# On derivatives of Boolean bent functions

**Shaporenko Alexander**

Sobolev Institute of Mathematics, Novosibirsk, Russia
Novosibirsk State University, Novosibirsk, Russia
Laboratory of cryptography JetBrains Research, Novosibirsk, Russia

CTCrypt 2021

# Boolean functions

Function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is called a **Boolean function in $n$ variables**.

A Boolean function $f$ is called **affine** if it can be represented as $\ell_{a,b}(x) = \langle a, x \rangle \oplus b$, where $a \in \mathbb{F}_2^n$ and $b \in \mathbb{F}_2$.

A Boolean function is called **balanced** if it takes values 0 and 1 equally often.

Every Boolean function $f$ in $n$ variables can be associated with its **support**:

$$\mathrm{supp}(f) = \{x \in \mathbb{F}_2^n : f(x) = 1\}.$$

# Boolean functions

Any Boolean function can be uniquely represented by its **algebraic normal form (ANF)**:

$$f(x_1, ..., x_n) = \bigoplus_{k=1}^{n} \bigoplus_{i_1,...,i_k} a_{i_1,...,i_k} x_{i_1} ... x_{i_k} \oplus a_0,$$

where for each $k$ indices $i_1, ..., i_k$ are pairwise distinct, sets $\{i_1, ..., i_k\}$ are exactly all different nonempty subsets of the set $\{1, ..., n\}$ and $a_{i_1,...,i_k}, a_0$ take values from $\mathbb{F}_2$.

For a Boolean function $f$, the number of variables in the longest item of its ANF is called the **algebraic degree** of a function (or briefly **degree**) and is denoted by $\deg f$.

# Bent functions

**The Hamming weight** $wt(f)$ of a Boolean function $f$ is the number of vectors $x \in \mathbb{F}_2^n$ such that $f(x) = 1$.

We denote by $dist(f, g)$ **the Hamming distance** between two Boolean functions $f$ and $g$; it is the number of positions in which their vectors of values differ:

$$dist(f, g) = |\{x \in \mathbb{F}_2^n : f(x) \neq g(x)\}|.$$

**The nonlinearity** of $f$, denoted by $N_f$, is the Hamming distance between $f$ and the set of affine functions.

A **bent function** is a Boolean function in an even number of variables that has the maximal nonlinearity, i.e., $N_f = 2^{n-1} - 2^{n/2-1}$.

# Applications of bent functions

Nonlinearity is an important property in cryptography. Ciphers using functions with high nonlinearity as components are more resistant to linear cryptanalysis (Matsui, 1994) because the greater the nonlinearity of function, the more difficult it is to approximate it by affine functions.

Bent functions were used in design of the block cipher CAST as coordinate functions of S-blocks (Adams, 1997). The nonlinear feedback polynomial of the NFSR (nonlinear feedback shift register) of the stream cipher Grain is constructed as the sum of a linear function and a bent function (Hell, 2006).

More information concerning applications of bent functions can be found in the monographs of N. Tokareva "**Bent functions: results and applications to cryptography**" (2015) and S. Mesnager "**Bent functions: Fundamentals and results**" (2016).

# Open problems

Some of open problems concerning bent functions are listed below:

- Number of bent functions for $n \geq 10$.
- Lower and upper bounds for the number of bent functions.
- Affine classification of bent functions for $n \geq 10$.
- New constructions of bent functions.

# Problem formulation

Bent function can also be defined as a Boolean function $f(x)$ in $n$ variables ($n$ is even) such that for any nonzero vector $y \in \mathbb{F}_2^n$ its derivative $D_y f(x) = f(x) \oplus f(x \oplus y)$ is balanced.

**Hypothesis 1**

*Any balanced function $f$ in $n$ variables of degree $\leq n/2 - 1$, such that $f(x) = f(x \oplus y)$ for every $x \in \mathbb{F}_2^n$ and some nonzero $y \in \mathbb{F}_2^n$, is a derivative of some bent function.*

This hypothesis was formulated by N. Tokareva in 2016. It was stated that this hypothesis holds for $n \leq 6$.

