Higher order functions and Walsh coefficients

M.Teresa Iglesias* Vicente S. Peñaranda Alain Verschoren

Abstract

In this note, we characterize the order of fitness functions in terms of their Walsh coefficients.

Introduction

Genetic algorithms (GAs) are a mathematical tool inspired upon the mechanisms of natural evolution and are mainly applied in the framework of function optimization. If a GA is unable to find an optimum of a function in a reasonable time, one says that this function is "(GA) hard". It remains an open problem to characterize these "difficult" functions.

In [9], Rawlins introduced the notion of *epistasis* (named after a related characteristic in genetics) as an estimator for function difficulty. In particular, Rawlins speaks of *minimal epistasis* when all genes are independent of the others and it appears that functions having this property are exactly what should be called "functions of order one". It has been shown in [6] that these functions may also be described by the fact that their Walsh coefficients w_t vanish for $t \notin \{0, 2^i\}$, a property which allows for efficient calculation of (normalized) epistasis, cf. loc. cit. The main purpose of this note is to show that a similar result holds for functions of higher order. *En passant*, we will show by some examples how order and difficulty are linked for several types of well-known functions.

1 Preliminaries

In order to apply a GA to the optimization of a function, one first has to codify the data to which this function will be applied. Our data will always be assumed

*Research partially supported by a grant of Xunta de Galicia, PGIDIT03PXIA10502PR
Received by the editors October 2004.
Communicated by A. Hoogewijs.

to be encoded as vectors (or strings) of length ℓ. GAs essentially act by processing schemata, i.e., the algorithm acts on certain particular subsets of the search space, ignoring others. Formally, if one considers the alphabet of alleles Σ (in our case $\Sigma = \{0,1\}$) and if one adds the symbol #, one obtains an extended alphabet Σ'. A *schema* is an hyperplane H in the set Ω_ℓ of binary strings of length ℓ and may be viewed as an element $H = h_{\ell-1} \dots h_0$ of $(\Sigma')^\ell$, where each place where the symbol # occurs may be filled in by 0 or 1. Actually, one says that a string $s \in \Omega_\ell$ "belongs" to H (and one writes $s \in H$), if it has the same structure as H, i.e., if $s_i = h_i$ whenever $h_i \neq \#$. For example, 100111 belongs to $H = 10\#\#1\#$, but 001011 does not. The number of positions not filled in by # is the *order* of the schema. The above schema is of order 3.

In what follows, we will consider for every $1 \leq p \leq \ell$, every set of indices $J = \{j_1, \dots, j_p\} \subset \{0, \dots, \ell-1\}$ and every $1 \leq n \leq p$ the set

$$\mathbf{P}_J = \{(\alpha_n, \beta_{p-n});\ \{\alpha_n, \beta_{p-n}\} \text{ is a partition of } J\} \subset \mathcal{P}(J) \times \mathcal{P}(J)$$

and the order p schema

$$H_{\alpha_n}^{\beta_{p-n}} = \{t \in \Omega_\ell; t_i = 0 \text{ if } i \in \alpha_n,\ t_j = 1 \text{ if } j \in \beta_{p-n}\} \in (\Sigma')^\ell,$$

where $(\alpha_n, \beta_{p-n}) \in \mathbf{P}_J$. Occasionally, it may be convenient to make explicit the loci which are defined in the schema: we will write

$$H_{i_1 \dots i_n}^{j_1 \dots j_{p-n}} = \{s \in \Omega_\ell;\ s_{i_1} = \dots = s_{i_n} = 0, s_{j_1} = \dots = s_{j_{p-n}} = 1\}.$$

On the other hand, we will denote the schema $\{s \in \Omega_\ell;\ s_{i_1} = \dots = s_{i_p} = 0\}$, resp. $\{s \in \Omega_\ell;\ s_{j_1} = \dots = s_{j_p} = 1\}$ by $H_{i_1 \dots i_p}$, resp. $H^{j_1 \dots j_p}$. One easily sees that the order 0 schema Ω_ℓ may be written as the disjoint union of two schemata of order 1. For example, considering the schemata $H_i = \{s \in \Omega_\ell; s_i = 0\}$ and $H^i = \{s \in \Omega_\ell; s_i = 1\}$ for fixed i, clearly $\{H_i, H^i\}$ defines a partition of Ω_ℓ. Moreover, if we denote for any fitness function $f : \Omega_\ell \to \mathbb{R}$ by $f(\Omega_\ell)$, $f(H_i)$ and $f(H^i)$ the average of f on respectively Ω_ℓ, H_i and H^i, then

$$f(\Omega_\ell) = \frac{1}{2^\ell} \sum_{s \in \Omega_\ell} f(s) = \frac{1}{2^\ell} \left[\sum_{s \in H_i} f(s) + \sum_{s \in H^i} f(s) \right] = \frac{1}{2} \left[f(H_i) + f(H^i) \right].$$

More generally, any order p schema may be written as the disjoint union of two schemata of order $p+1$. In particular, it then follows that

$$f(\Omega_\ell) = \frac{1}{4} \left[f(H_{ij}) + f(H_i^j) + f(H_j^i) + f(H^{ij}) \right],$$

for example. Iterating this process, we obtain the following result:

