On derivatives of polynomials over finite fields through integration

E. Pasalic[a,1,*], A. Muratović-Ribić[b,1], S. Hodzić[c,1], S. Gangopadhyay[d,1]

[a] *University of Primorska, FAMNIT & IAM, Glagoljaska 6, 6000 Koper, Slovenia*
[b] *University of Sarajevo, Department of Mathematics, Zmaja od Bosne 33-35, 71000 Sarajevo, Bosnia and Herzegovina*
[c] *University of Primorska, FAMNIT, Glagoljaska 6, 6000 Koper, Slovenia*
[d] *Department of Computer Science and Engineering, Indian Institute of Technology Roorkee, INDIA*

Abstract

In this note, using rather elementary technique and the derived formula that relates the coefficients of a polynomial over a finite field and its derivative, we deduce many interesting results related to derivatives of Boolean functions and derivatives of mappings over finite fields. For instance, we easily identify several infinite classes of polynomials which cannot possess linear structures. The same technique can be applied for deducing a nontrivial upper bound on the degree of so-called planar mappings.

Keywords: Finite fields, Boolean functions, Derivatives, Integration, Linear structures.

1. Introduction

Let $\mathbb{F}_q$ denote the Galois field of order $q = p^n$, and let the corresponding vector space be denoted as $\mathbb{F}_p^n$. For a given polynomial $F(x) \in \mathbb{F}_q[x]$ its derivative at $a \in \mathbb{F}_q^*$ is defined as $D_aF(x) = F(x+a) - F(x)$, where clearly $a = 0$ results in a trivial annihilation. In contrast to the standard notion of derivative, which is for instance useful for determination of multiple roots of F and which coincides to the derivation of polynomials over real numbers, this notion of derivatives is of great importance in cryptography and is directly related to differential properties of the mappings used in the substitution boxes. Indeed, when $p = 2$ the differential properties of F (that reflects the resistance to differential cryptanalysis [1]) are characterized by the number of solutions of $F(x+a) + F(x) = b$ for any $a \in \mathbb{F}_q^*$ and $b \in \mathbb{F}_q$. On the other hand, for fields of odd prime characteristic $p > 2$, if $F(x+a) - F(x)$ is a permutation for any nonzero a then F is called a planar function [6, 4, 5].

*Corresponding author

Email addresses: enes.pasalic6@gmail.com (E. Pasalic), amela@pmf.unsa.ba (A. Muratović-Ribić), samir_hodzic89@hotmail.com (S. Hodzić), gsugata@gmail.com (S. Gangopadhyay)

[1] The telephone number of the corresponding author: +386 (5) 611 75 70; the telefax number is: +386 (5) 611 75 71.

The concept of linear structures plays an important role in cryptographic applications. Certainly, for functions over finite fields (whose prime field is binary) the substitution boxes (S-boxes) identified as a polynomial $F(x) \in \mathbb{F}_{2^n}[x]$, represented as $F(x) = \sum_{i=0}^{q-1} b_i x^i$, should not contain linear structures a so that $F(x+a) + F(x) = b$ for some fixed $b \in \mathbb{F}_{2^n}$ and for all $x \in \mathbb{F}_{2^n}$. In this case a is called b-linear structure. A few general results are known about the form of polynomials $F(x)$ admitting linear structures. The same applies to the Boolean case when $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$ which again may be represented as $f(x) = \sum_{i=0}^{2^n-1} a_i x^i$ but the coefficients a_i must satisfy certain conditions, see Section 2. In [10], the properties of the set of differential functions defined as $\mathcal{DF}_q = \{D_a F(x) : F(x) \in F_q[x], a \in \mathbb{F}_q^*\}$ was investigated. One should notice that there exist polynomials in $\mathbb{F}_q[x]$ which are not derivatives of any polynomial, thus they do not belong to $\mathcal{DF}_q$. The main result in [10] concerning the existence of linear structures is that $F(x) \in \mathbb{F}_{2^n}[x]$ is a differential function (thus $F(x) \in \mathcal{DF}_q$) if and only if it has a 0-linear structure. This implies that the necessary condition to avoid linear structures is that $F(x) \notin \mathcal{DF}_q$, for $q = 2^n$. In [2], the authors investigated the existence of linear structures for the mappings of the form $F(x) = Tr(\delta x^s)$, where $F : \mathbb{F}_{p^n} \to \mathbb{F}_p$. For polynomials over finite fields a thorough treatment of binomials $F(x) = x^s + \alpha x^d$ was taken in [3].

A detailed study of the cryptanlytic significance of linear structures was initiated by Evertse [7] in which cryptanysis of DES like ciphers are discussed along with several possible extensions. Linear structures are also considered by Nyberg and Knudsen in a paper on provable security against a differential attack [8]. The connection between the existence of linear structures and the differential profile of functions over finite fields is an important area of investigation in the context of the designs of S-boxes. The relevance of this area has increased significantly due to the recent cryptographic need of development of S-boxes (vectorial Boolean functions) suitable for use in lightweight ciphers.

To sum up the critical technological impact of this area of research we refer to the foreword written by Bart Preneel in the recent book by Tokareva [9] which is entirely devoted to bent functions. Preneel writes: "Perhaps the largest impact on modern cryptography to date would be generated by the study of generalizations to vector Boolean functions that offer strong resistance against differential and linear attacks by Nyberg and others. This work resulted in the S-box used in the Advanced Encryption Standard (AES) that is today used in billions of devices." Incidentally bent functions are Boolean functions having no linear structures whose cryptographic applications include employment in the designs of CAST, Grain and HAVAL, as well as "non-cryptographic" uses in the designs of Hadamard matrices, strongly regular graphs, Kerdock codes and CDMA sequences.

