 X_3^2 = 1$ and $X_1 X_2^2 X_4 = 1$. Such a representation of the fraction is classically termed "multiplicative" notation. In our approach, it is not a question of notation or formalism, but rather the equations are actually defined on the complex field $\mathbb{C}$. As the recoding is a homomorphism from the additive group $\mathbb{Z}_3$ to the multiplicative group of $\mathbb{C}$, then the additive generating equations in $\mathbb{Z}_3$ (of the form $A + B + 2C = 0 \mod 3$ and $A + 2B + D = 0 \mod 3$) are mapped to the multiplicative equations in $\mathbb{C}$. In this case, the generating equations are binomial, i.e. polynomial with two terms.

In the following, we consider general subsets of the full factorial design and, as a consequence, no special form of the generating equations is assumed.

4.2 Responses defined on the fraction, indicator and counting functions

The indicator polynomial was first introduced in Fontana et al. (1997) to describe a fraction. In the two-level case, Ye (2003) suggested generalizing the idea of indicator function to fractions with replicates. However, the single replicate case has special features, mainly because, in such a case, the equivalent description with generating equations is available. For coherence with general mathematical terminology, we have maintained the indicator name, and introduced the new name, that is, *counting function* for the replicate case. The design with replicates associated to a counting function can be considered a multi-subset $\mathcal{F}$ of the design $\mathcal{D}$, or an array with repeated rows. In the following, we also use the name "fraction" in this extended sense.

Definition 1 (Indicator function and counting function) *The* counting function *R of a fraction $\mathcal{F}$ is a response defined on $\mathcal{D}$ so that for each $\zeta \in \mathcal{D}$, $R(\zeta)$ equals the number of appearances of ζ in the fraction.*

A 0-1 valued counting function is called indicator function *F of a single replicate fraction $\mathcal{F}$.*

We denote the coefficients of the representation of R on $\mathcal{D}$ using the monomial basis by b_α:

$$R(\zeta) = \sum_{\alpha \in L} b_\alpha \; X^\alpha(\zeta) \qquad \zeta \in \mathcal{D} \; .$$

A polynomial function R is a counting function of some fraction $\mathcal{F}$ with replicates up to r if, and only if, $R(R-1)\cdots(R-r) = 0$ on $\mathcal{D}$. In particular a function F is an indicator function if, and only if, $F^2 - F = 0$ on $\mathcal{D}$.

If F is the indicator function of the fraction $\mathcal{F}$, $F - 1 = 0$ is a set of generating equations of the same fraction.

As the counting function is real valued, we obtain $\overline{b_\alpha} = b_{[-\alpha]}$.

If f is a response on $\mathcal{D}$ then its *mean value on* $\mathcal{F}$, denoted by $E_\mathcal{F}(f)$, is:

$$E_\mathcal{F}(f) = \frac{1}{\#\mathcal{F}} \sum_{\zeta \in \mathcal{F}} f(\zeta) = \frac{\#\mathcal{D}}{\#\mathcal{F}} \; E_\mathcal{D}(R \; f)$$

where $\#\mathcal{F}$ is the total number of treatment combinations of the fraction, $\#\mathcal{F} = \sum_{\zeta \in \mathcal{D}} R(\zeta)$.

Proposition 1 *(1) The coefficients b_α of the the counting function of a fraction $\mathcal{F}$ are:*

$$b_\alpha = \frac{1}{\#\mathcal{D}} \sum_{\zeta \in \mathcal{F}} \overline{X^\alpha(\zeta)} \; ;$$

in particular, b_0 is the ratio between the number of points of the fraction and those of the design.

(2) In a single replicate fraction, the coefficients b_α of the indicator function are related according to:

$$b_\alpha = \sum_{\beta \in L} b_\beta \; b_{[\alpha - \beta]} \; .$$

(3) *If $\mathcal{F}$ and $\mathcal{F}'$ are complementary fractions without replications and b_α and b'_α are the coefficients of the respective indicator functions, $b_0 = 1 - b'_0$ and $b_\alpha = -b'_\alpha$.*

PROOF. Item (1) follows from :

$$\sum_{\zeta\in\mathcal{F}} \overline{X^\alpha(\zeta)} \sum_{\zeta\in\mathcal{D}} R\ \overline{X^\alpha(\zeta)} = \sum_{\zeta\in\mathcal{D}} \sum_{\beta\in L} b_\beta X^\beta(\zeta)\overline{X^\alpha(\zeta)} \sum_{\zeta\in\mathcal{D}} b_\alpha = \#\mathcal{D}\ b_\alpha\ .$$

Item (2) follows from relation $F = F^2$. In fact:

$$\begin{aligned}\sum_\alpha b_\alpha X^\alpha &= \sum_\beta b_\beta X^\beta \sum_\gamma b_\gamma X^\gamma = \sum_{\beta,\gamma} b_\beta b_\gamma X^{[\beta+\gamma]} = \\ &= \sum_\alpha \sum_{[\beta+\gamma]=\alpha} b_\beta b_\gamma X^\alpha = \sum_\alpha \sum_\beta b_\beta\ b_{[\alpha-\beta]} X^\alpha\ .\end{aligned}$$

Item (3) follows from $F' = 1 - F$. □

4.3 Orthogonal responses on a fraction

In this section, we discuss the general case of fractions $\mathcal{F}$ with or without replicates. As in the full design case, we say that a response f is *centered* on a fraction $\mathcal{F}$ if $E_\mathcal{F}(f) = E_\mathcal{D}(R\ f) = 0$ and we say that two responses f and g are *orthogonal* on $\mathcal{F}$ if $E_\mathcal{F}(f\ \overline{g}) = E_\mathcal{D}(R\ f\ \overline{g}) = 0$, i.e. the response $f\ \overline{g}$ is centered.

It should be noticed that the term "orthogonal" refers to vector orthogonality with respect to a given Hermitian product. The standard practise in orthogonal array literature, however, is to define an array as orthogonal when all the level combinations appear equally often in relevant subsets of columns, e.g. (Hedayat et al., 1999, Def. 1.1). Vector orthogonality is affected by the coding of the levels, while the definition of orthogonal array is purely combinatorial. A characterization of orthogonal arrays can be based on vector orthogonality of special responses. This section and the next one are devoted to discussing how the choice of complex coding makes such a characterization as straightforward as in the classical two-level case with coding -1,+1 .

Proposition 2 *Let $R = \sum_{\alpha\in L} b_\alpha X^\alpha$ be the counting function of a fraction $\mathcal{F}$.*

(1) *The term X^α is centered on $\mathcal{F}$ if, and only if, $b_\alpha = b_{[-\alpha]} = 0$.*
(2) *The terms X^α and X^β are orthogonal on $\mathcal{F}$ if, and only if, $b_{[\alpha-\beta]} = 0$;*
(3) *If X^α is centered then, for each β and γ such that $\alpha = [\beta - \gamma]$ or $\alpha = [\gamma - \beta]$, X^β is orthogonal to X^γ.*
(4) *A fraction $\mathcal{F}$ is self-conjugate, that is, $R(\zeta) = R(\overline{\zeta})$ for any $\zeta \in \mathcal{D}$, if, and only if, the coefficients b_α are real for all $\alpha \in L$.*

PROOF. The first three Items follow easily from Proposition 1.

