k-BENT FUNCTIONS AND QUADRATIC APPROXIMATIONS IN BLOCK CIPHERS *

Natalia N. Tokareva [1]

Abstract. We introduce the notion of *k-bent function*, here k is integer, $1 \leqslant k \leqslant m/2$, and m is an even number of variables. The parameter k makes the property of "bentness" more strong as k grows up. 1-Bent functions and bent functions coincide. The constructions of k-bent functions for any k are given. We study the cryptanalysis (of a block cipher) based on special quadratic approximations and prove that by using k-bent functions in a cipher it is possible to make it resistant to this attack.

Keywords. k-Bent functions, Hadamard codes, Walsh-Hadamard transform, quadratic approximations, block ciphers.

Introduction

Bent functions form a well-known topic in discrete mathematics and cryptology. A lot of papers deal with them: starting from papers of the pioneers O. Rothaus (sixties of XX) [15], J. F. Dillon (1972) [2], R. L. McFarland (1973) [12] and finishing for example with papers of A. Youssef and G. Gong (2001, 2007), C. Carlet and P. Gaborit (2006), A. S. Kuzmin et al. (2006, 2008) on generalized hyper-bent functions. We refer the interested reader to surveys in [3], [9] and [1].

* *This research was supported by the Siberian Branch of the Russian Academy of Sciences Integration project "Tree-like catalogue of mathematical Internet resources mathtree.ru"* (no. 35)*, by the Russian Foundation for Basic Research* (*projects* 07-01-00248, 08-01-00671) *and Russian Science Support Foundation.*

[1] Sobolev Institute of Mathematics, Pr. Koptyuga 4, 630090 Novosibirsk; Novosibirsk State University, St. Pirogova 2, 630090 Novosibirsk; Russian Federation. Phone: +7 383 3333 869 Fax: +7 383 333 25 98. email: tokareva@math.nsc.ru, web: www.math.nsc.ru/∼tokareva

In this paper we introduce the notion of *k-bent function*. It is the Boolean function with m variables $v_1, \ldots, v_m$ (m is even) that is on the maximal possible Hamming distance from the set of all Boolean functions $\langle \mathbf{u}, \mathbf{v} \rangle_j \oplus a$, where $j = 1, \ldots, k$, $\mathbf{u} \in \mathbb{Z}_2^m$, $a \in \mathbb{Z}_2$ and k is a fixed integer number, $1 \leqslant k \leqslant m/2$. Here $\langle \cdot, \cdot \rangle_k$ is the special type binary product similar with the inner product over $\mathbb{Z}_2$. Among functions $\langle \mathbf{u}, \mathbf{v} \rangle_k \oplus a$ (if k is fixed) there are exactly $2^{m-k+1}(k+1)$ affine functions and $2^{m-k+1}(2^k-k-1)$ functions of degree 2. The products $\langle \cdot, \cdot \rangle_k$ were defined in [19]. They are connected with the series of Hadamard-like codes constructed from $\mathbb{Z}_4$-linear Hadamard codes. The classification of $\mathbb{Z}_4$-linear Hadamard and perfect codes was given by D. S. Krotov [7], [8].

1-Bent functions and bent functions coincide. For $k > \ell$ the class of k-bent functions is a proper subclass of the class of ℓ-bent functions. We show how to construct k-bent functions for any k and study their properties.

For the application to cryptanalysis we study special quadratic approximations in block ciphers. We use the idea of the well-known method of linear cryptanalysis given in 1993 by M. Matsui [11]. There exist several nonlinear generalizations of the method's technique: such as the common nonlinear approach introduced by L. R. Knudsen and M. J. B. Robshaw (1996) [6] and developed by J. Nakahara et al. (2003) [13], J. M. E. Tapiador et al. (2007) [18] and some quadratic techniques for concrete ciphers for example as in the paper of T. Shimoyama and T. Kaneko (1998) [16] on DES. Some results on special nonlinear approximations based on using reduced representations of Boolean functions over $GF(2^m)$ are obtained by A. V. Ivanov (2007) [5]. However, the common nonlinear technique leaves many questions open.

We approximate Boolean functions by all functions $\langle \mathbf{u}, \pi(\mathbf{v}) \rangle_k$, where $\mathbf{u}$, k and permutation π on m variables $v_1, \ldots, v_m$ are arbitrary. The class of such functions includes 2^m (i. e. all) linear functions and not more than $2^{m(1+\log_2 m)}$ the special type quadratic functions. So, it is possible to look through all of them in order to find the best approximation. Properties of the product $\langle \cdot, \cdot \rangle_k$ and simple formulas for the Hamming distances between a Boolean function and the classes of functions $\langle \mathbf{u}, \pi(\mathbf{v}) \rangle_k$ (for fixed π, k) account for our choice of the functions $\langle \mathbf{u}, \pi(\mathbf{v}) \rangle_k$ for approximation. We show that by using k-bent-functions in a cipher it is possible to make the maximum absolute values of the biases be

minimal. In this way one can reach the extreme resistance of a cipher to these quadratic approximations. Also we consider several 4-bit permutations with the most high nonlinearity recommended for using in S-boxes of GOST 28147-89, DES, s^3DES in relation to the quadratic approximations. We show that for the all of them (excepting only one) there are special quadratic equalities on input and output bits with probability more high than any linear equality has.

We present our theoretical results in the concentrated form. All statements we give without proofs. The proofs for the facts from sections 1–2 one can find in [19], [20]; from sections 3–5 in [21]. In these papers we give also a detailed literature survey, which we omit here.

1. Basic definitions

1.1. Notions of coding theory

Let m be integer. Binary vector $(u_1, \ldots, u_m)$ we denote by $\mathbf{u}$. Consider the metric space $\langle \mathbb{Z}_2^m, d_H \rangle$ on the set of all binary vectors of length m with respect to the Hamming metric. The *Hamming distance* $d_H(\cdot, \cdot)$ between two binary vectors is the number of coordinates in which they differ. The *Hamming weight* $wt_H(\cdot)$ of a vector is given by the equality $wt_H(\mathbf{v}) = d_H(\mathbf{v}, \mathbf{0})$, where $\mathbf{0}$ is the all-zero-vector. By $\mathfrak{F}_m$ we denote the set of all Boolean functions with m variables. By the distance $\mathrm{dist}(\cdot, \cdot)$ between two Boolean functions we mean (as usual) the Hamming distance between their vectors of values. A subset C of $\mathbb{Z}_2^m$ is called a *code*; m is the *length* of C. The smallest distance between distinct elements of the code (i.e. *codewords*) is called the *minimum distance* of C.

1.2. Binary product $\langle \cdot, \cdot \rangle_k$

Let $\langle\, \mathbf{u}, \mathbf{v} \,\rangle = u_1 v_1 \oplus \ldots \oplus u_m v_m$ be the inner product of binary vectors $\mathbf{u}$ and $\mathbf{v}$ over $\mathbb{Z}_2$. For any integer k, $0 \leqslant k \leqslant m/2$, we define the binary operation $\langle \cdot, \cdot \rangle_k : \mathbb{Z}_2^m \times \mathbb{Z}_2^m \to \mathbb{Z}_2$ by the following rule:

$$\langle \mathbf{u}, \mathbf{v} \rangle_k = \Big(\bigoplus_{i=1}^{k} \bigoplus_{j=i}^{k} Y_i Y_j \Big) \oplus \langle \mathbf{u}, \mathbf{v} \rangle, \tag{1}$$

where $Y_i = (u_{2i-1} \oplus u_{2i})(v_{2i-1} \oplus v_{2i})$ for any integer i, $1 \leqslant i \leqslant m/2$. It is easy to see that $\langle \mathbf{u}, \mathbf{v} \rangle_0 = \langle \mathbf{u}, \mathbf{v} \rangle$ and $\langle \mathbf{u}, \mathbf{v} \rangle_1 = \langle \mathbf{u}, \hat{\mathbf{v}} \rangle$, where $\hat{\mathbf{v}}$ is obtained from $\mathbf{v}$ by changing v_1 and v_2 over. For example

in case $m = 4$ we have $\langle\, \mathbf{u}, \mathbf{v} \,\rangle_1 = u_2v_1 \oplus u_1v_2 \oplus u_3v_3 \oplus u_4v_4$ and $\langle\, \mathbf{u}, \mathbf{v} \,\rangle_2 = (u_1 \oplus u_2)(u_3 \oplus u_4)(v_1v_3 \oplus v_1v_4 \oplus v_2v_3 \oplus v_2v_4) \oplus u_2v_1 \oplus u_1v_2 \oplus u_4v_3 \oplus u_3v_4$. The operation $\langle \cdot, \cdot \rangle_k$ seems to be an analog of the standard inner product.