Lemma 1. *The space Ω_ℓ may be decomposed into the disjoint union of 2^p schemata of order p.*

Proof. Write $E_n^{p-n}(J)$ as the disjoint union $\bigcup_{\mathbf{P}_J} H_{\alpha_n}^{\beta_{p-n}}$ of $\binom{p}{n}$ schemata, then

$$\Omega_\ell = \bigcup_{0\le n\le p} E_n^{p-n}(J) = \bigcup_{0\le n\le p} \left(\bigcup_{\mathbf{P}_J} H_{\alpha_n}^{\beta_{p-n}} \right).$$

■

As an immediate corollary, note that the fitness $f(\Omega_\ell)$ of the space Ω_ℓ is the arithmetic average of the fitness values of the above schemata. Indeed, $f(E_n^{p-n}(J))$ is equal to

$$\frac{1}{2^{\ell-p}\binom{p}{n}} \sum_{s\in \bigcup_{\mathbf{P}_J} H_{\alpha_n}^{\beta_{p-n}}} f(s) = \frac{1}{2^{\ell-p}\binom{p}{n}} \sum_{\mathbf{P}_J} \sum_{s\in H_{\alpha_n}^{\beta_{p-n}}} f(s) = \frac{1}{\binom{p}{n}} \sum_{\mathbf{P}_J} f(H_{\alpha_n}^{\beta_{p-n}})$$

and hence

$$\begin{aligned} f(\Omega_\ell) &= \frac{1}{2^\ell} \sum_{n=0}^{p} \sum_{s\in E_n^{p-n}(J)} f(s) = \frac{1}{2^\ell} \sum_{n=0}^{p} 2^{\ell-p}\binom{p}{n} f(E_n^{p-n}(J)) \\ &= \frac{1}{2^p} \sum_{n=0}^{p} \sum_{\mathbf{P}_J} f\left(H_{\alpha_n}^{\beta_{p-n}}\right). \end{aligned}$$

As pointed out in the introduction, functions of minimal epistasis in the sense of [9] are exactly those fitness functions $f : \Omega_\ell \to \mathbb{R}$, for which there exist $g_i : \{0,1\} \to \mathbb{R}$ such that $f(s) = \sum_{i=0}^{\ell-1} g_i(s_i)$, i.e., they are described by components which individually only depend upon a single bit. These functions are usually referred to as being of order 1 and are "easy" to optimize with a genetic algorithm. This notion may be generalized by defining a function to be of order $k > 1$, if $f(s)$ may be written in the form

$$\sum_{0\le i<\ell} g_i(s_i) + \sum_{0\le i_1<i_2<\ell} g_{i_1 i_2}(s_{i_1}, s_{i_2}) + \cdots + \sum_{0\le i_1<\cdots<i_k<\ell} g_{i_1\cdots i_k}(s_{i_1}, \ldots, s_{i_k}) \qquad (1)$$

where $g_{i_1\cdots i_r}(s_{i_1}, \ldots, s_{i_r})$ essentially describes the interaction between the r alleles situated at the locations $i_1, i_2, \cdots, i_r$.

2 Order and difficulty: some examples

In this section, we will give some examples of functions, for which there is good correlation between GA hardness and the order of the function.

As a first, easy example, recall (from [10], e. g.) that a function f (defined on length ℓ strings) is said to be a *camel function* if for some $0 \le i < 2^\ell$ we have $\mathbf{f}_i = \mathbf{f}_{2^\ell-1-i} \neq 0$ and $\mathbf{f}_j = 0$ elsewhere. (We denote by $\mathbf{f}_k$ the k-th component of the associated vector $\mathbf{f}$ of f.)

It is obvious that camel functions have order ℓ. Note that camel functions are extremely hard to optimize and are exactly those functions whose (normalized) epistasis is maximal, cf. [10] for details. This already gives a first indication that the order of a function is connected to its difficulty of being optimized.

As another example, let us consider the so-called *Template functions*, defined by assigning to any string $s \in \Omega_\ell$ the number of times a certain substring t of length $n \leq \ell$ appears in it. For convenience's sake, we will always assume that $t = 1^n = 1...1$. These Template functions thus clearly depend upon two parameters only (ℓ and n) and will be denoted by T_ℓ^n. For example, $T_\ell^2(1^\ell) = T_\ell^2(1...1) = \ell - 1$ and $T_\ell^3(01110...011) = 1$.

Let us note:

Lemma 2. *The function T_ℓ^n has order n.*

Proof. It suffices to note that $T_\ell^n(s) = \sum_{j=0}^{\ell-n} \tau_j(s)$, where, for every $0 \leq j \leq \ell - n$, we have

$$\tau_j(s) = \begin{cases} 1 & \text{if } s_j = ... = s_{j+n-1} = 1 \\ 0 & \text{otherwise.} \end{cases}$$

■

As a last example, let us consider the so-called *generalized Royal Road functions*, defined as follows. We consider binary strings of length $\ell = 2^n$ and for every $m \leq n$ let us define the schema $\sigma_i^{n,m} = \#^{2^m i} 1^{2^m} \#^{2^n - 2^m(i+1)}$ of order 2^m, where $0 \leq i < 2^{n-m}$. As in [7] and inspired by previous work by Forrest and Mitchell [2], we then define the generalized Royal Road function $\Re_m^n$ by $\Re_m^n(s) = \sum_{s \in \sigma_i^{n,m}} 2^m$. Let us note:

Lemma 3. *The function $\Re_m^n$ has order 2^m.*

Proof. It suffices to note that $\Re_m^n(s) = \sum_{j=0}^{2^{n-m}-1} \rho_j(s)$, where, for each $0 \leq j < 2^{n-m}$, we have

$$\rho_j(s) = \begin{cases} 2^m & \text{if } s_{j \cdot 2^m} = \cdots = s_{(j+1) \cdot 2^m - 1} = 1 \\ 0 & \text{otherwise.} \end{cases}$$

■

With these definitions, one may verify the nice correlation between the order of the previous functions and their GA hardness. Indeed, in the table below we take a look at what happens over strings of length 64. We used ranking selection, one-point crossover with probability 0.7 and simple mutation with probability 0.001. We stop the GA when 50% of the population consists of the maximum and we use the number of generations needed to attain this as a measure for the difficulty of the function.

ord	1		2		4		8		16	
f	$\Re_0^6$	T_{64}^1	$\Re_1^6$	T_{64}^2	$\Re_2^6$	T_{64}^4	$\Re_3^6$	T_{64}^8	$\Re_4^6$	T_{64}^{16}
NG	37	41	68	46	164	63	> 1200	99	> 1200	> 1200

Note that jointly increasing the order of both functions, their difficulty increases as well. Nevertheless, for all practical purposes, the problem remains open to characterize in a simple way the order of a function. The order 1 case has been solved in [6], through the use of Walsh coefficients. In the present note, a similar approach will given for functions of higher order, inspired by ideas of Goldberg's, who successfully applied Walsh analysis in [3] to the study of the fitness of arbitrary schemata.

3 Walsh coefficients of functions of order k

For any $t \in \Omega_\ell$, define the associated Walsh function ψ_t by $\psi_t(s) = (-1)^{s\cdot t}$, where $s\cdot t$ denotes the scalar product of s and t. It is then well-known (cf. [3], for example) that the set $\{\psi_t,\ t \in \Omega_\ell\}$ forms a basis for the vector space of real-valued functions on Ω_ℓ.

Actually, considering the 2^ℓ-dimensional matrix $\mathbf{V}_\ell = (\psi_t(s))_{s,t\in\Omega_\ell} \in M_{2^\ell}(\mathbb{Z})$ and representing a function f by its associated vector $\mathbf{f} \in \mathbb{R}^{2^\ell}$, let us define the *Walsh transform* w of f by $\mathbf{w} = \mathbf{W}_\ell \mathbf{f}$, where $\mathbf{W}_\ell = 2^{-\ell/2}\mathbf{V}_\ell$. The components $w_i = w_i(f)$ are then said to be the *Walsh coefficients* of f and are, up to the factor $2^{-\ell/2}$, the coordinates of f with respect to the above basis. In practical situations, it is usually easier to work with the matrix $\mathbf{W}_\ell$ which satisfies the recursion formula

$$\mathbf{W}_{\ell+1} = 2^{-\frac{1}{2}} \begin{pmatrix} \mathbf{W}_\ell & \mathbf{W}_\ell \\ \mathbf{W}_\ell & -\mathbf{W}_\ell \end{pmatrix}.$$

In particular, $v_0 = 2^{-\ell/2} w_0 = 2^{-\ell/2}(\mathbf{W}\mathbf{f})_0 = 2^{-\ell}\sum_{i=0}^{2^\ell-1}\mathbf{f}_i = f(\Omega_\ell)$.

As $\mathbf{W}_\ell$ is idempotent, it easily follows that $\mathbf{f} = \mathbf{W}_\ell \mathbf{w}$ and that $\|\,\mathbf{f}\,\| = \|\mathbf{w}\|$.

On the other hand, let us also mention Goldberg's well-known *Hyperplane average theorem* (cf. [3], e. g.), which permits to calculate recursively the Walsh coefficients of order k:

$$w_{2^{i_1}+\cdots+2^{i_k}} = 2^{\ell/2} f(H_{i_1\cdots i_k}) - \left(\sum_{q=1}^{k-1}\ \sum_{1\le\lambda_1<\cdots<\lambda_q\le k} w_{2^{i_{\lambda_1}}+\cdots+2^{i_{\lambda_q}}}\right) - w_0, \qquad (2)$$

where $i_1,\ldots,i_k \in \{0,\ldots,\ell-1\}$. Given the fact that Walsh transforms respect the separation between the parts of order $1, 2, \ldots, k$ of the expression given in (1) for any function f of order k, it suffices to consider each of these components separately. Consider a *simple order k function*, i.e., an order $k \le \ell$ function $\mathcal{G}$ on Ω_ℓ, which has no components of lower order: $\mathcal{G}(s) = \mathcal{G}_{i_1\cdots i_k}(s_{i_1},\ldots,s_{i_k})$. We will prove that the Walsh coefficients of order higher than k all vanish and, using this, we will generalize this property to functions of arbitrary order given by the expression (1).

