Counting Boolean Functions with Faster Points

Ana Sălăgean[1] and Ferruh Özbudak[2]

[1] Department of Computer Science, Loughborough University, Loughborough, UK
`e-mail: a.m.salagean@lboro.ac.uk`
[2] Department of Mathematics and Institute of Applied Mathematics, Middle East Technical University, Ankara, Turkey `e-mail: ozbudak@metu.edu.tr`

Abstract. Duan and Lai introduced the notion of "fast point" for a Boolean function f as being a direction a so that the algebraic degree of the derivative of f in direction a is strictly lower than the expected $\deg(f) - 1$. Their study was motivated by the fact that the existence of fast points makes many cryptographic differential attacks (such as the cube and AIDA attack) more efficient. The number of functions with fast points was determined by Duan et. al. in some special cases and by Sălăgean and Mandache in the general case.
We generalise the notion of fast point, defining a fast points of order ℓ as being a fast point a so that the degree of the derivative of f in direction a is lower by at least ℓ than the expected degree. We determine the number of functions of degree d in n variables which have fast points of order ℓ. We also determine the number of functions which have a given space U as their space of fast points of order ℓ. As an application, we compute the number of functions which admit linear structures (i.e. their derivative in a certain direction is constant); such functions have a long history of being used in the analysis of symmetric ciphers.

Keywords: Boolean Functions · Differential attacks · Linear structures

1 Introduction

Boolean functions used in cryptography are usually required to resist a range of attacks. They need to have a sufficiently high algebraic degree (i.e. the degree of the function written in its algebraic normal form) in order to resist algebraic attacks. Differential attacks on cryptographic functions typically exploit properties of the discrete derivative. The discrete derivative of a function f in the direction a is defined as $D_a f(x) = f(x + a) - f(x)$. The derivatives should also have a high degree; the highest that can be achieved is one less than the degree of the original function, i.e. $\deg(D_a f) \leq \deg(f) - 1$. For example, the cube attack of Dinur and Shamir [2] and the AIDA attack of Vielhaber [7], as well as many further variants of these attacks, exploit the situation where a higher order derivative of the function has a very low degree (usually 1 or 2). Higher order derivatives are obtained by repeatedly differentiating in several directions.

Motivated by these applications, Duan and Lai [3] introduced the notion of "fast point" for a cryptographic function: a is a fast point for a function f if the degree of $D_a f$ drops more than expected, i.e. the degree is strictly lower than $\deg(f)-1$. The fast points of a function f form a linear space. In [4] they started computing the number of functions that admit fast points; explicit formulae were obtained for small degrees and very large degrees, and exhaustive search results were obtained for small numbers of variables.

In [6] Sălăgean and Mandache obtained a recurrence relation as well as an explicit formula for the number of functions that admit fast points, for any number of variables n and any degree d, and also for any given dimension of the space of fast points.

In this paper we define "faster points" i.e. points where the degree of the derivative drops by at least 2 more than expected. More generally, a fast point of order ℓ for a function f will be a point where the degree of $D_a f$ is at most $\deg(f) - 1 - \ell$, i.e. it dropped ℓ more than expected. The fast points of order ℓ of a function f form a linear space.

We will count the number of functions of degree d in n variables which have a given space U as their space of fast points of order ℓ. This number does not depend on the space itself, only on its dimension, so this allows us to count, for each fixed k, the number of functions which have exactly 2^k fast points of order ℓ; also the number of functions which have no fast points of order ℓ. For all these numbers we give both recurrence relations and explicit formulae, see Theorem 2 for fast points of order 2 and Theorem 3 for arbitrary order. The proofs use some techniques similar to the ones in [6], but also some different techniques, see Lemma 2.

As an application of these counting results, we can determine the number of functions which have linear structures. The notion of linear structure was introduced by Chaum and Evertse in 1985 in [1] and has since been used widely in the analysis of cryptographic primitives. An element a is a *linear structure* for a function f if $D_a f$ is a constant function. With our definition, a linear structure for f is a fast point of order $\deg(f) - 1$, so we can apply our results directly to compute the number of functions which have linear structures for each degree d and n variables.