In this article we firstly derive the relationship between the coefficients b_i of $F(x) = \sum_{i=0}^{q-1} b_i x^i$ and the coefficients c_i of its derivative $G(x) = F(x+a) - F(x) = \sum_{i=0}^{q-2} c_i x^i$. This connection can be efficiently used for specifying conditions regarding the existence of linear structures for either Boolean functions or for mappings over finite fields. Though the approach is quite elementary it leads to several important results in this direction. For instance, it is sufficient that $F(x)$ contains the highest polynomial degree term x^{q-1}

so that F does not admit linear structures, which when translated into the domain of Boolean functions corresponds to a class of functions of highest algebraic degree. Noticing that any n-variable Boolean function can also be represented as a univariate polynomial $f(x) = \sum_{i=0}^{q-1} b_i x^i \in \mathbb{F}_{2^n}[x]$, where the coefficients b_i satisfy certain conditions, we apply the same technique to either mappings over finite fields or to Boolean mappings. While the linear structures of monomials and binomials are quite easy to handle, in general the existence of linear structures for arbitrary polynomials is harder to analyze. Nevertheless, we provide a few interesting results in this direction covering also some particular cases when F contains an arbitrary number of terms. Finally, using the same technique we provide a nontrivial upper bound on the degree of planar mappings.

This article is organized as follows. Some basic definitions and notions are given in Section 2. In Section 3, some general results (based on the derived connection between a given function and its derivative) related to the existence of linear structures for polynomials over finite fields and for Boolean functions are presented. In Section 4, a nontrivial upper bound on the degree of planar mappings is derived. Some concluding remarks are given in Section 5.

2. Preliminaries

Let $\mathbb{F}_2 = \{0, 1\}$ denote the binary field of characteristic two. Furthermore, let $\mathbb{F}_{2^n}$ denote the Galois field of order 2^n and $\mathbb{F}_2^n$ be its corresponding vector space (once the basis is fixed). Any function from $\mathbb{F}_2^n$ or $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ is called an n variable Boolean function, and the set of all Boolean functions in n variables is denoted by $\mathcal{B}_n$. The algebraic normal form (ANF) of a Boolean function, f on $\mathbb{F}_2^n$ is a multivariate polynomial in $x_1, \ldots, x_n$,

$$f(x_1, \ldots, x_n) = \sum_{\mathbf{a} \in \mathbb{F}_2^n} \mu_{\mathbf{a}} \prod_{i=1}^{n} x_i^{a_i}, \text{ where } \mu_{\mathbf{a}} \in \mathbb{F}_2.$$

The *algebraic degree* of $f \in \mathcal{B}_n$, denoted by $\deg(f)$, is defined as $\max\{wt(\mathbf{a}) : \mu_{\mathbf{a}} \neq 0, \mathbf{a} \in \mathbb{F}_{2^n}\}$, where $wt(\mathbf{a})$ denotes the Hamming weight of a binary vector $\mathbf{a}$.

For the purpose of this paper another equivalent representation of Boolean functions is also of interest. The univariate representation of Boolean functions $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$ is given as,

$$f(x) = \sum_{i=0}^{2^n-1} a_i x^i, \quad a_i \in \mathbb{F}_{2^n}, \tag{1}$$

where the coefficients $a_i \in \mathbb{F}_{2^n}$ satisfy the following (Boolean conditions): $a_0, a_{2^n-1} \in \mathbb{F}_2$ and $a_{2i \pmod{2^n-1}} = a_i^2$ for $i = 1, \ldots, 2^n - 2$, due to the condition $f(x)^2 \equiv f(x) \pmod{x^{2^n} - x}$. Consequently, using the univariate representation we formally do not distinguish between $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ and a Boolean mapping $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$. Also, the polynomial degree of $F(x) = \sum_{i=0}^{q-1} b_i x^i$ is the largest i for which $b_i \neq 0$.

The *derivative* of $f \in \mathfrak{B}_n$ at $a \in \mathbb{F}_{2^n}$, denoted by $D_a f$, is a Boolean function defined by

$$D_a f(x) = f(x + a) + f(x), \text{ for all } x \in \mathbb{F}_{2^n}.$$

Accordingly, an element $a \in \mathbb{F}_{2^n}^*$ is called a linear structure of f if $f(x+a)+f(x) = const. \in \mathbb{F}_2$, for any $x \in \mathbb{F}_{2^n}$.

3. Linear structures and derivatives

Throughout this article we write $F(x) = \sum_{i=0}^{q-1} b_i x^i$ and $D_{F,a}(x) = F(x+a)-F(x) = G(x) = \sum_{i=0}^{q-2} c_i x^i$, where $b_i, c_i \in \mathbb{F}_q$ and $a \in \mathbb{F}_q^*$, for $q = p^n$. Thus, given $D_{F,a}(x)$ specified by the known coefficients c_i our goal is to recover the values of b_i (or possibly a set of different polynomials $\{F\}$) so that the derivative of F at a corresponds to $G(x)$. For convenience, we sometimes write,

$$F(x) = \sum_{i=0}^{q-1} b_i x^i = \sum_{\substack{i=1 \\ i \neq p^j; 0 \leq j < n-1}}^{q-1} b_i x^i + \left(b_0 + \sum_{j=0}^{n-1} b_{p^j} x^{p^j}\right) = F^*(x) + A(x), \tag{2}$$

where $A(x) = b_0 + \sum_{j=0}^{n-1} b_{p^j} x^{p^j}$ denotes an affine polynomial in $\mathbb{F}_q[x]$. Also, $A(x) = b_0 + L(x)$, where L is a linearized polynomial. Furthermore, denote by $\mathcal{L}_q$ and $\mathcal{A}_q$ the sets of all linearized and affine polynomials over $\mathbb{F}_q$, respectively, where $q = p^n$ and $p > 2$. Since for any $G, H \in \mathbb{F}_q[x]$ we have $D_{G+H,a}(x) = D_{G,a}(x) + D_{H,a}(x)$, then $D_{F,a}(x) = D_{F^*,a}(x) + D_{A,a}(x) = D_{F^*,a}(x) + L(a)$ due to the fact that $D_{A,a}(x) = L(a)$.