In order to realize this, let us fix $k \ge 1$ and a set of indices $I = \{i_1,\ldots,i_k\} \subset \{0,\ldots,\ell-1\}$. For any set of indices $J \subset \{0,\ldots,\ell-1\}$ and any partition of J formed by, say, Q and $J-Q$, we then define the order p schema $H^{\beta_{q-m}\widetilde{\beta}_{p-q-(n-m)}}_{\alpha_m\widetilde{\alpha}_{n-m}}$ as

$$\{t \in \Omega_\ell;\ t_j = 0, j \in \alpha_m \cup \widetilde{\alpha}_{n-m},\ t_j = 1, j \in \beta_{q-m} \cup \widetilde{\beta}_{p-q-(n-m)}\},$$

where $\big((\alpha_m,\beta_{q-m}),(\widetilde{\alpha}_{n-m},\widetilde{\beta}_{p-q-(n-m)})\big) \in \mathbf{P}_Q \times \mathbf{P}_{J-Q}$.

As before, subindices refer to the respective cardinalities of the components of the partition. Moreover, for non-proper partitions, one eliminates the corresponding index in the notation.

Example 1. *If $\ell = 4$, $J = \{0,2,3\}$ and $Q = \{0,2\}$, then we have H_{023}, H_{02}^{3} if $m = 2$, we have H_{03}^{2}, H_{0}^{23}, H_{2}^{03}, H_{23}^{0} if $m = 1$ and H_{3}^{02}, H^{023} if $m = 0$.*

In a similar way, let us denote by $\widetilde{H}_{\alpha_m\widetilde{\alpha}_{n-m}}^{\beta_{q-m}\widetilde{\beta}_{p-q-(n-m)}}$ the schema in $(\Sigma')^k$ obtained from $H_{\alpha_m\widetilde{\alpha}_{n-m}}^{\beta_{q-m}\widetilde{\beta}_{p-q-(n-m)}}$ by only considering the bits $t_{i_1}, \ldots, t_{i_k}$.

We will need the following technical result:

Lemma 4. *For every $p > 0$ and any $0 \le q, n \le p$, we have:*

$$\sum_{q=0}^{p-1}\left(\sum_{j=0}^{q}(-1)^{q-j}\binom{n}{j}\binom{p-n}{q-j}\right) = (-1)^{p-n+1}.$$

Proof.

$$\begin{aligned}
\sum_{q=0}^{p-1}\left(\sum_{j=0}^{q}(-1)^{q-j}\binom{n}{j}\binom{p-n}{q-j}\right) &= \sum_{j=0}^{p-1}\left(\sum_{q=j}^{p-1}(-1)^{q-j}\binom{n}{j}\binom{p-n}{q-j}\right)\\
&= \sum_{j=0}^{p-1}\left(\sum_{z=0}^{p-j-1}(-1)^{z}\binom{n}{j}\binom{p-n}{z}\right)\\
&= \sum_{z=0}^{p-1}\left(\sum_{j=0}^{p-z-1}(-1)^{z}\binom{n}{j}\binom{p-n}{z}\right)\\
&= \sum_{z=0}^{p-n}(-1)^{z}\binom{p-n}{z}\sum_{j=0}^{p-z-1}\binom{n}{j}.
\end{aligned}$$

But now, as

$$\sum_{j=0}^{p-z-1}\binom{n}{j} = \sum_{j=0}^{n}\binom{n}{j} = 2^n,$$

for $p - z - 1 \ge n$ and

$$\sum_{j=0}^{p-z-1}\binom{n}{j} = \sum_{j=0}^{n-1}\binom{n}{j} = 2^n - 1$$

if $z = p - n$, it follows that

$$\begin{aligned}
\sum_{z=0}^{p-n}(-1)^{z}\binom{p-n}{z}\sum_{j=0}^{p-z-1}\binom{n}{j} &= \sum_{z=0}^{p-n-1}(-1)^{z}\binom{p-n}{z}\sum_{j=0}^{p-z-1}\binom{n}{j} + (-1)^{p-n}\sum_{j=0}^{n-1}\binom{n}{j}\\
&= 2^n\sum_{z=0}^{p-n-1}(-1)^{z}\binom{p-n}{z} + (-1)^{p-n}\left(2^n - 1\right)\\
&= (-1)^{p-n+1},
\end{aligned}$$

as claimed. ∎

Using this, one may prove:

Proposition 1. *For every set* $J = \{j_1, \ldots, j_p\} \subset \{0, \ldots, \ell-1\}$*, the value of the coefficient* $w_{2^{j_1}+\cdots+2^{j_p}}$ *of a simple order* k *function* $\mathcal{G}$ *is given by*

$$w_{2^{j_1}+\ldots+2^{j_p}} = \begin{cases} \frac{2^{\ell/2}}{2^p} \sum_{n=0}^{p} (-1)^{p-n} \binom{p}{n} \ \mathcal{G}(E_n^{p-n}(J)) & \text{if } J \subset I \\ 0 & \text{if } J \not\subset I. \end{cases}$$

Proof. We will use induction on p. For simplicity's sake, let us calculate the values $v_{2^j} = 2^{-\ell/2} w_{2^j}$. For $p = 1$, we find $J = \{j\}$ and for $J \in I$

$$v_{2^j} = \mathcal{G}(H_j) - v_0 = \mathcal{G}(H_j) - \mathcal{G}(\Omega_\ell)$$

is clearly equal to

$$\frac{1}{2}\left[\mathcal{G}(H_j) - \mathcal{G}(H^j)\right] = \frac{1}{2}\left[-\mathcal{G}(E_0^1(J)) + \mathcal{G}(E_1^0(J))\right].$$

If $j \notin I$, then, since the function $\mathcal{G}$ is simple of order k, clearly $\mathcal{G}(H_j) = \mathcal{G}(H^j)$ and hence $v_{2^j} = 0$.