Proposition 1.1. *For any integer m, k, such that $0 \leqslant k \leqslant m/2$, for any $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{Z}_2^m$ it holds*

(i) $\langle \mathbf{u}, \mathbf{v} \rangle_k = \langle \mathbf{v}, \mathbf{u} \rangle_k$;

(ii) $\langle a\mathbf{u}, \mathbf{v} \rangle_k = a\langle \mathbf{u}, \mathbf{v} \rangle_k$ *for arbitrary* $a \in \mathbb{Z}_2$;

(iii) $\sum\limits_{\mathbf{v} \in \mathbb{Z}_2^m} (-1)^{\langle \mathbf{u}, \mathbf{v} \rangle_k \oplus \langle \mathbf{w}, \mathbf{v} \rangle_k} = \begin{cases} 2^m, & \text{if } \mathbf{u} = \mathbf{w}, \\ 0 & \text{else.} \end{cases}$

For more complicated properties of $\langle \cdot, \cdot \rangle_k$ see sections 1.4 and 1.5.

1.3. k-Affine functions $\langle \mathbf{u}, \mathbf{v} \rangle_k \oplus a$

For a fixed k, $0 \leqslant k \leqslant m/2$, we define the following class of Boolean functions on m variables $v_1, \ldots, v_m$:

$$\mathfrak{A}_m^k = \{\, \langle \mathbf{u}, \mathbf{v} \rangle_k \oplus a \mid \mathbf{u} \in \mathbb{Z}_2^m, a \in \mathbb{Z}_2 \,\}. \tag{2}$$

A function from $\mathfrak{A}_m^k$ we call *k-affine function*. It is easy to see that size of $\mathfrak{A}_m^k$ (denote it by $|\mathfrak{A}_m^k|$) is equal to 2^{m+1}. We note that classes $\mathfrak{A}_m^0$ and $\mathfrak{A}_m^1$ coincide and give us the class of affine functions. That is why in what follows we often consider k, such that $1 \leqslant k \leqslant m/2$. Note that for the functions from $\mathfrak{A}_m^k$ the first $2k$ and the last $m - 2k$ variables play distinct roles. The partition into pairs is also important. So, there is "inequality of rights" for variables. If $g_{\mathbf{u}}^{(k)}(\mathbf{v}) = \langle\, \mathbf{u}, \mathbf{v} \,\rangle_k$, then it is true

$$\operatorname{dist}(f, g_{\mathbf{u}}^{(k)}) = 2^{m-1} - \tfrac{1}{2} W_f^{(k)}(\mathbf{u}),$$
$$\operatorname{dist}(f, g_{\mathbf{u}}^{(k)} \oplus 1) = 2^{m-1} + \tfrac{1}{2} W_f^{(k)}(\mathbf{u}),$$

where W_f^k is the k-Walsh-Hadamard transform and will be defined in 1.5. As we know the *degree* of a Boolean function is the number of variables in the longest item of its algebraic normal form.

Proposition 1.2. *For any integer m, k, $0 \leqslant k \leqslant m/2$, the class $\mathfrak{A}_m^k$ consists of $2^{m-k+1}(k+1)$ functions of degree 1 and $2^{m-k+1}(2^k - k - 1)$ functions of degree 2.*

1.4. From k-affine functions to Hadamard-like code A_m^k

At first the definition of $\langle \cdot, \cdot \rangle_k$ was given in terms of coding theory [19]. Binary vectors of values for all functions $\langle \mathbf{u}, \mathbf{v} \rangle_k \oplus a$

form the binary Hadamard-like code A_m^k of length 2^m, size 2^{m+1} and minimum distance 2^{m-1}. Let us present its construction and properties. Here m and k are any integer numbers, $0 \leqslant k \leqslant m/2$.

Let $\mathbf{G}_m^k$ be the $(m-k) \times 2^m$-matrix over $\mathbb{Z}_4$ that consists of the lexicographically ordered columns $\mathbf{z}^T$, where $\mathbf{z}$ runs through $\mathbb{Z}_4^k \times (2\mathbb{Z}_4)^{m-2k}$. Matrices of this type at first were considered by D. S. Krotov, see [7], [8] for the classification of $\mathbb{Z}_4$-linear Hadamard-like and perfect codes. For example,

$$\mathbf{G}_4^1 = \begin{pmatrix} 0000111122223333 \\ 0022002200220022 \\ 0202020202020202 \end{pmatrix}, \mathbf{G}_4^2 = \begin{pmatrix} 0000111122223333 \\ 0123012301230123 \end{pmatrix}.$$

We use the following notations. Let $\beta, \gamma : \mathbb{Z}_4^i \to \mathbb{Z}_2^i$ be coordinate-wise extensions of the mappings $\beta : 0,1 \to 0;\ 2,3 \to 1$ and $\gamma : 0,3 \to 0;\ 1,2 \to 1$. Let $\varphi : \mathbb{Z}_4^i \to \mathbb{Z}_2^{2i}$ be the coordinate-wise extension of the *Gray mapping*: $\varphi : a \to (\beta(a), \gamma(a))$ for any $a \in \mathbb{Z}_4$. Let the mapping $\varphi_k : \mathbb{Z}_4^k \times \mathbb{Z}_2^{m-2k} \to \mathbb{Z}_2^m$ be given by $\varphi_k : (\mathbf{u}', \mathbf{u}'') \to (\varphi(\mathbf{u}'), \mathbf{u}'')$ for any $\mathbf{u}' \in \mathbb{Z}_4^k, \mathbf{u}'' \in \mathbb{Z}_2^{m-2k}$. Consider the vector $\mathbf{h}^{\mathbf{u}} = \varphi_k^{-1}(\mathbf{u}) \cdot \mathbf{G}_m^k$ of length 2^m over $\mathbb{Z}_4$ for any $\mathbf{u} \in \mathbb{Z}_2^m$. Let $\mathbf{C}_m^k = (c_{\mathbf{u},\mathbf{v}}^k)$, $\mathbf{u}, \mathbf{v} \in \mathbb{Z}_2^m$, be the $2^m \times 2^m$-matrix over $\mathbb{Z}_4$ with the rows $\mathbf{h}^{\mathbf{u}}$. These rows we arrange in the lexicographical order of vectors $\varphi_k^{-1}(\mathbf{u})$. We will numerate the columns of $\mathbf{C}_m^k$ by the vectors $\mathbf{v}$ in the lexicographical order of vectors $\varphi_k^{-1}(\mathbf{v})$. For example, see matrices $\mathbf{C}_4^1$ and $\mathbf{C}_4^2$ with the required numbering of lines by vectors $\mathbf{u}, \mathbf{v}$. Here for the binary vector (u_1, u_2, u_3, u_4) we write the number $8u_1 + 4u_2 + 2u_3 + u_4$ in the hexadecimal system; e. g. we write 7 for (0111), B for (1011), etc.