In general, for a given $a \in \mathbb{F}_q^*$ and $G(x)$ the coefficients b_i such that

$$F(x+a) - F(x) = G(x) \qquad \text{for all } x \in \mathbb{F}_q,$$

can be easily derived. Namely, using

$$F(x+a) - F(x) = \sum_{i=0}^{q-1} b_i[(x+a)^i - x^i] =$$

$$\sum_{i=0}^{q-1} b_i \left[\sum_{t=0}^{i} \binom{i}{t} x^t a^{i-t} - x^i\right] = \sum_{i=0}^{q-1} b_i \left[\sum_{t=0}^{i-1} \binom{i}{t} a^{i-t} x^t\right] =$$

$$\sum_{t=0}^{q-2} \left[\sum_{i=t+1}^{q-1} \binom{i}{t} a^{i-t} b_i\right] x^t,$$

the following equations relating a, b_i and c_t is valid

$$c_t = \sum_{i=t+1}^{q-1} \binom{i}{t} a^{i-t} b_i, \quad \text{for } t = 0, 1, \ldots, q-2. \tag{3}$$

The set of equations can be written as

$$\begin{array}{lllll}
\binom{1}{0} a b_1 + & \binom{2}{0} a^2 b_2 + & \ldots + & \binom{q-1}{0} a^{q-1} b_{q-1} & = c_0 \\
& \binom{2}{1} a b_2 + & \ldots + & \binom{q-1}{1} a^{q-2} b_{q-1} & = c_1 \\
& & \ddots & \vdots & \vdots \\
& & \binom{q-2}{q-3} a b_{q-2} + & \binom{q-1}{q-3} a^2 b_{q-1} & = c_{q-3} \\
& & & \binom{q-1}{q-2} a b_{q-1} & = c_{q-2}.
\end{array} \tag{4}$$

In particular, if $q = p$ then all the diagonal coefficients are of the form $\binom{k}{k-1}a = ka$, for $k = 1, 2, \ldots, p-1$, and since these are nonzero the system has a unique solution.

For $q = p^n$ and $n > 1$, we have $\binom{p^u}{t} \equiv 0$, for all $t \neq 0, p^u$. Furthermore, on the main diagonal we have the coefficients $\binom{k}{k-1}a = ka \equiv 0 \mod p$, for all $k = ps$, where $s = 0, 1, \ldots, \frac{q}{p} - 1$. The last p equations of the above system are of the form:

$$\begin{array}{lllll}
\binom{q-p}{q-p-1}ab_{q-p} & +\binom{q-p+1}{q-p-1}a^2b_{q-p+1} & +\binom{q-p+2}{q-p-1}a^3b_{q-p+2}+ & \cdots & +\binom{q-1}{q-p-1}a^pb_{q-1} = c_{q-p-1} \\
\binom{q-p+1}{q-p}ab_{q-p+1} & +\binom{q-p+2}{q-p}a^2b_{q-p+2}+ & \cdots & +\binom{q-1}{q-p}a^{p-1}b_{q-1} = c_{q-p} \\
 & \ddots & & \vdots & \vdots \\
 & & \binom{q-2}{q-3}ab_{q-2} & +\binom{q-1}{q-3}a^2b_{q-1} & = c_{q-3} \\
 & & & \binom{q-1}{q-2}ab_{q-1} & = c_{q-2}
\end{array}$$

The last $p-1$ equations can be uniquely solved for $b_{q-1}, \ldots, b_{q-p+1}$ recursively, but the first equation has to be a linear combination of the last $p-1$ equations, as $\binom{q-p}{q-p-1} \equiv 0 \pmod p$. Therefore, the coefficient c_{q-p-1} depends on a and on the coefficients $c_{q-2}, \ldots, c_{q-p}$ and furthermore b_{q-p} is free due to the fact that $\binom{q-p}{q-p-1} \equiv 0 \pmod p$. This also implies that the derivative $G(x)$ cannot be arbitrary due to this restriction on c_{q-p-1}. Similarly, by considering the last $2p$ equations of the system, the fact that $\binom{q-2p}{q-2p-1} \equiv 0$ implies that b_{q-2p} is free. Since the diagonal coefficient with b_{q-p} is zero, we can choose b_{q-p} to be arbitrary but fixed and evaluate uniquely the coefficients $b_{q-p-1} \ldots, b_{q-2p+1}$, but again c_{q-2p-1} will depend on a and on $c_{q-2}, \ldots, c_{q-2p}$. The same reasoning applies if we take p more equations.

In general, on the diagonal we have $\binom{sp}{sp-1} \equiv 0$, for $s = 0, 1, \ldots, \frac{q}{p} - 1$, and thus the coefficients b_{sp} are free (can be chosen arbitrary) but the corresponding equations are linear combinations of the equations below so the coefficient c_{sp-1} is not arbitrary but it is determined by this linear combination, i.e., with a and c_k where $k > sp - 1$. Note that the system has q/p free coefficients and therefore $q^{(q/p)}$ distinct solutions $F(x)$. On the other hand, given arbitrary $G(x)$ there may not exist any function $F(x)$ such that $G(x)$ is its derivative for some $a \in \mathbb{F}_q$. The reason for this is that q/p coefficients in $G(x)$ are determined by other coefficients.

3.1. Some preliminary results using integration formula

It is of interest to investigate whether the differentiation of two polynomials whose difference is not an affine polynomial can give rise to same derivatives for different values of a. We firstly treat the case when the derivative a is fixed and show that for a given derivative $G(x)$ there is a unique polynomial (up to addition of affine polynomial) $F(x)$.

Proposition 1. *Let $G(x) = \sum_{i=0}^{q-2} c_i x^i$ be a given derivative of some $F(x) \in \mathbb{F}_q[x]$ for a fixed $a \in \mathbb{F}_q^*$. Let $F^*(x) \in \mathbb{F}_q[x]$ be one solution to $F(x+a) - F(x) = G(x)$ and denote by*

$$\mathcal{F}(A) = \{F^*(x) + A(x) : A(x) \in \mathcal{A}_n\}.$$

Now for any $F_1(x) \notin \mathcal{F}(A)$, we have $F_1(x+a) - F_1(x) \neq G(x)$.