For the Item (4), we obtain:

$$R(\zeta) = \sum_{\alpha \in L} b_\alpha X^\alpha(\zeta) = \sum_{\alpha \in L} b_{[-\alpha]} X^{[-\alpha]}(\zeta) = \sum_{\alpha \in L} \overline{b_\alpha} X^{[-\alpha]}(\zeta)$$
$$R(\overline{\zeta}) = \sum_{\alpha \in L} b_\alpha X^\alpha(\overline{\zeta}) = \sum_{\alpha \in L} b_\alpha X^{[-\alpha]}(\zeta) \ .$$

Therefore $R(\zeta) = R(\overline{\zeta})$ if, and only if, $b_\alpha = \overline{b_\alpha}$. It should be noticed that the same applies to all real valued responses. □

Interest in self-conjugate fractions concerns the existence of a real valued linear basis of the response space, as explained in (Kobilinsky, 1990, Prop. 3.1). It follows that it is possible to fit a real linear model on such a fraction, even though the levels have complex coding.

An important property of the centered responses follows from the structure of the roots of the unity as a cyclical group. This connects the combinatorial properties to the coefficients b_α's through the following two basic properties which hold true for the full design $\mathcal{D}$.

P-1 Let X_i be a simple term with level set Ω_n. Let us define $s = n/\mathrm{gcd}(r,n)$ and let Ω_s be the set of the s-th roots of the unity. The term X_i^r takes all the values of Ω_s equally often.

P-2 Let $X^\alpha = X_{j_1}^{\alpha_{j_1}} \cdots X_{j_k}^{\alpha_{j_k}}$ be an interaction term of order k where $X_{j_i}^{\alpha_{j_i}}$ takes values in $\Omega_{s_{j_i}}$. Let us define $s = \mathrm{lcm}\{s_{j_1}, \ldots, s_{j_k}\}$. The term X^α takes values in Ω_s equally often.

Let X^α be a term with level set Ω_s on the design $\mathcal{D}$. Let r_k be the number of times X^α takes the value ω_k on $\mathcal{F}$, $k = 0, \ldots, s-1$. The polynomial $P(\zeta)$ is associated to the sequence $(r_k)_{k=0,\ldots,s-1}$ so that:

$$P(\zeta) = \sum_{k=0}^{s-1} r_k \zeta^k \qquad \text{with } \zeta \in \mathbb{C} \ .$$

It should be noticed that

$$E_{\mathcal{F}}(X^\alpha) = \frac{1}{\#\mathcal{F}} \sum_{k=0}^{s-1} r_k \omega_k = \frac{1}{\#\mathcal{F}} P(\omega_1)$$

See Lang (1965) and the Appendix for a review of the properties of such a polynomial P.

Proposition 3 *Let X^α be a term with level set Ω_s on full design $\mathcal{D}$.*

(1) *X^α is centered on $\mathcal{F}$ if, and only if,*

$$P(\zeta) = \Phi_s(\zeta)\Psi(\zeta)$$

where Φ_s is the cyclotomic polynomial of the s-roots of the unity and Ψ is a suitable polynomial with integer coefficients.

(2) *Let s be prime. Therefore, the term X^α is centered on $\mathcal{F}$ if, and only if, its s levels appear equally often:*

$$r_0 = \cdots = r_{s-1} = r$$

(3) *Let $s = p_1^{h_1} \cdot \cdots \cdot p_d^{h_d}$, with p_i prime, for $i = 1, \ldots, d$. The term X^α is centered on $\mathcal{F}$ if, and only if, the following equivalent conditions are satisfied.*

(a) *The remainder*

$$H(\zeta) = P(\zeta) \mod \Phi_s(\zeta) \ ,$$

whose coefficients are integer combination of r_k, $k = 0, \ldots, s-1$, is identically zero.

(b) *The polynomial of degree s*

$$\tilde{P}(\zeta) = P(\zeta) \prod_{d|s} \Phi_d(\zeta) \mod (\zeta^s - 1) \ ,$$

whose coefficients are integer combination of the replicates r_k, $k = 0, \ldots, s-1$, is identically zero. The indices of the product are the d's that divide s.

(4) *Let g_i be an indicator of a subgroup or of a lateral of a subgroup of Ω_s; i.e.: $g_i = (g_{i1}, \ldots, g_{ij}, \ldots, g_{is})$, $g_{ij} \in \{0, 1\}$, such that $\{k \ : \ g_{ik} = 1\}$ is a subgroup or a lateral of a subgroup of Ω_s.*

If the vector of level replicates $(r_0, r_1, \ldots, r_{s-1})$ is a combination with positive weights of g_i:

$$(r_0, r_1, \ldots, r_{s-1}) = \sum a_i \ g_i \qquad \text{with } a_i \in \mathbb{N}$$

X^α is centered.

Proof. (1) As $\omega_k = \omega_1^k$, the assumption $\sum_k r_k \omega_k = 0$ is equivalent to $P(\omega_1) = 0$. From Section 8, Items 4 and 5, we know that this implies that $P(\omega) = 0$ for all primitive s-roots of the unity, that is, $P(\zeta)$ is divisible by the cyclotomic polynomial Φ_s.

(2) If s is a prime number, the cyclotomic polynomial is $\Phi_s(\zeta) = \sum_{k=0}^{s-1} \zeta^k$. The polynomial $P(\zeta)$ is divided by the cyclotomic polynomial, and $P(\zeta)$ and $\Phi_s(\zeta)$ have the same degree, therefore $r_{s-1} > 0$ and $P(\zeta) = r_{s-1}\Phi(\zeta)$, so that $r_0 = \cdots = r_{s-1}$.

(3) The divisibility shown in Item 1 is equivalent to the condition of null remainder. Such a remainder is easily computed as the reduction of the

polynomial $P(\zeta) \mod \Phi_s(\zeta)$. According to the same condition and Equation (8), we obtain that $\tilde{P}(\zeta)$ is divisible by $\zeta^s - 1$, therefore it also equals $0 \mod \zeta^s - 1$.

(4) If Ω_p is a prime subgroup of Ω_s, then $\sum_{\omega \in \Omega_p} \omega = 0$. Now let us assume that the replicates on a primitive subgroup Ω_{p_i} are 1. Therefore $\sum_{\omega \in \Omega_{p_i}} \omega = 0$ according the equation in Item (3). The same occurs in the case of the laterals and the sum of such cases.

