This paper appeared in Discrete Appl. Math. **160** (2012), 383–390.

THE ARITY GAP OF ORDER-PRESERVING FUNCTIONS AND EXTENSIONS OF PSEUDO-BOOLEAN FUNCTIONS

MIGUEL COUCEIRO, ERKKO LEHTONEN, AND TAMÁS WALDHAUSER

Abstract. The aim of this paper is to classify order-preserving functions according to their arity gap. Noteworthy examples of order-preserving functions are the so-called aggregation functions. We first explicitly classify the Lovász extensions of pseudo-Boolean functions according to their arity gap. Then we consider the class of order-preserving functions between partially ordered sets, and establish a similar explicit classification for this function class.

1. Introduction

In this paper, we study the arity gap of functions of several variables. Essentially, the arity gap of a function $f\colon A^n \to B$ ($n \geq 2$) that depends on all of its variables can be defined as the minimum decrease in the number of essential variables when variables of f are identified. Salomaa [18] showed that the arity gap of any Boolean function is at most 2. This result was extended to functions defined on arbitrary finite domains by Willard [21], who showed that the same upper bound holds for the arity gap of any function $f\colon A^n \to B$, provided that $n > \max(|A|, 3)$. In fact, he showed that if the arity gap of such a function f equals 2, then f is totally symmetric. This line of research culminated into a complete classification of functions $f\colon A^n \to B$ according to their arity gap (see Theorem 2.5), originally presented in [4] in the setting of functions with finite domains; in [6] it was observed that this result holds for functions with arbitrary, possibly infinite domains.

Salomaa's [18] result on the upper bound for the arity gap of Boolean functions mentioned above was strengthened in [3], where Boolean functions were completely classified according to their arity gap. Using tools provided by Berman and Kisielewicz [1] and Willard [21], in [4] a similar explicit classification was established for all pseudo-Boolean functions, i.e., functions $f\colon \{0,1\}^n \to \mathbb{R}$. As it turns out, this leads to analogous classifications of wider classes of functions. In [5], this result on pseudo-Boolean functions was the key step in showing that among lattice polynomial functions only truncated ternary medians have arity gap 2; all the others have arity gap 1.

Similar techniques are used in Section 3 to derive explicit descriptions of the arity gap of well-known extensions of pseudo-Boolean functions to the whole real line, namely, Owen and Lovász extensions.

In Section 4 we consider the arity gap of order-preserving functions. To this extent, we present a complete classification of functions over arbitrary domains according to their arity gap (originally established in [4] for functions over finite domains), which is then used to derive a dichotomy theorem based on the arity gap (and the so-called quasi-arity), and to explicitly determine those order-preserving functions that have arity gap 1 and those that have arity gap 2.

Aggregation functions became a widely studied class of order-preserving functions. Thus, as a by-product of our general results, we obtain an explicit classification of these functions according to their arity gap, which we present in the end of Section 4.

2010 *Mathematics Subject Classification.* 08A40.

Key words and phrases. arity gap, order-preserving function, aggregation function, Owen extension, Lovász extension.

2. Preliminaries: arity gap and the simple minor relation

Throughout this paper, let A and B be arbitrary sets with at least two elements. A *B-valued function (of several variables) on A* is a mapping $f\colon A^n \to B$ for some positive integer n, called the *arity* of f. The A-valued functions on A are called *operations on A.* Operations on $\{0, 1\}$ are called *Boolean functions.* We denote the set of real numbers by $\mathbb{R}$. Functions $f\colon \{0, 1\}^n \to \mathbb{R}$ are referred to as *pseudo-Boolean functions.* For a natural number $n \geq 1$, we denote $[n] = \{1, \dots, n\}$.

The i-th variable is said to be *essential* in $f\colon A^n \to B$, or f is said to *depend* on x_i, if there is a pair

$$((a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n), (a_1, \dots, a_{i-1}, b_i, a_{i+1}, \dots, a_n)) \in A^n \times A^n,$$

called a *witness of essentiality* of x_i in f, such that

$$f(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n) \neq f(a_1, \dots, a_{i-1}, b_i, a_{i+1}, \dots, a_n).$$

The number of essential variables in f is called the *essential arity* of f, and it is denoted by $\operatorname{ess} f$. If $\operatorname{ess} f = m$, we say that f is *essentially m-ary.*

For $n \geq 2$, define

$$A^n_= := \{(a_1, \dots, a_n) \in A^n : a_i = a_j \text{ for some } i \neq j\}.$$

We also define $A^1_= := A$. Note that if A has less than n elements, then $A^n_= = A^n$.

Consider $f\colon A^n \to B$. Any function $g\colon A^n \to B$ satisfying $f|_{A^n_=} = g|_{A^n_=}$ is called a *support* of f. The *quasi-arity* of f, denoted $\operatorname{qa} f$, is defined as the minimum of the essential arities of the supports of f, i.e., $\operatorname{qa} f = \min_g \operatorname{ess} g$, where g ranges over the set of all supports of f. If $\operatorname{qa} f = m$, we say that f is *quasi-m-ary.*

A function $f\colon A^n \to B$ is said to be obtained from $g\colon A^m \to B$ by *simple variable substitution,* or f is a *simple minor* of g, if there is a mapping $\sigma\colon \{1, \dots, m\} \to \{1, \dots, n\}$ such that

$$f(x_1, \dots, x_n) = g(x_{\sigma(1)}, \dots, x_{\sigma(m)}) \qquad \text{for all } (x_1, \dots, x_n) \in A^n.$$

The simple minor relation constitutes a quasi-order $\leq$ on the set of all B-valued functions of several variables on A which is given by the following rule: $f \leq g$ if and only if f is obtained from g by simple variable substitution. If $f \leq g$ and $g \leq f$, we say that f and g are *equivalent,* denoted $f \equiv g$. If $f \leq g$ but $g \not\leq f$, we denote $f < g$. It can be easily observed that if $f \leq g$ then $\operatorname{ess} f \leq \operatorname{ess} g$, with equality if and only if $f \equiv g$. For background, extensions and variants of the simple minor relation, see, e.g., [2, 7, 8, 9, 12, 13, 17, 20, 22].