2 Preliminaries

We denote by $\mathbb{F}_2$ the binary field. A boolean function f with n-bits input and one bit output can be viewed as a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$. Any such function can be represented in algebraic form as a polynomial function of degree at most one in each variable. (More precisely, because $x^2 = x$ when $x \in \mathbb{F}_2$, each function corresponds to an element in $\mathbb{F}_2[x_1, \ldots, x_n]/\langle x_1^2 + x_1, \ldots, x_n^2 + x_n\rangle$; we identify each coset with its unique representative multivariate polynomial which has degree at most one in each variable.)

The (algebraic) degree of f, denoted $\deg(f)$, is the total degree of the polynomial, with the usual convention that the degree of the zero function is $-\infty$. We will denote by $BF(n)$ the set of Boolean functions in n variables, and by

$BF(n,d)$ the set of Boolean functions in n variables of degree exactly d, where $0 \leq d \leq n$.

Let $f \in BF(n,d)$ and $a = (a_1, \ldots, a_n) \in \mathbb{F}_2^n$. The *derivative* $D_a f$ of f is defined as the Boolean function in n variables $x \mapsto f(x+a) - f(x)$. The degree of $D_a f$ is lower than or equal to $d-1$ (see [5]), i.e. differentiation decreases the degree of the function by at least 1. Vectors a for which the degree of $D_a f$ is strictly lower than $d-1$ (i.e. the degree drops more than expected) are called "fast points" of f (see [3]). The set of fast points of f is a linear subspace of $\mathbb{F}_2^n$ of dimension at most $n-d$ (see [3]).

Note that only the monomials of f of degree d matter when determining whether the degree of $D_a f$ is equal to $d-1$ or strictly lower, and therefore determining whether a is a fast point. In other words, $a \in \mathbb{F}_2^n$ is a fast point of f if and only if a is a fast point of $f+g$, where g is an arbitrary Boolean function in n variables of degree at most $d-1$. Hence for counting Boolean functions having fast points, it is natural to define the following equivalence on $BF(n)$: for $f_1, f_2 \in BF(n)$

$$f_1 \overset{(i)}{\sim} f_2 \iff \deg(f_1 - f_2) \leq i.$$

Then $[f]^{(i)}$ denotes the equivalence class of f with respect to the equivalence relation $\overset{(i)}{\sim}$. For deciding whether f has fast points, it suffices to consider f up to the equivalence $\overset{(d-1)}{\sim}$.

Next we formally define "faster points".

Definition 1. *Let $f \in B(n,d)$ and $1 \leq \ell \leq d$. An element $a \in \mathbb{F}_2^n$ is called a* fast point of order ℓ *for f if* $\deg(D_a f) \leq d - 1 - \ell$. *The set of fast points of order ℓ of f is denoted*

$$\mathrm{FP}^{(\ell)}(f) = \{a \in \mathbb{F}_2^n : \deg(D_a f) \leq d - 1 - \ell\}. \tag{1}$$

Note the usual fast points are fast points of order 1. If $a_1, a_2 \in \mathrm{FP}^{(\ell)}(f)$, then

$$\begin{aligned} D_{a_1+a_2} f(x) &= f(x + a_1 + a_2) - f(x) \\ &= f(x + a_1 + a_2) - f(x + a_1) + f(x + a_1) - f(x) \\ &= D_{a_2} f(x + a_1) + D_{a_1} f(x). \end{aligned}$$

Hence $\deg(D_{a_1+a_2} f) \leq \max\{\deg(D_{a_1} f), \deg(D_{a_2} f)\}$. This proves that $\mathrm{FP}^{(\ell)}(f)$ is a linear subspace of $\mathbb{F}_2^n$. The dimension of this space is at most $n-d$, since a fast point of order ℓ is also a fast point of order $\ell - 1$ and $\dim(\mathrm{FP}^{(1)}(f)) \leq n - d$. We have a filtration of linear subspaces:

$$\mathbb{F}_2^n \supseteq \mathrm{FP}^{(1)}(f) \supseteq \mathrm{FP}^{(2)}(f) \supseteq \cdots \supseteq \mathrm{FP}^{(d-1)}(f) \supseteq \mathrm{FP}^{(d)}(f) \supseteq \{0\}.$$

When determining whether a function f has fast points of order ℓ only the monomials of degree $d, d-1, \ldots, d-\ell+1$ matter, as they are the only ones that can produce polynomials of degree strictly above $d - 1 - \ell$ after differentiation; so we only need to consider the function f up to the equivalence $\overset{(d-\ell)}{\sim}$. The set of functions (up to the suitable equivalence) which have their space of fast points of order ℓ equal to a given subspace $U \subseteq \mathbb{F}_2^n$ will be denoted:

$$\mathrm{F}^{(\ell)}(n,d;U) = \{[f]^{(d-\ell)} : f \in BF(n,d) \text{ and } \mathrm{FP}^{(\ell)}(f) = U\}.$$

