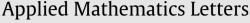
Applied Mathematics Letters 25 (2012) 739-744



Contents lists available at SciVerse ScienceDirect



journal homepage: www.elsevier.com/locate/aml



(2)

# On the boolean partial derivatives and their composition

# A. Martín del Rey<sup>a,\*</sup>, G. Rodríguez Sánchez<sup>b</sup>, A. de la Villa Cuenca<sup>c</sup>

<sup>a</sup> Department of Applied Mathematics, E.P.S. de Ávila, Universidad de Salamanca, C/ Hornos Caleros 50, 05003-Ávila, Spain

<sup>b</sup> Department of Applied Mathematics, E.P.S. de Zamora, Universidad de Salamanca, Avda. Requejo 33, 49022-Zamora, Spain

<sup>c</sup> Department of Applied Mathematics and Computation, E.T.S.I. (ICAI) Universidad Pontificia Comillas, C/ Alberto Aguilera 23, 28015-Madrid, Spain

#### ARTICLE INFO

Article history: Received 26 May 2011 Received in revised form 12 October 2011 Accepted 13 October 2011

*Keywords:* Boolean functions Boolean derivative

#### 1. Introduction and preliminaries

Let  $\mathbb{F}_2^n$  be the *n*-dimensional vector space over the Galois field  $\mathbb{F}_2 = \{0, 1\}$ , and set  $\{e_1, \ldots, e_n\}$  as its standard basis, that is,

$$e_1 = (1, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1).$$
<sup>(1)</sup>

For two vectors  $x = (x_1, ..., x_n) \in \mathbb{F}_2^n$  and  $y = (y_1, ..., y_n) \in \mathbb{F}_2^n$ , we can define the XOR addition operation as follows:

$$x \oplus y = (x_1 \oplus y_1, \ldots, x_n \oplus y_n) \in \mathbb{F}_2^n$$
.

An *n*-variable boolean function is a map of the form  $f : \mathbb{F}_2^n \to \mathbb{F}_2$ . The set of all *n*-variable boolean functions is denoted by  $\mathscr{BF}_n$  and its cardinality is  $|\mathscr{BF}_n| = 2^{2^n}$ . The vector

$$t_f = (f(v_0), f(v_1), \dots, f(v_{2^n - 1})) \in \mathbb{F}_2^{2^n},$$
(3)

where  $v_0 = (0, ..., 0)$ ,  $v_1 = (0, ..., 0, 1)$ , ...,  $v_{2^n-1} = (1, ..., 1)$ , is called the truth table of f. Note that for  $1 \le i \le 2^n - 1$ ,  $v_i$  is the binary representation of i written as a vector of length  $2^n$ .

The usual representation of a boolean function f is by means of its algebraic normal form (ANF for short) which is the n-variable polynomial representation over  $\mathbb{F}_2$ , that is,

$$f(x_1, \dots, x_n) = a_0 \oplus \bigoplus_{\substack{1 \le k \le n \\ 1 \le i_1, i_2, \dots, i_k \le n}} a_{i_1, i_2, \dots, i_k}, x_{i_1}, x_{i_2}, \dots, x_{i_k},$$
(4)

where  $a_0, a_{i_1,...,i_k} \in \mathbb{F}_2$ . The degree of the ANF is the algebraic degree of the function. The simplest boolean functions, considering their ANF, are the affine boolean functions:  $f(x_1, ..., x_n) = a_0 \oplus a_1 x_1 \oplus a_2 x_2 \oplus \cdots \oplus a_n x_n$ , where  $a_0, a_1, ..., a_n \in \mathbb{F}_2$ . If  $a_0 = 0$ , we have the linear boolean functions and they are denoted by  $l_a(x)$  with  $a = (a_1, ..., a_n) \in \mathbb{F}_2^n$ .

#### ABSTRACT

The main goal of this work is to introduce the relation between the partial boolean derivatives of an *n*-variable boolean function and their directional boolean derivatives.  $\$   $\$  2011 Elsevier Ltd. All rights reserved.

<sup>\*</sup> Corresponding author. Tel.: +34 920 353500x3785; fax: +34 920 353501.

E-mail addresses: delrey@usal.es (A. Martín del Rey), gerardo@usal.es (G. Rodríguez Sánchez), avilla@dmc.icai.upcomillas.es (A. de la Villa Cuenca).