PROOF. Let $F^*(x) = \sum_{i=0}^{q-1} b_i x^i$ and $F_1(x) = \sum_{i=0}^{q-1} b'_i x^i \notin \mathcal{F}(A)$. Since $F_1(x) \notin \mathcal{F}(A)$ then there exists b'_i such that $i \not\equiv 0 \pmod p$ and $b_i \neq b'_i$. Let i be the largest such integer satisfying $b_i \neq b'_i$. Then, we necessarily have $b_j = b'_j$ for all $i < j \leq q-1$ and by cancelling the equal terms in the triangular system above, the condition $\binom{i}{i-1} a b_i = c_{i-1} = \binom{i}{i-1} a b'_i$ implies $b_i = b'_i$, a contradiction. □

In our analysis we have assumed that a is known, but if this is not the case the same analysis can be preformed for any $a \in \mathbb{F}_q^*$ thus providing $(q/p-1)q$ many solutions for each a. Providing that $p \not| \deg(F)$, these solutions are however distinct as shown below.

Proposition 2. *For a given function $F(x)$ such that $p \not| \deg(F)$, the condition $F(x+a) - F(x) = F(x+a') - F(x)$ implies $a = a'$, unless $b_i = 0$ for all $i \not\equiv 0 \pmod p$.*

PROOF. Assume $\deg(F) = m$. The largest nonzero coefficient of both $D_{F,a}(x)$ and $D_{F,a'}(x)$ being c_{m-1}, we have

$$\binom{m}{m-1} a b_m = c_{m-1} = \binom{m}{m-1} a' b_m.$$

Since $\binom{m}{m-1} = m \not\equiv 0$ it immediately follows $a = a'$. Now assuming that $F(x+a) - F(x) = F(x+a') - F(x)$ for $a \neq a'$, then $b_m = 0$ and in general $b_i = 0$ for all $i \not\equiv 0 \pmod p$. □

Corollary 1. *If the field is of prime order then $F(x+a) - F(x) = F(x+a') - F(x)$, that is, $F(x+a) = F(x+a')$, implies $a = a'$.*

Notice that in the case $q = p$ the system has a unique solution for any $a \in \mathbb{F}_p^*$, thus Corollary 1 implies that all the solutions are distinct.

Corollary 2. *Let $L(x)$ be a linearized polynomial over $\mathbb{F}_q$ such that $L(a) = L(a')$. Then $L(x+a) - L(x) = L(x+a') - L(x)$. In particular if $L(x)$ is a permutation over $\mathbb{F}_q$ then $L(x+a) - L(x) = L(x+a') - L(x)$ if and only of $a = a'$.*

Open Problem 1. *It would be of interest to show whether for two polynomials $F(x)$ and $F'(x)$, related through $F(x) \neq F'(x) + A(x)$, we may have $D_{F,a}(x) = D_{F',b}(x)$ for some $b \neq a$.*

In what follows we use a well-known result concerning the parity due to James W. L. Glaisher (also referred to as Luca's theorem).

Theorem 1. *Let n and k be two non-negative integers. Then,*

$$\binom{n}{k} \equiv \begin{cases} 0 \mod 2 & \text{if } n \text{ is even and } k \text{ is odd} \\ \binom{\lfloor n/2 \rfloor}{\lfloor k/2 \rfloor} \mod 2 & \text{otherwise.} \end{cases} \tag{5}$$

In general, $\binom{n}{k} \equiv 0 \pmod p$ as soon as $n_i < d_i$ for at least one i, so that

$$\binom{n}{k} \not\equiv 0 \pmod p \text{ if and only if } k \preccurlyeq n.$$

3.2. Linear structures of mappings over finite fields and Boolean functions

Obviously, the easiest way of applying the above result in the context of determining the existence of linear structures is to study sparse polynomials over finite fields. Notice that a linear structure $a \in \mathbb{F}_{2^n}$ of $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$ means that $F(x+a) + F(x) = \gamma$ for all $x \in \mathbb{F}_{2^n}$ and some constant element $\gamma \in \mathbb{F}_{2^n}$. Furthermore, using the above notation, it is equivalent to saying that $c_i = 0$ for all $i \in [1, 2^n - 1]$ and $c_0 = \gamma$. In the case of monomials of the form $F(x) = b_r x^r$ we have the following result.

Theorem 2. *Let $F(x) = b_r x^r$ be a non-zero monomial, where $F(x) \in \mathbb{F}_{2^n}[x]$ and $1 \leq r \leq 2^n - 1$. Then, a is a non-zero linear structure of F if and only if $r = 2^i$ for some $i \in [0, n-1]$.*

Proof. If $r = 2^i$ so that $F(x) = b_{2^i} x^{2^i}$, then $F(x+a) + F(x) = b_{2^i} a^{2^i}$. Thus, any a is a linear structure of F. Conversely, assume that a is a linear structure of $F(x) = b_r x^r$ and consider

$$c_{r-k} = \binom{r-k+1}{r-k} a b_{r-k+1} + \ldots + \binom{r}{r-k} a^k b_r,$$

for some $1 \leq k \leq r$. Since b_r is the only nonzero b_i, we have $c_{r-k} = \binom{r}{r-k} a^k b_r$. Now if a is a linear structure, then $c_{r-k} = 0$ for all $k \in [1, r-1]$. Consequently, $\binom{r}{r-k} \equiv 0 \mod 2$ for these values of k. Especially, for $k = 1$ we have $\binom{r}{r-1} \equiv 0 \mod 2$ implying that r is even. Then, assuming $r > 2$, the condition that $\binom{r}{r-2} \equiv \binom{r/2}{r/2-1} \equiv 0 \mod 2$ (corresponding to c_{r-2}) implies that $r/2$ is even. Continuing this way, for any $k = 2^i$ we necessarily have that $r/2^i$ is even. Let $r = \sum_{j=0}^{n-1} r_j 2^j$ be the 2-adic representation of r and assume that v is the largest j such $r_j = 1$, thus $r_v = 1$ and $r_j = 0$ for $j > v$. Since k ranges from 1 to r, taking $k = 2^{v-1}$ implies that $r/2^{v-1}$ is also even. It means that $2^v \mid r$ and therefore r is of the form 2^v. □

A similar analysis can be performed for the case of binomials of the form $F(x) = x^d + ux^e$, but this has already been done in [3]where it was proved that $F(x)$ cannot have linear structures unless F is affine.