$c^1_{\mathbf{u},\mathbf{v}}$	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**C**	**D**	**E**	**F**	**8**	**9**	**A**	**B**
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	2	0	2	0	2	0	2	0	2	0	2	0	2	0	2
2	0	0	2	2	0	0	2	2	0	0	2	2	0	0	2	2
3	0	2	2	0	0	2	2	0	0	2	2	0	0	2	2	0
4	0	0	0	0	1	1	1	1	2	2	2	2	3	3	3	3
5	0	2	0	2	1	3	1	3	2	0	2	0	3	1	3	1
6	0	0	2	2	1	1	3	3	2	2	0	0	3	3	1	1
7	0	2	2	0	1	3	3	1	2	0	0	2	3	1	1	3
C	0	0	0	0	2	2	2	2	0	0	0	0	2	2	2	2
D	0	2	0	2	2	0	2	0	0	2	0	2	2	0	2	0
E	0	0	2	2	2	2	0	0	0	0	2	2	2	2	0	0
F	0	2	2	0	2	0	0	2	0	2	2	0	2	0	0	2
8	0	0	0	0	3	3	3	3	2	2	2	2	1	1	1	1
9	0	2	0	2	3	1	3	1	2	0	2	0	1	3	1	3
A	0	0	2	2	3	3	1	1	2	2	0	0	1	1	3	3
B	0	2	2	0	3	1	1	3	2	0	0	2	1	3	3	1

$c^2_{\mathbf{u},\mathbf{v}}$	**0**	**1**	**3**	**2**	**4**	**5**	**7**	**6**	**C**	**D**	**F**	**E**	**8**	**9**	**B**	**A**
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
3	0	2	0	2	0	2	0	2	0	2	0	2	0	2	0	2
2	0	3	2	1	0	3	2	1	0	3	2	1	0	3	2	1
4	0	0	0	0	1	1	1	1	2	2	2	2	3	3	3	3
5	0	1	2	3	1	2	3	0	2	3	0	1	3	0	1	2
7	0	2	0	2	1	3	1	3	2	0	2	0	3	1	3	1
6	0	3	2	1	1	0	3	2	2	1	0	3	3	2	1	0
C	0	0	0	0	2	2	2	2	0	0	0	0	2	2	2	2
D	0	1	2	3	2	3	0	1	0	1	2	3	2	3	0	1
F	0	2	0	2	2	0	2	0	0	2	0	2	2	0	2	0
E	0	3	2	1	2	1	0	3	0	3	2	1	2	1	0	3
8	0	0	0	0	3	3	3	3	2	2	2	2	1	1	1	1
9	0	1	2	3	3	0	1	2	2	3	0	1	1	2	3	0
B	0	2	0	2	3	1	3	1	2	0	2	0	1	3	1	3
A	0	3	2	1	3	2	1	0	2	1	0	3	1	0	3	2

All matrices $\mathbf{C}_m^k$ are symmetric and can be constructed iteratively: $\mathbf{C}_{m+1}^k = (\mathbf{C}_m^k \otimes \mathbf{J}_2) + (\mathbf{J}_{2^m} \otimes \mathbf{C}_1^0)$ and $\mathbf{C}_{m+2}^{k+1} = (\mathbf{J}_4 \otimes \mathbf{C}_m^k) + (\mathbf{C}_2^1 \otimes \mathbf{J}_{2^m})$, where $\mathbf{J}_s$ is the all-one-matrix of order s, $\mathbf{C}_1^0 = \binom{00}{02}$, $A \otimes B$ is the Kronecker product $\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \dots & a_{1p}\mathbf{B} \\ \dots & \dots & \dots \\ a_{p1}\mathbf{B} & \dots & a_{pp}\mathbf{B} \end{pmatrix}$ for the $p \times p$-matrix $\mathbf{A}$ and $q \times q$-matrix $\mathbf{B}$, $+$ is taken over $\mathbb{Z}_4$.

Let $\mathbf{2}$ be the all-two-vector. The rows of matrices $\mathbf{C}_m^k$ and $\mathbf{C}_m^k + 2\,\mathbf{J}_{2^m}$ (i. e. all vectors $\mathbf{h^u}$ and $\mathbf{h^u} + \mathbf{2}$) form a group code $\mathcal{A}_m^k$ with operation $+$ over $\mathbb{Z}_4$, i. e. if $\mathbf{x}, \mathbf{y} \in \mathcal{A}_m^k$ then $\mathbf{x} + \mathbf{y}$ belongs to $\mathcal{A}_m^k$ too. Denote by A_m^k the binary code obtained from $\mathcal{A}_m^k$ by an action of β in every position of its codewords, $A_m^k = \beta(\mathcal{A}_m^k)$. Although the map $\beta : \mathbb{Z}_4^{2^m} \to \mathbb{Z}_2^{2^m}$ is not one-to-one, it is possible to inverse it on the set A_m^k (one can easily prove it). On codewords of A_m^k define the binary operation $\bullet : A_m^k \times A_m^k \to A_m^k$ connected with $+$ on the set $\mathcal{A}_m^k$. Let

$$\mathbf{x} \bullet \mathbf{y} = \beta(\beta^{-1}(\mathbf{x}) + \beta^{-1}(\mathbf{y})) \text{ for any vectors } \mathbf{x}, \mathbf{y} \in A_m^k.$$

It is easy to see that $(A_m^k, \bullet)$ is an Abelian group. Moreover the operation $\bullet$ is coordinated with the Hamming metric, i. e.

Proposition 1.3. *For any codewords $\mathbf{x}, \mathbf{y} \in A_m^k$ it holds*

$$d_H(\mathbf{x}, \mathbf{y}) = wt_H(\mathbf{x} \bullet \mathbf{y}^{-1}),$$

where $\mathbf{y}^{-1} \in A_m^k$ is the codeword such that $\mathbf{y} \bullet \mathbf{y}^{-1} = \mathbf{0}$.

Our products $\langle \cdot, \cdot \rangle_k$ can be defined like this.

Proposition 1.4. *For any $\mathbf{u}, \mathbf{v} \in \mathbb{Z}_2^m$ it holds $\langle \mathbf{u}, \mathbf{v} \rangle_k = \beta(c_{\mathbf{u},\mathbf{v}}^k)$.*

The good product's properties that we have are caused by this "geometric" definition of $\langle \cdot, \cdot \rangle_k$. With respect to the operation $\bullet$ we can say that the nonlinear product $\langle \cdot, \cdot \rangle_k$ has the "linear-like" properties coordinated with Hamming metric. This fact can be developed in details in the future.

1.5. k-Walsh-Hadamard transform

Now for k, $0 \leqslant k \leqslant m/2$, we define the function $W_f^{(k)} : \mathbb{Z}_2^m \to \mathbb{Z}$ by the rule

$$W_f^{(k)}(\mathbf{v}) = \sum_{\mathbf{u} \in \mathbb{Z}_2^m} (-1)^{\langle \mathbf{u}, \mathbf{v} \rangle_k \oplus f(\mathbf{u})} \text{ for any } \mathbf{v} \in \mathbb{Z}_2^m,$$

and call it *k-Walsh-Hadamard transform* of $f \in \mathfrak{F}_m$. It is obvious that $W_f^{(0)}$ is the standard Walsh-Hadamard transform of f and $W_f^{(0)}(\mathbf{v}) = W_f^{(1)}(\hat{\mathbf{v}})$, here $\hat{\mathbf{v}}$ is the vector defined in 1.2. For the k-Walsh-Hadamard transform of any $f \in \mathfrak{F}_m$ the *inversion formula* holds

$$(-1)^{f(\mathbf{u})} = \frac{1}{2^m} \sum_{\mathbf{v} \in \mathbb{Z}_2^m} (-1)^{\langle \mathbf{u}, \mathbf{v} \rangle_k} W_f^{(k)}(\mathbf{v}) \quad \text{for any } \mathbf{u} \in \mathbb{Z}_2^m,$$

Theorem 1.5. *For any m, k, $1 \leqslant k \leqslant m/2$, for any $f \in \mathfrak{F}_m$ the analog of Parseval's equality holds* $\sum\limits_{\mathbf{v} \in \mathbb{Z}_2^m} \left(W_f^{(k)}(\mathbf{v}) \right)^2 = 2^{2m}$.