□

Example
Let us consider the case $s = 6$. This situation occurs in the case of mixed-level factorial designs with both three-level factors and two-level factors. In this case, the cyclotomic polynomial is $\Phi_6(\zeta) = \zeta^2 - \zeta + 1$ whose roots are ω_1 and ω_5. The remainder is

$$\begin{aligned} H(\zeta) &= \sum_{k=0}^{5} r_k \zeta^k \mod \Phi_6(\zeta) \\ &= r_0 + r_1\zeta + r_2\zeta^2 + r_3\zeta^3 + r_4\zeta^4 + r_5\zeta^5 \mod \left(\zeta^2 - \zeta + 1\right) \\ &= (r_1 + r_2 - r_4 - r_5)\zeta + (r_0 - r_2 - r_3 + r_5) \end{aligned}$$

The condition $H(\zeta) = 0$ implies the following relations concerning the numbers of replicates: $r_0 + r_1 = r_3 + r_4$, $r_1 + r_2 = r_4 + r_5$, $r_2 + r_3 = r_0 + r_5$, where the first one follows by summing of the second with the third one. Equivalently:

$$r_0 - r_3 = r_4 - r_1 = r_2 - r_5 \ . \tag{4}$$

Let us consider the replicates corresponding to the sub-group $\{\omega_0, \omega_2, \omega_4\}$ and denote the $\min\{r_0, r_2, r_4\}$ by m_1. We then consider the replicates corresponding to the lateral of the previous sub-group $\{\omega_1, \omega_3, \omega_5\}$ and we denote by m_2 the $\min\{r_1, r_3, r_5\}$. We consider the new vector of the replicates:

$$\begin{aligned} r' &= (r'_0, r'_1, r'_2, r'_3, r'_4, r'_5) \\ &= (r_0 - m_1, r_1 - m_2, r_2 - m_1, r_3 - m_2, r_4 - m_1, r_5 - m_2) \\ &= r - m_1(1, 0, 1, 0, 1, 0) - m_2(0, 1, 0, 1, 0, 1) \end{aligned}$$

The vector r' satisfies Equation (4).

As at least r'_0, r'_2 or r'_4 is zero, the common value in Equation (4) is zero or negative. Moreover, as at least r'_1, r'_3 or r'_5 is zero, the common value in Equations (4) is zero or positive. The common value is therefore zero and $r'_0 = r'_3$, $r'_1 = r'_4$, $r'_2 = r'_5$ and

$$r' = r'_0(1, 0, 0, 1, 0, 0) + r'_1(0, 1, 0, 0, 1, 0) + r'_2(0, 0, 1, 0, 0, 1)$$

A term is therefore centered if the vector of the replicates is of the form:

$$\begin{aligned}(r_0, \ldots, r_5) = a_1(1,0,0,1,0,0) + a_2(0,1,0,0,1,0) \\ + a_3(0,0,1,0,0,1) + a_4(1,0,1,0,1,0) + a_5(0,1,0,1,0,1)\end{aligned}$$

with a_i non negative integers. There are 5 generating integer vectors of the replicate vector.

It should be noticed that if the number of levels of X^α is not prime, $E_{\mathcal{F}}(X^\alpha) = 0$ does not imply $E_{\mathcal{F}}(X^{r\alpha}) = 0$. In the previous six-level example, if X^α is centered, the vector of replicates of $X^{2\alpha}$ is of the form $(2a_1 + a_4 + a_5, 0, 2a_2 + a_4 + a_5, 0, 2a_3 + a_4 + a_5, 0)$ and $X^{2\alpha}$ is centered only if $a_1 = a_2 = a_3$.

5 Orthogonal arrays

In this sectionn we discuss the relations between the coefficients b_α, $\alpha \in L$, of the counting function and the property of being an orthogonal array. Let

$$\mathrm{OA}(n, s_1^{p_1}, \ldots, s_m^{p_m}, t)$$

be a mixed-level orthogonal array with n rows and m columns, $m = p_1 + \cdots + p_m$, in which p_1 columns have s_1 symbols, ..., p_k columns have s_m symbols, and with strength t, as defined e.g. in (Wu and Hamada, 2000, p. 260). Strength t means that, for any t columns of the matrix design, all possible combinations of symbols appear equally often in the matrix.

Definition 2 *Let I be a non-empty subset of $\{1, \ldots, m\}$, and let J be its complement set, $J = I^c$. Let $\mathcal{D}_I$ and $\mathcal{D}_J$ be the corresponding full factorial designs over the I-factors and the J-factors, so that $\mathcal{D} = \mathcal{D}_I \times \mathcal{D}_J$. Let $\mathcal{F}$ be a fraction of $\mathcal{D}$ and let $\mathcal{F}_I$ and $\mathcal{F}_J$ be its projections.*

(1) A fraction $\mathcal{F}$ factorially projects *on the I-factors if $\mathcal{F}_I = s\ \mathcal{D}_I$, that is, the projection is a full factorial design where each point appears s times.*
(2) A fraction $\mathcal{F}$ is a mixed orthogonal array *of strength t if it factorially projects on any I-factors with $\#I = t$.*

Using the notations of Definition 2, for each point ζ of a complex coded fraction $\mathcal{F}$, we consider the decomposition $\zeta = (\zeta_I, \zeta_J)$ and we denote the counting function restricted to the I-factors of a fraction by R_I, i.e. $R_I(\zeta_I)$ is the number of points in $\mathcal{F}$ whose projection on the I-factors is ζ_I.
We denote the sub-set of the exponents restricted to the I-factors by L_I and an element of L_I by α_I:

$$L_I = \{\alpha_I = (\alpha_1, \ldots, \alpha_j, \ldots, \alpha_m), \qquad \alpha_j = 0 \text{ if } j \in J\} \ .$$

Therefore, for each $\alpha \in L$ and $\zeta \in \mathcal{D}$: $\alpha = \alpha_I + \alpha_J$ and $X^\alpha(\zeta) = X^{\alpha_I}(\zeta_I) X^{\alpha_J}(\zeta_J)$. We denote the cardinalities of the projected designs by $\#\mathcal{D}_I$ and $\#\mathcal{D}_J$.

Proposition 4

(1) *The number of replicates of the points of a fraction projected onto the I-factors is:*

$$R_I(\zeta_I) = \#\mathcal{D}_J \sum_{\alpha_I} b_{\alpha_I} \; X^{\alpha_I}(\zeta_I) \; .$$

(2) *A fraction* factorially projects onto the I-factors *if, and only if,*

$$R_I(\zeta_I) = \#\mathcal{D}_J \; b_0 = \frac{\#\mathcal{F}}{\#\mathcal{D}_I} \quad \textit{for all} \;\; \zeta_I \; .$$

This is equivalent to all the coefficients of the counting function involving only the I-factors being 0:

$$b_{\alpha_I} = 0 \quad \textit{with} \;\; \alpha_I \in L_I, \;\; \alpha_I \neq (0, 0, \ldots, 0) \; .$$

In such a case, the levels of a factor X_i, $i \in I$, appear equally often in $\mathcal{F}$.