Remark 1. *For the Boolean case, when $p = 2$ and $F, G : \mathbb{F}_{2^n} \to \mathbb{F}_2$, the coefficients of both F and G must satisfy the Boolean conditions mentioned in the introduction.*

In what follows, we derive some interesting results regarding the polynomial form of $F : \mathbb{F}_{2^n} \to \mathbb{F}_2$ regarding linear structures.

It is well-known that the presence of the highest degree term in the ANF of F, corresponding to the term x^{2^n-1}, implies unbalancedness of F (the converse is of course not true). This means that specifying $b_{2^n-1} = 1$ the function F is unbalanced and we show that in this particular case any such F cannot have linear structures. Assuming that a is a nonzero linear structure of F satisfying $b_{2^n-1} = 1$, then $c_{2^n-2} = 0$ and (3) gives for $t = 2^n - 2$,

$$c_{2^n-2} = \binom{q-1}{q-2} a b_{q-1} = a \cdot 1 = 0,$$

which then implies $a = 0$, a contradiction.

Theorem 3. *Let $F(x) = \sum_{i=0}^{q-1} b_i x^i$, $q = 2^n$, where $F : \mathbb{F}_2^n \to \mathbb{F}_2$ so that the coefficients of F satisfy the Boolean conditions. If $b_{q-1} = 1$ so that F is necessarily unbalanced, since its ANF contains the term $x_1 x_2 \cdots x_n$, then any such F does not admit linear structures.*

The importance of this result lies in the fact that any balanced Boolean function with good cryptographic properties apart from possibly having linear structures can easily be transformed into (just slightly) unbalanced function which does not possess linear structures. Moreover, the algebraic degree is then optimized.

Remark 2. *It is known that if a is all-one linear structure, that is $F(x+a)+F(x) = 1$, then F (which is Boolean) is necessarily balanced since the relation $F(x+a) = F(x)+1$ means that F takes an equal number of ones and zeros. Nevertheless, the unbalancedness of F in Theorem 3, through the term x^{2^n-1}, also excludes all-zero linear structures.*

Let us proceed our investigation for the special case of potentially balanced functions F, thus requiring that $b_{q-1} = 0$. In this case, $ab_{q-1} = c_{q-2} = 0$ does not lead to a contradiction. Then, computing the next few relations between b_i and c_j from (3) (and constantly using $\binom{k}{k-1} = k \equiv 0 \mod 2$, for all $k = 2s$ where s is a positive integer) gives for $q = 2^n \geq 8$ the following

$$\begin{aligned} c_{q-3} &= 0 = \binom{q-2}{q-3} ab_{q-2} + \binom{q-1}{q-3} a^2 b_{q-1} = a^2 b_{q-1} = a^2 \cdot 0 \\ c_{q-4} &= 0 = \binom{q-3}{q-4} ab_{q-3} + \binom{q-2}{q-4} a^2 b_{q-2} + \binom{q-1}{q-4} a^3 b_{q-1} \\ &= \binom{q-3}{q-4} ab_{q-3} + \binom{q-2}{q-4} a^2 b_{q-2}. \end{aligned}$$

The first equation gives us no condition on b_{q-2}, it can be chosen arbitrary (since $\binom{q-2}{q-3} \equiv 0$) though if F is Boolean we must also have $b_{q/2-1}^2 = b_{q-2}$. The second equation depends on the parity of $\binom{q-3}{q-4}$ and $\binom{q-2}{q-4}$. Now, obviously $\binom{q-3}{q-4} = q-3 \equiv 1 \mod 2$, whereas $\binom{q-2}{q-4} \equiv \binom{q/2-1}{q/2-2} = q/2-1 \equiv 1 \mod 2$. This implies that the second equation above yields $b_{q-3} = ab_{q-2}$. In particular, since $b_i \in \mathbb{F}_2$ then assuming that either $b_{q-2} = 1$ or $b_{q-3} = 1$ we necessarily have that $a = 1$.

The expression for c_{q-5} given by

$$c_{q-5} = 0 = \binom{q-4}{q-5} ab_{q-4} + \binom{q-3}{q-5} a^2 b_{q-3} + \binom{q-2}{q-5} a^3 b_{q-2} + \binom{q-1}{q-5} a^4 b_{q-1},$$

requires again the analysis of the coefficients $\binom{q-3}{q-5}$ and $\binom{q-2}{q-5}$. Clearly $\binom{q-2}{q-5} \equiv 0 \mod 2$, since $q-2$ is even and $q-5$ is odd. Similarly, $\binom{q-3}{q-5} \equiv \binom{q/2-2}{q/2-3} = q/2-2 \equiv 0 \mod 2$. Thus, since also $\binom{q-4}{q-5} \equiv 0 \mod 2$, implies that b_{q-4} is arbitrary and at the same time

$b_{q/2-2}^2 = b_{q-4}$. Similarly, computing

$$\begin{aligned} c_{q-6} &= 0 = \binom{q-5}{q-6} a b_{q-5} + \binom{q-4}{q-6} a^2 b_{q-4} + \binom{q-3}{q-6} a^3 b_{q-3} + \binom{q-2}{q-6} a^4 b_{q-2} \\ &\quad + \binom{q-1}{q-6} a^5 b_{q-1} = a b_{q-5} + a^4 b_{q-2}. \end{aligned}$$

implies that $b_{q-5} = a^3 b_{q-2}$ and also $b_{q-9} = (a^3 b_{q-2})^2$ using $b_{2i}^2 = b_i$.