The inversion formula and Parseval's equality for $W_f^{(k)}$ can be easily derived from Proposition 1.1. It follows from Theorem 1.5 that for any $f \in \mathfrak{F}_m$ it holds

$$\max_{\mathbf{v} \in \mathbb{Z}_2^m} |W_f^{(k)}(\mathbf{v})| \geqslant 2^{m/2}. \tag{3}$$

1.6. k-Bent functions

Let m, k be arbitrary, such that $1 \leqslant k \leqslant m/2$. Let *$k$-nonlinearity* of $f \in \mathfrak{F}_m$ be defined as $N_f^{(k)} = \min_{g \in \mathfrak{L}_m^k} \operatorname{dist}(f, g)$.

Proposition 1.6. *It holds* $N_f^{(k)} = 2^{m-1} - \frac{1}{2} \max\limits_{\mathbf{v} \in \mathbb{Z}_2^m} |W_f^{(k)}(\mathbf{v})|$.

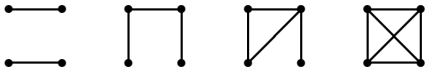
Thus, we have inequality $N_f^{(k)} \leqslant 2^{m-1} - 2^{(m/2)-1}$. Boolean function $f \in \mathfrak{F}_m$ we call *maximal k-nonlinear*, if each parameter $N_f^{(j)}$, $j = 1, \ldots, k$, gets the maximal possible value. For even m Boolean function $f \in \mathfrak{F}_m$ we call *k-bent function*, if all coefficients $W_f^{(j)}(\mathbf{v})$, $j = 1, \ldots, k$, are equal to $\pm 2^{m/2}$ (and hence $N_f^{(j)} = 2^{m-1} - 2^{(m/2)-1}$ for $j = 1, \ldots, k$). By constructing k-bent functions we show that for any even m two definitions given above do coincide. The class of all k-bent functions on m variables (m is even) we denote by $\mathfrak{B}_m^k$. We see that $\mathfrak{B}_m^1 \supseteq \cdots \supseteq \mathfrak{B}_m^{m/2}$, and $\mathfrak{B}_m^1$ coincides with the class of bent functions $\mathfrak{B}_m$. Further we show that here any symbol $\supseteq$ can be changed to $\supset$ and any set $\mathfrak{B}_m^k$ is not empty.

2. How to construct k-bent functions

2.1. Small values of m

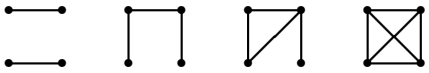
For $m = 2$ it is known that $\mathfrak{B}_2^1$ includes any Boolean function f such that its vector $\mathbf{f}$ of values has an odd Hamming weight. Thus, $|\mathfrak{B}_2^1| = 8$.

Let $m = 4$. We have $|\mathfrak{B}_4^1| = 896$ and any function from $\mathfrak{B}_4^1$ is quadratic (i. e. of degree 2). It is well known (see [10]) that the set $\mathfrak{B}_4^1$ of all 1-bent functions can be divided into 28 classes. Each class contains 32 Boolean functions with the same quadratic part and all possible linear parts. These 28 different types of quadratic parts can be given with the diagram:



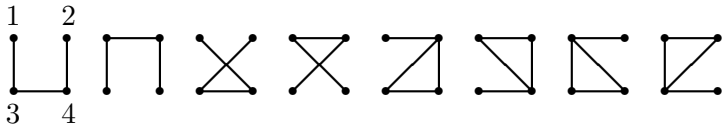
(3) (12) (12) (1)

Here vertices correspond to variables, edges correspond to the quadratic items in algebraic normal form of a function. We have proven that $|\mathfrak{B}_m^2| = 384$. The class $\mathfrak{B}_m^2$ consists of 12 classes. Every class again contains 32 Boolean functions distinct from each other only by a linear part. These 12 quadratic parts are of the following types



(3) (4) (4) (1)

In details, this series of graphs is obtained from the previous one by taking only those graphs for which the number of non horizontal edges is *even.* The ordering of vertices is important. We should numerate them from left to right and then from up to down. Below one can see these special 8 graphs with 3 or 4 edges:



For example, the function $\xi(v_1, v_2, v_3, v_4) = v_1v_2 \oplus v_2v_3 \oplus v_3v_4$ is 1-bent, but not 2-bent.

2.2. Iterative construction

Now we give simple iterative constructions for k-bent functions.

Proposition 2.1. *Let m, r be even, k be integer, $1 \leqslant k \leqslant m/2$. Let a function $f \in \mathfrak{F}_{m+r}$ be represented in the form*

$$f(\mathbf{u}', \mathbf{u}'') = p(\mathbf{u}') \oplus q(\mathbf{u}'') \text{ for any } \mathbf{u}' \in \mathbb{Z}_2^m, \mathbf{u}'' \in \mathbb{Z}_2^r,$$

where $p \in \mathfrak{F}_m$, $q \in \mathfrak{F}_r$ are functions with nonintersecting sets of variables. Then $f \in \mathfrak{B}_{m+r}^k$ if and only if $p \in \mathfrak{B}_m^k$, $q \in \mathfrak{B}_r^1$.

We know that a Boolean function $s \in \mathfrak{F}_m$ is called *symmetric* if it holds $s(\pi(\mathbf{v})) = s(\mathbf{v})$ for any permutation π on m variables. The set of all symmetric functions on 2 variables we denote by $\mathfrak{F}_2^1$.

Proposition 2.2. *Let m be even, k be integer, $1 \leqslant k \leqslant m/2$. Let a function $f \in \mathfrak{F}_{m+2}$ be represented in the form*

$$f(a, a', \mathbf{u}) = s(a, a') \oplus p(\mathbf{u}) \text{ for any } a, a' \in \mathbb{Z}_2, \mathbf{u} \in \mathbb{Z}_2^m,$$

where $s \in \mathfrak{F}_2^1$, $p \in \mathfrak{F}_m$ are functions with nonintersecting sets of variables. Then $f \in \mathfrak{B}_{m+2}^{k+1}$ if and only if $s \in \mathfrak{B}_2^1$, $p \in \mathfrak{B}_m^k$.

Summarizing both Propositions 2.1 and 2.2 we have

Theorem 2.3. *Let $m, r \geqslant 0$ be even, $j \geqslant 0, k \geqslant 1$ be integer, $k \leqslant m/2$. Let a function $f \in \mathfrak{F}_{2j+m+r}$ be represented in the form*

$$f(a_1, a_1', \ldots, a_j, a_j', \mathbf{u}', \mathbf{u}'') = \left(\bigoplus_{i=1}^{j} s_i(a_i, a_i') \right) \oplus p(\mathbf{u}') \oplus q(\mathbf{u}''),$$

where $s_1, \ldots, s_j \in \mathfrak{F}_2^1, p \in \mathfrak{F}_m$ and $q \in \mathfrak{F}_r$ are functions with nonintersecting sets of variables. Then $f \in \mathfrak{B}_{2j+m+r}^{j+k}$ if and only if $s_1, \ldots, s_j \in \mathfrak{B}_2^1$, $p \in \mathfrak{B}_m^k$ and $q \in \mathfrak{B}_r^1$.

Let m be even ($m \geqslant 2$), k be any integer ($1 \leqslant k \leqslant m/2$). Then

Corollary 2.4. *The class $\mathfrak{B}_m^k$ is not empty.*

Corollary 2.5. *It holds $\mathfrak{B}_m^1 \supset \cdots \supset \mathfrak{B}_m^{m/2}$.*

Corollary 2.6. *If $m \geqslant 4$ then for any integer d, $2 \leqslant d \leqslant \max\{2, (m/2) - k + 1\}$, there exists k-bent function of degree d.*

Corollary 2.7. *It is true that $|\mathfrak{B}_m^k| \geqslant 2^{2k-2} |\mathfrak{B}_{m-2k+2}^1|$.*

But the last bound seems to be very rough.

