Higher order functions and Walsh coefficients revisited

M. T. Iglesias C. Vidal* A. Verschoren

Abstract

The main purpose of this note is to present an alternative, more transparent treatment of the results obtained in [4], which link the epistasis of a function to its Walsh coefficients and its order.

1 Introduction

Throughout, we will denote by $\Omega_\ell = \{0,1\}^\ell$ the set of binary strings of length ℓ and use binary representation to identify Ω_ℓ with the set of integers $0, \ldots, 2^\ell - 1$. As the results of this note should be viewed in the context of genetic algorithms (cf. [4], for details), we will usually refer to real-valued functions on Ω_ℓ as *fitness functions*. It is easy to see that any fitness function f may be written in polynomial form as

$$f(x_0, \ldots, x_{\ell-1}) = \sum_{i_0 \ldots i_{\ell-1} \in \{0,1\}} a_{i_0 \ldots i_{\ell-1}} x_0^{i_0} \ldots x_{\ell-1}^{i_{\ell-1}}. \tag{1}$$

We say that f is of *order* (at most) k, if it may be written as

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

for some functions $g_{i_1 \cdots i_r}(s_{i_1}, \ldots, s_{i_r})$ on Ω_r, which essentially describe the interaction between the bits situated at the locations $i_1, i_2, \cdots, i_r$. It is not difficult to see that

*Research partially supported by the Xunta de Galicia (Spain) project REGACA 2006/38

Received by the editors April 2007.
Communicated by B. Hoogewijs.
1991 *Mathematics Subject Classification :* 68T20, 68W05.
Key words and phrases : Genetic algorithm, function order, epistasis, Walsh coefficient.

this is also equivalent to $a_{i_0 \dots i_{\ell-1}} = 0$ for $u(i_0 \dots i_{\ell-1}) > k$, in the above form (1) of f. Here, $u(s)$ denotes the *weight* of the string s, i.e., the number of ones in s.

The Walsh coefficients of a fitness function f are easy to calculate by using its vector representation ${}^t\mathbf{f} = (\mathrm{f}_0, \dots, \mathrm{f}_{2^\ell - 1}) \in \mathbb{R}^{2^\ell}$, where f_i is the value of f on the binary string $i_0 \dots i_{\ell-1}$ representing $0 \leq i \leq 2^\ell - 1$.

Recall that for any string $t \in \Omega_\ell$, the associated Walsh function ψ_t is defined 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. [2], 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 $\mathbf{V}_\ell = (\psi_t(s))_{s,t \in \Omega_\ell} \in M_{2^\ell}(\mathbb{Z})$ which satisfies the recursion formula

$$\mathbf{V}_{\ell+1} = \begin{pmatrix} \mathbf{V}_\ell & \mathbf{V}_\ell \\ \mathbf{V}_\ell & -\mathbf{V}_\ell \end{pmatrix},$$

with

$$\mathbf{V}_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

and putting

$$\mathbf{v}_\ell = (v_i) = 2^{-\ell} \mathbf{V}_\ell \, \mathbf{f},$$

it is easy to check that $v_i = v_i(f)$ is the $i-$th coordinate of $\mathbf{f}$ with respect to the above basis.

The vector $\mathbf{w} = \mathbf{W}_\ell \mathbf{f}$, where $\mathbf{W}_\ell = 2^{-\ell/2} \mathbf{V}_\ell$ defines the *Walsh transform* w of f and its components

$$w_i = 2^{-\ell/2} v_i$$

are the *Walsh coefficients* of f.

In [4], we proved the following result, which links the order of a fitness function to its Walsh coefficients:

Theorem 1.1. *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 $t \in \Omega_\ell$ with $u(t) > k$.*

The proof of the implication 2) $\Rightarrow$ 1) is easy. Indeed, as

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

for any $s \in \Omega_\ell$, we have

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

with

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

for every $0 \leq i_1 < \cdots < i_j < \ell$.

The other implication was also proven in [4], the proof being very technical, however. In the next section, we will present a surprisingly straightforward alternative to it.

Note 1.1. In [4], we linked the previous result to the notion of "higher epistasis", showing that the previous statements are also equivalent to asserting that $\varepsilon_k^*(f) = 0$, where $\varepsilon_k^*(f)$ denotes the normalized k-epistasis of f - we refer to [3, 4, 5] for definitions and details. Specializing to the order one (or *linear*) case, the result thus says that a fitness function is linear if and only if it has zero "standard" normalized epistasis, and that this is also equivalent to all of its higher Walsh coefficients vanishing, cf. [3] for details and the proper interpretation of this result in the framework of GA hardness.