(3) *If there exists a subset J of $\{1, \ldots, m\}$ such that the J-factors appear in all the non null elements of the counting function, the fraction* factorially projects onto the I-factors, *with $I = J^c$.*

(4) *A fraction is an* orthogonal array of strength t *if, and only if, all the coefficients of the counting function up to the order t are zero:*

$$b_\alpha = 0 \quad \forall \; \alpha \;\; \textit{of order up to} \; t, \;\; \alpha \neq (0, 0, \ldots, 0) \; .$$

Proof. (1) We obtain:

$$\begin{aligned} R_I(\zeta_I) &= \sum_{\zeta_J \in \mathcal{D}_J} R(\zeta_I, \zeta_J) = \sum_{\zeta_J \in \mathcal{D}_J} \sum_{\alpha \in L} b_\alpha \; X^\alpha(\zeta_I, \zeta_J) \\ &= \sum_{\zeta_J \in \mathcal{D}_J} \sum_{\alpha \in L} b_\alpha \; X^{\alpha_I}(\zeta_I) X^{\alpha_J}(\zeta_J) \\ &= \sum_{\alpha_I \in L_I} b_{\alpha_I} X^{\alpha_I}(\zeta_I) \sum_{\zeta_J \in \mathcal{D}_J} X^{\alpha_J}(\zeta) + \sum_{\alpha \notin L_I} b_\alpha X^{\alpha_I}(\zeta) \sum_{\zeta_J \in \mathcal{D}_J} X^{\alpha_J}(\zeta) \; . \end{aligned}$$

The thesis follows from $\sum_{\zeta_J \in \mathcal{D}_J} X^{\alpha_J}(\zeta_J) = 0$ if $\alpha_J \neq (0, 0, \ldots, 0)$ and $\sum_{\zeta_J \in \mathcal{D}_J} X^{\alpha_J}(\zeta_J) = \#\mathcal{D}_J$ if $\alpha_J = (0, 0, \ldots, 0)$.

(2) The number of replicates of the points of the fraction projected onto the I-factors, $R_I(\zeta_I) = \#\mathcal{D}_J \sum_{\alpha_I} b_{\alpha_I} \; X^{\alpha_I}(\zeta_I)$, is a polynomial and it is a constant if all the coefficients b_{α_I}, with $\alpha_I \neq (0, 0, \ldots, 0)$, are zero.

(3) This condition implies that the b_{α_I}'s are zero, if $\alpha_I \neq (0, 0, \ldots, 0)$, and the thesis follows from the previous item.

(4) This item follows from the previous items and the definition.

□

Remarks

(1) If a fraction factorially projects onto the I-factors, its cardinality must be equal to, or a multiple of the cardinality of $\mathcal{D}_I$.
(2) If the number of levels of each factors is a prime, the condition $b_{\alpha_i} = 0$ for each $i \in I$ and $0 < \alpha_i \leq n_i - 1$ in Items (2) and (3) of the previous Proposition, simplify to $E_{\mathcal{F}}(X_i) = 0$, according to Item (2) of Proposition 3.

6 Regular fractions: a partial generalization to mixed-level design

A short review of the theory of regular fractions is here made from the view point of the present paper. Various definitions of regular fraction appear in literature, e.g. in the books by (Raktoe et al., 1981, p. 123), (Collombier, 1996, p. 125), (Kobilinsky, 1997, p. 70), (Dey and Mukerjee, 1999, p. 164), (Wu and Hamada, 2000, p. 305). To our knowledge, all the definitions are known to be equivalent if all the factors have the same prime number of levels, $n = p$. The definition based on Galois Field computations is given for $n = p^s$ power of a prime number. All definitions assume symmetric factorial design, i.e. all the factors have the same number of levels.

Regular fraction designs are usually considered for qualitative factors, where the coding of the levels is arbitrary. The integer coding, the GF(p^s) coding, and the roots of the unity coding, as introduced by Bailey (1982) and used extensively in this paper, can all be used. Each of those codings is associated to specific ways of characterizing a fraction, and even more important for us, to a specific basis for the responses. One of the possible definitions of a regular fraction refers to the property of non-existence of partial confounding of simple and interaction terms, and this property has to be associated to a specific basis, as explicitly pointed out in Wu and Hamada (2000).

In our approach, we use polynomial algebra with complex coefficients, the n-roots of the unit coding, and the idea of indicator polynomial function, and we make no assumption about the number of levels. In the specific coding we use, the indicator polynomial is actually a linear combination of monomial terms which are centered and orthogonal on the full factorial design. We refer to such a basis to state the no-partial confounding property.

The definition of the regular fraction is hereafter generalized in the symmetric case with a prime number of levels. The new setting includes asymmetric design with any number of levels. Proposition 5 below does not include regular fractions defined in GF(p^s). A full discussion of this point shall be published elsewhere.

We consider a fraction *without replicates*. Let $n = \mathrm{lcm}\{n_1, \ldots, n_m\}$. It should be recalled that Ω_n is the set of the n-th roots of the unity, $\Omega_n = \{\omega_0, \ldots, \omega_{n-1}\}$. Let $\mathcal{L}$ be a subset of exponents, $\mathcal{L} \subset L = \mathbb{Z}_1 \times \cdots \times \mathbb{Z}_m$, containing $(0, \ldots, 0)$ and let l be its cardinality ($l > 0$). Let e be a map from $\mathcal{L}$ to Ω_n, $e : \mathcal{L} \to \Omega_n$.

Definition 3 *A fraction $\mathcal{F}$ is* regular *if*

(1) $\mathcal{L}$ is a sub-group of L,
(2) e is a group homomorphism, $e([\alpha + \beta]) = e(\alpha)\ e(\beta)$ for each $\alpha, \beta \in \mathcal{L}$,
(3) the equations

$$X^\alpha = e(\alpha) \quad , \qquad \alpha \in \mathcal{L} \tag{5}$$

define the fraction $\mathcal{F}$, i.e. they are a set of generating equations, according to Section 4.1. Equations (5) are also called the defining *equations of $\mathcal{F}$.*

If $\mathcal{H}$ is a minimal generator of the group $\mathcal{L}$, Equations $X^\alpha = e(\alpha)$, $\alpha \in \mathcal{H} \subset \mathcal{L}$, are called a minimal set of generating *equations.*

It should be noticed that our situation is general because the values $e(\alpha)$ can be different from 1. From items (1) and (2) it follows that a necessary condition is that the $e(\alpha)$'s must belong to the subgroup spanned by the values of X^α. For example, for $n_1 = n_2 = n = 6$, an equation such as $X_1^3 X_2^3 = \omega_2$ cannot be a defining equation.

For example, in the fraction of Section 4.1, we have: $\mathcal{H} = \{(1,1,2,0), (1,2,0,1)\}$ and $e(1,1,2,0) = e(1,2,0,1) = \omega_0 = 1$. The set $\mathcal{L}$ is: $\{(0,0,0,0), (0,1,1,2),$ $(0,2,2,1), (1,1,2,0), (2,2,1,0), (1,2,0,1), (2,1,0,2), (1,0,1,1), (2,0,2,2)\}$.