2.3. Several properties of k-bent functions

Let S_m be the symmetric group on m elements. Denote by $S_{m,k}$ the subgroup of S_m generated by all transpositions $(1,2)$, $(3,4)$, $\dots$, $(2k-1,2k)$. Let $\mathfrak{F}_m^k$ be the set of all Boolean functions from $\mathfrak{F}_m$ that are constant on the each orbit of $\mathbb{Z}_2^m$ by an action of $S_{m,k}$. The number of such orbits is equal to $3^k 2^{m-2k}$, and hence we have $|\mathfrak{F}_m^k| = 2^{3^k 2^{m-2k}} = 2^{2^{m-k\log_2\frac{4}{3}}}$.

Theorem 2.8. *For any even m, any integer k, $1 \leqslant k \leqslant m/2$, it is true that $\mathfrak{F}_m^k \cap \mathfrak{B}_m^k = \mathfrak{F}_m^k \cap \mathfrak{B}_m^1$.*

Theorem 2.8 helps to study known bent functions in order to determine their "k-bentness". It seems to be an interesting question.

3. The class Δ_m of functions for approximation

Let m be even, k be integer, $1 \leqslant k \leqslant m/2$. For a permutation $\pi \in S_m$ define the following class of functions on variables $v_1, \dots, v_m$: $\mathfrak{A}_{m,0}^k(\pi) = \{\, \langle\, \mathbf{u}, \pi(\mathbf{v}) \,\rangle_k \,|\, \mathbf{u} \in \mathbb{Z}_2^m \,\}$. Note that for any π the set $\mathfrak{A}_{m,0}^1(\pi)$ consists of all linear functions. We will approximate Boolean functions in ciphers with functions from the class

$$\Delta_m = \bigcup_{1\leqslant k\leqslant m/2} \bigcup_{\pi\in S_m} \mathfrak{A}_{m,0}^k(\pi).$$

We call it the *class of approximating functions*. To say informally by using distinct permutations π here we make all variables $v_1, \dots, v_m$ be "equal in rights" in this set of functions (see the remark in 1.3).

Let $\mathrm{ANF}(f)$ be the set of all conjunctions in *algebraic normal form* of a function $f \in \mathfrak{F}_m$. For $\pi \in S_m$ by f^π we denote the Boolean function given by the equality $f^\pi(\mathbf{v}) = f(\pi(\mathbf{v}))$. Denote by $\mathrm{Act}(f)$ the subset of $\{1, 2, \dots, m/2\}$ with the maximal possible size such that for any distinct elements i, j from $\mathrm{Act}(f)$ (we call them *active*) conjunctions $v_{2i-1}v_{2j-1}$, $v_{2i-1}v_{2j}$, $v_{2i}v_{2j-1}$ and $v_{2i}v_{2j}$ belong to $\mathrm{ANF}(f)$. Note that $|\mathrm{Act}(f)| = 0$ or $|\mathrm{Act}(f)| \geqslant 2$ for any f (the case $|\mathrm{Act}(f)| = 1$ has no sense). By ρ we denote a permutation on m elements such that $|\mathrm{Act}(f^\rho)| = \max_{\pi\in S_m} |\mathrm{Act}(f^\pi)|$; it can be chosen in several ways.

E. g. if $\mathrm{ANF}(f) = \{v_1v_2,\ v_1v_3,\ v_1v_4,\ v_2v_3,\ v_3v_4,\ v_2,\ v_3,\ 1\}$ then $\mathrm{Act}(f) = \varnothing$, but $\mathrm{Act}(f^{(1,3,2,4)}) = \{1, 2\}$.

Let us characterize Δ_m in terms of ANFs.

Theorem 3.1. *Boolean function $f \in \mathfrak{F}_m$ of degree less or equal to 2, such that $f(\mathbf{0}) = 0$, belongs to Δ_m if and only if it holds*

1) *conjunctions $v_{2i-1}v_{2j-1}$, $v_{2i-1}v_{2j}$, $v_{2i}v_{2j-1}$, $v_{2i}v_{2j}$ belong (or not) to $\mathrm{ANF}(f^\rho)$ simultaneously for any i, j $(1 \leqslant i \neq j \leqslant m/2)$;*

2) *there are no conjunctions of the form $v_{2i-1}v_{2i}$ in $\mathrm{ANF}(f^\rho)$;*

3) *if $i \in \mathrm{Act}(f^\rho)$ then exactly one variable among v_{2i-1}, v_{2i} belongs to $\mathrm{ANF}(f^\rho)$.*

Corollary 3.2. *For any even m it is true that*

$$|\Delta_m| = 2^m \left(1 + \sum_{k=2}^{m/2} \binom{m}{2k} \frac{(2k-1)!!}{2^k}\right).$$

Here $(2k-1)!!$ is equal to $1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2k-1)$. For instance, $|\Delta_4| = 28$, $|\Delta_6| = 904$, $|\Delta_8| = 28816$. But the numbers of linear functions in these classes are 16, 64 and 256. Theorem 3.1 tells us how to run through Δ_m without repetitions. At first consider all linear functions. Then for any k, $2 \leqslant k \leqslant m/2$, take all distinct k-subsets of pairwise nonintersecting 2-subsets of $\{1, \ldots, m\}$. Each k-subset of pairs uniquely determines the quadratic part of a function. For any chosen subset take all possible (see Theorem 3.1) linear parts.

4. Quadratic approximations

In this section we give our modification for the well known method of linear cryptanalysis (M. Matsui, 1993). The main idea of it: to use (linear and quadratic) Boolean functions from Δ_m in approximations.

Consider a block cipher with r rounds of ciphering. We use the following notations:

$m = m_{\text{text}}$ — an even length of a plaintext and a ciphertext;

P — a plaintext, $P \in \mathbb{Z}_2^m$;

m_{key} — an even length of a key;

K — a key for ciphering, $K \in \mathbb{Z}_2^{m_{\text{key}}}$;

$F : \mathbb{Z}_2^m \times \mathbb{Z}_2^{m_{\text{key}}} \to \mathbb{Z}_2^m$ — a transform that is one-to-one if we fix the second argument;

$C = F(P, K)$ — a ciphertext, $C \in \mathbb{Z}_2^m$;

m'_{key} — an even length of a subkey for a round;

$K^{(i)}$ — a subkey for the i-th round of ciphering, $K^{(i)} \in \mathbb{Z}_2^{m'_{\text{key}}}$, $1 \leqslant i \leqslant r$, determined by the key K;

$F_i : \mathbb{Z}_2^m \times \mathbb{Z}_2^{m'_{\text{key}}} \to \mathbb{Z}_2^m$ — a transform for the i-th round of ciphering, $1 \leqslant i \leqslant r$; it is one-to-one for a fixed second argument;

$C^{(0)} = P$; $C^{(i)} = F_i(C^{(i-1)}, K^{(i)})$ — an intermediate ciphertext, $C^{(i)} \in \mathbb{Z}_2^m$, $1 \leqslant i \leqslant r$; $C = C^{(r)}$ — the ciphertext;

Suppose that all plaintexts P (and keys K) are equiprobable.