The set of functions (up to the suitable equivalence) for which the space of fast points of order ℓ has a given dimension k will be denoted:

$$\mathrm{F}^{(\ell)}(n,d;k) = \{[f]^{(d-\ell)} : f \in BF(n,d) \text{ and } \dim(\mathrm{FP}^{(\ell)}(f)) = k\}.$$

For integers $0 \le k \le n$ the Gaussian binomial coefficients (or q-binomial coefficients) are defined as

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{(q^n-1)(q^{n-1}-1)\cdots(q^{n-k+1}-1)}{(q^k-1)(q^{k-1}-1)\cdots(q-1)}. \tag{2}$$

Recall that $\begin{bmatrix} n \\ k \end{bmatrix}_q$ is the number of k dimensional $\mathbb{F}_q$-linear subspaces of $\mathbb{F}_q^n$. We will mostly use these coefficients for $q = 2$, and in this case we will omit the index and simply denote $\begin{bmatrix} n \\ k \end{bmatrix} = \begin{bmatrix} n \\ k \end{bmatrix}_2$.

The cardinality of $\mathrm{F}^{(1)}(n,d;k)$ was computed explicitly by Sălăgean and Mandache-Sălăgean:

Theorem 1. *([6, Corollary 3]) For integers $1 \le d \le n$ and $0 \le k \le n-d$ we have*

$$|\mathrm{F}^{(1)}(n,d;k)| = \begin{bmatrix} n \\ k \end{bmatrix} \sum_{i=0}^{n-k-d} (-1)^i 2^{\frac{i(i-1)}{2}} \begin{bmatrix} n-k \\ i \end{bmatrix} \left(2^{\binom{n-k-i}{d}} - 1\right).$$

For $1 \le i \le n$ let $e_i = (0,\ldots,0,1,0,\ldots,0) \in \mathbb{F}_2^n$ be the vector which has 1 in its i-th position and zeroes elsewhere. The set $\{e_1,\ldots,e_n\}$ forms the standard basis of $\mathbb{F}_2^n$ over $\mathbb{F}_2$. Linear changes of variables (changes of coordinates) will be useful, so we collect a number of straightforward results:

Lemma 1. *Let $f, f_1, f_2 : \mathbb{F}_2^n \to \mathbb{F}_2$ and $a \in \mathbb{F}_2^n$. Let $\varphi : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be an invertible linear map.*
(i) $\deg(f) = \deg(f \circ \varphi)$.
(ii) $(D_a f) \circ \varphi = D_{\varphi^{-1}(a)}(f \circ \varphi)$.
(iii) a is a fast point of order ℓ for f iff $\varphi^{-1}(a)$ is a fast point of order ℓ for $f \circ \varphi$.
(iv) e_i is a fast point of order ℓ for f iff x_i does not appear in any of the monomials of degree $d, d-1, \ldots, d-\ell+1$ of f.
(v) $f_1 \overset{(i)}{\sim} f_2$ iff $f_1 \circ \varphi \overset{(i)}{\sim} f_2 \circ \varphi$.

3 Counting Faster Points

We will make extensive use of the following result, which might be known but so far we have not found it in the literature.

Lemma 2. *Let $S, T : \mathbb{N} \to \mathbb{C}$ be functions. Then*

$$S(n) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q T(k) \text{ for all } n \geq 0 \tag{3}$$

if and only if

$$T(n) = \sum_{k=0}^{n} (-1)^k q^{\frac{k(k-1)}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q S(n-k) \text{ for all } n \geq 0 \tag{4}$$

Proof. Recall the analogue of the binomial formula for Gaussian binomial coefficients:

$$\prod_{k=0}^{n-1} (1 + q^k t) = \sum_{k=0}^{n} q^{\frac{k(k-1)}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q t^k.$$

which, for $t = -1$, gives:

$$\sum_{k=0}^{n} (-1)^k q^{\frac{k(k-1)}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{cases} 1 & \text{for } \ n = 0 \\ 0 & \text{for } \ n \geq 1 \end{cases} \tag{5}$$

