Annales Univ. Sci. Budapest., Sect. Comp. 23 (2004) 179-192

CONSTRUCTION OF HIGHLY NONLINEAR FUNCTIONS

I. Licskó (Budapest, Hungary)

1. Abstract

The highly nonlinear odd-dimensional Boolean functions have many applications in cryptographic practice, that is why the research of this class of functions is important. This study focusses on some types of functions in the class of highly nonlinear odd-dimensional Boolean functions which have special characteristics. Upper bound is given for the number of non-zero linear structures of such functions and regarding them as mappings some functional relations are proved. From the results one can gain an algorithm to construct highly nonlinear odd-dimensional Boolean functions with special characteristics by the use of functions having the same characteristics.

2. Introduction

The functions used for the purposes of cryptographic applications basically determine the strength of the cipher. The information to be coded is generally stored in binary format, that is why one applies the Boolean functions. There are two basic models of encryption: the stream cipher and the block cipher. The standard model of the stream cipher combines the outputs of several independent linear feedback shift registers using a nonlinear Boolean function. The model of the different block ciphers uses a so-called round-function for combining the round-key with the text to be coded. To be cryptographically strong, the Boolean function used in the process must have high nonlinearity. Moreover, such functions should be balanced, correlation immune and have high algebraic degree.

As we have seen the use of nonlinear functions is advisable but their construction often leads to much trouble.

3. Background

The mapping $f: \{0,1\}^n \to \{0,1\}$ is called a *Boolean function*. Sometimes the $\overline{f}: \{0,1\}^n \to \{-1,1\}$ mapping is used instead of the Boolean function. The relation between f and $\overline{f}$ can be described as follows:

$$\overline{f}(x) = (-1)^{f(x)}$$

or

$$\overline{f}(x) = 1 - 2f(x).$$

In the following we use the notation $f(x)$ to denote the whole $\{-1,1\}$ sequence generated by $f(x)$, that means $f(x)$ can be regarded as a vector having 2^n elements.

The elements of the set $\{0,1\}^n$ can be regarded as vectors. In this case $\{0,1\}^n$ is a vector space called Boolean space. As the coordinates of these vectors are the numbers 0 and 1, the vector can be regarded as an integer written in binary form. We can refer to a vector by an integer and in this case the components of the vector show the binary representation of the integer. It is also possible to refer to a vector by an indexed name, where the index is the integer corresponding to the vector.

The weight of a function $f(x)$ is the number of 1-s in its truth table: $w(f) = \sum_{x \in \{0,1\}^n} f(x)$.

The function $f(x)$ is called balanced if the number of 1-s and the number of 0-s in its truth table are equal, that is $\sum_{x \in \{0,1\}^n} \overline{f}(x) = 0$, and the weight of a balanced function is $w(f) = 2^{n-1}$.

Ordinary *operations* are defined by components in GF(2). If $a, b \in \{0,1\}^n$, $a = (a_0, a_1, \ldots, a_{n-1})$ and $b = (b_0, b_1, \ldots, b_{n-1})$, then $a \oplus b = (a_0 \oplus b_0, a_1 \oplus \oplus b_1, \ldots, a_{n-1} \oplus b_{n-1})$ is the *sum* and $ab = \bigoplus_{i=0}^{n-1} a_i b_i$ is the *scalar product* of the two vectors.

When the Zhegalkin polynomial of a Boolean function is

$$f(x_0, x_1, \ldots, x_{n-1}) = a_0 x_0 \oplus a_1 x_1 \oplus \ldots \oplus a_{n-1} x_{n-1} \oplus c,$$

where $c,\ a_i \in \{0,1\}$ for $0 \le i \le n-1$, the function is called *affine* and *linear* in the special case if $c = 0$. A linear function can be regarded as the scalar product of the constant $a \in \{0,1\}^n$ and the variable x. Function f satisfies the

propagation criterion regarding an element $a \in \{0,1\}^n,\ a \neq 0$ if $f(x) \oplus f(x \oplus a)$ is balanced. The distance between the functions f and g is

$$d(f,g) = w(f \oplus g) = \sum_{x \in \{0,1\}^n} (f(x) \oplus g(x)),$$

while their correlation is specified as

$$c(f,g) = \sum_{x \in \{0,1\}^n} \overline{f}(x)\overline{g}(x) = \sum_{x \in \{0,1\}^n} (-1)^{f(x) \oplus g(x)} = 2^n - 2d(f,g).$$

The *autocorrelational* function $r_f(a)$ of f is defined as

$$r_f(a) = \sum_{x \in \{0,1\}^n} \overline{f}(x)\overline{f}(x+a).$$

This expression is in principle the scalar product of the 2^n-dimensional $(-1,1)$ vectors $\xi(0), \xi(a)$, where $\xi(0)$ is given by $\overline{f}(x)$ and $\xi(a)$ is given by $\overline{f}(x+a)$.