For $f\colon A^n \to B$, $i, j \in \{1, \dots, n\}$, $i \neq j$, we define $f_{i\leftarrow j}\colon A^n \to B$ to be the simple minor of f given by the substitution of x_j for x_i, that is,

$$f_{i\leftarrow j}(x_1, \dots, x_n) = f(x_1, \dots, x_{i-1}, x_j, x_{i+1}, \dots, x_n).$$

Note that on the right-hand side of the above equality, x_j occurs twice, namely both at the i-th and the j-th positions. We denote

$$\operatorname{ess}^< f = \max_{g<f} \operatorname{ess} g,$$

and we define the *arity gap* of f by $\operatorname{gap} f = \operatorname{ess} f - \operatorname{ess}^< f$. It is easily observed that

$$\operatorname{gap} f = \min_{i \neq j} (\operatorname{ess} f - \operatorname{ess} f_{i\leftarrow j}),$$

where i and j range over the set of indices of essential variables of f.

In the sequel, whenever we consider the arity gap of some function f, we will assume that all variables of f are essential. This is not a significant restriction, because every nonconstant function is equivalent to a function with no inessential variables and equivalent functions have the same arity gap.

Salomaa [18] proved that the arity gap of every Boolean function with at least two essential variables is at most 2. This result was generalized by Willard [21, Lemma 1.2] in the following theorem.

Theorem 2.1. *Let A be a finite set. Suppose $f\colon A^n \to B$ depends on all of its variables. If $n > \max(|A|, 3)$, then* $\operatorname{gap} f \leq 2$.

In [3], Salomaa's result was strengthened into an explicit classification of Boolean functions in terms of arity gap.

Theorem 2.2. *Assume that $f\colon \{0,1\}^n \to \{0,1\}$ depends on all of its variables. We have $\operatorname{gap} f = 2$ if and only if f is equivalent to one of the following Boolean functions:*

- $x_1 \oplus x_2 \oplus \cdots \oplus x_n \oplus c$,
- $x_1 x_2 \oplus x_1 \oplus c$,
- $x_1 x_2 \oplus x_1 x_3 \oplus x_2 x_3 \oplus c$,
- $x_1 x_2 \oplus x_1 x_3 \oplus x_2 x_3 \oplus x_1 \oplus x_2 \oplus c$,

where $\oplus$ denotes addition modulo 2 and $c \in \{0,1\}$. Otherwise $\operatorname{gap} f = 1$.

Based on this, a complete classification of pseudo-Boolean functions according to their arity gap was presented in [4].

Theorem 2.3. *For a pseudo-Boolean function $f\colon \{0,1\}^n \to \mathbb{R}$ which depends on all of its variables, $\operatorname{gap} f = 2$ if and only if f satisfies one of the following conditions:*

- *$n = 2$ and f is a nonconstant function satisfying $f(0,0) = f(1,1)$,*
- *$f = g \circ h$, where $g\colon \{0,1\} \to \mathbb{R}$ is injective and $h\colon \{0,1\}^n \to \{0,1\}$ is a Boolean function with $\operatorname{gap} h = 2$, as listed in Theorem 2.2.*

Otherwise $\operatorname{gap} f = 1$.

Remark 2.4. It is noteworthy that there is a complete one-to-one correspondence between pseudo-Boolean functions and *set functions,* i.e., functions $v\colon 2^{[n]} \to \mathbb{R}$ for some $n \geq 1$. This correspondence is based on the natural order-isomorphism between $\{0,1\}^n$ and the power set $2^{[n]}$ of $[n]$. For a pseudo-Boolean function $f\colon \{0,1\}^n \to \mathbb{R}$ we can associate a set function $v_f\colon 2^{[n]} \to \mathbb{R}$ given by $v_f(T) = f(\mathbf{e}_T)$, where $\mathbf{e}_T$ denotes the characteristic vector of $T \subseteq [n]$. Conversely, for a set function $v\colon 2^{[n]} \to \mathbb{R}$, let $f_v\colon \{0,1\}^n \to \mathbb{R}$ be the pseudo-Boolean function defined by $f_v(\mathbf{e}_T) = v(T)$. Clearly, $f_{v_f} = f$ and $v_{f_v} = v$ for every pseudo-Boolean function $f\colon \{0,1\}^n \to \mathbb{R}$ and every set function $v\colon 2^{[n]} \to \mathbb{R}$.

The study of the arity gap of functions $A^n \to B$ culminated into the characterization presented in Theorem 2.5, originally proved in [4]. We need to introduce some terminology to state the result.