Now assume the equation (3) holds, and we want to prove equation (4). We evaluate the right hand side of (4), substituting $S(n-k)$ by the expression given in (3):

$$\begin{aligned}
&\sum_{k=0}^{n} (-1)^k q^{\frac{k(k-1)}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q S(n-k) = \\
&= \sum_{k=0}^{n} (-1)^k q^{\frac{k(k-1)}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q \sum_{i=0}^{n-k} \begin{bmatrix} n-k \\ i \end{bmatrix}_q T(i) \\
&= \sum_{i=0}^{n} \sum_{k=0}^{n-i} (-1)^k q^{\frac{k(k-1)}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q \begin{bmatrix} n-k \\ i \end{bmatrix}_q T(i) \\
&= \sum_{i=0}^{n} \sum_{k=0}^{n-i} (-1)^k q^{\frac{k(k-1)}{2}} \begin{bmatrix} n \\ i \end{bmatrix}_q \begin{bmatrix} n-i \\ k \end{bmatrix}_q T(i) \\
&= \sum_{i=0}^{n} \left(\sum_{k=0}^{n-i} (-1)^k q^{\frac{k(k-1)}{2}} \begin{bmatrix} n-i \\ k \end{bmatrix}_q \right) \begin{bmatrix} n \\ i \end{bmatrix}_q T(i) \\
&= T(n).
\end{aligned}$$

We used the identity

$$\begin{bmatrix} n \\ k \end{bmatrix}_q \begin{bmatrix} n-k \\ i \end{bmatrix}_q = \begin{bmatrix} n \\ i \end{bmatrix}_q \begin{bmatrix} n-i \\ k \end{bmatrix}_q$$

which can easily be verified directly from the definition of Gaussian binomial coefficients (2); for the last equality we used (5) above.

Proving that (4) implies (3) is similar.

We will also need the following lemma which exploits invariance to invertible linear changes of coordinates.

Lemma 3. *Let U be a space of dimension k. Then*

$$|\mathrm{F}^{(\ell)}(n,d;U)| = |\mathrm{F}^{(\ell)}(n,d;\langle e_{n-k+1},\dots,e_n\rangle)| = |\mathrm{F}^{(\ell)}(n-k,d;\{0\})|$$

Proof. Let $a_1,\dots,a_k$ be a basis for U. Let φ be an invertible linear map so that $\varphi(e_{n-i+1}) = a_i$. We will use Lemma 1. A function f is such that $[f]^{(d-\ell)} \in \mathrm{F}^{(\ell)}(n,d;U)$ iff U is the space of fast points of order ℓ of f iff $\varphi^{-1}(U) = \langle e_{n-k+1},\dots,e_n\rangle$ is the space of fast points of order ℓ of $f\circ\varphi$ iff $f\circ\varphi$ is such that $[f\circ\varphi]^{(d-\ell)} \in \mathrm{F}^{(\ell)}(n,d;\langle e_{n-k+1},\dots,e_n\rangle)$. This proves the first equality.

For the second equality, note that for the equivalence class $[g]^{(d-\ell)}$ we can pick a representative which only contains monomials of degree $d, d-1,\dots,d-\ell+1$, where $d=\deg(g)$. Using Lemma 1(iv), we see that $[g]^{(d-\ell)} \in \mathrm{F}^{(\ell)}(n,d;\langle e_{n-k+1},\dots,e_n\rangle)$ iff the representative g which only contains monomials of degree $d, d-1,\dots,d-\ell+1$ does not depend on any of the variables $x_{n-k+1},\dots,x_n$; in other words, g is a polynomial in the $n-k$ variables $x_1,\dots,x_{n-k}$. Since the space of fast points of order ℓ of g is $\langle e_{n-k+1},\dots,e_n\rangle$, this means that g has no non-trivial fast points when viewed as a function in $n-k$ variables, so we have $[g]^{(d-\ell)} \in \mathrm{F}^{(\ell)}(n-k,d;\{0\})$. Conversely, any function in $n-k$ variables in $\mathrm{F}^{(\ell)}(n-k,d;\{0\})$ can be viewed as a function in n variables and it has the required fast points to be in $\mathrm{F}^{(\ell)}(n,d;\langle e_{n-k+1},\dots,e_n\rangle)$.

We are now ready to count the functions which have a given space U (or any space of given dimension k) as their space of fast points of order 2.