Proposition 5 *Let $\mathcal{F}$ be a fraction. The following statements are equivalent:*

(1) Fraction $\mathcal{F}$ is regular according to Definition 3.
(2) The indicator function of the fraction has the form

$$F(\zeta) = \frac{1}{l} \sum_{\alpha \in \mathcal{L}} \overline{e(\alpha)}\ X^\alpha(\zeta) \qquad \zeta \in \mathcal{D}$$

where $\mathcal{L}$ is a given subset of L and $e : \mathcal{L} \to \Omega_n$ is a given mapping.
(3) For each $\alpha, \beta \in L$, the parametric functions represented on $\mathcal{F}$ by the terms X^α and X^β are either orthogonal or totally confounded.

Proof. First we prove the equivalence between (1) and (2).

(1) $\Rightarrow$ (2).
Let $\mathcal{F}$ be a regular fraction and let $X^\alpha = e(\alpha)$ be its defining equations with $\alpha \in \mathcal{L}$, $\mathcal{L}$ a sub-group of L and e a homomorphism.

If, and only if, $\zeta \in \mathcal{F}$:

$$
\begin{aligned}
0 &= \sum_{\alpha \in \mathcal{L}} |X^\alpha(\zeta) - e(\alpha)|^2 = \sum_{\alpha \in \mathcal{L}} \left(X^\alpha(\zeta) - e(\alpha)\right) \overline{\left(X^\alpha(\zeta) - e(\alpha)\right)} \\
&= \sum_{\alpha \in \mathcal{L}} \left(X^\alpha(\zeta)\overline{X^\alpha(\zeta)} + e(\alpha)\overline{e(\alpha)} - e(\alpha)\overline{X^\alpha(\zeta)} - \overline{e(\alpha)}X^\alpha(\zeta) \right) \\
&= 2\ l - \sum_{\alpha \in \mathcal{L}} \overline{\overline{e(\alpha)}X^\alpha(\zeta)} - \sum_{\alpha \in \mathcal{L}} \overline{e(\alpha)}\ X^\alpha(\zeta) = 2\left(l - \sum_{\alpha \in \mathcal{L}} \overline{e(\alpha)}\ X^\alpha(\zeta) \right)
\end{aligned}
$$

therefore

$$
\frac{1}{l} \sum_{\alpha \in \mathcal{L}} \overline{e(\alpha)}\ X^\alpha(\zeta) - 1 = 0 \quad \text{if, and only if,} \quad \zeta \in \mathcal{F}\ .
$$

The function $F = \frac{1}{l} \sum_{\alpha \in \mathcal{L}} \overline{e(\alpha)}\ X^\alpha$ is an indicator function, as it can be shown that $F = F^2$ on $\mathcal{D}$. In fact, $\mathcal{L}$ is a sub-group of L and e is a homomorphism; therefore:

$$
\begin{aligned}
F^2 &= \frac{1}{l^2} \sum_{\alpha \in \mathcal{L}} \sum_{\beta \in \mathcal{L}} \overline{e(\alpha)\ e(\beta)}\ X^{[\alpha+\beta]} = \frac{1}{l^2} \sum_{\alpha \in \mathcal{L}} \sum_{\beta \in \mathcal{L}} \overline{e([\alpha+\beta])}\ X^{[\alpha+\beta]} = \\
&= \frac{1}{l^2} \sum_{\gamma \in \mathcal{L}} l\ \overline{e(\gamma)}\ X^\gamma = F
\end{aligned}
.
$$

It follows that F is the indicator function of $\mathcal{F}$, and $b_\alpha = \frac{\overline{e(\alpha)}}{l}$, for all $\alpha \in \mathcal{L}$.

$(2) \Rightarrow (1)$.
It should be noticed that an indicator function is real valued, therefore $\overline{F} = F$.

$$
\frac{1}{l} \sum_{\alpha \in \mathcal{L}} |X^\alpha(\zeta) - e(\alpha)|^2 = 2 - \overline{F(\zeta)} - F(\zeta) = 2 - 2F(\zeta) = \begin{cases} 0 & \text{on } \mathcal{F} \\ 2 & \text{on } \mathcal{D} \setminus \mathcal{F}\ . \end{cases}
$$

Equations $X^\alpha = e(\alpha)$, with $\alpha \in \mathcal{L}$, define the fraction $\mathcal{F}$ as the generating equations of a regular fraction. It is easy to see that $\mathcal{L}$ is a group. In fact, if $\gamma = [\alpha + \beta] \notin \mathcal{L}$, there exists one ζ such that $X^\gamma(\zeta) = X^\alpha(\zeta)X^\beta(\zeta) = e(\alpha)e(\beta) \subset \Omega_n$ and the value $e(\alpha)e(\beta)$ only depends on γ. By repeating the previous proof, the uniqueness of the polynomial representation of the indicator function leads a contradiction.

Now we prove the equivalence between (2) and (3).

$(2) \Rightarrow (3)$
The non-zero coefficients of the indicator function are of the form $b_\alpha = \overline{e(\alpha)}/l$.

We consider two terms X^α and X^β with $\alpha, \beta \in L$. If $[\alpha - \beta] \notin \mathcal{L}$ then X^α and X^β are orthogonal on $\mathcal{F}$ as the coefficient $b_{[\alpha-\beta]}$ of the indicator function equals

0. If $[\alpha - \beta] \in \mathcal{L}$ then X^α and X^β are confounded because $X^{[\alpha-\beta]} = e([\alpha-\beta])$; therefore $X^\alpha = e([\alpha - \beta])\ X^\beta$.

(3) $\Rightarrow$ (2).
Let $\mathcal{L}$ be the set of exponents of the terms confounded with a constant:

$$\mathcal{L} = \{\alpha \in L : X^\alpha = \text{constant} = e(\alpha), \quad e(\alpha) \in \Omega_n\} \ .$$

For each $\alpha \in \mathcal{L}$, $b_\alpha = \overline{e(\alpha)}\ b_0$. For each $\alpha \notin \mathcal{L}$, because of the assumption, X^α is orthogonal to X^0, therefore $b_\alpha = 0$. $\square$

Corollary 1 *Let $\mathcal{F}$ be a regular fraction with $X^\alpha = 1$ for all the defining equations. $\mathcal{F}$ is therefore self-conjugate and a multiplicative subgroup of $\mathcal{D}$.*

Proof. It follows from Prop. 2 Item 4. $\square$

The following proposition extends a result presented in Fontana et al. (2000) for the two level case.