Let 2^A be the power set of A, and define $\operatorname{oddsupp}\colon \bigcup_{n\geq 1} A^n \to 2^A$ by

$$\operatorname{oddsupp}(a_1, \ldots, a_n) = \big\{a_i : |\{j \in [n] : a_j = a_i\}| \text{ is odd}\big\}.$$

A partial function $f\colon S \to B$, $S \subseteq A^n$, is said to be *determined by* oddsupp if $f = f^* \circ \operatorname{oddsupp}|_S$ for some function $f^*\colon 2^A \to B$.

Theorem 2.5. *Suppose that $f\colon A^n \to B$, $n \geq 2$, depends on all of its variables.*

(i) *For $3 \leq p \leq n$, $\operatorname{gap} f = p$ if and only if $\operatorname{qa} f = n - p$.*

(ii) *For $n \neq 3$, $\operatorname{gap} f = 2$ if and only if $\operatorname{qa} f = n - 2$ or $\operatorname{qa} f = n$ and $f|_{A^n_=}$ is determined by* oddsupp.

(iii) *For $n = 3$, $\operatorname{gap} f = 2$ if and only if there is a nonconstant unary function $h\colon A \to B$ and $i_1, i_2, i_3 \in \{0,1\}$ such that*

$$f(x_1, x_0, x_0) = h(x_{i_1}),$$
$$f(x_0, x_1, x_0) = h(x_{i_2}),$$
$$f(x_0, x_0, x_1) = h(x_{i_3}).$$

(iv) *Otherwise* $\operatorname{gap} f = 1$.

Remark 2.6. The notion of a function's being determined by oddsupp is due to Berman and Kisielewicz [1]. Willard [21] showed that if $f\colon A^n \to B$ where A is finite, $n > \max(|A|, 3)$ and $\operatorname{gap} f = 2$, then f is determined by oddsupp.

Remark 2.7. While Theorem 2.5 was originally stated and proved in the setting of functions with finite domains, its proof presented in [4] does not make use of any assumption on the cardinalities of the domain and codomain – as long as they contain at least two elements. Hence the theorem immediately generalizes for functions with arbitrary domains.

3. The arity gap of Lovász and Owen extensions

In this section, we consider well-known extensions of pseudo-Boolean functions and generalize Theorem 2.3 accordingly. For further background on pseudo-Boolean functions, we refer the reader to Hammer and Rudeanu [11].

As is well-known, every pseudo-Boolean function can be uniquely represented by a multilinear polynomial expression. A common way to construct such representations makes use of the notion of "Möbius transform".

Let $v\colon 2^{[n]} \to \mathbb{R}$ be a set function. The *Möbius transform* (or *Möbius inverse*) of v is the map $m_v\colon 2^{[n]} \to \mathbb{R}$ given by

$$m_v(S) = \sum_{T \subseteq S} (-1)^{|S|-|T|} v(T), \quad \text{for all } S \subseteq [n].$$

In view of Remark 2.4, we say that $m\colon 2^{[n]} \to \mathbb{R}$ is the Möbius transform of $f\colon \{0,1\}^n \to \mathbb{R}$ if $m = m_{v_f}$.

Theorem 3.1 ([11]). *Let $f\colon \{0,1\}^n \to \mathbb{R}$ be a pseudo-Boolean function. Then*

$$f(\mathbf{x}) = \sum_{S \subseteq [n]} m_{v_f}(S) \prod_{i \in S} x_i, \quad \textit{for all } \mathbf{x} \in \{0,1\}^n. \tag{1}$$

Remark 3.2. Theorem 3.1 motivates the terminology "Möbius inverse of v" since it implies in particular that for every $S \subseteq [n]$, $v(S) = \sum_{T \subseteq S} m_v(T)$.

The following result is well known and easy to verify (see, e.g., [15] for the case of order-preserving pseudo-Boolean functions).

Lemma 3.3. *Let $f\colon \{0,1\}^n \to \mathbb{R}$ be a pseudo-Boolean function and consider its corresponding set function v_f. If x_i is inessential in f, then $m_{v_f}(S) = 0$ whenever $i \in S$. In particular, f depends on x_i if and only if x_i appears in the multilinear polynomial representation* (1) *of f.*

There are several ways of extending a pseudo-Boolean function $f\colon \{0,1\}^n \to \mathbb{R}$ to a function on $\mathbb{R}$. Perhaps the most natural is the multilinear polynomial extension. The *Owen extension* [16] (or *multilinear extension*) of a pseudo-Boolean function $f\colon \{0,1\}^n \to \mathbb{R}$ is the mapping $P_f\colon \mathbb{R}^n \to \mathbb{R}$ defined by

$$P_f(\mathbf{x}) = \sum_{S \subseteq [n]} m_{v_f}(S) \prod_{i \in S} x_i, \quad \text{for all } \mathbf{x} \in \mathbb{R}^n.$$

Clearly, f coincides with the restriction of P_f to $\{0,1\}^n$.

Another extension of pseudo-Boolean functions to functions on $\mathbb{R}$ is the so-called "Lovász extension". This terminology is due to Singer [19] who refined a result by Lovász [14] concerning convex functions. The *Lovász extension* of a pseudo-Boolean function $f\colon \{0,1\}^n \to \mathbb{R}$ is the mapping $F_f\colon \mathbb{R}^n \to \mathbb{R}$ defined by

$$F_f(\mathbf{x}) = \sum_{S \subseteq [n]} m_{v_f}(S) \bigwedge_{i \in S} x_i, \quad \text{for all } \mathbf{x} \in \mathbb{R}^n.$$

Observe that the Lovász extension of a pseudo-Boolean function f is the unique extension of f which is linear on the "standard simplices"

$$\mathbb{R}^n_\sigma = \{\mathbf{x} \in \mathbb{R}^n : x_{\sigma(1)} \leq x_{\sigma(2)} \leq \cdots \leq x_{\sigma(n)}\},$$

for any permutation σ on $[n]$ (see [10]).