<sup>0893-9659/\$ –</sup> see front matter s 2011 Elsevier Ltd. All rights reserved. doi:10.1016/j.aml.2011.10.013

The Hamming weight of a boolean vector x is denoted by wt(x) and is defined as the number of ones in the vector x. In this sense, the Hamming weight of a boolean function f is the Hamming weight of its truth table  $t_f$ . An n-variable boolean function f is said to be balanced if its weight is exactly  $2^{n-1}$ , that is, if the number of ones equals the number of zeros of its truth table.

# 2. The partial derivative of a boolean function

The notion of the boolean derivative was introduced by Vichniac (see [1]) and it is defined as follows:

**Definition 1.** The partial derivative of an *n*-variable boolean function f with respect to the *i*th variable  $x_i$  is another *n*-variable boolean function,  $D_i f$ , defined as follows:

$$D_i f: \mathbb{F}_2^n \to \mathbb{F}_2$$
  
$$x \mapsto D_i f(x) = f(x) \oplus f(x \oplus e_i), \qquad (5)$$

that is,

$$D_{i}f(x) = f(x_{1}, \dots, x_{i}, \dots, x_{n}) \oplus f(x_{1}, \dots, x_{i} \oplus 1, \dots, x_{n}).$$
(6)

The notion of the boolean derivative is very important and useful in, for example, cryptography (see [2]).

**Example 1.** Let us consider the four-variable boolean function whose ANF is  $f(x_1, x_2, x_3, x_4) = 1 \oplus x_3 \oplus x_1x_2 \oplus x_2x_3 \oplus x_1x_2x_3 \oplus x_2x_3x_4$ ; then, as simple CALCULATIONS show,  $D_1f(x_1, x_2, x_3, x_4) = x_2 \oplus x_2x_3$ .

This definition allows one to state a derivation rule similar to the derivation rule for multivariate polynomials over real numbers:

**Lemma 1.** Let f be an n-variable boolean function whose ANF is (4). Then for each variable  $x_i$  we have

$$f(x) = g_i(x \oplus x_i e_i) \oplus x_i h_i(x \oplus x_i e_i),$$
(7)

where  $h_i$  and  $g_i$  are (n - 1)-variable boolean functions which do not depend on the variable  $x_i$ . Moreover, if f does not depend on the variable  $x_i$  then  $h_i = 0$ .

**Proof.** Set  $1 \le i \le n$ ; then the *n*-variable boolean function *f* can be factored by taking the common factor  $x_i$ , and consequently *f* can be written as follows:

$$f(\mathbf{x}) = g_i(\mathbf{x}_1, \dots, \widehat{\mathbf{x}_i}, \dots, \mathbf{x}_n) \oplus \mathbf{x}_i h_i(\mathbf{x}_1, \dots, \widehat{\mathbf{x}_i}, \dots, \mathbf{x}_n),$$
(8)

where  $g_i$  and  $h_i$  are (n - 1)-variable boolean functions which do not depend on  $x_i$ . For the sake of simplicity we set  $x \oplus x_i e_i = (x_1, \ldots, \hat{x_i}, \ldots, x_n)$ ; then,

$$f(\mathbf{x}) = g_i(\mathbf{x} \oplus \mathbf{x}_i e_i) \oplus \mathbf{x}_i h_i(\mathbf{x} \oplus \mathbf{x}_i e_i),$$
(9)

thus finishing the proof.  $\Box$ 

**Example 2.** Let us consider the four-variable boolean function whose ANF is  $f(x_1, x_2, x_3, x_4) = x_1 \oplus x_2 x_3 \oplus x_3 x_4 \oplus x_2 x_3 x_4$ . Then,

$$f(x_1, x_2, x_3, x_4) = x_2 x_3 \oplus x_3 x_4 \oplus x_2 x_3 x_4 \oplus x_1$$
  
=  $x_1 \oplus x_3 x_4 \oplus x_2 (x_3 \oplus x_3 x_4)$   
=  $x_1 \oplus x_3 (x_2 \oplus x_4 \oplus x_2 x_4)$   
=  $x_1 \oplus x_2 x_3 \oplus x_4 (x_3 \oplus x_2 x_3)$ . (10)