Proposition 6 *Let $\mathcal{F}$ be a fraction with indicator function F. We denote the set of the exponents α such that $\frac{b_\alpha}{b_0} = \overline{e(\alpha)} \in \Omega_n$ by $\mathcal{L}$. The indicator function can be written as*

$$F(\zeta) = b_0 \sum_{\alpha \in \mathcal{L}} \overline{e(\alpha)}\ X^\alpha(\zeta) + \sum_{\beta \in \mathcal{K}} b_\beta\ X^\beta(\zeta) \qquad \zeta \in \mathcal{F}\ ,\ \mathcal{L} \cap \mathcal{K} = \emptyset\ .$$

It follows that $\mathcal{L}$ is a subgroup and the equations $X^\alpha = e(\alpha)$, with $\alpha \in \mathcal{L}$, are the defining equations of the smallest regular fraction $\mathcal{F}_r$ containing $\mathcal{F}$ restricted to the factors involved in the $\mathcal{L}$-exponents.

Proof. The coefficients b_α, $\alpha \in \mathcal{L}$, of the indicator function F are of the form $b_0 \overline{e(\alpha)}$. Therefore, from the extremality of n-th roots of the unity, $X^\alpha(\zeta) = e(\alpha)$ if $\zeta \in \mathcal{F}$ and $X^\alpha(\zeta)F(\zeta) = e(\alpha)F(\zeta)$ for each $\zeta \in \mathcal{D}$ and $\mathcal{L}$ is a group.

We denote the indicator function of $\mathcal{F}_r$ by F_r. For each $\zeta \in \mathcal{D}$ we have:

$$\begin{aligned} F(\zeta)F_r(\zeta) &= \frac{1}{l} F(\zeta) \sum_{\alpha \in \mathcal{L}} \overline{e(\alpha)}\ X^\alpha(\zeta) = \frac{1}{l} \sum_{\alpha \in \mathcal{L}} \overline{e(\alpha)}\ X^\alpha(\zeta)\ F(\zeta) = \\ &= \frac{1}{l} \sum_{\alpha \in \mathcal{L}} \overline{e(\alpha)}\ e(\alpha)\ F(\zeta) = \frac{1}{l}\ l\ F(\zeta)\ . \end{aligned}$$

The relation $F(\zeta)F_r(\zeta) = F(\zeta)$ implies $\mathcal{F} \subseteq \mathcal{F}_r$. The fraction $\mathcal{F}_r$ is minimal because we have collected all the terms confounded with a constant. $\square$

Remark

Given generating equations $X^{\alpha_1} = 1, \ldots, X^{\alpha_h} = 1$, with $\{\alpha_1, \ldots, \alpha_h\} = \mathcal{H} \subset \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_m}$, the corresponding fraction $\mathcal{F}$ is a subgroup of $\Omega_{n_1} \times \cdots \times \Omega_{n_m}$. If the same fraction is represented in the additive notation, such a set of treatment combinations is the principal block of a single replicate generalized cyclic design, see John and Dean (1975), Dean and John (1975) and Lewis (1979). A complex vector of the form

$$\left(e^{\mathrm{i}\, 2\pi \frac{k_1}{n_1}}, e^{\mathrm{i}\, 2\pi \frac{k_2}{n_2}}, \ldots, e^{\mathrm{i}\, 2\pi \frac{k_m}{n_m}}\right) \qquad \text{with } \ 0 \leq k_i < n_i,\ i = 1, \ldots, m \ ,$$

is in fact a solution of the generating equations if, and only if,

$$\begin{cases} \sum_{j=1}^{m} \alpha_{ij}\ \gamma_j\ k_j = 0 \mod s \\ 0 \leq k_j < n_j \end{cases} \tag{6}$$

with $s = \operatorname{lcm}\{n_1, \ldots, n_m\}$ and $\gamma_j = \frac{s}{n_j}$.

A set of generators can be computed from Equation (6). It should be noticed that the following equivalent integer linear programming problem does not involve computation mod s, see Schrijver (2002)

$$\begin{cases} \sum_{j=1}^{m} \alpha_{ij}\ \gamma_j\ k_j - sq = 0 \\ 0 \leq k_j < n_j\ ,\ q \geq 0 \end{cases} \tag{7}$$

In Lewis (1982), the monomial part of our defining equations is called defining contrast, according to Bailey et al. (1977). The paper contains extensive tables of the generator subgroups of the treatment combinations and the corresponding defining contrasts.

Viceversa, given a set of generators of the treatment combinations,

$$\{b_1, \ldots, b_r \mid b_i = (b_{i1}, \ldots, b_{im})\}\ ,$$

Equation (6) with indeterminates α_i

$$\begin{cases} \sum_{j=1}^{m} \gamma_j\ b_{ij}\ \alpha_j = 0 \mod s \\ 0 \leq \alpha_j < n_j \end{cases}$$

produces generating equations for the fraction.

7 Examples

A regular fractions with $\mathbf{n = 3}$.
Let us consider the classical 3^{4-2} fraction of Section 4.1. Its indicator function is:

$$F = \frac{1}{9}\Big(1 + X_2X_3X_4 + X_2^2X_3^2X_4^2 + X_1X_2X_3^2 + X_1^2X_2^2X_3 \\ + X_1X_2^2X_4 + X_1^2X_2X_4^2 + X_1X_3X_4^2 + X_1^2X_3^2X_4\Big) \ .$$

We can observe that the coefficients are all equal to $\frac{1}{9}$. The minimum order of interactions that appear in the indicator function is 3, therefore the fraction is an orthogonal array of strength 2. All the defining equations are of the form $X^\alpha = 1$, therefore the fraction is self-conjugate.

A regular fraction with $\mathbf{n = 6}$.
Let us consider a 6^3 design. From property [P-2] of Section 4.3, the terms X^α take values either in Ω_6 or in one of the two subgroups either $\{1, \omega_3\}$ or $\{1, \omega_2, \omega_4\}$.
Let $\mathcal{F}$ be a fraction whose generating equations are: $X_1^3X_2^3X_3^3 = \omega_3$ and $X_2^4X_2^4X_3^2 = \omega_2$. In this case we have: $\mathcal{H} = \{(3,3,3),(4,4,2)\}$ and $e(3,3,3) = \omega_3$, $e(4,4,2) = \omega_2$. The set $\mathcal{L}$ is: $\{(0,0,0),(3,3,3),(4,4,2),(2,2,4),(1,1,5),(5,1,1)\}$. The full factorial design has 216 points and the fraction has 36 points. The indicator function is:

$$F = \frac{1}{6}\Big(1 + \omega_3X_1^3X_2^3X_3^3 + \omega_4X_1^4X_2^4X_3^2 + \omega_2X_1^2X_2^2X_3^4 + \omega_1X_1X_2X_3^5 + \omega_5X_1^5X_2X_3\Big)$$

It should be noticed that this fraction is an $OA(36, 6^3, 2)$.