Remark 3.4. The defining expressions of Owen and Lovász extensions differ only in the fact that the connecting operations between variables are the product and the minimum, respectively. In the sequel, this observation can be used to translate the results concerning Lovász extensions into analogous results about Owen extensions.

Remark 3.5. Every function $F\colon \mathbb{R}^n \to \mathbb{R}$ of the form

$$F(\mathbf{x}) = \sum_{S\subseteq[n]} m(S) \bigwedge_{i\in S} x_i, \tag{2}$$

where $m\colon 2^{[n]} \to \mathbb{R}$ is the Lovász extension of a unique pseudo-Boolean function, namely, $f = F|_{\{0,1\}^n}$. Therefore, we shall refer to any map of the form (2) as a *Lovász extension.*

Theorem 3.6. *Let $f\colon \{0,1\}^n \to \mathbb{R}$ be a pseudo-Boolean function. Then the i-th variable is essential in f if and only if the i-th variable is essential in F_f.*

Proof. As observed, f coincides with F_f on $\{0,1\}^n$, and thus if the i-th variable is inessential in F_f, then the i-th variable is inessential in f.

Conversely, if the i-th variable is inessential in f, then by Lemma 3.3 it follows that x_i does not appear in the defining expression of F_f. Hence, the i-th variable is inessential in F_f. □

Corollary 3.7. *Let $f\colon \{0,1\}^n \to \mathbb{R}$ be a pseudo-Boolean function. Then $\operatorname{gap} f = \operatorname{gap} F_f$. In particular, $\operatorname{gap} F_f \leq 2$.*

Using Theorems 2.2 and 2.3, we obtain the following explicit descriptions of those Lovász extensions that have arity gap 2.

Theorem 3.8. *Assume that $F\colon \mathbb{R}^n \to \mathbb{R}$ is a Lovász extension that depends on all of its variables. Then $\operatorname{gap} F = 2$ if and only if F is of one of the following forms:*

(i) $F \equiv \dfrac{a-b}{2} \displaystyle\sum_{S\subseteq[n]} \big((-2)^{|S|} \cdot \bigwedge_{i\in S} x_i\big)$,
(ii) $F \equiv a + (b-a)x_1 + (a-b)(x_1 \wedge x_2)$,
(iii) $F \equiv a + (b-a)\big((x_1 \wedge x_2) + (x_1 \wedge x_3) + (x_2 \wedge x_3)\big) + 2(a-b)(x_1 \wedge x_2 \wedge x_3)$,
(iv) $F \equiv a + (b-a)(x_1 + x_2) + (a-b)\big((x_1 \wedge x_2) + (x_1 \wedge x_3) + (x_2 \wedge x_3)\big)$ $+ 2(b-a)(x_1 \wedge x_2 \wedge x_3)$,
(v) $F \equiv a + (b-a)x_1 + (c-a)x_2 + (2a-b-c)(x_1 \wedge x_2)$,

for some $a,b,c \in \mathbb{R}$. Otherwise $\operatorname{gap} F = 1$.

Note that since F is assumed to depend on all of its variables, for functions of the form (i)–(iv) it holds that $a \neq b$, and for functions of the form (v) it holds that $\{a,b,c\} \neq \{a\}$.

Proof. Let $f\colon \{0,1\}^n \to \mathbb{R}$ be the pseudo-Boolean function determined by F. By Theorems 2.2 and 2.3, $\operatorname{gap} f = 2$ if and only if

(i) $f \equiv (b-a)(x_1 \oplus \cdots \oplus x_n) + a$,
(ii) $f \equiv (b-a)(x_1x_2 \oplus x_1) + a$,
(iii) $f \equiv (b-a)(x_1x_2 \oplus x_1x_3 \oplus x_2x_3) + a$,
(iv) $f \equiv (b-a)(x_1x_2 \oplus x_1x_3 \oplus x_2x_3 \oplus x_1 \oplus x_2) + a$, or
(v) $f\colon \{0,1\}^2 \to \mathbb{R}$ is nonconstant such that $f(0,0) = f(1,1)$, say, $f(0,0) = f(1,1) = a$, $f(1,0) = b$ and $f(0,1) = c$,

where $\oplus$ denotes addition modulo 2, and $a,b,c \in \mathbb{R}$. The theorem now follows by computing the Möbius transform of v_f in each possible case. □

Corollary 3.9. *A nondecreasing Lovász extension $F\colon \mathbb{R}^n \to \mathbb{R}$ has arity gap 2 if and only if*

$$F \equiv a + (b-a)\big((x_1 \wedge x_2) + (x_1 \wedge x_3) + (x_2 \wedge x_3)\big) + 2(a-b)(x_1 \wedge x_2 \wedge x_3). \tag{3}$$

Otherwise $\operatorname{gap} F = 1$.