A matrix $H = (h_{ij})$, where $h_{ij} \in \{-1,1\}$ for $i,j = 0,1,\ldots,m-1$, is called Hadamard matrix if $HH^T = mI$. H^T means the transpose of H, and I is the identity matrix of order m. The 2^n-order Hadamard matrix is denoted by H_n and it can be generated by the following recursive process:

$$H_0 = 1, \quad H_n = \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}, \quad n = 1,2,\ldots$$

The rows of H_n are denoted by $l_i,\ i = 0,1,\ldots,2^n-1$. l_i can be regarded as a $\{-1,1\}$ sequence generated by the linear function ix. The $\{-1,1\}$ sequences of all linear functions in $\{0,1\}^n$ appear in the rows of H_n.

The *Walsh transform* of the function f at $a = (a_0, a_1, \ldots, a_{n-1}) \in \{0,1\}^n$ is

$$F(a) = \sum_{x \in \{0,1\}^n} f(x)(-1)^{ax},$$

while the *Walsh transform* of $\overline{f}$ is specified as

$$\overline{F}(a) = \sum_{x \in \{0,1\}^n} (-1)^{f(x) \oplus xa}.$$

The value $\overline{F}(a)$ is the correlation of function f with the linear function ax and it can be written as the scalar product $(\overline{f}(x) l_a)$.

A vector $a \in \{0,1\}^n$ is called the *linear structure* of the function f if $f(x) \oplus f(x \oplus a)$ is a constant function, that is $\sum_{x \in \{0,1\}^n} (-1)^{f(x) \oplus f(x \oplus a)} = \pm 2^n$.

A Boolean function f is called *bent function* if for all $a \in \{0,1\}^n$ the correlation of f with the linear function ax has a constant absolute value

$$|\overline{F}(a)| = \left| \sum_{x \in \{0,1\}^n} (-1)^{f(x) \oplus xa} \right| = 2^{\frac{n}{2}}.$$

The Wiener-Khintchine theorem represents the relation between the Hadamard matrix H_n, the autocorrelational function of f and its correlation with the linear functions

$$(\overline{f}(x)\overline{f}(x \oplus a_0), \ldots, \overline{f}(x)\overline{f}(x \oplus a_{2^n-1}))H_n = ((\overline{f}(x)l_0)^2, \ldots, (\overline{f}(x)l_{2^n-1})^2),$$

where l_i is the i-th row of H_n (see [8]).

The *non-linearity* of a function f is the distance between f and the set of affine functions denoted by N_f,

$$N_f = 2^{n-1} - \frac{1}{2} \max_{i=0,1,\ldots,2^n-1} (|\overline{f}(x)l_i|).$$

The minimal value of the non-linearity is 0, which is the non-linearity of the affine functions. The maximum of it is generated if $|\overline{f}(x)l_i|$ is constant and its value is $2^{\frac{n}{2}}$. The maximal value of the non-linearity can be realized by bent functions, whose non-linearity is $N_f = 2^{n-1} - 2^{\frac{n}{2}-1}$. Such functions are interpreted in a space of even dimensions only. Despite their favourable non-linearity characteristics they are not preferred for cryptographic applications, because these functions are never balanced, correlation-immune, etc.

It is useful therefore to find other functions, which are in all probable ways maximally nonlinear, balanced or at least can easily be arranged to have these properties.

4. Characteristics of highly nonlinear functions

The behaviour of the correlation of a given Boolean function with the linear functions determines its nonlinearity.