Thus, in order to deduce stronger conditions on the coefficients we need to assume further restrictions on the form of F. Indeed, by requesting that $b_{q-2} = 0$ we necessarily have $b_{q-3} = b_{q-5} = 0$. Then, checking the expression for c_{q-7} which is given by,

$$\begin{aligned} c_{q-7} &= 0 = \binom{q-6}{q-7} a b_{q-6} + \binom{q-5}{q-7} a^2 b_{q-5} + \binom{q-4}{q-7} a^3 b_{q-4} + \binom{q-3}{q-7} a^4 b_{q-3} \\ &\quad + \binom{q-2}{q-7} a^5 b_{q-2} + \binom{q-1}{q-7} a^6 b_{q-1} = \binom{q-4}{q-7} a^3 b_{q-4} = a^3 b_{q-4}, \end{aligned}$$

taking into account that $b_{q-1} = b_{q-2} = b_{q-3} = b_{q-5} = 0$ and that $\binom{q-6}{q-7} \equiv 0 \mod 2$. But, since b_{q-4} is arbitrary then assuming it is non-zero leads to a contradiction $c_{q-7} = a^3 b_{q-4} \neq 0$. Therefore, assuming that $b_{q-1} = b_{q-2} = b_{q-3} = b_{q-5} = 0$ and $b_{q-4} \neq 0$ implies that such an $F : \mathbb{F}_{2^n} \to \mathbb{F}_2$ cannot have linear structures. Notice that the same reasoning is also valid for $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ since the Boolean conditions are actually irrelevant in the above derivation.

Theorem 4. *Let $F(x) = \sum_{i=0}^{q-1} b_i x^i$, $q = 2^n$, where $F : \mathbb{F}_{2^n} \to \mathbb{F}_{2^n}$. If $b_{q-1} = b_{q-2} = b_{q-3} = b_{q-5} = 0$ and $b_{q-4} \neq 0$, then F cannot have linear structure. Furthermore, if the coefficients of F satisfy the Boolean conditions the same condition implies that $F : \mathbb{F}_{2^n} \to \mathbb{F}_2^n$ does not have linear structures.*

Remark 3. *The above result appears to be rather peculiar in the context of linear structures. There is no obvious reason why the above condition ensures the non-existence of linear structures. Certainly, there are other possibilities of specifying the coefficients b_i (for instance without forcing that $b_{q-2} = 0$) for the same purpose, though we do not explore this further.*

Example 1. *Let $q = 2^4$ and assume that $b_{15} = b_{14} = b_{13} = b_{11} = 0$ and $b_{12} \neq 0$. Then, $F(x) = \sum_{i=0}^{15} b_i x^i$, $b_i \in \mathbb{F}_{2^4}$, does not posses linear structures for arbitrary choice of $b_0, \ldots, b_{11}$. If the remaining coefficients satisfy the Boolean conditions then $F : \mathbb{F}_{2^4} \to \mathbb{F}_2$ does not admit linear structures.*

A similar analysis also implies the following result.

Theorem 5. *Let $F(x) = \sum_{i=0}^{2^n-1} b_i x^i$, such that $b_i = 0$ for $i > d$ and $b_d \neq 0$, for some $d \in [1, 2^n - 1]$. Then,*

(i) *If d is odd and $d > 1$, then F has no linear structures.*
(ii) *If d is even such that $4 \not| d$ and $b_{d-1} = 0$, then F has no linear structures.*
(iii) *If d is even such that $4 \mid d$ and $b_{d-1} = 1$, then F cannot have linear structures.*

PROOF. (i) The case when $d = 2^n - 1$ follows from Theorem 3, regardless whether b_{q-2} is zero or not. Thus, let $d < 2^n - 1$, where d is odd and $b_d \neq 0$. Since $c_i = 0$ for $i > d-1$ let us consider

$$c_{d-1} = \binom{d}{d-1} ab_d = dab_d \neq 0,$$

because $d \not\equiv 0$. Since $c_{d-1} \neq 0$ and $d - 1 > 0$, F does not have linear structures.
(ii) If $b_{d-1} = 0$ and d is even such that $4 \not| d$, then

$$c_{d-2} = \binom{d-1}{d-2} ab_{d-1} + \binom{d}{d-2} a^2 b_d = a^2 b_d \neq 0,$$

and F cannot have linear structures.
(iii) If $b_{d-1} = 1$ and d is even such that $4 \mid d$, then

$$c_{d-2} = \binom{d-1}{d-2} ab_{d-1} + \binom{d}{d-2} a^2 b_d = ab_{d-1} = a,$$

thus F cannot have linear structures in this case. □

Notice that the above result covers a large class of polynomials, having arbitrary number of terms, without linear structures. For instance, the main result in [3] was to establish the fact that binomials $F(x) = x^e + \alpha x^d$ cannot have linear structures unless F is affine. The result in Theorem 5 and a further simple analysis would lead to the same conclusion as already stated in [3].

4. Upper bounds on degree of planar mappings

In this section we will apply formulas for the integration of the polynomials to the planar mappings and consequently we deduce a nontrivial upper bound on the polynomial degree of these mappings. Assume p is odd and that $F(x) = \sum_{i=0}^{q-1} b_i x^i$ is a planar polynomial, thus $p > 2$. Then, for all $a \in \mathbb{F}_q^*$, the polynomial $G(x) = F(x+a) - F(x) = \sum_{i=1}^{q-2} c_i x^i$ is a permutation, where the connection between the coefficients c_i and b_i has been established in the previous section.

Theorem 6. *Let $F(x) = \sum_{i=0}^{q-1} b_i x^i$ be a planar polynomial over $\mathbb{F}_q$, where the prime field of $\mathbb{F}_q$ is of odd characteristic. Then, the polynomial degree of F is less than $q - 1 - \frac{p+1}{2}$.*

PROOF. For $F(x) = \sum_{i=0}^{q-1} b_i x^i$, if $G(x) = F(x+a) - F(x) = \sum_{i=1}^{q-2} c_i x^i$ is a permutation, then by Hermite's criterion $G^n(x) \pmod{x^q - x}$ has the coefficients with x^{q-1} equal to zero, for all $n = 1, 2, \ldots, q-2$. The case $n = 1$ implies $c_{q-1} = 0$.