4.1. **The first algorithm**

We introduce a modification of the Matsui's algorithm for the one key bit determination. It is based on the equality

$$\langle\, \mathbf{a}, \pi(P) \,\rangle_i \,\oplus\, \langle\, \mathbf{b}, \sigma(C) \,\rangle_j = \langle\, \mathbf{d}, \tau(K) \,\rangle_k, \tag{4}$$

where $\mathbf{a}, \mathbf{b} \in \mathbb{Z}_2^m$, $\mathbf{d} \in \mathbb{Z}_2^{m_{\text{key}}}$ — chosen vectors; $\pi, \sigma \in S_m$, $\tau \in S_{m_{\text{key}}}$ — fixed permutations; i, j, k — integer numbers such that $1 \leqslant i, j \leqslant m/2$, $1 \leqslant k \leqslant m_{\text{key}}/2$. Assume that (4) holds with probability $p = \frac{1}{2} + \varepsilon$, where $0 < |\varepsilon| \leqslant 1/2$. The number ε we call the *bias* of (4). To choose parameters $\mathbf{a}, \mathbf{b}, \mathbf{d}, \pi, \sigma, \tau, i, j, k$ for the value $|\varepsilon|$ to be maximal — it is the next task for a concrete algorithm of ciphering. We do not consider this task here. Let us note that if the parameter i, j or k is equal to 1, then dependence on bits of the corresponding block P, C or K in (4) is linear.

Let us fix a key K. Consider the set $\{\, (P_t, C_t) \mid t = 1, \ldots, N \,\}$ of known pairs (plaintext,ciphertext), where $C_t = F(P_t, K)$. The algorithm (as in the linear case) is based on the principle of maximum likelihood.

Algorithm 1

- define $N_0 = |\, \{\, t \,:\, \langle\, \mathbf{a}, \pi(P_t) \,\rangle_i \,\oplus\, \langle\, \mathbf{b}, \sigma(C_t) \,\rangle_j = 0 \,\}\,|$;
- guess $\langle\, \mathbf{d}, \tau(K) \,\rangle_k = \begin{cases} 0, & \text{if } (N_0 - \frac{N}{2}) \cdot \varepsilon > 0; \\ 1, & \text{else;} \end{cases}$
- try to find K using the correlation obtained.

The end

The reliability ξ_0 of Algorithm 1 (the mathematical expectation of its correct work) can be estimated in the same way as in the case of linear cryptanalysis:

$$\xi_0 \simeq \int_{-2|\varepsilon|\sqrt{N}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy.$$

For a fixed key K, for any integer i, j, such that $1 \leqslant i, j \leqslant m/2$, for any permutations $\pi, \sigma \in S_m$ denote by $\varepsilon^{i,\mathbf{a},\pi}_{j,\mathbf{b},\sigma}(K;0)$ the real number between $-1/2$ and $1/2$ such that the probability of $\langle\, \mathbf{a}, \pi(P)\,\rangle_i \oplus \langle\, \mathbf{b}, \sigma(F(P,K))\,\rangle_j = 0$ equals $1/2 + \varepsilon^{i,\mathbf{a},\pi}_{j,\mathbf{b},\sigma}(K;0)$.

Proposition 4.1. *For any map $F(\cdot, K) : \mathbb{Z}_2^m \to \mathbb{Z}_2^m$, any $i, j, \mathbf{a}, \mathbf{b}, \pi$ and σ it holds*

$$2^{m+1} \cdot \varepsilon^{i,\mathbf{a},\pi}_{j,\mathbf{b},\sigma}(K;0) = W^{(i)}_{\langle\, \mathbf{b}, \sigma(F(\pi^{-1}(\cdot),K))\,\rangle_j}(\mathbf{a}).$$

This proposition gives us the strict connection between the bias $\varepsilon^{i,\mathbf{a},\pi}_{j,\mathbf{b},\sigma}(K;0)$ and k-Walsh-Hadamard coefficients of the Boolean function $\langle\, \mathbf{b}, \sigma(F(\pi^{-1}(\cdot), K))\,\rangle_j$.

Let $\varepsilon(K)$ be a bias of (4) for a fixed key K. Note that for any k, $\mathbf{d}$ and τ it is true $|\varepsilon(K)| = |\varepsilon^{i,\mathbf{a},\pi}_{j,\mathbf{b},\sigma}(K;0)|$.

Theorem 4.2. *Let $K \in \mathbb{Z}_2^{m_{\text{key}}}$ be fixed. Let $\mathbf{b} \in \mathbb{Z}_2^m$, $\pi, \sigma \in S_m$ and j, $1 \leqslant j \leqslant m/2$, be such that the function*

$$\langle\, \mathbf{b}, \sigma(F(\pi^{-1}(\cdot), K))\,\rangle_j : \mathbb{Z}_2^m \to \mathbb{Z}_2^m$$

is $(m/2)$-bent. Then we have

$$\max_{i,k,\mathbf{a},\mathbf{d},\tau} |\varepsilon(K)| = \min_{i,k,\mathbf{a},\mathbf{d},\tau} |\varepsilon(K)| = 2^{-(m/2)-1}.$$

It follows from the inequality (3) that $2^{-(m/2)-1}$ is the minimal possible value for $\max_{\mathbf{a}\in\mathbb{Z}_2^m} |\varepsilon(K)|$ when $i, j, k, \mathbf{b}, \mathbf{d}, \pi, \sigma$ and τ are arbitrary. Theorem 4.2 tells us that in order to achieve this minimum we should use $(m/2)$-bent functions for constructing F. But to make it possible in practice seems to be a hard problem.

Note that "linear-like" properties of $\langle \cdot, \cdot \rangle_k$ (see 1.4) can be used in approximations too. But it requires for a separate investigation.

For $\mathbf{u} = (u_1, \ldots, u_m)$ let $\bar{\mathbf{u}}^k = (u_1 \oplus u_2, \ldots, u_{2k-1} \oplus u_{2k})$ be a vector of length k. By $*$ denote the componentwise multiplication of vectors. In a concrete cipher the following fact can be helpful for the joint approximation to be made.

Proposition 4.3. *For any $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{Z}_2^m$, $1 \leqslant k \leqslant m/2$, it holds*

$$\langle \mathbf{u} \oplus \mathbf{v}, \mathbf{w} \rangle_k = \langle \mathbf{u}, \mathbf{w} \rangle_k \oplus \langle \mathbf{v}, \mathbf{w} \rangle_k \oplus \langle \bar{\mathbf{u}}^k, \bar{\mathbf{w}}^k \rangle \cdot \langle \bar{\mathbf{v}}^k, \bar{\mathbf{w}}^k \rangle \oplus \bar{\mathbf{u}}^k * \bar{\mathbf{v}}^k * \bar{\mathbf{w}}^k.$$

4.2. **The second algorithm**

Here we introduce a modification of the improved Matsui's algorithm (see [11]). Let s_1, s_2, $0 \leqslant s_1 < s_2 \leqslant r$, be integers.

Consider the equality

$$\langle\, \mathbf{a}, \pi(C^{(s_1)})\,\rangle_i \oplus \langle\, \mathbf{b}, \sigma(C^{(s_2)})\,\rangle_j = \langle\, \tau(\mathbf{d}), K\,\rangle_k, \qquad (5)$$

where $\mathbf{a}, \mathbf{b} \in \mathbb{Z}_2^m$, $\mathbf{d} \in \mathbb{Z}_2^{m_{\text{key}}}$ are fixed vectors; $\pi, \sigma \in S_m$, $\tau \in S_{m_{\text{key}}}$ are given permutations; i, j, k are integer numbers such that $1 \leqslant i, j \leqslant m/2$, $1 \leqslant k \leqslant m_{\text{key}}/2$. Assume that (5) holds with probability $\tilde{p} = \frac{1}{2} + \tilde{\varepsilon}$, where $0 < |\tilde{\varepsilon}| \leqslant 1/2$. Suppose that vectors P and C are known. Denote by $\tilde{K}$ those bits of a key K, which we need to know for the values $\langle\, \mathbf{a}, \pi(C^{(s_1)})\,\rangle_i$ and $\langle\, \mathbf{b}, \sigma(C^{(s_2)})\,\rangle_j$ to be determined. Let m_{s_1,s_2} be the number of bits in $\tilde{K}$.