An $\mathbf{OA(18, 2^1 3^7, 2)}$.
We consider the fraction of a 2×3^7 design with 18 runs, taken from (Wu and Hamada, 2000, Table 7C.2) and recoded with complex levels. Here X_1 takes values in Ω_2, X_i, with $i = 2, \ldots, 8$, and their interactions take values in Ω_3, and the interactions involving X_1 take values in Ω_6.
All the 4374 X^α terms of the fraction have been computed in SAS using $\mathbb{Z}_2$, $\mathbb{Z}_3$ and $\mathbb{Z}_6$ arithmetic. The replicates of the values in the relevant $\mathbb{Z}_k$ have then been computed for each terms. We found:

(1) 3303 centered responses. These are characterized by Proposition 3. The replicates are of the type: $(9,9)$, $(6,6,6)$, $(3,3,3,3,3,3)$ and $(9,0,0,9,0,0)$. We have:
   (a) the two-level simple term and 1728 terms involving only the three-level factors (14 of order 1, 84 of order 2, 198 of order 3, 422 of order

4, 564 of order 5, 342 of order 6 and 104 of order 7);
   (b) 1574 terms involving both the two-level factor and the three-level factors (14 of order 2, 66 of order 3, 188 of order 4, 398 of order 5, 492 of order 6, 324 of order 7 and 92 of order 8).
(2) 9 terms with corresponding b_α coefficients equal to $b_0 = \frac{18}{2 \times 3^7} = 3^{-5}$;
(3) 1062 terms with corresponding coefficients different from zero and b_0: 450 terms involving only the three-level factors (80 of order 3, 138 of order 4, 108 of order 5, 100 of order 6 and 24 of order 7) and 612 terms involving both the two-level factor and the three-level factors (18 of order 3, 92 of order 4, 162 of order 5, 180 of order 6, 124 of order 7 and 36 of order 8).

Some statistical properties of the fraction are:

(1) Analyzing the centered responses we can observe that:
   (a) All the 15 simple terms are centered.
   All the 98 interactions of order 2 (84 involving only the three-level factors and 14 also involving the two-level factor) are centered. This implies that both the "linear" terms and the "quadratic" terms of the three-level factors are mutually orthogonal and they are orthogonal to the two-level factor.
   The fraction is a mixed orthogonal array of strength 2.
   (b) The fraction factorially projects onto the following factor subsets:

$$\begin{array}{l} \{X_1, X_2, X_3\},\ \{X_1, X_2, X_4\},\ \{X_1, X_2, X_5\},\ \{X_1, X_2, X_6\}, \\ \{X_1, X_3, X_6\},\ \{X_1, X_3, X_7\},\ \{X_1, X_4, X_5\},\ \{X_1, X_4, X_8\}, \\ \{X_1, X_5, X_8\},\ \{X_1, X_6, X_7\},\ \{X_1, X_6, X_8\}\ . \end{array}$$

   All the terms of order 1, 2 and 3 involving the same set of factors are in fact centered.
   (c) The minimal regular fraction containing our fraction restricted to the three-level factors has the following defining relations:

$$\begin{array}{l} X_2^2 X_4^2 X_5 = 1\ ,\ X_2 X_4 X_5^2 = 1\ , \\ X_2 X_3 X_4^2 X_6 X_7 X_8 = 1\ ,\ X_2^2 X_3^2 X_4 X_6^2 X_7^2 X_8^2 = 1\ , \\ X_2^2 X_3 X_5^2 X_6 X_7 X_8 = 1\ ,\ X_2 X_3^2 X_5 X_6^2 X_7^2 X_8^2 = 1\ , \\ X_3 X_4 X_5 X_6 X_7 X_8 = 1\ ,\ X_3^2 X_4^2 X_5^2 X_6^2 X_7^2 X_8^2 = 1\ . \end{array}$$

   (d) The non centered terms have levels in Ω_6 and in Ω_3.

Acknowledgments

We wish to thank many colleagues for their helpful and interesting comments, especially G.-F. Casnati, R. Notari, L. Robbiano, E. Riccomagno, H.P. Wynn and K.Q. Ye. Last but not least, we extensively used the comments and suggestions made by the anonymous referees of the previous versions. We regret we are unable to thank them by name.

8 Appendix: Algebra of the n-th roots of the unity.

We hereafter list some facts concerning the algebra of the complex n-th roots of the unity, for ease of reference.

(1) The conjugate of a n-th root of the unity equals its inverse: $\overline{\omega_k} = \omega_k^{-1} = \omega_{[-k]}$ for all $\omega_k \in \Omega_n$.

(2) If $\zeta \neq \omega_m$, we obtain: $\prod_{k=0\ k\neq m}^{n-1}(\zeta - \omega_k) = \frac{\zeta^n - 1}{\zeta - \omega_m} = \sum_{h=0}^{n-1} \omega_m^{n-h-1}\ \zeta^h$ where the last equality follows from algebraic computation. Therefore, for $\zeta = \omega_m$:

$$\prod_{k=0\ k\neq m}^{n-1} (\omega_m - \omega_k) = \sum_{h=0}^{n-1} \omega_m^{n-h-1}\ \omega_m^h = n\ \omega_m{}^{n-1} = n\ \overline{\omega_m}$$

and especially: $\prod_{k=1}^{n-1}(1 - \omega_k) = n$.

(3) We have: $\zeta^n - 1 = (\zeta - \omega_0) \cdots (\zeta - \omega_{n-1}) = \sum_{k=0}^{n-1}(-1)^{n-k} S_{n-k}(\omega_0, \ldots, \omega_{n-1})\, \zeta^k$ where $S_{n-k}(x_0, \ldots, x_{n-1})$ is the elementary symmetric polynomial of order $n-k$. We therefore obtain the following notable cases:
- $S_1(\omega_0, \ldots, \omega_{n-1}) = \sum_k \omega_k = 0$
- $S_2(\omega_0, \ldots, \omega_{n-1}) = \sum_{\ell < m} \omega_\ell\ \omega_m = 0$
- $S_n(\omega_0, \ldots, \omega_{n-1}) = \prod_k \omega_k = (-1)^{n+1}$

where the indices of the sums and products are from 0 to $n-1$.

(4) Let ω be a primitive n-th root of the unity, that is, a generator of Ω_n as a cyclic group: $\{1, \omega, \omega^2, \ldots, \omega^{n-1}\} = \Omega_n$.
The root $\omega_p \in \Omega_n$ is primitive if p is relatively prime with n. In particular, ω_1 is a primitive root and, for $\omega_k \in \Omega_n$, we obtain: $\omega_k = (\omega_1)^k$. If n is a prime number, all the roots of the unity, except 1, are primitive roots. The number of the primitive n-th roots of the unity is denoted by $\phi(n)$.