Whether it holds for every even $n$ is an open problem.

**Remark 1**

We should consider only balanced functions of degree $\leq n/2 - 1$, since the maximum possible degree of bent function is $n/2$ and it's known that $deg\, D_y f(x) \leq deg\, f(x) - 1$.

**Remark 2**

Not every Boolean fucntion can be a derivative. The following lemma gives the necessary and sufficient condition for a Boolean function to be a derivative of some Boolean function.

**Lemma 1**

A Boolean function $f$ in $n$ variables is a derivative of some Boolean function in $n$ variables in nonzero direction $y$ if and only if $f(x) \oplus f(x \oplus y) = 0$ for all $x \in \mathbb{F}_2^n$.

# Affine derivatives of bent functions

It is known that any nonconstant affine function is balanced.

**Lemma 2**

*Let $\ell_{a,b}(x) = \langle a, x\rangle \oplus b$, where $a \in \mathbb{F}_2^n$, $a$ is nonzero, and $b \in \mathbb{F}_2$. There are $2^{n-1} - 1$ nonzero directions for which $\ell_{a,b}$ is a derivative of a some Boolean function. Namely, these directions are exactly those nonzero vectors $y$ such that $\langle a, y\rangle = 0$.*

**Lemma 3**

*Let $\ell$ be a nonconstant affine function that is a derivative of bent functions $g$ and $g'$ in distinct nonzero directions $y$ and $y'$, respectively. Then $g \neq g'$.*

Suppose that $D_{y'}g(x) = \ell_{a,b}(x)$.

It follows from Lemma 1 that $\ell_{a,b}(x) = l_{a,b}(x \oplus y')$ and hence

$$x \in \mathrm{supp}(l_{a,b}) \iff x \oplus y' \in \mathrm{supp}(l_{a,b}). \tag{1}$$

Note that for any Boolean function $g$ in $n$ variables that has $\ell_{a,b}$ as its derivative in the direction $y'$ it holds that

$$g(x) \oplus g(x \oplus y') = l_{a,b}(x). \tag{2}$$

Let $i$ be the number of the first nonzero coordinate of $y'$ and $j$ be the number such that $j \neq i$ and $x_j$ is an essential variable for $\ell_{a,b}$. It can be shown that such $j$ always exists.

Without loss of generality, let $i = 1$ and $j = 2$.

It follows from (1) and (2) that any Boolean function $g$, such that $D_{y'}g(x) = \ell_{a,b}(x)$, has the following representation

$$
\begin{aligned}
&g(0, x_2, \bar{x}) = f_0(\bar{x}), && (0, x_2, \bar{x}) \in \mathrm{supp}(\ell_{a,b}),\\
&g(1, x_2 \oplus y'_2, \bar{x} \oplus \bar{y'}) = f_0(\bar{x}) \oplus 1, && (1, x_2 \oplus y'_2, \bar{x} \oplus \bar{y'}) \in \mathrm{supp}(\ell_{a,b}),\\
&g(0, x_2, \bar{x}) = f_1(\bar{x}), && (0, x_2, \bar{x}) \notin \mathrm{supp}(\ell_{a,b}),\\
&g(1, x_2 \oplus y'_2, \bar{x} \oplus \bar{y'}) = f_1(\bar{x}), && (1, x_2 \oplus y'_2, \bar{x} \oplus \bar{y'}) \notin \mathrm{supp}(\ell_{a,b}),
\end{aligned}
$$

where

$$\bar{z} = (z_3, \ldots, z_n), \text{for } z_k \in \mathbb{F}_2$$

and $f_0, f_1$ are arbitrary Boolean functions in $n - 2$ variables.

# Affine derivatives of bent functions

It was shown that $g$ is bent if and only if $f_0$ and $f_1$ are bent. Therefore, according to Lemma 3 we get the following result

**Theorem 1**

*Let $\ell_{a,b}$ be a nonconstant affine function in $n$ variables. Then $\ell_{a,b}$ is a derivative of $(2^{n-1}-1)|\mathcal{B}_{n-2}|^2$ bent functions in $n$ variables.*

We can use this way of constructing bent function $g$ as iterative construction of bent functions.