Algorithm 2

- for any $\tilde{K} \in \mathbb{Z}_2^{m_{s_1,s_2}}$ determine

$$N_0(\tilde{K}) = |\,\{\, t \,:\, \langle\, \mathbf{a}, \pi(C_t^{(s_1)})\,\rangle_i \oplus \langle\, \mathbf{b}, \sigma(C_t^{(s_2)})\,\rangle_j = 0\,\}\,|;$$

- arrange all vectors from $\mathbb{Z}_2^{m_{s_1,s_2}}$: $\tilde{K}_1, \ldots, \tilde{K}_{2^{m_{s_1,s_2}}}$, in such a way that $\left|\frac{N}{2} - N_0(\tilde{K}_1)\right| \geqslant \ldots \geqslant \left|\frac{N}{2} - N_0(\tilde{K}_{2^{m_{s_1,s_2}}})\right|$;

- for any q from 1 to $2^{m_{s_1,s_2}}$ do
  - ▷ guess $\langle\, \mathbf{d}, \tau(K)\,\rangle_k = \begin{cases} 0, & \text{if } (N_0(\tilde{K}_q) - \frac{N}{2}) \cdot \tilde{\varepsilon} > 0; \\ 1, & \text{else;} \end{cases}$
  - ▷ try to find K using the correlation obtained.

The end

The same facts on $\tilde{\varepsilon}$ as in the case of the first algorithm can be obtained. We omit them here.

5. Permutations in S-boxes

For a binary vector $\mathbf{x} = (x_1, x_2, x_3, x_4)$ denote by $\tilde{x}$ the number $8x_1 + 4x_2 + 2x_3 + x_4$ from 0 to 15. Let p_1, p_2, p_3, p_4 and c_1, c_2, c_3, c_4 be respectively binary inputs and outputs of a 4-bit permutation $S : P \to C$, i.e. $S(\tilde{P}) = \tilde{C}$. Using Δ_4 we find the most probable equalities on input and output bits of S. According to Corollary 3.2 we have $|\Delta_4| = 28$. There are 16 linear functions. 12 quadratic functions we can list in the way: $\langle \mathbf{u}, (v_1, v_2, v_3, v_4)\rangle_2$, $\langle \mathbf{u}, (v_1, v_3, v_2, v_4)\rangle_2$, $\langle \mathbf{u}, (v_1, v_4, v_2, v_3)\rangle_2$, where $\tilde{u}$ gets values 5, 6, 9 and 10. Consider equalities

$$\langle\, \mathbf{a}, \pi(P)\,\rangle_i \,\oplus\, \langle\, \mathbf{b}, \sigma(C)\,\rangle_j = 0, \tag{6}$$

where if $i = 1$ we take vectors $\mathbf{a}$ with $\tilde{a}$ equal to $0, \ldots, 15$ and identity permutation π; if $i = 2$ we take vectors $\mathbf{a}$ with $\tilde{a} = 5, 6, 9, 10$ and permutations $\pi = \mathrm{id}, (1,3,2,4), (1,3,4,2)$ (the same we do for $\mathbf{b}$ and σ in the cases $j = 1$, $j = 2$). With these conditions functions $\langle\, \mathbf{a}, \pi(\cdot)\,\rangle_i$ and $\langle\, \mathbf{b}, \sigma(\cdot)\,\rangle_j$ run trough Δ_4 without repetitions. For S consider the table with rows numbered by triples $(i, \tilde{a}, \pi)$ and columns numbered by triples $(j, \tilde{b}, \sigma)$, such that an element of the table is the bias of the corresponding equality (6) multiplied by 16. In other words an element of the table is the number of times when (6) holds minus 8. Note that one can construct this table for any ℓ-bit permutation, $\ell > 4$.

Let us define the *non-quadratic characteristic* of a permutation S in the following way:

$$NQ(S) = \min_{i,j} \min_{\mathbf{a}\neq\mathbf{0},\mathbf{b}\neq\mathbf{0}} \min_{\delta\in\mathbb{Z}_2,\pi,\sigma} |\,\{P : \langle\, \mathbf{a}, \pi(P)\,\rangle_i \,\oplus\, \langle\, \mathbf{b}, \sigma(C)\,\rangle_j \neq \delta\}\,|.$$

In other words $NQ(S)$ equals the difference of 8 and the maximal absolute value of an element in the table (excepting elements for zero combinations of bits).

By the *nonlinearity* of a permutation S we as usual mean

$$NL(S) = \min_{\mathbf{a}\neq\mathbf{0},\mathbf{b}\neq\mathbf{0}} \min_{\delta,\pi,\sigma} |\,\{P : \langle\, \mathbf{a}, \pi(P)\,\rangle_1 \,\oplus\, \langle\, \mathbf{b}, \sigma(C)\,\rangle_1 \neq \delta\}\,|.$$

The part of the table given by $i = j = 1$ corresponds only to linear equalities of bits. So, the difference of 8 and the maximal absolute value of an element in this part (excepting elements of the first lines) is equal to $NL(S)$. Obviously, $NQ(S) \leqslant NL(S)$. It is known fact [4] that for any 4-bit permutation S it holds $NL(S) \leqslant 4$.

For instance, for $S' = (0, 1, 9, 14, 13, 11, 7, 6, 15, 2, 12, 5, 10, 4, 3, 8)$ we have the Table 1. We see that $NQ(S') = 2$ but $NL(S') = 4$. There are 7 quadratic equations on input and output bits of S' with probability 7/8, although any linear equality has probability not higher than 3/4. Among these 7 equalities there is for example $c_3 = \langle (0110), P\rangle_2 = p_1p_3 \oplus p_1p_4 \oplus p_2p_3 \oplus p_2p_4 \oplus p_1 \oplus p_4$ that is linear with respect to output bits.

We test [21] permutations with the most high nonlinearity $NL = 4$ recommended for using in S-boxes of GOST 28147-89, DES, s^3DES (see for them [14], [17]) and found that for all of them (excepting one) there are special quadratic relations on input and output bits with probability 7/8 whereas any linear equality has

probability not more then $3/4$. And only for one permutation $S'' = (13, 10, 0, 7, 3, 9, 14, 4, 2, 15, 12, 1, 5, 6, 11, 8)$ we have $NL(S'') = 4$ and $NQ(S'') = 4$; in this case we can not deliver any additional information in comparison with linear approximations.