Theorem 2. *Let $0\le d\le n$ and $0\le k\le n-d$ be integers. For the cardinality of $\mathrm{F}^{(2)}(.,d,0)$ we have the recurrence formula*

$$\sum_{k=0}^{n-d}\begin{bmatrix}n\\k\end{bmatrix}|\mathrm{F}^{(2)}(n-k,d;0)| = (2^{\binom{n}{d}}-1)2^{\binom{n}{d-1}} \tag{6}$$

and the explicit formula

$$|\mathrm{F}^{(2)}(n,d;0)| = \sum_{i=0}^{n-d}(-1)^i 2^{\frac{i(i-1)}{2}}\begin{bmatrix}n\\i\end{bmatrix}\left(2^{\binom{n-i}{d}}-1\right)2^{\binom{n-i}{d-1}} \tag{7}$$

We also have, for any space U of dimension k with $0\le k\le n-d$:

$$|\mathrm{F}^{(2)}(n,d;U)| = \sum_{i=0}^{n-k-d}(-1)^i 2^{\frac{i(i-1)}{2}}\begin{bmatrix}n-k\\i\end{bmatrix}\left(2^{\binom{n-k-i}{d}}-1\right)2^{\binom{n-k-i}{d-1}} \tag{8}$$

$$|\mathrm{F}^{(2)}(n,d;k)| = \begin{bmatrix}n\\k\end{bmatrix}\sum_{i=0}^{n-k-d}(-1)^i 2^{\frac{i(i-1)}{2}}\begin{bmatrix}n-k\\i\end{bmatrix}\left(2^{\binom{n-k-i}{d}}-1\right)2^{\binom{n-k-i}{d-1}}. \tag{9}$$

Proof. For the recurrence relation, note that

$$\{[f]^{(d-2)} | f \in B(n,d)\} = \bigcup_U \mathrm{F}^{(2)}(n,d;U)$$

where the union is disjoint and U ranges over all subspaces of $\mathbb{F}_2^n$. The cardinality of the set on the left hand side is $(2^{\binom{n}{d}} - 1)2^{\binom{n}{d-1}}$, since for any class we can pick the representative which only has monomials of degree d and $d-1$; there are $\binom{n}{d}$ monomials of degree d, and at least one of them should have non-zero coefficient; there are $\binom{n}{d-1}$ monomials of degree $d-1$. For the set on the right hand side we use Lemma 3 and the fact that there are $\begin{bmatrix} n \\ k \end{bmatrix}$ spaces of each dimension k. This completes the proof of the recurrence formula (6).

For the proof of the first explicit formula we rewrite (6) as

$$\begin{aligned}\sum_{k=0}^{n-d} \begin{bmatrix} n \\ k \end{bmatrix} |\mathrm{F}^{(2)}(n-k,d;0)| &= \sum_{k=0}^{n} \begin{bmatrix} n \\ n-k \end{bmatrix} |\mathrm{F}^{(2)}(n-k,d;0)| \\ &= \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} |\mathrm{F}^{(2)}(k,d;0)| \\ &= (2^{\binom{n}{d}} - 1)2^{\binom{n}{d-1}}\end{aligned}$$

using the fact that $|\mathrm{F}^{(2)}(k,d;0)| = 0$ when $k < d$. This recurrence relation is of the type of equation (3) in Lemma 2, viewing d as fixed and putting $S(n) = (2^{\binom{n}{d}} - 1)2^{\binom{n}{d-1}}$ and $T(n) = |\mathrm{F}^{(2)}(n,d;0)|$. Therefore, equation (4) in Lemma 2 gives the first explicit formula (7) in the theorem statement (with the summation going up to n, but then note that $S(n-i) = 0$ for $n-d < i \leq n$).

Alternatively, (7) could also be proven using the technique from the proof of [6, Theorem 6].

For the next explicit formula, (8) we use Lemma 3. Finally for the final formula (9) we use the fact that $\mathrm{F}^{(2)}(n,d;k) = \cup_U \mathrm{F}^{(2)}(n,d;U)$ where U ranges over all the $\begin{bmatrix} n \\ k \end{bmatrix}$ spaces of dimension k in $\mathbb{F}_2^n$ and the sets in the union are disjoint.

The Theorem above can be generalised to counting the functions which have a given space U (or any space of given dimension k) as their space of fast points of order ℓ.