Like any construction it can be used in order to find new quivalence classes of bent functions.

# Affine derivatives of bent functions

**Lemma 4**

*Let $g$ be bent. If $D_y g(x) = \ell(x)$, then $D_{y'} g(x) \neq \ell(x) \oplus 1$ for every nonzero $y' \neq y$.*

Theorem 1 and Lemma 4 give us the following iterative lower bound.

**Theorem 2**

*For even $n \geq 4$ it holds $|\mathcal{B}_{n+2}| \geqslant (2^{n+2} - 2)|\mathcal{B}_n|^2$.*

| Variables | 4 | 6 | 8 | 10 |
|---|---|---|---|---|
| Bent | 896 | 5 425 430 528 | $\approx 2^{106.29}$ | ? |
| Theorem 2 | 896 | 49 774 592 | $\approx 2^{72.6}$ | $\approx 2^{222.5}$ |
| Tokareva, 2011 | 512 | 322 961 408 | $\approx 2^{87.35}$ | $\approx 2^{262.16}$ |

**Table 1:** Number of bent functions constructed with different methods

Iterative lower bound from Theorem 2 is not better than one intriduced by N. Tokareva in 2011 when $n \geq 6$. But it theoretically can be improved if we consider more then two affine functions $\ell$ and $\ell \oplus 1$. Unfortunately, it is hard to keep track of bent functions that were already counted because it is possible that $D_y g(x) = \ell_1(x)$ and $D_{y'} g(x) = \ell_2(x)$, where $\ell_2 \neq \ell_1$, $\ell_2 \neq \ell_1 \oplus 1$ and $y \neq y'$. We also can consider bent functions that do not have affine derivatives. Such functions of degree 3 were studied by A. Canteaut and P. Charpin in 2003. Although, the number of such functions was not presented.

# What's next

- Improvment of iterative lower bound from Theorem 1.
- Search for new affine equivalence classes using construction from Theorem 1.
- ***Proof of Hypothesis 1.***

# Bent sum decomposition problem

In 2011 the following hypothesis was raised.

**Hypothesis 2**

*Let $n$ be an even positive integer. Then any Boolean function in $n$ variables with degree $\leq n/2$ can be expressed as the sum of two bent functions in $n$ variables.*

It was shown (Qu L., 2014) that if Hypothesis 2 is true, then one can obtain the following lower bound for the number of bent functions

$$|\mathcal{B}_n| \geq 2^{2^{n-2}+\frac{1}{4}\binom{n}{n/2}+\frac{n+1}{2}},$$

which will be the best known.

# Connection between two open problems

**Proposition 1**

*A Boolean function $f = f_1(\bar{x})x_1 \oplus f_0(\bar{x}) \oplus x_2$ in $n$ variables is a derivative of a bent function in the direction $y$ $(y_1 = 1)$ if and only if a Boolean function $f_0(\bar{x})$ in $n-2$ variables can be represented as the sum of two bent functions in $n-2$ variables.*

**Remark 3**

Any balanced Boolean function $f$ in $n$ variables of degree $d$ that depends linearly on variable $x_2$ which is a derivative of some Boolean function in the direction $y$ $(y_1 = 1)$ can be expressed as $f = (f_0(\bar{x}) \oplus f_0(\bar{x} \oplus \bar{y}))x_1 \oplus f_0(\bar{x}) \oplus x_2$.

**Remark 4**

It can be shown that $\deg f_0 = deg\, f(x)$.

# Connection between two open problems

**Corollary 1**

*Every balanced Boolean function $f$ in $n$ variables of degree $d \leqslant \frac{n}{2} - 1$ that depends linearly on at least one of its variables is a derivative of a bent function if and only if every Boolean function in $n-2$ variables of degree $d$ can be represented as the sum of two bent functions in $n-2$ variables.*

Therefore, by proving Hypothesis 1 one can verify Hypothesis 2 which will at least lead us to new lower bound for the number of bent functions.

Thanks for attention!