2 The "other" implication

As we just mentioned, in this section we will give an alternative proof of the implication 1) ⇒ 2) in the above theorem. For $s_0 \in \{0,1\}$, define $f_{s_0\#\ldots\#} : \Omega_{\ell-1} \to \mathbb{R}$ by

$$f_{s_0\#\ldots\#}(s_1, \ldots, s_{\ell-1}) = f(s_0, s_1, \ldots, s_{\ell-1}).$$

From

$$f_{0\#\ldots\#}(x_1, \ldots, x_{\ell-1}) = \sum_{i_1 \ldots i_{\ell-1} \in \{0,1\}} c_{i_1 \ldots i_{\ell-1}} \, x_1^{i_1} \ldots x_{\ell-1}^{i_{\ell}-1}$$

and

$$f_{1\#\ldots\#}(x_1, \ldots, x_{\ell-1}) = \sum_{i_1 \ldots i_{\ell-1} \in \{0,1\}} b_{i_1 \ldots i_{\ell-1}} \, x_1^{i_1} \ldots x_{\ell-1}^{i_{\ell}-1},$$

it follows that

$$a_{0 i_1 \ldots i_{\ell-1}} = c_{i_1 \ldots i_{\ell-1}}$$

and

$$a_{1 i_1 \ldots i_{\ell-1}} = b_{i_1 \ldots i_{\ell-1}} - c_{i_1 \ldots i_{\ell-1}}. \tag{2}$$

Indeed, the very definition of $f_{0\#\ldots\#}$ yields

$$\begin{aligned} f_{0\#\ldots\#}(x_1, \ldots, x_{\ell-1}) &= \sum_{i_0 \ldots i_{\ell-1} \in \{0,1\}} a_{i_0 \ldots i_{\ell-1}} \, 0^{i_0} x_1^{i_1} \ldots x_{\ell-1}^{i_{\ell-1}} \\ &= \sum_{i_1 \ldots i_{\ell-1} \in \{0,1\}} a_{0 i_1 \ldots i_{\ell-1}} \, x_1^{i_1} \ldots x_{\ell-1}^{i_{\ell-1}} \\ &= \sum_{i_1 \ldots i_{\ell-1} \in \{0,1\}} c_{i_1 \ldots i_{\ell-1}} x_1^{i_1} \ldots x_{\ell-1}^{i_{\ell-1}}. \end{aligned}$$

The second relation may be derived similarly.

Let us denote by $v^{s_0}_{i_1 \ldots i_{\ell-1}}$ the Walsh coefficients $v_{s_0 i_1 \ldots i_{\ell-1}}(f_{s_0\#\ldots\#})$ of $f_{s_0\#\ldots\#}$.

We will need the following lemmas:

Lemma 2.1. *With notations as before, we have*

$$v_{s_0 i_1 \dots i_{\ell-1}}(f) = 2^{-1}\left(v^0_{i_1\dots i_{\ell-1}} + (-1)^{s_0} v^1_{i_1\dots i_{\ell-1}}\right)$$

resp.

$$v^{s_0}_{i_1\dots i_{\ell-1}} = v_{0 i_1\dots i_{\ell-1}}(f) + (-1)^{s_0} v_{1 i_1 \dots i_{\ell-1}}(f).$$

Proof. By a straightforward recursion argument, the result easily follows from:

$$\begin{aligned} \mathbf{v}_\ell &= 2^{-\ell}\mathbf{V}_\ell \mathbf{f} = 2^{-\ell}\begin{pmatrix} \mathbf{V}_{\ell-1} & \mathbf{V}_{\ell-1} \\ \mathbf{V}_{\ell-1} & -\mathbf{V}_{\ell-1}\end{pmatrix}\begin{pmatrix} f_{0\#\dots\#} \\ f_{1\#\dots\#}\end{pmatrix} = \\ &= 2^{-\ell}\begin{pmatrix} 2^{\ell-1}\left[\mathbf{v}_{\ell-1}(f_{0\#\dots\#}) + \mathbf{v}_{\ell-1}(f_{1\#\dots\#})\right] \\ 2^{\ell-1}\left[\mathbf{v}_{\ell-1}(f_{0\#\dots\#}) - \mathbf{v}_{\ell-1}(f_{1\#\dots\#})\right]\end{pmatrix}. \end{aligned}$$

■

Lemma 2.2. *For any fitness function*

$$f(x_0,\dots,x_{\ell-1}) = \sum_{i_0\dots i_{\ell-1}\in\{0,1\}} a_{i_0\dots i_{\ell-1}} x_0^{i_0}\dots x_{\ell-1}^{i_{\ell-1}}$$

on Ω_ℓ *and any* $s = s_0\dots s_{\ell-1}\in\Omega_\ell$, *we have*

$$a_s = (-2)^{u(s)}\sum_{t\in J(s)} v_t$$

resp.

$$v_s = (-1)^{u(s)}\sum_{t\in J(s)} 2^{-u(t)} a_t,$$

where $J(s) = \{t\in\Omega_\ell;\ s_i = 1 \Rightarrow t_i = 1\}$.