Theorem 3. *Let $1 \leq \ell \leq d$ and let U be a space of dimension k. Then*

$$|\mathrm{F}^{(\ell)}(n,d;U)| = \sum_{i=0}^{n-k-d} (-1)^i 2^{\frac{i(i-1)}{2}} \begin{bmatrix} n-k \\ i \end{bmatrix} \left(2^{\binom{n-k-i}{d}} - 1\right) 2^{\sum_{j=1}^{\ell-1} \binom{n-k-i}{d-j}} \quad (10)$$

$$|\mathrm{F}^{(\ell)}(n,d;k)| = \begin{bmatrix} n \\ k \end{bmatrix} \sum_{i=0}^{n-k-d} (-1)^i 2^{\frac{i(i-1)}{2}} \begin{bmatrix} n-k \\ i \end{bmatrix} \left(2^{\binom{n-k-i}{d}} - 1\right) 2^{\sum_{j=1}^{\ell-1} \binom{n-k-i}{d-j}} \quad (11)$$

Furthermore, the number of functions which have fast points of order ℓ (any number of non-trivial fast points of order ℓ) is:

$$|\{[f]^{(d-\ell)} : f \in B(n,d), \mathrm{FP}^{(\ell)}(f) \neq \{0\}\}| = \tag{12}$$

$$= \sum_{i=1}^{n-d} (-1)^{i-1} 2^{\frac{i(i-1)}{2}} \begin{bmatrix} n \\ i \end{bmatrix} \left(2^{\binom{n-i}{d}} - 1\right) 2^{\sum_{j=1}^{\ell-1} \binom{n-i}{d-j}}. \tag{13}$$

Proof. As in the proof of Theorem 2, we have

$$\{[f]^{(d-\ell)} | f \in B(n,d)\} = \bigcup_U \mathrm{F}^{(\ell)}(n,d;U).$$

For the left hand side we have:

$$|\{[f]^{(d-\ell)} | f \in B(n,d)\}| = \left(2^{\binom{n}{d}} - 1\right) 2^{\sum_{j=1}^{\ell-1} \binom{n}{d-j}}.$$

and using Lemma 3 we obtain the recurrence relation

$$\left(2^{\binom{n}{d}} - 1\right) 2^{\sum_{j=1}^{\ell-1} \binom{n}{d-j}} = \sum_{k=0}^{n-d} \begin{bmatrix} n \\ k \end{bmatrix} |\mathrm{F}^{(\ell)}(n-k,d;0)|$$

which we solve using Lemma 2 to obtain

$$|\mathrm{F}^{(\ell)}(n,d;0)| = \sum_{i=0}^{n-d} (-1)^{i} 2^{\frac{i(i-1)}{2}} \begin{bmatrix} n \\ i \end{bmatrix} \left(2^{\binom{n-i}{d}} - 1\right) 2^{\sum_{j=1}^{\ell-1} \binom{n-i}{d-j}}.$$

Using this formula and Lemma 3, we then obtain (10) and (11).

For Equation (12) we have:

$$\begin{aligned}
&|\{[f]^{(d-\ell)} : f \in B(n,d), \mathrm{FP}^{(\ell)}(f) \neq \{0\}\}| = \\
&= |\{[f]^{(d-\ell)} | f \in B(n,d)\} \setminus \mathrm{F}^{(\ell)}(n,d;0)| \\
&= \left(2^{\binom{n}{d}} - 1\right) 2^{\sum_{j=1}^{\ell-1} \binom{n}{d-j}} - \sum_{i=0}^{n-d} (-1)^{i} 2^{\frac{i(i-1)}{2}} \begin{bmatrix} n \\ i \end{bmatrix} \left(2^{\binom{n-i}{d}} - 1\right) 2^{\sum_{j=1}^{\ell-1} \binom{n-i}{d-j}} \\
&= \sum_{i=1}^{n-d} (-1)^{i-1} 2^{\frac{i(i-1)}{2}} \begin{bmatrix} n \\ i \end{bmatrix} \left(2^{\binom{n-i}{d}} - 1\right) 2^{\sum_{j=1}^{\ell-1} \binom{n-i}{d-j}}.
\end{aligned}$$

As an application of these counting results, we can determine the number of functions which have linear structures. An element $a \in \mathbb{F}_2^n \setminus \{0\}$ is a *linear structure* for a function f if $D_a f$ is a constant function. With our definition, a linear structure for f is a fast point of order $\deg(f) - 1$. Therefore:

Corollary 1. *The number of functions of degree d in n variables which have linear structures is:*

$$\sum_{i=1}^{n-d} (-1)^{i-1} 2^{\frac{i(i-1)}{2}} \begin{bmatrix} n \\ i \end{bmatrix} \left(2^{\binom{n-i}{d}} - 1\right) 2^{\sum_{j=1}^{d-2} \binom{n-i}{d-j}}.$$

where the functions are counted up to addition of an affine function.