Let us now assume that the result holds true up to $p-1$ and let us prove it for p. We will distinguish two cases.

Case 1: $J \subset I$.

Since $\mathcal{G}$ is simple of order $k \geq p$, it is easy to prove that $\mathcal{G}(H_{j_1 \cdots j_p}) = \mathcal{G}(\widetilde{H}_{j_1 \cdots j_p})$. Indeed,

$$\begin{aligned} \mathcal{G}(H_{j_1 \cdots j_p}) &= \frac{1}{2^{\ell-p}} \sum_{s \in H_{j_1 \cdots j_p}} \mathcal{G}(s) = \frac{1}{2^{\ell-p}} \sum_{s \in \widetilde{H}_{j_1 \cdots j_p}} 2^{\ell-k} \mathcal{G}(s) \\ &= \frac{1}{2^{k-p}} \sum_{s \in \widetilde{H}_{j_1 \cdots j_p}} \mathcal{G}(s) = \mathcal{G}(\widetilde{H}_{j_1 \cdots j_p}). \end{aligned}$$

On the other hand, let $Q = \{j_{\lambda_1}, \cdots, j_{\lambda_q}\}$ be a subset of J of cardinality $1 \leq q < p$ and let (α_m, β_{q-m}) be an element of $\mathbf{P}_Q$. Then, by the induction hypothesis,

$$\begin{aligned} v_{2^{j_{\lambda_1}}+\cdots+2^{j_{\lambda_q}}} &= \frac{1}{2^q} \sum_{m=0}^{q} (-1)^{q-m} \binom{q}{m} \mathcal{G}\left(E_m^{q-m}(Q)\right) = \frac{1}{2^q} \sum_{m=0}^{q} (-1)^{q-m} \sum_{\mathbf{P}_Q} \mathcal{G}\left(\widetilde{H}_{\alpha_m}^{\beta_{q-m}}\right) \\ &= \frac{1}{2^p} \sum_{n=0}^{p} \sum_{\mathbf{P}_Q \times \mathbf{P}_{J-Q}} (-1)^{q-m} \mathcal{G}\left(\widetilde{H}_{\alpha_m \widetilde{\alpha}_{n-m}}^{\beta_{q-m} \widetilde{\beta}_{p-n-(q-m)}}\right), \end{aligned}$$

and $v_0 = \mathcal{G}(\Omega_\ell) = \mathcal{G}(\Omega_k) = \frac{1}{2^p} \sum_{n=0}^{p} \sum_{\mathbf{P}_J} \mathcal{G}\left(\widetilde{H}_{\mu_n}^{\tau_{p-n}}\right)$. Now, in order to apply (2) to calculate $v_{2^{j_1}+\cdots+2^{j_p}}$, we first need to calculate the value of

$$\Gamma = v_0 + \sum_{q=1}^{p-1} \sum_{1 \leq \lambda_1 < \cdots < \lambda_q \leq p} v_{2^{j_{\lambda_1}}+\cdots+2^{j_{\lambda_q}}}$$

as follows

$$\begin{aligned}\Gamma &= \frac{1}{2^p}\sum_{n=0}^{p}\sum_{\mathbf{P}_J}\mathcal{G}\left(\widetilde{H}^{\tau_{p-n}}_{\mu_n}\right)\\ &\quad+\sum_{q=1}^{p-1}\sum_{1\le\lambda_1<\cdots<\lambda_q\le p}\left[\frac{1}{2^p}\sum_{n=0}^{p}\sum_{\mathbf{P}_Q\times\mathbf{P}_{J-Q}}(-1)^{q-m}\mathcal{G}\left(\widetilde{H}^{\beta_{q-m}\widetilde{\beta}_{p-n-(q-m)}}_{\alpha_m\widetilde{\alpha}_{n-m}}\right)\right]\\ &= \frac{1}{2^p}\left(\mathcal{G}\left(\widetilde{H}_{j_1\cdots j_p}\right)+\sum_{n=0}^{p-1}\sum_{\mathbf{P}_J}\mathcal{G}\left(\widetilde{H}^{\tau_{p-n}}_{\mu_n}\right)\right)\\ &\quad+\frac{1}{2^p}\sum_{n=0}^{p}\sum_{\mathbf{P}_Q\times\mathbf{P}_{J-Q}}\left\{\sum_{q=1}^{p-1}\sum_{1\le\lambda_1<\cdots<\lambda_q\le p}(-1)^{q-m}\mathcal{G}\left(\widetilde{H}^{\beta_{q-m}\widetilde{\beta}_{p-n-(q-m)}}_{\alpha_m\widetilde{\alpha}_{n-m}}\right)\right\}.\end{aligned}$$