Techniques similar to those developed in this section were successfully used in [5] to classify the class of lattice polynomial functions, i.e., functions which can be obtained as compositions of the lattice operations and variables (projections) and constants. A well-known example of a lattice polynomial function on a distributive lattice A is the *median function* $\mathrm{med}\colon A^3 \to A$ given by

$$\begin{aligned}\mathrm{med}(x_1, x_2, x_3) &= (x_1 \wedge x_2) \vee (x_1 \wedge x_3) \vee (x_2 \wedge x_3)\\ &= (x_1 \vee x_2) \wedge (x_1 \vee x_3) \wedge (x_2 \vee x_3).\end{aligned}$$

As shown in [5], lattice polynomial functions with arity gap 2 are exactly the truncated median functions.

Theorem 3.10 ([5])**.** *Let $f\colon A^n \to A$ be a lattice polynomial function on a bounded distributive lattice A. Then* $\operatorname{gap} f = 2$ *if and only if*

$$f \equiv (a \vee \mathrm{med}(x_1, x_2, x_3)) \wedge b,$$

for some $a, b \in A$, $a < b$. Otherwise $\operatorname{gap} f = 1$.

In the next section, we extend these results to the more general class of order-preserving maps between possibly different ordered sets A and B.

4. The arity gap of order-preserving functions

Let $(A; \leq)$ be a partially ordered set. We say that $(A; \leq)$ is

- *upwards directed* if every pair of elements of A has an upper bound,
- *downwards directed* if every pair of elements of A has a lower bound,
- *bidirected* if $(A; \leq)$ is both upwards directed and downwards directed,
- *pseudo-directed* if every pair of elements of A has an upper bound or a lower bound.

Remark 4.1. In the above definitions, existence of a least upper bound or a greatest lower bound is not stipulated. Therefore, an upwards (or downwards) directed poset is not the same thing as a semilattice, nor is a bidirected poset the same thing as a lattice. However, every semilattice is either upwards or downwards directed, and every lattice and every bounded poset is bidirected. Moreover, every upwards directed or downwards directed poset is pseudo-directed.

Let $(A; \leq_A)$ and $(B; \leq_B)$ be partially ordered sets. A function $f\colon A^n \to B$ is said to be *order-preserving* (with respect to the partial orders $\leq_A$ and $\leq_B$) if for all $\mathbf{a}, \mathbf{b} \in A^n$, $f(\mathbf{a}) \leq_B f(\mathbf{b})$ whenever $\mathbf{a} \leq_A \mathbf{b}$, where $\mathbf{a} \leq_A \mathbf{b}$ denotes the componentwise ordering of tuples, i.e., $\mathbf{a} \leq_A \mathbf{b}$ if and only if $a_i \leq_A b_i$ for all $i \in \{1, \dots, n\}$.

Lemma 4.2. *Let $(A; \leq_A)$ be a pseudo-directed poset, and let $f\colon A^n \to B$ be a function. If x_i is essential in f then there are elements $a_1, \dots, a_n, b_i \in A$ such that $a_i <_A b_i$ and*

$$f(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n) \neq f(a_1, \dots, a_{i-1}, b_i, a_{i+1}, \dots, a_n).$$

Moreover, if B is partially ordered by $\leq_B$ and f is order-preserving with respect to $\leq_A$ and $\leq_B$, then

$$f(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n) <_B f(a_1, \dots, a_{i-1}, b_i, a_{i+1}, \dots, a_n).$$

Proof. Since x_i is essential in f, there exist elements $a_1, \dots, a_{i-1}, a', b', a_{i+1}, \dots, a_n \in A$ such that

$$f(a_1, \dots, a_{i-1}, a', a_{i+1}, \dots, a_n) \neq f(a_1, \dots, a_{i-1}, b', a_{i+1}, \dots, a_n).$$

By the assumption that $(A; \leq)$ is pseudo-directed, a' and b' have an upper bound or a lower bound. Assume first that a' and b' have an upper bound c. We clearly have that

$$f(a_1, \dots, a_{i-1}, a', a_{i+1}, \dots, a_n) \neq f(a_1, \dots, a_{i-1}, c, a_{i+1}, \dots, a_n) \quad \text{or} \tag{4}$$

$$f(a_1, \dots, a_{i-1}, b', a_{i+1}, \dots, a_n) \neq f(a_1, \dots, a_{i-1}, c, a_{i+1}, \dots, a_n). \tag{5}$$

The claim thus follows by choosing $b_i := c$ and $a_i := a'$ if (4) holds or $a_i := b'$ if (5) holds.

Otherwise a' and b' have a lower bound, and a similar argument shows that the claim holds also in this case.

If f is order-preserving with respect to $\leq_A$ and $\leq_B$, then we have in fact that

$$f(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n) <_B f(a_1, \dots, a_{i-1}, b_i, a_{i+1}, \dots, a_n). \qquad \square$$

Lemma 4.3. *Let $(A; \leq_A)$ be a bidirected poset, let $(B; \leq_B)$ be any poset, and let $f\colon A^n \to B$ $(n \geq 2)$ be an order-preserving function that depends on all of its variables. Then, for all $i, j \in \{1, \dots, n\}$ $(i \neq j)$, x_j is essential in $f_{i \leftarrow j}$. Furthermore, if $i < j$, then there exist elements $c, d, a_1, \dots, a_n \in A$ such that $c <_A d$ and*

$$(6) \quad f(a_1, \dots, a_{i-1}, c, a_{i+1}, \dots, a_{j-1}, c, a_{j+1}, \dots, a_n) <_B f(a_1, \dots, a_{i-1}, d, a_{i+1}, \dots, a_{j-1}, d, a_{j+1}, \dots, a_n).$$