**Proposition 1.** Let f be an n-variable boolean function. Then,

$$D_{if}(x) = h_{i}(x \oplus x_{i}e_{i}).$$
<sup>(11)</sup>

**Proof.** By definition,  $D_i f(x) = f(x) \oplus f(x \oplus e_i)$ , and taking into account the last lemma, this yields

$$D_{i}f(x) = g_{i}(x \oplus x_{i}e_{i}) \oplus x_{i}h_{i}(x \oplus x_{i}e_{i}) \oplus g_{i}(x \oplus x_{i}e_{i}) \oplus (x_{i} \oplus 1)h_{i}(x \oplus x_{i}e_{i}) = h_{i}(x \oplus x_{i}e_{i}),$$
(12)

thus finishing the proof.  $\Box$ 

As a consequence, the partial derivative (with respect to one variable) reduces the algebraic degree of the boolean function by 1.

**Example 3.** The partial derivatives of the four-variable boolean function  $f(x_1, x_2, x_3, x_4) = x_1 \oplus x_2 x_3 \oplus x_3 x_4 \oplus x_2 x_3 x_4$  with respect to its four variables are

$$D_1 f(x) = 1,$$

$$D_2 f(x) = x_2 \oplus x_2 x_3.$$
(13)
(14)

$$D_{2j}(x) = x_3 \oplus x_3 x_4,$$

$$D_{3f}(x) = x_2 \oplus x_4 \oplus x_2 x_4,$$
(15)

$$D_4f(x) = x_3 \oplus x_2 x_3. \tag{16}$$

**Definition 2.** The composition of partial derivatives of an n-variable boolean function f with respect to the *i*th and *j*th variables is defined as follows:

$$D_{i} \circ D_{j} f(x) = D_{i} (D_{j}f) (x) = D_{j}f(x) \oplus D_{j}f(x \oplus e_{i})$$
  
=  $f(x) \oplus f(x \oplus e_{j}) \oplus f(x \oplus e_{i}) \oplus f(x \oplus e_{i} \oplus e_{j}).$  (17)

Furthermore,

(

$$\left(D_{i_1} \circ D_{i_2} \circ \cdots \circ D_{i_k}\right) f(x) = D_{i_1}\left(D_{i_2}\left(\cdots D_{i_k}\left(f\right)\right)\right)(x).$$

$$\tag{18}$$

Consequently, the following result holds:

Proposition 2. The composition of boolean partial derivatives commutes:

$$\left(D_i \circ D_j\right) f\left(x\right) = \left(D_j \circ D_i\right) f\left(x\right).$$
(19)

**Example 4.** Set the four-variable boolean function  $f(x_1, x_2, x_3, x_4) = x_1 \oplus x_2 x_3 \oplus x_3 x_4 \oplus x_2 x_3 x_4$ ; then,

$$(D_2 \circ D_3) f(x) = D_2 (x_2 \oplus x_4 \oplus x_2 x_4) = 1 \oplus x_4,$$

$$(20)$$

$$(D_3 \circ D_2)f(x) = D_3(x_3 \oplus x_3 x_4) = 1 \oplus x_4.$$
(21)

We can extend the notion of a partial derivative to that of a directional derivative as follows:

**Definition 3.** The directional derivative of the *n*-variable boolean function f with respect to  $b \in \mathbb{F}_2^n$  is another *n*-variable boolean function defined as follows:

$$D_b f: \mathbb{F}_2^n \to \mathbb{F}_2$$
  
$$x \mapsto D_b f(x) = f(x) \oplus f(x \oplus b).$$
(22)

Note that if wt(b) = k then  $b \in \mathbb{F}_2^n$  has k non-zero coefficients placed at positions  $1 \le i_1 < \cdots < i_k \le n$ ; consequently  $b = e_{i_1} \oplus \cdots \oplus e_{i_k} \in \mathbb{F}_2^n$ . In this sense, for the sake of simplicity, we take

$$D_{b}f(x) = D_{e_{i_{1}} \oplus \dots \oplus e_{i_{k}}}f(x) = D_{i_{1},\dots,i_{k}}f(x).$$
(23)