If $n=p$, then $q=m$ and

$$\widetilde{H}^{\beta_{q-m}\widetilde{\beta}_{p-n-(q-m)}}_{\alpha_m\widetilde{\alpha}_{n-m}}=\widetilde{H}_{\alpha_q\widetilde{\alpha}_{p-q}}=\widetilde{H}_{j_1\cdots j_p},$$

as $\alpha_q=Q$ and $\widetilde{\alpha}_{p-q}=J-Q$. It thus follows that

$$\begin{aligned}\sum_{q=1}^{p-1}\sum_{1\le\lambda_1<\cdots<\lambda_q\le p}(-1)^{q-m}\mathcal{G}\left(\widetilde{H}^{\beta_{q-m}\widetilde{\beta}_{p-n-(q-m)}}_{\alpha_m\widetilde{\alpha}_{n-m}}\right) &= \sum_{q=1}^{p-1}\sum_{1\le\lambda_1<\cdots<\lambda_q\le p}\mathcal{G}\left(\widetilde{H}_{j_1\cdots j_p}\right)\\ &= \sum_{q=1}^{p-1}\binom{p}{q}\mathcal{G}\left(\widetilde{H}_{j_1\cdots j_p}\right)\\ &= (2^p-2)\mathcal{G}\left(\widetilde{H}_{j_1\cdots j_p}\right).\end{aligned}$$

On the other hand, if $n<p$ a rather technical calculation (see [8] for details) shows that

$$\sum_{\mathbf{P}_Q\times\mathbf{P}_{J-Q}}\left\{\sum_{q=1}^{p-1}\sum_{1\le\lambda_1<\cdots<\lambda_q\le p}(-1)^{q-m}\mathcal{G}\left(\widetilde{H}^{\beta_{q-m}\widetilde{\beta}_{p-n-(q-m)}}_{\alpha_m\widetilde{\alpha}_{n-m}}\right)\right\}=\sum_{\mathbf{P}_J}\gamma_{p,n}\mathcal{G}\left(\widetilde{H}^{\tau_{p-n}}_{\mu_n}\right)$$

whence

$$\gamma_{p,n}=\sum_{q=1}^{p-1}\sum_{j=0}^{q}(-1)^{q-j}\binom{n}{j}\binom{p-n}{q-j}.$$

Applying lemma 4, it now follows that

$$\begin{aligned}&\sum_{\mathbf{P}_Q\times\mathbf{P}_{J-Q}}\left\{\sum_{q=1}^{p-1}\sum_{1\le\lambda_1<\cdots<\lambda_q\le p}(-1)^{q-m}\mathcal{G}\left(\widetilde{H}^{\beta_{q-m}\widetilde{\beta}_{p-n-(q-m)}}_{\alpha_m\widetilde{\alpha}_{n-m}}\right)\right\}\\ &=\sum_{\mathbf{P}_J}\gamma_{p,n}\mathcal{G}\left(\widetilde{H}^{\tau_{p-n}}_{\mu_n}\right)=\sum_{\mathbf{P}_J}\left((-1)^{p-n+1}-1\right)\mathcal{G}\left(\widetilde{H}^{\tau_{p-n}}_{\mu_n}\right)\end{aligned}$$

and we finally obtain

$$\begin{aligned}
v_{2^{j_1}+\cdots+2^{j_p}} &= \mathcal{G}\left(\widetilde{H}_{j_1\cdots j_p}\right) - \Gamma = \mathcal{G}\left(\widetilde{H}_{j_1\cdots j_p}\right) - \frac{2^p-1}{2^p}\mathcal{G}\left(\widetilde{H}_{j_1\cdots j_p}\right) \\
&\quad -\frac{1}{2^p}\sum_{n=0}^{p-1}\sum_{\mathbf{P}_J}(1+\gamma_{p,n})\mathcal{G}\left(\widetilde{H}_{\mu_n}^{\tau_{p-n}}\right) \\
&= \frac{1}{2^p}\mathcal{G}\left(\widetilde{H}_{j_1\cdots j_p}\right) - \frac{1}{2^p}\sum_{n=0}^{p-1}\sum_{\mathbf{P}_J}(-1)^{p-n+1}\mathcal{G}\left(\widetilde{H}_{\mu_n}^{\tau_{p-n}}\right) \\
&= \frac{1}{2^p}\sum_{n=0}^{p}(-1)^{p-n}\sum_{\mathbf{P}_J}\mathcal{G}\left(\widetilde{H}_{\mu_n}^{\tau_{p-n}}\right) = \frac{1}{2^p}\sum_{n=0}^{p}(-1)^{p-n}\binom{p}{n}\mathcal{G}\left(E_n^{p-n}(J)\right).
\end{aligned}$$

Case 2: $J = \{j_1, \dots, j_p\} \not\subset I$.

Without loss of generality, assume that $R = \{j_1, \dots, j_r\} \subset I$, $\{j_{r+1}, \dots, j_p\} \cap I = \emptyset$. It is straightforward to check that $\mathcal{G}\left(H_{j_1\cdots j_p}\right) = \mathcal{G}\left(\widetilde{H}_{j_1\cdots j_r}\right)$.