(5) Given an algebraic number x, the unique irreducible monic polynomial of the smallest degree with rational coefficients P such that $P(x) = 0$ and whose leading coefficient is 1, is called the minimal polynomial of x. The minimal polynomial of a primitive n-th root of the unity is called the cyclotomic polynomial $\Phi_n(\zeta)$ and its degree is $\phi(n)$:

$$\Phi_n(\zeta) = \prod_p (\zeta - \omega_p)\ , \quad \zeta \in \mathbb{C},\ \omega_p \in \Omega_n \text{ primitive } n\text{-th root of the unity.}$$

If n is prime, the minimal polynomial of a primitive n-th root of the unity is $\Phi_n(\zeta) = \zeta^{n-1} + \zeta^{n-2} + \cdots + 1$. Moreover:

$$\zeta^n - 1 = \Phi_n(\zeta) \cdot \cdots \cdot \Phi_d(\zeta) \cdot \cdots \cdot \Phi_1(\zeta) \qquad \text{where } d \text{ divides } n \ . \tag{8}$$

(6) The recoding in Equation (2) is a polynomial function of degree $n-1$ and complex coefficients in both directions:

$$\begin{aligned} \omega_k &= \sum_{s=0}^{n-1} \omega_s \frac{\prod_{h=0,h\neq s}^{n-1}(x-h)}{\prod_{h=0,h\neq s}^{n-1}(s-h)} \ , \qquad x = k \in \{0,\ldots,n-1\} \\ k &= \frac{1}{n}\sum_{h=0}^{n-1} \zeta^h \sum_{s=1}^{n-1} s\ \omega_{[s-sh]} \qquad , \qquad \zeta = \omega_k \in \Omega_n \ . \end{aligned} \tag{9}$$

The last Equation follows from

$$k = \sum_{s=1}^{n-1} s \frac{\prod_{h=0,h\neq s}^{n-1}(\zeta-\omega_h)}{\prod_{h=0,h\neq s}^{n-1}(\omega_s-\omega_h)} \quad , \qquad \zeta = \omega_k \in \Omega_n$$

and from the properties of the n-th roots of the unity, see Item 2.

References

Bailey, R. A., 1982. The decomposition of treatment degrees of freedom in quantitative factorial experiments. J. R. Statist. Soc., B 44 (1), 63–70.

Bailey, R. A., Gilchrist, F. H. L., Patterson, H. D., 1977. Identification of effects and confounding patterns in factorial designs. Biometrika 64, 347–354.

Bose, R. C., 1947. Mathematical theory of the symmetrical factorial designs. Sankhyā 8, 107–166.

Cheng, S.-W., Ye, K. Q., 2004. Geometric isomorphism and minimum aberration for factorial designs with quantitative factorss. The Annals of Statistics 32 (5).

Collombier, D., 1996. Plans D'Expérience Factoriels. Construction et propriétés des fractions de plans. No. 21 in Mathématiques et Applications. Springer, Paris.

Dean, A. M., John, J. A., 1975. Single replicate factorial experiments in generalized cyclic designs. II. Asymmetrical arrangements. J. Roy. Statist. Soc. Ser. B 37, 72–76.

Dey, A., Mukerjee, R., 1999. Fractional Factorial Plans. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York.

Edmondson, R. N., 1994. Fractional factorial designs for factors with a prime number of quantitative levels. J. R. Statist. Soc., B 56 (4), 611–622.

Fontana, R., Pistone, G., Rogantin, M.-P., 1997. Algebraic analysis and generation of two-levels designs. Statistica Applicata 9 (1), 15–29.

Fontana, R., Pistone, G., Rogantin, M. P., 2000. Classification of two-level factorial fractions. J. Statist. Plann. Inference 87 (1), 149–172.

Galetto, F., Pistone, G., Rogantin, M. P., 2003. Confounding revisited with commutative computational algebra. J. Statist. Plann. Inference 117 (2), 345–363.

Hedayat, A. S., Sloane, N. J. A., Stufken, J., 1999. Orthogonal arrays. Theory and applications, With a foreword by C. R. Rao. Springer-Verlag, New York.

John, J. A., Dean, A. M., 1975. Single replicate factorial experiments in generalized cyclic designs. I. Symmetrical arrangements. J. Roy. Statist. Soc. Ser. B 37, 63–71.

Kobilinsky, A., 1990. Complex linear model and cyclic designs. Linear Algebra and its Applications 127, 227–282.

Kobilinsky, A., 1997. Les Plans Factoriels. ASU–SSdF. Éditions Technip, Ch. 3, pp. 69–209.

Kobilinsky, A., Monod, H., 1991. Experimental design generated by group morphism: An introduction. Scand. J. Statist. 18, 119–134.

Kobilinsky, A., Monod, H., 1995. Juxtaposition of regular factorial designs and the complex linear model. Scand. J. Statist. 22, 223–254.

Lang, S., 1965. Algebra. Addison Wesley, Reading, Mass.

Lewis, S. M., 1979. The construction of resolution III fractions from generalized cyclic designs. J. Roy. Statist. Soc. Ser. B 41 (3), 352–357.

Lewis, S. M., 1982. Generators for asymmetrical factorial experiments. J. Statist. Plann. Inference 6 (1), 59–64.

Pistone, G., Riccomagno, E., Wynn, H. P., 2001. Algebraic Statistics: Computational Commutative Algebra in Statistics. Chapman&Hall, Boca Raton.

Pistone, G., Rogantin, M., 2005. Indicator function and different codings for fractional factorial designs. Tech. rep., Dipartimento di Matematica, Politecnico di Torino.

Pistone, G., Wynn, H. P., 1996. Generalised confounding with Gröbner bases. Biometrika 83 (3), 653–666.

Raktoe, B. L., Hedayat, A., Federer, W. T., 1981. Factorial designs. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York.

Robbiano, L., 1998. Gröbner bases and statistics. In: Buchberger, B., Winkler, F. (Eds.), Gröbner Bases and Applications (Proc. of the Conf. 33 Years of Gröbner Bases). Vol. 251 of London Mathematical Society Lecture Notes. Cambridge University Press, pp. 179–204.

Robbiano, L., Rogantin, M.-P., 1998. Full factorial designs and distracted fractions. In: Buchberger, B., Winkler, F. (Eds.), Gröbner Bases and Applications (Proc. of the Conf. 33 Years of Gröbner Bases). Vol. 251 of London Mathematical Society Lecture Notes Series. Cambridge University Press, pp. 473–482.

Schrijver, A., 2002. Theory of linear and integer programming. Wiley-Interscience Series in Discrete Mathematics. Wiley-Interscience [John Wiley & Sons], New York.

Tang, B., 2001. Theory of J-characteristics for fractional factorial designs and projection justification of minimum G_2-aberration. Biometrika 88 (2), 401–407.

Tang, B., Deng, L. Y., 1999. Minimum G_2-aberration for nonregular fractinal factorial designs. The Annals of Statistics 27 (6), 1914–1926.

Wu, C. F. J., Hamada, M., 2000. Experiments. John Wiley & Sons Inc., New York.

Xu, H., Wu, C. F. J., 2001. Generalized minimum aberration for asymmetrical fractional factorial designs. Ann. Statist. 29 (4), 1066–1077.

Ye, K. Q., 2003. Indicator function and its application in two-level factorial designs. The Annals of Statistics 31 (3), 984–994.

Ye, K. Q., 2004. A note on regular fractional factorial designs. Statistica sinica 14 (4), 1069–1074.