**Example 5.** Set the four-variable boolean function  $f(x_1, x_2, x_3, x_4) = x_1 \oplus x_2 x_3 \oplus x_3 x_4 \oplus x_2 x_3 x_4$  and b = (0, 1, 1, 0); then,

$$D_{b}f(x) = D_{23}f(x) = f(x) \oplus f(x \oplus b)$$
  
=  $x_{1} \oplus x_{2}x_{3} \oplus x_{3}x_{4} \oplus x_{2}x_{3}x_{4} \oplus x_{1}(x_{2} \oplus 1)(x_{3} \oplus 1) \oplus (x_{3} \oplus 1)x_{4} \oplus (x_{2} \oplus 1)(x_{3} \oplus 1)x_{4}$   
=  $1 \oplus x_{2} \oplus x_{3} \oplus x_{2}x_{4} \oplus x_{3}x_{4}.$  (24)

#### 3. The relation between partial and directional boolean derivatives

The relation between partial boolean derivatives and directional boolean derivatives is stated in the following.

**Theorem 1.** Let *f* be an *n*-variable boolean function and set

$$1 \le i_1 < i_2 < \dots < i_k \le n \tag{25}$$

with  $k \leq n$ ; then,

$$(D_{i_1} \circ \dots \circ D_{i_k}) f(x) = \bigoplus_{\substack{1 \le l \le k \\ j_1 < \dots < j_l \\ j_1, \dots, j_l \in \{i_1, \dots, i_k\}}} D_{j_1, \dots, j_l} f(x) .$$
 (26)

**Proof.** We prove the statement by induction on *k*. For k = 2 the formula is true:

$$(D_{i_1} \circ D_{i_2})f(x) = D_{i_2}f(x) \oplus D_{i_2}f(x \oplus e_{i_1})$$

$$= f(x) \oplus f(x \oplus e_{i_2}) \oplus f(x \oplus e_{i_1}) \oplus f(x \oplus e_{i_2} \oplus e_{i_1})$$

$$= f(x) \oplus (D_{i_2}f(x) \oplus f(x)) \oplus (D_{i_1}f(x) \oplus f(x)) \oplus (D_{e_{i_1} \oplus e_{i_2}}f(x) \oplus f(x))$$

$$= D_{i_1}f(x) \oplus D_{i_2}f(x) \oplus D_{i_1,i_2}f(x) .$$

$$(27)$$

For k = 3, the statement of the proposition is also true:

$$\begin{pmatrix} D_{i_1} \circ D_{i_2} \circ D_{i_3} \end{pmatrix} f(x) = \begin{pmatrix} D_{i_1} \circ (D_{i_2} \circ D_{i_3}) \end{pmatrix} f(x) = \begin{pmatrix} D_{i_2} \circ D_{i_3} \end{pmatrix} f(x) \oplus \begin{pmatrix} D_{i_2} \circ D_{i_3} \end{pmatrix} f(x \oplus e_{i_1}) = D_{i_2}f(x) \oplus D_{i_3}f(x) \oplus D_{i_2,i_3}f(x) \oplus D_{i_2}f(x \oplus e_{i_1}) \oplus D_{i_3}f(x \oplus e_{i_1}) \oplus D_{i_2,i_3}f(x \oplus e_{i_1}) \\ = D_{i_2}f(x) \oplus D_{i_3}f(x) \oplus D_{i_2,i_3}f(x) \oplus f(x \oplus e_{i_1}) \oplus f(x \oplus e_{i_1} \oplus e_{i_2}) \\ \oplus f(x \oplus e_{i_1}) \oplus f(x \oplus e_{i_1} \oplus e_{i_3}) \oplus f(x \oplus e_{i_1}) \oplus f(x \oplus e_{i_1} \oplus e_{i_2} \oplus e_{i_3}) \\ = D_{i_2}f(x) \oplus D_{i_3}f(x) \oplus D_{i_2,i_3}f(x) \oplus D_{i_1}f(x) \oplus D_{i_1,i_2}f(x) \oplus D_{i_1}f(x) \oplus D_{i_1,i_3}f(x) \\ \oplus D_{i_1}f(x) \oplus D_{i_1,i_2,i_3}f(x) \\ = D_{i_1}f(x) \oplus D_{i_2}f(x) \oplus D_{i_3}f(x) \oplus D_{i_1,i_2}f(x) \oplus D_{i_1,i_3}f(x) \oplus D_{i_1,i_2,i_3}f(x).$$
 (28)