Proof. First, consider $s = 1s_1\dots s_{\ell-1} = 1\hat{s}$, then $u(s) = u(\widehat{s}) + 1$. Using (2), lemma 2.1 and an easy induction argument, it follows that:

$$\begin{aligned} a_{1\widehat{s}} &= b_{\widehat{s}} - c_{\widehat{s}} \\ &= (-2)^{u(\widehat{s})}\sum_{t_1\dots t_{\ell-1}\in J(\widehat{s})} v^1_{t_1\dots t_{\ell-1}} - v^0_{t_1\dots t_{\ell-1}} \\ &= (-2)^{u(\widehat{s})}\sum_{t_1\dots t_{\ell-1}\in J(\widehat{s})} (-2) v_{1t_1\dots t_{\ell-1}}(f) \\ &= (-2)^{u(\widehat{s})+1}\sum_{t_1\dots t_{\ell-1}\in J(\widehat{s})} v_{1t_1\dots t_{\ell-1}}(f) \\ &= (-2)^{u(s)}\sum_{t\in J(s)} v_t. \end{aligned}$$

The corresponding expression of $v_{1\hat{s}}$ in terms of a_t may be derived similarly.

Next, consider $s = 0\hat{s}$. In this case $u(s) = u(\hat{s})$, and it follows, in a similar way, that

$$\begin{aligned} v_{0\widehat{s}} &= 2^{-1}[v^0_{\widehat{s}} + v^1_{\widehat{s}}] \\ &= 2^{-1}(-1)^{u(\widehat{s})}\sum_{\widehat{t}\in J(\widehat{s})} 2^{-u(\widehat{t})}\left(c_{\widehat{t}} + b_{\widehat{t}}\right) \\ &= (-1)^{u(s)}\sum_{\widehat{t}\in J(\widehat{s})} 2^{-u(\widehat{t})} c_{\widehat{t}} + (-1)^{u(s)}\sum_{\widehat{t}\in J(\widehat{s})} 2^{-u(\widehat{t})-1}\left(b_{\widehat{t}} - c_{\widehat{t}}\right) \\ &= (-1)^{u(s)}\sum_{\widehat{t}\in J(\widehat{s})} 2^{-u(\widehat{t})} a_{0\widehat{t}} + (-1)^{u(s)}\sum_{\widehat{t}\in J(\widehat{s})} 2^{(-u(\widehat{t})+1)} a_{1\widehat{t}} \\ &= (-1)^{u(s)}\sum_{t\in J(s)} 2^{-u(t)} a_t. \end{aligned}$$

The corresponding expression of $a_{0\hat{s}}$ in terms of v_t may be obtained in a similar way. ■

The following corollary now clearly proves the implication 1) ⇒ 2) in the above theorem:

Corollary 2.3. *For any fitness function $f : \Omega_\ell \to \mathbb{R}$ with polynomial form*

$$f(x_0, \ldots, x_{\ell-1}) = \sum_{i_0 \ldots i_{\ell-1} \in \{0,1\}} a_{i_0 \ldots i_{\ell-1}} x_0^{i_0} \ldots x_{\ell-1}^{i_{\ell-1}},$$

and $1 \le k \le \ell$, the following assertions are equivalent:

1. *$a_s = 0$ for any $s \in \Omega_\ell$ with $u(s) > k$;*
2. *$v_s = 0$ for any $s \in \Omega_\ell$ with $u(s) > k$.*

Proof. This is an immediate consequence of 2.2 and the fact that for any $t \in J(s)$, we have $u(t) \ge u(s)$. ■

Corollary 2.4. *With notations as before, the following assertions are equivalent:*

1. *f is "strictly" of order k, i.e., $a_s = 0$ for any $s \in \Omega_\ell$ with $u(s) > k$ and there exists $t \in \Omega_\ell$ with $u(t) = k$ and $a_t \ne 0$;*
2. *$v_s = 0$ for any $s \in \Omega_\ell$ with $u(s) > k$ and there exists $t \in \Omega_\ell$ with $u(t) = k$ and $v_t \ne 0$.*

Proof. If we suppose 1) with $a_t \ne 0$ and $u(t) = k$, we can use 2.2 and put:

$$a_t = (-2)^k \sum_{z \in J(t)} v_z = (-2)^k \sum_{z \in J(t)\,:\,u(z)=k} v_z,$$

so there exists $z \in J(t)$ with $u(z) = k$ and $v_z \ne 0$, which yields 2). The other implication may be proved similarly. ■

References

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

[2] Goldberg, D., *Genetic Algorithms and Walsh Functions: Part I, A Gentle Introduction,* Complex Systems **3**, 129–152, 1989.

[3] Iglesias, M.T., Naudts, B., Verschoren, A. and Vidal, C. Foundations of Generic Optimization. Volume 1: A Combinatorial approach to epistasis, Springer, Dordrecht, 2005.

[4] Iglesias, M.T., Peñaranda, V. S. and Verschoren, A., *Higher order functions and Walsh coefficients*, Bulletin of the Belgian Mathematical Society Simon Stevin **13**, 633-643, 2006.

[5] Suárez Peñaranda, V.: Epistasis superior, Ph.D. Thesis, Universidade da Coruña, 2006.

Universidade da Coruña
A Coruña, Spain