On the other hand, by the induction hypothesis,

$$\begin{aligned}
v_{2^{j_1}+\cdots+2^{j_p}} &= \mathcal{G}(H_{j_1\cdots j_p}) - \left(v_0 + \sum_{q=1}^{p-1}\sum_{1\le\lambda_1<\cdots<\lambda_q\le p} v_{2^{j_{\lambda_1}}+\cdots+2^{j_{\lambda_q}}}\right) \\
&= \mathcal{G}(H_{j_1\cdots j_p}) - \left(v_0 + \sum_{q=1}^{r}\sum_{\substack{1\le\lambda_1<\cdots<\lambda_q\le r\\ j_{\lambda_i}\in R}} v_{2^{j_{\lambda_1}}+\cdots+2^{j_{\lambda_q}}}\right) \\
&= \mathcal{G}(H_{j_1\cdots j_p}) - \Gamma.
\end{aligned}$$

The set $Q = \{j_{\lambda_1}, \dots, j_{\lambda_q}\} \subset R$ has cardinality q, hence, again by induction and arguing as in case 1, it follows that

$$v_{2^{j_{\lambda_1}}+\cdots+2^{j_{\lambda_q}}} = \frac{1}{2^r}\sum_{n=0}^{r}(-1)^{q-m}\sum_{\mathbf{P}_Q\times\mathbf{P}_{R-Q}}\mathcal{G}\left(\widetilde{H}_{\alpha_m\widetilde{\alpha}_{n-m}}^{\beta_{q-m}\widetilde{\beta}_{r-n-(q-m)}}\right),$$

whence, as before, $(\alpha_m, \beta_{q-m}) \in \mathbf{P}_Q$. So,

$$\begin{aligned}
\Gamma &= \frac{1}{2^r}\sum_{n=0}^{r}\sum_{\mathbf{P}_R}\mathcal{G}\left(\widetilde{H}_{\mu_n}^{\tau_{r-n}}\right) \\
&\quad +\sum_{q=1}^{r}\sum_{1\le\lambda_1<\cdots<\lambda_q\le r}\frac{1}{2^r}\left(\sum_{n=0}^{r}\sum_{\mathbf{P}_Q\times\mathbf{P}_{R-Q}}(-1)^{q-m}\mathcal{G}\left(\widetilde{H}_{\alpha_m\widetilde{\alpha}_{n-m}}^{\beta_{q-m}\widetilde{\beta}_{r-n-(q-m)}}\right)\right) \\
&= \frac{1}{2^r}\left(\mathcal{G}\left(\widetilde{H}_{j_1\cdots j_r}\right) + \sum_{n=0}^{r-1}\sum_{\mathbf{P}_R}\mathcal{G}\left(\widetilde{H}_{\mu_n}^{\tau_{r-n}}\right)\right) \\
&\quad +\frac{1}{2^r}\sum_{n=0}^{r}\sum_{\mathbf{P}_Q\times\mathbf{P}_{R-Q}}\left(\sum_{q=1}^{r}\sum_{1\le\lambda_1<\cdots<\lambda_q\le r}(-1)^{q-m}\right)\mathcal{G}\left(\widetilde{H}_{\alpha_m\widetilde{\alpha}_{n-m}}^{\beta_{q-m}\widetilde{\beta}_{r-n-(q-m)}}\right) \\
&= \frac{1}{2^r}\left[2^r + \sum_{n=0}^{r-1}\sum_{\mathbf{P}_R}\left(1 + \sum_{q=1}^{r}\sum_{1\le\lambda_1<\cdots<\lambda_q\le r}(-1)^{q-m}\right)\right]\mathcal{G}\left(\widetilde{H}_{\alpha_m\widetilde{\alpha}_{n-m}}^{\beta_{q-m}\widetilde{\beta}_{r-n-(q-m)}}\right) \\
&= \mathcal{G}(\widetilde{H}_{j_1\cdots j_r}).
\end{aligned}$$

Indeed, for $n = r$ we have $q = m$ and $\widetilde{H}^{\beta_{q-m}\widetilde{\beta}_{r-n-(q-m)}}_{\alpha_m\widetilde{\alpha}_{n-m}} = \widetilde{H}_{\alpha_q\widetilde{\alpha}_{r-q}} = \widetilde{H}_{j_1\cdots j_r}$ (as $\alpha_q = Q$ and $\widetilde{\alpha}_{r-q} = R - Q$) hence

$$\begin{aligned}\sum_{q=1}^{r}\sum_{1\le\lambda_1<\cdots<\lambda_q\le r}(-1)^{q-m}\mathcal{G}\left(\widetilde{H}^{\beta_{q-m}\widetilde{\beta}_{r-n-(q-m)}}_{\alpha_m\widetilde{\alpha}_{n-m}}\right) &= \sum_{q=1}^{r}\sum_{1\le\lambda_1<\cdots<\lambda_q\le r}\mathcal{G}\left(\widetilde{H}_{j_1\cdots j_r}\right)\\ &= (2^r-1)\,\mathcal{G}\left(\widetilde{H}_{j_1\cdots j_r}\right).\end{aligned}$$