Proof. Assume, without loss of generality, that $i = 1$, $j = 2$. Since x_1 is essential in f, by Lemma 4.2 there exist elements $a_1, \dots, a_n, b_1 \in A$ such that $a_1 <_A b_1$ and $f(a_1, a_2, \dots, a_n) <_B f(b_1, a_2, \dots, a_n)$. By the assumption that $(A; \leq)$ is bidirected, there exist a lower bound c of a_1 and a_2 and an upper bound d of b_1 and a_2. Again, by the monotonicity of f,

$$f_{1 \leftarrow 2}(a_1, c, a_3, \dots, a_n) = f(c, c, a_3, \dots, a_n) \leq_B f(a_1, a_2, a_3, \dots, a_n) <_B f(b_1, a_2, a_3, \dots, a_n) \leq_B f(d, d, a_3, \dots, a_n) = f_{1 \leftarrow 2}(a_1, d, a_3, \dots, a_n),$$

which shows that x_2 is essential in $f_{1 \leftarrow 2}$ and inequality (6) holds. $\square$

Proposition 4.4. *Let $(A; \leq_A)$ be a bidirected poset, let $(B; \leq_B)$ be any poset, and let $f\colon A^n \to B$ $(n \geq 2)$ be an order-preserving function that depends on all of its variables. Then* $\operatorname{qa} f \geq n - 1$ *and $f|_{A^{n}_{=}}$ is not determined by* oddsupp.

Proof. Suppose first, on the contrary, that $\operatorname{qa} f = n - p$ for some $p \geq 2$. Let g be a support of f with essential arity $n - p$. Then g has at least two inessential variables, say x_i and x_j, and these variables are clearly inessential in $g_{i \leftarrow j}$ as well. But, since $f_{i \leftarrow j} = g_{i \leftarrow j}$, this constitutes a contradiction to Lemma 4.3 which asserts that x_j is essential in $f_{i \leftarrow j}$.

Suppose then, on the contrary, that $f|_{A^{n}_{=}}$ is determined by oddsupp. Then $f|_{A^{n}_{=}} = f^* \circ \operatorname{oddsupp}$ for some $f^*\colon 2^A \to B$. We clearly have that for all c, $d, a_3, \dots, a_n \in A$, $\operatorname{oddsupp}(c, c, a_3, \dots, a_n) = \operatorname{oddsupp}(d, d, a_3, \dots, a_n)$ (note that $(c, c, a_3, \dots, a_n), (d, d, a_3, \dots, a_n) \in A^{n}_{=}$); hence $f(c, c, a_3, \dots, a_n) = f(d, d, a_3, \dots, a_n)$. This contradicts Lemma 4.3. $\square$

Proposition 4.5. *Let $(A; \leq_A)$ be a bidirected poset, let $(B; \leq_B)$ be any poset, and let $f\colon A^3 \to B$ be an order-preserving function that depends on all of its variables. Then* $\operatorname{gap} f = 2$ *if and only if there is a nonconstant order-preserving unary function $h\colon A \to B$ such that*

$$f(x_1, x_0, x_0) = f(x_0, x_1, x_0) = f(x_0, x_0, x_1) = h(x_0).$$

Proof. By Theorem 2.5, the condition is sufficient. For necessity, assume that $\operatorname{gap} f = 2$. Then, by Theorem 2.5, there is a nonconstant unary function $h\colon A \to B$ and $i_1, i_2, i_3 \in \{0, 1\}$ such that

$$f(x_1, x_0, x_0) = h(x_{i_1}), \quad f(x_0, x_1, x_0) = h(x_{i_2}), \quad f(x_0, x_0, x_1) = h(x_{i_3}).$$

We claim that $i_1 = i_2 = i_3 = 0$. Suppose, on the contrary, that $i_1 = 1$. By Lemma 4.3, there exist elements $a, b, c \in A$ such that $b <_A c$ and $f(a, b, b) <_B f(a, c, c)$, but this is a contradiction to $f(a, b, b) = h(a) = f(a, c, c)$. Similarly, we can derive a contradiction from the assumption that $i_2 = 1$ or $i_3 = 1$.

The monotonicity of h follows from the monotonicity of f. For, if $a \leq_A b$, then

$$h(a) = f(a, a, a) \leq_B f(b, b, b) = h(b). \qquad \square$$

Theorem 4.6. *Let $(A; \leq_A)$ be a bidirected poset, let $(B; \leq_B)$ be any poset, and let $f\colon A^n \to B$ ($n \geq 2$) be an order-preserving function that depends on all of its variables. Then* $\operatorname{gap} f = 2$ *if and only if $n = 3$ and there is a nonconstant order-preserving unary function $h\colon A \to B$ such that*

$$f(x_1, x_0, x_0) = f(x_0, x_1, x_0) = f(x_0, x_0, x_1) = h(x_0).$$

Otherwise $\operatorname{gap} f = 1$.

Proof. Immediate consequence of Theorem 2.5 and Propositions 4.4 and 4.5. □

By imposing stronger assumptions on the underlying posets, we obtain more stringent descriptions of order-preserving functions with arity gap 2.

Lemma 4.7. *Let $(A; \leq_A)$ and $(B; \leq_B)$ be lattices, and let $h\colon A \to B$ be a lattice homomorphism. Let $f\colon A^3 \to B$ be an order-preserving function such that*

$$f(x_1, x_0, x_0) = f(x_0, x_1, x_0) = f(x_0, x_0, x_1) = h(x_0).$$

If the homomorphic image of $(A; \leq_A)$ by h is a distributive sublattice of $(B; \leq_B)$, then $f = \operatorname{med}\bigl(h(x_1), h(x_2), h(x_3)\bigr)$, where med *denotes the ternary median function on* $\operatorname{Im} h$.