Assume that the statement is true for  $k \le m$ ; then for k = m + 1,

$$\left(D_{i_1}\circ\cdots\circ D_{i_{m+1}}\right)f(x)=\left(D_{i_1}\circ\left(D_{i_2}\circ\cdots\circ D_{i_{m+1}}\right)\right)f(x).$$
(29)

Now, by induction, this yields

$$\left(D_{i_1}\circ\cdots\circ D_{i_{m+1}}\right)f(x) = \left(D_{i_2}\circ\cdots\circ D_{i_{m+1}}\right)f(x) \oplus \left(D_{i_2}\circ\cdots\circ D_{i_{m+1}}\right)f(x\oplus e_{i_1}). \tag{30}$$

As

$$(D_{i_2} \circ \dots \circ D_{i_{m+1}}) f (x \oplus e_{i_1}) = \bigoplus_{\substack{1 \le l \le m \\ j_1 < \dots < j_l \\ j_1 \dots , j_l \in \{l_2, \dots, l_{m+1}\}}} D_{j_1, \dots, j_l} f (x \oplus e_{i_1})$$

$$= \bigoplus_{\substack{1 \le l \le m \\ j_1 \dots , j_l \in \{l_2, \dots, l_{m+1}\}}} f (x \oplus e_{i_1}) \oplus f \left( x \oplus e_{i_1} \oplus e_{i_{j_1}} \oplus \dots \oplus e_{i_{j_l}} \right)$$

$$= \bigoplus_{\substack{1 \le l \le m \\ j_1 \dots , j_l \in \{l_2, \dots, l_{m+1}\}}} D_{i_1} f (x) \oplus D_{i_1, j_1, \dots, j_l} f (x) ,$$

then

$$\begin{pmatrix} D_{i_1} \circ \dots \circ D_{i_{m+1}} \end{pmatrix} f(x) = (D_{i_2} \circ \dots \circ D_{i_{m+1}}) f(x) \oplus (D_{i_2} \circ \dots \circ D_{i_{m+1}}) f(x \oplus e_{i_1}) \\ = \bigoplus_{\substack{1 \le l \le m \\ j_1 < \dots < j_l \\ j_1 < \dots < j_l \in [l_2, \dots, l_{m+1}]}} D_{j_1, \dots, j_l} f(x) \oplus \bigoplus_{\substack{1 \le l \le m \\ j_1 < \dots < j_l \\ j_1 < \dots < j_l \in [l_2, \dots, l_{m+1}]}} D_{i_1} f(x) \oplus D_{i_1, j_1, \dots, j_l} f(x)$$

$$= \bigoplus_{\substack{1 \le l \le m+1 \\ j_1 < \dots < j_l \\ j_1 < \dots < j_l \in [l_1, \dots, l_m]}} D_{j_1, \dots, j_l} f(x) ,$$

$$(31)$$

thus finishing the proof.  $\Box$ 

**Example 6.** Let us consider the four-variable boolean function  $f(x_1, x_2, x_3, x_4) = x_1 \oplus x_2 x_3 \oplus x_3 x_4 \oplus x_2 x_3 x_4$ ; then,

$$(D_2 \circ D_3) f (x) = D_2 (x) \oplus D_3 (x) \oplus D_{23} (x) = (x_3 \oplus x_3 x_4) \oplus (x_2 \oplus x_4 \oplus x_2 x_4) \oplus (1 \oplus x_2 \oplus x_3 \oplus x_2 x_4 \oplus x_3 x_4) = 1 \oplus x_4.$$
(32)

Corollary 1. Taking into account Theorem 1, simple calculus shows the following results:

1. The directional derivative can be given in terms of the composition of partial derivatives as follows:

$$D_{i_1,\dots,i_k}f(\mathbf{x}) = \bigoplus_{\substack{1 \le l \le k \\ j_1 < \dots < j_l \\ j_1,\dots,j_l \in \{i_1,\dots,i_k\}}} \left( D_{j_1} \circ \dots \circ D_{j_l} \right) f(\mathbf{x}) \,.$$

$$(33)$$

2. The directional derivative reduces the algebraic degree of the boolean function to be applied by, at least, 1.

3. If k = n, then

$$(D_1 \circ \dots \circ D_n) f(x) = \bigoplus_{b \in \mathbb{F}_2^n} D_b f(x) .$$
(34)