Consider now $n = 2$. Squaring $G(x)$ we have that the coefficient d with x^{q-1} equals to $d = c_1c_{q-2}+c_2c_{q-3}+\ldots+c^2_{\frac{q-1}{2}}+\ldots+c_{q-2}c_1 = \sum_{t=1}^{q-2} c_t c_{q-1-t}$. Using $c_t = \sum_{i=t+1}^{q-1} \binom{i}{t} a^{i-t} b_i$ and substituting in d we obtain

$$d = \sum_{t=1}^{q-2} \left(\sum_{i=t+1}^{q-1} b_i \binom{i}{t} a^{i-t} \right) \left(\sum_{j=q-t}^{q-1} b_j \binom{j}{q-1-t} a^{j-q+1+t} \right).$$

Let us use a new variable $s = i+j+1-q$. If $j = q-t$ then $s = i+1-t$ and for $j = q-1$ we have that $s = i$. Therefore,

$$d = \sum_{t=1}^{q-2} \sum_{i=t+1}^{q-1} \sum_{s=i+1-t}^{i} \binom{i}{t} \binom{s+q-1-i}{q-1-t} b_i b_{s+q-1-i} a^s.$$

By changing the order of summation we obtain

$$d = \sum_{t=1}^{q-2} \sum_{s=2}^{q-1} \left(\sum_{i=\max\{s,t+1\}}^{\min\{s+t-1,q-1\}} \binom{i}{t} \binom{s+q-1-i}{q-1-t} b_i b_{s+q-1-i} \right) a^s =$$

$$\sum_{s=2}^{q-1} \left(\sum_{t=1}^{q-2} \sum_{i=\max\{s,t+1\}}^{\min\{s+t-1,q-1\}} \binom{i}{t} \binom{s+q-1-i}{q-1-t} b_i b_{s+q-1-i} \right) a^s.$$

We have that $d = 0$, for all $a \in \mathbb{F}_q^*$. Note that this is a polynomial in a, which is identically equal to zero for all $a \in \mathbb{F}_q^*$ and its degree is $q-1$. Thus, all the coefficients with a^s, for $s = 2, 3, \ldots, q-1$, are equal to zero.

The coefficient with a^{q-1}, i.e., for $s = q-1$, equals to

$$\sum_{t=1}^{q-2} \binom{q-1}{t} \binom{q-1}{q-1-t} b_{q-1}^2,$$

since $i = q-1$. The binomial formula implies

$$\sum_{t=0}^{2(q-1)} \binom{2(q-1)}{t} y^t = (y+1)^{2(q-1)} = (y+1)^{q-1}(y+1)^{q-1} = \sum_{i=0}^{q-1} \binom{q-1}{i} y^i \cdot \sum_{j=0}^{q-1} \binom{q-1}{j} y^j.$$

Equalling the coefficient with y^{q-1} we obtain the equality

$$\binom{2(q-1)}{q-1} = \sum_{i=0}^{q-1} \binom{q-1}{i} \binom{q-1}{q-1-i}.$$

Using this identity we obtain a simpler expression for the coefficient with a^{q-1} (note that the summation in the formula for the coefficient starts with 1)

$$\Big(\binom{2(q-1)}{q-1}-1\Big)b_{q-1}^2=\Big(\frac{2(q-1)\dots q}{(q-1)!}-1\Big)b_{q-1}^2=-2b_{q-1}^2.$$

Since this coefficient is equal to zero we have $b_{q-1}=0$.

Assume now that $b_{q-1}=\ldots=b_{q-u}=0$, with $u<p$. Let us evaluate the coefficient with a^{q-1-2u}. Since $s=q-1-2u$, $\max\{q-1-2u,t+1\}=q-1-2u$, for $t\le q-2-2u$ and similarly $\max\{q-1-2u,t+1\}=t+1$, for $t>q-2-2u$. Also, $\min\{q-1-2u+t-1,q-1\}=q-1$ if $t\ge 2u+1$ and $\min\{q-1-2u+t-1,q-1\}=q-1-2u+t-1$, for $t<2u+1$. The coefficient with a^{q-1-2u} is

$$\sum_{t=1}^{q-2}\sum_{i=\max\{s,t+1\}}^{\min\{s+t-1,q-1\}}\binom{i}{t}\binom{s+q-1-i}{q-1-t}b_ib_{2(q-1)-2u-i}=$$

$$\sum_{t=1}^{2u}\sum_{i=q-1-2u}^{q-1-2u+t-1}\binom{i}{t}\binom{2(q-1)-2u-i}{q-1-t}b_ib_{2(q-1)-2u-i}+$$
$$\sum_{t=2u+1}^{q-2-2u}\sum_{i=q-1-2u}^{q-1}\binom{i}{t}\binom{2(q-1)-2u-i}{q-1-t}b_ib_{2(q-1)-2u-i}+$$
$$\sum_{t=q-1-2u}^{q-2}\sum_{i=t+1}^{q-1}\binom{i}{t}\binom{2(q-1)-2u-i}{q-1-t}b_ib_{2(q-1)-2u-i}.$$

Consider now the sum in the middle. If $i=q-1-2u,q-1-2u+1,\ldots,q-1-2u+(u-1)=q-2-u$ then $b_{2(q-1)-2u-i}$ equals to $b_{q-1}=b_{q-2}=\ldots=b_{q-u}=0$. If $i=q-u,\ldots,q-1$ then $b_i=0$ by assumption. For $i=q-1-u$ we have that $b_ib_{2(q-1)-2u-i}=b_{q-1-u}^2$. Therefore, the inner sum equals to

$$\sum_{t=2u+1}^{q-2-2u}\binom{q-1-u}{t}\binom{q-1-u}{q-1-t}b_{q-1-u}^2.$$

Consider now the first sum. Here, $q-2\ge i\ge q-2-2u$. Similarly, the product $b_ib_{2(q-1)-2u-i}\neq 0$ only if $i=q-1-u\le q-1-2u+t-1$. There are nonzero terms only for $t\ge u+1$ and first sum equals to the