Proof. By the monotonicity of f and the assumption that A is a lattice, we have that for all $a_1, a_2, a_3 \in A$,

$$\begin{aligned} h(a_1 \wedge a_2) = f(a_1 \wedge a_2, a_1 \wedge a_2, a_3) &\leq f(a_1, a_2, a_3) \\ &\leq f(a_1 \vee a_2, a_1 \vee a_2, a_3) = h(a_1 \vee a_2). \end{aligned}$$

A similar argument shows that for all $i, j \in \{1, 2, 3\}$, we have

$$h(a_i \wedge a_j) \leq f(a_1, a_2, a_3) \leq h(a_i \vee a_j).$$

By the assumption that B is a lattice, it follows from the above inequalities that

$$\begin{aligned} h(a_1 \wedge a_2) \vee h(a_2 \wedge a_3) \vee h(a_1 \wedge a_3) &\leq f(a_1, a_2, a_3) \\ &\leq h(a_1 \vee a_2) \wedge h(a_2 \vee a_3) \wedge h(a_1 \vee a_3). \end{aligned}$$

Since h is a lattice homomorphism, we have that

$$\begin{aligned} &h(a_1 \wedge a_2) \vee h(a_2 \wedge a_3) \vee h(a_1 \wedge a_3) \\ &\quad = \bigl(h(a_1) \wedge h(a_2)\bigr) \vee \bigl(h(a_2) \wedge h(a_3)\bigr) \vee \bigl(h(a_1) \wedge h(a_3)\bigr), \end{aligned} \tag{7}$$

$$\begin{aligned} &h(a_1 \vee a_2) \wedge h(a_2 \vee a_3) \wedge h(a_1 \vee a_3) \\ &\quad = \bigl(h(a_1) \vee h(a_2)\bigr) \wedge \bigl(h(a_2) \vee h(a_3)\bigr) \wedge \bigl(h(a_1) \vee h(a_3)\bigr). \end{aligned} \tag{8}$$

By the assumption that $\operatorname{Im} h$ is a distributive sublattice of B, the right-hand sides of (7) and (8) are equal, and they are actually equal to $\operatorname{med}\bigl(h(a_1), h(a_2), h(a_3)\bigr)$. We conclude that $f(a_1, a_2, a_3) = \operatorname{med}\bigl(h(a_1), h(a_2), h(a_3)\bigr)$. □

Corollary 4.8. *Let $(A; \leq_A)$ be a chain and let $(B; \leq_B)$ be any lattice. Let $f\colon A^n \to B$ be an order-preserving function. Then* $\operatorname{gap} f = 2$ *if and only if $n = 3$ and $f = \operatorname{med}\bigl(h(x_1), h(x_2), h(x_3)\bigr)$ for some nonconstant order-preserving unary function $h\colon A \to B$ (here* med *denotes the median function on* $\operatorname{Im} h$*). Otherwise* $\operatorname{gap} f = 1$.

Proof. If $f = \operatorname{med}\bigl(h(x_1), h(x_2), h(x_3)\bigr)$, where h is as described in the statement, then clearly $\operatorname{gap} f = 2$. For the converse implication, assume that $\operatorname{gap} f = 2$. By Theorem 4.6, $n = 3$ and there is a nonconstant order-preserving unary function $h\colon A \to B$ such that

$$f(x_1, x_0, x_0) = f(x_0, x_1, x_0) = f(x_0, x_0, x_1) = h(x_0).$$

Since every order-preserving function h is a lattice homomorphism from a chain A to any lattice B and the homomorphic image of A by h is a chain and hence a distributive sublattice of B, it follows from Lemma 4.7 that $f = \operatorname{med}(h(x_1), h(x_2), h(x_3))$.

The last claim follows from Theorem 4.6, which asserts that $\operatorname{gap} f \leq 2$. □

To illustrate the use of the results obtained in this section, we present an alternative proof of Theorem 3.10.

Proof of Theorem 3.10. It is well-known that lattice polynomial functions are order-preserving. Therefore Theorem 4.6 applies, and $\operatorname{gap} f \leq 2$. Assume, without loss of generality, that $\operatorname{ess} f = n$. Suppose that $\operatorname{gap} f = 2$. Then, by Theorem 4.6, $n = 3$ and there is a nonconstant order-preserving unary function $h \colon A \to A$ such that

$$f(x_1, x_0, x_0) = f(x_0, x_1, x_0) = f(x_0, x_0, x_1) = h(x_0).$$

Since f is a polynomial function, h is a polynomial function as well, and hence $h(x) = (a \vee x) \wedge b$ for some $a, b \in A$, $a < b$. In particular, h is a lattice homomorphism. Since A is a distributive lattice, $\operatorname{Im} h$ is a distributive sublattice of A, and Lemma 4.7 then implies that

$$f = \operatorname{med}\bigl(h(x_1), h(x_2), h(x_3)\bigr) = h\bigl(\operatorname{med}(x_1, x_2, x_3)\bigr).$$

Clearly, if f has the above form, then $\operatorname{gap} f = 2$. Since $\operatorname{gap} f \leq 2$, the last claim of the theorem follows. ☐

As mentioned, the class of order-preserving functions includes the noteworthy class of aggregation functions. Traditionally, an *aggregation function* on a closed real interval $[a, b] \subseteq \mathbb{R}$ is defined as a mapping $M \colon [a, b]^n \to [a, b]$ which is nondecreasing and fulfills the boundary conditions $M(a, \dots, a) = a$ and $M(b, \dots, b) = b$. From Corollary 4.8, we obtain the following.