4. If  $\sigma$  is a permutation of *n* elements, then

$$(D_1 \circ \dots \circ D_n) f(x) = (D_{\sigma(1)} \circ \dots \circ D_{\sigma(n)}) f(x).$$
(35)

### 4. Cryptographic applications

The boolean derivative plays an important role in the study of cryptographic properties of boolean functions such as the strict avalanche criterion (SAC for short) and its generalization: the propagation criterion (PC for brevity), and the non-existence of non-linear structures.

It is well-known (see, for example, [2]) that  $f \in \mathscr{BF}_2^n$  is SAC if and only if f(x) changes with probability 0.5 whenever a single input bit of x is complemented, that is, if changing only one bit in the input yields the output being changed for exactly  $2^{n-2}$  vectors with the changed input bit. The characterization of SAC is terms of boolean derivatives was introduced in [3] and it is as follows: f is SAC if  $D_b f$  is a balanced n-variable boolean function for every  $b \in \mathbb{F}_2^n$  such that wt(b) = 1, that is, using the notation introduced in Section  $1, f \in \mathscr{BF}_2^n$  is SAC if  $D_b f$  is a balanced function for  $1 \le i \le n$ .

The notion of SAC can be generalized to the concept of a propagation criterion as follows: f satisfies the propagation criterion of degree k (f is PC(k) for short) if f (x) changes with a probability of 0.5 whenever i ( $1 \le i \le k$ ) bits of the input x are complemented, *i.e.*, if changing any i ( $1 \le i \le k$ ) of the n bits in the input results in the output being changed for exactly  $2^{n-2}$  vectors with the changed input bit. As in the previous case, a characterization in terms of derivatives was stated as follows: A boolean function  $f \in \mathscr{BF}_2^n$  is PC(k) if and only if  $D_{i_1,...,i_m}f$  are balanced functions for  $1 \le m \le k$ .

Furthermore, the following are very interesting results (see [2]):

Lemma 2. Every affine boolean function is a balanced boolean function.

**Lemma 3.** If f is an n-variable boolean function which is PC(k) for some k,  $1 \le k \le n$ , then so is  $f \oplus g$ , where g is any affine function in n variables.

An *n*-variable boolean function is said to be a bent function if it is PC(n). Bent functions have been extensively studied for their applications in cryptography since they are as different as possible from all linear and affine functions. Such functions have also been applied in areas such as the spread spectrum, coding theory, and combinatorial design.

Then, taking into account Theorem 1 and the result 1 from Corollary 1, we obtain:

Proposition 3. An n-variable boolean function is a bent function if

$$\bigoplus_{\substack{1 \le l \le k \\ j_1 < \cdots < j_l \\ j_1, \cdots, j_l \in \{i_1, \dots, i_k\}}} (D_{j_1} \circ \cdots \circ D_{j_l}) f(x)$$

are balanced functions for  $1 \le k \le n$ .

#### 5. Conclusions

In this work the notion of a boolean derivative is revisited, stating a derivation rule similar to the traditional derivation rule for a multivariate polynomial over  $\mathbb{R}$ . Moreover, the relation between the composition of boolean derivatives with respect to one variable and the directional derivatives with respect to a suitable boolean vector is shown. In this sense, some applications of this theoretical result to the characterization of cryptographic boolean functions are stated. Future work will aim at studying more detailed cryptographic implications of this result in relation to the strict avalanche criterion, propagation criterion and extended propagation criterion.

(36)

#### Acknowledgment

This work was supported by Ministerio de Ciencia e Innovación (Spain) under grant MTM2008-02773.

## References

- G.Y. Vichniac, Boolean derivatives on cellular automata, Physica D 45 (1999) 63–74.
   T.W. Cusick, P. Stanica, Cryptographic Boolean Functions and Applications, Academic Press, Oxford, 2009.
   A.F. Webster, S.E. Tavares, On the design of S-boxes, in: Advances in Cryptology, Proc. of Crypto'85, in: Lect. Notes Comput. Sci., vol. 219, 1985, pp. 523–534.