$$\sum_{t=u+1}^{2u}\binom{q-1-u}{t}\binom{q-1-u}{q-1-t}b_{q-1-u}^2.$$

Finally, let us consider the third sum. Here, i takes values $q-2u,q-2u+1,\ldots,q-1$. As already mentioned, $b_ib_{2(q-1)-2u-i}=0$ for all values of i except for $i=q-1-u\ge t+1$ and thus $t\le q-u-2$. Therefore, the third sum equals to

$$\sum_{t=q-1-2u}^{q-u-2}\binom{q-1-u}{t}\binom{q-1-u}{q-1-t}b_{q-1-u}^2.$$

The coefficient now is

$$\sum_{t=u+1}^{q-u-2} \binom{i}{t}\binom{q-1-u}{q-1-t} b_{q-1-u}^2 = b_{q-1-u}^2 \sum_{t=u+1}^{q-1-u-1} \binom{q-1-u}{t}\binom{q-1-u}{q-1-t}.$$

In order to simplify this expression consider

$$\sum_{t=0}^{2(q-1)-2u} \binom{2(q-1)-2u}{t} y^t = (y+1)^{2(q-1)-2u} = (y+1)^{q-1-u}(y+1)^{q-1-u} =$$

$$\sum_{i=0}^{q-1-u} \binom{q-1-u}{i} y^i \sum_{j=0}^{q-1-u} \binom{q-1-u}{j} y^j.$$

Equalling the coefficient with y^{q-1} on both sides ($j = q-1-i$) we obtain equality

$$\sum_{i=u}^{q-1-u} \binom{q-1-u}{i}\binom{q-1-u}{q-1-i} = \binom{2(q-1)-2u}{q-1}.$$

Now we have that

$$\sum_{t=u+1}^{q-1-u-1} \binom{q-1-u}{t}\binom{q-1-u}{q-1-t} =$$

$$\binom{2(q-1)-2u}{q-1} - \binom{q-1-u}{u}\binom{q-1-u}{q-1-u} - \binom{q-1-u}{q-1-u}\binom{q-1-u}{u} =$$

$$\frac{(2(q-1)-2u)\cdots q\cdots(q-2u)}{(q-1)!} - 2\binom{q-1-u}{q-1-u}\binom{q-1-u}{u} \equiv -2\binom{q-1-u}{u} \pmod{p}.$$

Therefore, the coefficient is now

$$-2b_{q-1-u}^2 \binom{q-1-u}{u}.$$

Note that

$$\binom{q-1-u}{u} = \frac{(q-1-u)(q-1-u-1)\cdots(q-2u)}{u!} \neq 0 \pmod{p}$$

if $q-2u > q-p$, i.e., for $u < \frac{p}{2}$. Therefore, if $u < \frac{p}{2}$ we can conclude that $b_{q-1-u} = 0$. Inductively, we have that

$$b_{q-1} = b_{q-2} = \ldots = b_{q-\frac{p+1}{2}} = 0$$

for planar polynomials. □

If $q = p$ in the previous proof, successively considering $s = q-1$, $s = q-2$, we can show that $b_i = 0$ for all $i > \frac{q-1}{2}$. Applying the same idea to $(G(x))^3$, …, $(G(x))^{p-2}$ it can be shown that the only planar polynomials over a prime field are quadratic, which is a well-known and established fact.

Corollary 3. *Assume that* $f(x) = \sum_{i=0}^{q-1} b_i x^i$ *is a planar polynomial. If there exists* $1 \leq s \leq n-1$ *where* $q = p^n$ *such that* $q - \frac{p+1}{2} \leq kp^s \mod (q-1) \leq q-1$ *then* $b_k = 0$.

PROOF. If $f(x)$ is planar then $f(x^{p^s})$ is also planar where the coefficient with $x^{kp^s} \mod (x^q - x) = x^{kp^s \mod (q-1)}$ is b_k. Since the degree of f is less than $q - \frac{p+1}{2}$ we have that $b_k = 0$ if $kp^s \mod (q-1) \geq q - \frac{p+1}{2}$. □

5. Conclusions

In this article we have derived the relation between the derivatives and the original polynomial. These results are then proved useful for establishing various results related to the existence of linear structures and in particular a nontrivial upper bound on the degree of planar mappings has been deduced.

[1] E. BIHAM AND A. SHAMIR. Differential cryptanalysis of DES-like cryptosystems. *Journal of Cryptology*, vol. 4(1):3–72, 1991.

[2] P. CHARPIN, G. KYUREGHYAN. Monomial functions with linear structure and permutation polynomials. *Contemporary Mathematics*, vol. 518: 99–111, 2010.

[3] P. CHARPIN, S. SARKAR. Polynomials with linear structure and Maiorana–McFarland construction. *IEEE Trans. on Inform. Theory*, IT-57(6):3796–3804, 2011.

[4] R. S. COULTER AND R. W. MATTHEWS. Planar functions and planes of Lenz-Barlotti class ii. *Des. Codes Cryptogr.*, vol. 10:167–184, 1997.

[5] R. S. COULTER AND R. W. MATTHEWS. Dembowski-ostrom polynomials from Dickson polynomials. *Finite Fields and Their Applications*, pages 369 – 379, 2010.

[6] P. DEMBOWSKI AND T. G. OSTROM. Planes of order n with collineation groups of order n^2. *Mathematische Zeitschrift*, 103(3):239–258, 1968.

[7] J. H. EVERTSE. Linear structures in block ciphers. In *Advances in Cryptology—EUROCRYPT 1987*, volume LNCS 304, pages 249–266. Springer-Verlag, 1988.

[8] K. NYBERG AND L. R. KNUDSEN. Provable security against a differential attack. *J. Cryptology*, 8(1): 27–37 (1995).

[9] N. TOKAREVA. *Bent functions: results and applications to cryptography.* Academic Press – Elsevier, 2015.

[10] H. XIONG, L. QU, C. LI, AND Y LI. Some results on the differential functions over finite fields. *Applicable Algebra in Engeneering, Communication and Computing*, Vol. 25(3):189–195, 2014.