Corollary 4.9. *Let $M \colon [a, b]^n \to [a, b]$ be an aggregation function on a real interval $[a, b]$. Then $\operatorname{gap} M = 2$ if and only if $n = 3$ and*

$$M = \operatorname{med}\bigl(h(x_1), h(x_2), h(x_3)\bigr)$$

for some nonconstant order-preserving unary function $h \colon [a, b] \to [a, b]$ satisfying $h(a) = a$, $h(b) = b$. Otherwise $\operatorname{gap} f = 1$.

Acknowledgements

The third author acknowledges that the present project is supported by the National Research Fund, Luxembourg, and cofunded under the Marie Curie Actions of the European Commission (FP7-COFUND), and supported by the Hungarian National Foundation for Scientific Research under grant no. K77409.

We would like to thank the anonymous reviewer for constructive remarks which helped in improving the present manuscript.

References

[1] J. Berman, A. Kisielewicz, On the number of operations in a clone, Proc. Amer. Math. Soc. 122 (1994) 359–369.

[2] M. Couceiro, On the lattice of equational classes of Boolean functions and its closed intervals, J. Mult.-Valued Logic Soft Comput. 18 (2008) 81–104.

[3] M. Couceiro, E. Lehtonen, On the effect of variable identification on the essential arity of functions on finite sets, Int. J. Found. Comput. Sci. 18 (2007) 975–986.

[4] M. Couceiro, E. Lehtonen, Generalizations of Świerczkowski's lemma and the arity gap of finite functions, Discrete Math. 309 (2009) 5905–5912.

[5] M. Couceiro, E. Lehtonen, The arity gap of polynomial functions over bounded distributive lattices, 40th IEEE International Symposium on Multiple-Valued Logic (ISMVL 2010), IEEE Computer Society, Los Alamitos, 2010, pp. 113–116.

[6] M. Couceiro, E. Lehtonen, T. Waldhauser, Decompositions of functions based on arity gap, arXiv:1003.1294.

[7] M. Couceiro, M. Pouzet, On a quasi-ordering on Boolean functions, Theoret. Comput. Sci. 396 (2008) 71–87.

[8] O. Ekin, S. Foldes, P.L. Hammer, L. Hellerstein, Equational characterizations of Boolean function classes, Discrete Math. 211 (2000) 27–51.

[9] A. Feigelson, L. Hellerstein, The forbidden projections of unate functions, Discrete Appl. Math. 77 (1997) 221–236.

[10] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, Aggregation Functions, Encyclopedia of Mathematics and Its Applications, vol. 127, Cambridge University Press, Cambridge, 2009.

[11] P.L. Hammer, S. Rudeanu, Boolean Methods in Operations Research and Related Areas, Springer-Verlag, Berlin, 1968.
[12] E. Lehtonen, Descending chains and antichains of the unary, linear, and monotone subfunction relations, Order 23 (2006) 129–142.
[13] E. Lehtonen, Á. Szendrei, Equivalence of operations with respect to discriminator clones, Discrete Math. 309 (2009) 673–685.
[14] L. Lovász, Submodular function and convexity. In: A. Bachem, M. Grötschel, B. Korte (eds.), Mathematical programming. The state of the art. Bonn 1982, Springer-Verlag, Berlin–Heidelberg–New York–Tokyo, 1983, pp. 235–257.
[15] T. Murofushi, M. Sugeno, A theory of fuzzy measures: representations, the Choquet integral, and null sets, J. Math. Anal. Appl. 159 (1991) 532–549.
[16] G. Owen, Multilinear extensions of games, Management Science 18 (1972) 64–79.
[17] N. Pippenger, Galois theory for minors of finite functions, Discrete Math. 254 (2002) 405–419.
[18] A. Salomaa, On essential variables of functions, especially in the algebra of logic, Ann. Acad. Sci. Fenn. Ser. A I. Math. 339 (1963) 3–11.
[19] I. Singer, Extensions of functions of 0–1 variables and applications to combinatorial optimization, Numer. Funct. Anal. Optim. 7 (1985) 23–62.
[20] C. Wang, Boolean minors, Discrete Math. 141 (1991) 237–258.
[21] R. Willard, Essential arities of term operations in finite algebras, Discrete Math. 149 (1996) 239–259.
[22] I.E. Zverovich, Characterizations of closed classes of Boolean functions in terms of forbidden subfunctions and Post classes, Discrete Appl. Math. 149 (2005) 200–218.

(M. Couceiro) Mathematics Research Unit, University of Luxembourg, 6, rue Richard Coudenhove-Kalergi, L–1359 Luxembourg, Luxembourg
E-mail address: `miguel.couceiro@uni.lu`

(E. Lehtonen) Computer Science and Communications Research Unit, University of Luxembourg, 6, rue Richard Coudenhove-Kalergi, L–1359 Luxembourg, Luxembourg
E-mail address: `erkko.lehtonen@uni.lu`

(T. Waldhauser) University of Luxembourg, 6, rue Richard Coudenhove-Kalergi, L–1359 Luxembourg, Luxembourg and Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H–6720 Szeged, Hungary
E-mail address: `twaldha@math.u-szeged.hu`