$16 \cdot \varepsilon^{i,\mathbf{a},\pi}_{j,\mathbf{b},\sigma}$		$j=1$ id																$j=2$ id				$j=2$ (1,3,2,4)				$j=2$ (1,3,4,2)			
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	5	6	9	10	5	6	9	10	5	6	9	10
$i=1$ id	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	-2	0	2	-2	0	-2	**4**	0	2	0	-2	2	0	2	**4**	-2	0	0	2	0	2	2	0	0	0	**4**	0
	2	0	0	0	0	0	**4**	**4**	0	2	2	-2	-2	2	-2	2	-2	2	(6)	0	0	2	**4**	-2	0	**4**	2	0	2
	3	0	2	0	-2	-2	0	2	0	-2	**4**	2	**4**	0	-2	0	2	0	2	**4**	2	2	-2	0	0	**4**	-2	0	2
	4	0	-2	0	-2	0	2	0	2	0	-2	**4**	2	0	2	**4**	-2	**4**	-2	0	2	**4**	-2	0	2	0	0	-2	-2
	5	0	0	**4**	0	2	2	-2	2	0	**4**	0	0	-2	2	-2	-2	-2	2	0	**4**	0	0	-2	2	0	**-4**	2	2
	6	0	-2	**4**	2	0	-2	0	-2	2	0	-2	**4**	2	0	2	0	-2	0	0	-2	-2	2	2	-2	0	2	2	0
	7	0	**4**	0	0	2	2	-2	2	-2	-2	-2	2	**4**	0	0	0	0	0	**-4**	0	2	**4**	0	-2	0	2	2	**-4**
	8	0	0	2	-2	0	0	2	-2	0	0	2	-2	**4**	**4**	-2	2	2	0	2	0	**4**	0	2	-2	**4**	2	0	-2
	9	0	2	2	**4**	-2	**4**	0	-2	0	-2	2	0	-2	0	0	2	**4**	0	-2	2	0	2	0	(6)	**-4**	2	0	2
	10	0	0	-2	2	**4**	0	-2	-2	2	2	**4**	0	2	-2	0	0	0	-2	2	**4**	2	0	**4**	2	0	0	**4**	0
	11	0	-2	-2	**4**	2	0	**4**	2	-2	0	0	2	0	2	-2	0	2	2	2	-2	-2	2	2	2	0	**4**	0	**4**
	12	0	2	2	0	0	-2	2	**4**	**4**	-2	2	0	0	-2	-2	0	2	-2	2	-2	0	-2	2	0	0	2	-2	0
	13	0	0	-2	-2	2	2	0	0	**4**	0	-2	2	-2	2	2	0	**4**	0	2	-2	0	0	0	**-4**	0	0	-2	-2
	14	0	2	2	0	**4**	-2	2	0	-2	0	0	-2	-2	0	**4**	2	0	0	2	-2	-2	-2	0	0	0	0	-2	2
	15	0	**4**	-2	2	-2	-2	0	0	2	2	0	0	0	**4**	2	-2	-2	0	2	0	-2	0	2	0	0	0	2	2
$i=2$ id	5	0	-2	2	0	**4**	-2	-2	0	2	**4**	0	2	2	0	0	-2	**-4**	0	2	2	0	0	2	-2	2	-2	**4**	0
	6	0	0	(6)	2	-2	2	0	0	0	0	-2	2	-2	2	0	0	0	2	-2	0	-2	2	-2	2	-2	0	0	2
	9	0	0	0	**4**	0	0	0	**-4**	2	-2	2	2	2	-2	2	2	2	-2	0	0	0	2	**4**	2	-2	**4**	2	0
	10	0	2	0	2	2	**4**	-2	0	0	2	**4**	-2	-2	0	-2	0	2	0	0	(6)	2	0	0	(6)	-2	-2	2	2
$i=2$ (1,3,2,4)	5	0	0	2	2	**4**	0	-2	2	**4**	0	2	-2	0	0	-2	-2	0	-2	0	2	0	0	2	2	-2	0	2	0
	6	0	2	**4**	-2	-2	0	2	**4**	0	2	0	2	-2	0	-2	0	0	2	2	0	0	-2	-2	0	2	-2	-2	2
	9	0	2	-2	0	0	-2	2	0	2	0	**4**	2	2	**-4**	0	2	2	-2	**4**	0	2	-2	**4**	0	2	2	0	0
	10	0	0	0	0	2	2	-2	-2	-2	(6)	2	2	0	0	0	0	-2	2	2	(6)	2	0	0	2	2	**-4**	**4**	2
$i=2$ (1,3,4,2)	5	0	0	**4**	**4**	0	0	0	0	**4**	**-4**	0	0	0	0	0	0	2	-2	-2	-2	-2	2	2	2	**-4**	**4**	0	0
	6	0	0	2	-2	0	**-4**	2	2	2	2	0	**4**	2	-2	0	0	-2	0	**4**	-2	0	-2	2	**-4**	**4**	0	0	0
	9	0	**4**	0	0	-2	2	2	2	0	0	**4**	0	-2	-2	-2	2	**4**	0	2	2	2	-2	0	**4**	0	0	-2	2
	10	0	0	0	2	2	-2	2	0	**-4**	-2	2	0	**4**	0	0	2	2	0	2	0	2	0	2	0	2	0	0	2	2

Table 1. The biases for the permutation S'.

References

[1] Carlet C. *Boolean Functions for Cryptography and Error Correcting Codes*, chap. of the monogr. *Boolean methods and models*, Cambridge Univ. Press (P. Hammer, Y. Crama eds.), to appear. Prelim. version is available at `www-rocq.inria.fr/secret/Claude.Carlet/chap-fcts-Bool.pdf`.

[2] Dillon J. F. *A survey of bent functions* Special Issue of the *NSA Technical Journal*, 1972. P. 191–215.

[3] Dobbertin H., Leander G. *A survey of some recent results on bent functions* Proc. of the Int. conf. on *Sequences and their applications* (Seul, Korea. October 24–28, 2004). Berlin: Springer, 2005. P. 1–29 (LNCS 3486).

[4] Heys H. M., Tavares S. E. *Substitution-permutation networks resistant to differential and linear cryptanalysis*, volume 9 of *J. Cryptology*, 1996. N 1. P. 1–19.

[5] Ivanov A. V. *The use of the reduced representation of the Boolean functions in their nonlinear approximations' construction*, Bulletin of Tomsk State University. Supplement. 2007. N 23. P. 31–35.

[6] Knudsen L. R., Robshaw M. J. B. *Non-linear approximation in linear cryptanalysis* Proc. of EUROCRYPT'96. (Saragossa, Spain. May 12–16, 1996). P. 224–236 (LNCS 1070).

[7] Krotov D. S. $\mathbb{Z}_4$*-linear perfect codes*, volume 7 of *Discrete Analysis and Operation Research*, 2000. N. 4. P. 78–90 (in Russian). English translation is available at `http://arxiv.org/abs/0710.0198`.

[8] Krotov D. S. $\mathbb{Z}_4$*-linear Hadamard and extended perfect codes*, Proc. of the *Int. Workshop on Coding and Cryptography* (Paris, France. January 8–12, 2001). P. 329–334.

[9] Logachev O. A., Sal'nikov A. A., Yashenko V. V. *Boolean functions in coding theory and cryptology.* Moscow center for the uninterrupted mathematical education, 2004 (in Russian).

[10] MacWilliams F. J., Sloane N. J. A. *Theory of error correcting codes.* North-Holland: Amsterdam, 1977.

[11] Matsui M. *Linear cryptanalysis method for DES cipher* Proc. of EUROCRYPT'93. (Lofthus, Norway. May 23–27, 1993) P. 386–397 (LNCS 765).

[12] McFarland R. L. *A family of difference sets in non-cyclic groups*, volume 15 of *Journal of Combin. Theory*, Ser. A. 1973. N 1. P. 1–10.

[13] Nakahara J., Preneel B., Vandewalle J. *Experimental Non-Linear Cryptanalysis*, COSIC Internal Report. Katholieke Univ. Leuven. 2003. 17 p.

[14] Rostovtsev A., Mahovenko E. *Introduction to the theory of iterative ciphers*, Saint-Petersburg, 2003 (in Russian).

[15] Rothaus O. *On bent functions*, volume 20 of *Journal of Combin. Theory*, Ser. A. 1976. N 3. P. 300–305.

[16] Shimoyama T., Kaneko T. *Quadratic relation of S-box and its application to the linear attack of full round DES*, Proc. of CRYPTO'98. (Santa Barbara, California. USA. August 23–27, 1998) P. 200–211 (LNCS 1462).

[17] Shneier B. *Applied cryptography. Protocols, algorithms and source code in C.* 2002.

[18] Tapiador J. M. E., Clark J. A., Hernandez-Castro J. C. *Non-linear Cryptanalysis Revisited: Heuristic Search for Approximations to S-Boxes*, IMA Int. conf. (Cirencester, UK. Dec. 18–20, 2007) P. 99–117 (LNCS 4887).

[19] Tokareva N. N. *Bent functions with stronger nonlinear properties: k-bent functions*, volume 14 of *Discrete Analysis and Operation Research*, 2007. N 4. P. 76–102 (in Russian). Engl. trans. will be available soon in *J. Applied and Industrial Mathematics* and at `www.math.nsc.ru/~tokareva`.

[20] Tokareva N. N. *Description of k-bent functions in four variables*, 2008. submitted to *Discrete Analysis and Operation Research* (in Russian).

[21] Tokareva N. N. *On quadratic approximations in block ciphers*, 2008. submitted to *Problems of Information Transmission.*