Example 1. Let $n = 7, d = 3$. We compute the number of functions of degree 3 in 7 variables which have fast points of order 2, i.e. they have linear structures. Using Corollary 1, this number is:

$$\sum_{i=1}^{4}(-1)^{i-1}2^{\frac{i(i-1)}{2}}\begin{bmatrix}7\\i\end{bmatrix}\left(2^{\binom{7-i}{3}}-1\right)2^{\binom{7-i}{2}} = 4\ 358\ 179\ 630\ 080.$$

Note the counting is done up to equivalence $\overset{(1)}{\sim}$ i.e. up to addition of affine functions. The remaining

$$\begin{aligned}\left(2^{\binom{7}{3}}-1\right)2^{\binom{7}{2}} - 4\ 358\ 179\ 630\ 080 &= 72\ 057\ 594\ 035\ 830\ 784 - 4\ 358\ 179\ 630\ 080\\ &= 72\ 053\ 235\ 856\ 200\ 704\end{aligned}$$

functions have no linear structures. We can also compute the number of functions with no fast points of order 2 directly using Theorem 2:

$$|\mathrm{F}^{(2)}(7,3;\{0\})| = \sum_{i=0}^{4}(-1)^{i}2^{\frac{i(i-1)}{2}}\begin{bmatrix}7\\i\end{bmatrix}\left(2^{\binom{7-i}{3}}-1\right)2^{\binom{7-i}{2}} = 72053235856200704.$$

The proportion of functions which have linear structures out of all functions of degree 3 in 7 variables is

$$\frac{4358179630080}{72057594035830784} \approx 0.00006.$$

4 Conclusion

Motivated by the properties of cryptographic functions exploited by differential attacks, Duan and Lai [3] introduced the notion of Boolean functions that admit “fast points”. We generalised this notion, defining functions f which have “fast points of order ℓ” i.e. the degree of at least one of the discrete derivatives of f is lower by ℓ than the expected value (i.e. it is $d-1-\ell$ or less, instead of the expected $d-1$). We obtained explicit formulae for the number of such functions of degree d in n variables. As an important particular case, this allowed us to compute the number of functions which admit a linear structure.

Acknowledgement

The authors thank to Royal Society for their support through the Newton Mobility Grant NI170158. The second author thanks also Loughborough University, Department of Computer Science for hospitality.

References

1. D. Chaum and J.-H. Evertse. Cryptanalysis of DES with a reduced number of rounds. In *Proceedings of Crypto '85*, pages 192–211, 1985.
2. Itai Dinur and Adi Shamir. Cube attacks on tweakable black box polynomials. In *EUROCRYPT*, pages 278–299, 2009.
3. Ming Duan and Xuejia Lai. Higher order differential cryptanalysis framework and its applications. In *International Conference on Information Science and Technology (ICIST)*, pages 291–297, 2011.
4. Ming Duan, Mohan Yang, Xiaorui Sun, Bo Zhu, and Xuejia Lai. Distinguishing properties and applications of higher order derivatives of boolean functions. *Information Sciences*, 271:224–235, 2014.
5. Xuejia Lai. Higher order derivatives and differential cryptanalysis. In Richard E. Blahut, Daniel J. Costello, Jr., Ueli Maurer, and Thomas Mittelholzer, editors, *Communications and Cryptography*, volume 276 of *The Springer International Series in Engineering and Computer Science*, pages 227–233. Springer, 1994.
6. Ana Sălăgean and Matei Mandache-Sălăgean. Counting and characterising functions with fast points for differential attacks. *Cryptography and Communications*, pages 1–23, 2015.
7. M. Vielhaber. Breaking ONE.FIVIUM by AIDA an algebraic IV differential attack. Cryptology ePrint Archive, Report 2007/413, 2007. http://eprint.iacr.org/.