On the other hand, if $n < r$, it has been proved in [8] that

$$\sum_{\mathbf{P}_Q\times\mathbf{P}_{R-Q}}\sum_{q=1}^{r}\sum_{1\le\lambda_1<\cdots<\lambda_q\le r}(-1)^{q-m}\mathcal{G}\left(\widetilde{H}^{\beta_{q-m}\widetilde{\beta}_{r-n-(q-m)}}_{\alpha_m\widetilde{\alpha}_{n-m}}\right) = \sum_{\mathbf{P}_R}\zeta_{r,n}\mathcal{G}\left(\widetilde{H}^{\tau_{r-n}}_{\mu_n}\right)$$

with

$$\zeta_{r,n} = \sum_{q=1}^{r}\sum_{j=0}^{q}(-1)^{q-j}\binom{n}{j}\binom{r-n}{q-j}$$

equal to

$$\gamma_{r,n} + \sum_{j=0}^{r}(-1)^{r-j}\binom{n}{j}\binom{r-n}{r-j} = [-1+(-1)^{r-n+1}] + (-1)^{r-n} = -1.$$

From (2), it now immediately follows that $v_{2^{j_1}+\cdots+2^{j_p}} = \mathcal{G}\left(H_{j_1\cdots j_p}\right) - \Gamma = 0$, which finishes the proof. ∎

Corollary 1. *If the function $\mathcal{G} : \Omega_\ell \to \mathbb{R}$ has order k, then its Walsh coefficients of order $p > k$ all vanish.*

We thus finally obtain:

Theorem 1. *For any function $f : \Omega_\ell \to \mathbb{R}$ with Walsh coefficients w_t, the following statements are equivalent:*

1. *f has order k;*
2. *$w_t = 0$ for all $t \notin \{0,\ 2^{i_1}+\cdots+2^{i_j};\ 1\le j\le k,\ 0\le i_1<\cdots<i_k<\ell\}$.*

Proof. In view of the previous results, it only remains to prove the converse of the previous corollary. As

$$f(s) \;=\; (\mathbf{W}_\ell\mathbf{w})_s = 2^{-\ell/2}w_0 + 2^{-\ell/2}\sum_{j=1}^{k}\sum_{0\le i_1<\cdots<i_j<\ell}(-1)^{(s_{i_1}+\cdots+s_{i_j})}w_{2^{i_1}+\cdots+2^{i_j}},$$

it suffices to define

$$g_{i_1\cdots i_j}(s) = 2^{-\ell/2}\left(\frac{w_0}{k\binom{\ell}{j}} + (-1)^{(s_{i_1}+\cdots+s_{i_j})}\ w_{2^{i_1}+\cdots+2^{i_j}}\right),$$

for every $0 \le i_1 < \cdots < i_j < \ell$ and to note that

$$f(s) = \sum_{j=1}^{k}\sum_{0\le i_1<\cdots<i_j<\ell} g_{i_1\cdots i_j}(s).$$

∎

References

[1] Davidor, Y.: Epistasis and Variance: A Viewpoint on GA-Hardness. In: Rawlins, G. J. E. (Ed.): Foundations of Genetic Algorithms 1, Morgan Kaufmann Publishers, 1991, pp. 23–25.

[2] Forrest, S. and Mitchell, M.: Relative Building-Block Fitness and the Building-Block Hypotesis, Foundations of Genetic Algorithms 2. Ed. L.D. Whitley, Morgan Kaufmann Publishers, San Mateo, 1993.

[3] Goldberg, D. E.: Genetic Algorithms and Walsh functions: Part I, A Gentle Introduction. Complex Systems, Vol. 3, 1989, pp. 129–152.

[4] Iglesias, M. T., Naudts, B., Verschoren, A. and Vidal, C.: A Combinatorial Approach to Epistasis, Springer, Dordrecht, 1995.

[5] Iglesias, M. T., Verschoren, A. and Vidal, C.: Template Functions and their Epistasis. In: Proceedings of MS'2000 International Conference on Modelling and Simulation, Las Palmas de Gran Canaria, 2000, pp. 539-546.

[6] Naudts, B., Suys, D. and Verschoren, A.: Epistasis, Deceptivity and Walsh Transforms. In: Proceedings of the International ICSC Symposium on Engineering of Intelligent Systems (EIS'98), V1: Fuzzy Logic/Genetic Algorithms, ICSC Academic Press, 1998, pp. 210–216.

[7] Naudts, B., Suys, D. and Verschoren, A.: Generalized Royal Road functions and their Epistasis, Artificial Intelligence, Vol. 19, 2000, pp. 317-334.

[8] Peñaranda, V. S.: Epistasis superior, Ph.D. Thesis, Universidade da Coruña, in preparation.

[9] Rawlins, G. J. E.: Foundations of Genetic Algorithms, Morgan Kaufmann Publishers, San Mateo, 1991.

[10] Suys, D. and Verschoren, A.: Extreme Epistasis. In: International Conference on Intelligent Technologies in Human-Relates Sciences (ITHURS 96), Vol. II, León, 1996, pp. 251–258.

[11] Van Hove, H.: Representational Issues in Genetic Algorithms, Ph.D. Thesis, University of Antwerp, 1995.

M.Teresa Iglesias, Vicente S. Peñaranda
Departamento de Matemáticas
Universidade da Coruña
A Coruña, Spain
email : totero@udc.es, vesp@udc.es

Alain Verschoren
Department of Mathematics and Computer Science
University of Antwerp
Antwerp, Belgium
email : alain.verschoren@ua.ac.be