Lemma 1. *Let $f:\ \{0,1\}^n \to \{0,1\}$ be such a function that its correlation with the linear functions $\overline{F}(i) = \overline{f}(x)l_i = \sum\limits_{x\in\{0,1\}^n} (-1)^{f(x)\oplus ix}$ takes only the values 0 and $\pm C$ and the number of the linear functions having non-zero correlation with it is 2^{2m}. In this case $C = \pm 2^{\frac{2n-2m}{2}} = \pm 2^{n-m}$.*

Proof. According to the Parseval equality

$$\sum_{a\in\{0,1\}^n} \overline{F}(a)^2 = 2^n \sum_{x\in\{0,1\}^n} \overline{f}(x)^2 \Rightarrow \sum_{a,F(a)\neq 0} C^2 + \sum_{a,F(a)=0} 0 = 2^{2n} \Rightarrow$$

$$\Rightarrow 2^{2m}C^2 \Rightarrow C = \pm 2^{\frac{2n-2m}{2}} = \pm 2^{n-m}$$

and this is the proof of the proposition.

An upper bound can be set for the number of nonzero linear structures of functions having the characteristics shown in Lemma 1.

Theorem 1. *If $f:\ \{0,1\}^n \to \{0,1\}$ is a function, whose correlation with the linear functions takes only the values 0 and $\pm C$ and the number of linear functions having nonzero correlation with f is 2^{2m}, then the number of nonzero linear structures of f is maximum $L = 2^{n-2m} - 1$.*

Proof. Let us use the Wiener-Khintchine form

$$(\overline{f}(x)\overline{f}(x\oplus a_0),\ldots,\overline{f}(x)\overline{f}(x\oplus a_{2^n-1}))H_n = ((\overline{f}(x)l_0)^2,\ldots,(\overline{f}(x)l_{2^n-1})^2).$$

The result of the multiplication on the left side of the equation

$$(\overline{f}(x)\overline{f}(x\oplus a_0),\ldots,\overline{f}(x)\overline{f}(x\oplus a_{2^n-1}))H_n =$$

$$= \left(\sum_{a=0}^{2^n-1} \overline{f}(x)\overline{f}(x\oplus a)(-1)^{aa_0},\ldots,\sum_{a=0}^{2^n-1} \overline{f}(x)\overline{f}(x\oplus a)(-1)^{aa_{2^n-1}}\right).$$

If we square both sides of the Wiener-Khintchine equation, we get

$$\sum_{y=0}^{2^n-1}\left(\left(\sum_{a=0}^{2^n-1} \overline{f}(x)\overline{f}(x\oplus a)(-1)^{ay}\right)\left(\sum_{b=0}^{2^n-1} \overline{f}(x)\overline{f}(x\oplus b)(-1)^{by}\right)\right) =$$

$$= \sum_{i=0}^{2^n-1} (\overline{f}(x)l_i)^4.$$

With further reductions on the left side of the equation we receive

$$2^n \sum_{a=0}^{2^n-1} \left(\overline{f}(x)\overline{f}(x \oplus a)\right)^2 = \sum_{i=0}^{2^n-1} (\overline{f}(x) l_i)^4.$$

The left-side sum can be divided into two parts, taking $a = 0$ the value of the autocorrelation is $\pm 2^n$ and its square is 2^{2n}. On the right side we may check the following considerations: f is a function whose correlation with linear functions is $\pm 2^{n-m}$ for 2^{2m} linear functions and is 0 for all other cases. In this way the above equation will have the form

$$2^n \left(2^{2n} + \sum_{a=1}^{2^n-1} \left(\overline{f}(x)\overline{f}(x \oplus a)\right) \right) = 2^{2m} 2^{4n-4m} = 2^{4n-2m}.$$

A further transformation thereof results in

$$\sum_{a=1}^{2^n-1} \left(\overline{f}(x)\overline{f}(x \oplus a)\right)^2 = \frac{2^n - 2^{2m}}{2^{2m}} 2^{2n} = \left(2^{n-2m} - 1\right) 2^{2n},$$

by which the proposition for the maximal number of the nonzero linear structures is proved.

On the basis of the above result we can state the following with respect to the functions interesting for us:

1. in the case of bent functions $2m = n$ and $c(f, ax) = \pm 2^{\frac{n}{2}}$ for each a, and the maximum number of nonzero linear structures is $L = 2^{n-2m} - 1 = = 2^0 - 1 = 0$;
2. for highly nonlinear functions of odd dimension $2m = n - 1$ and $c(f, ax) = \begin{cases} \pm 2^{\frac{n+1}{2}} \\ 0 \end{cases}$. The maximal number of nonzero linear structures is $L = 2^{n-2m} - 1 = 2^1 - 1 = 1$;
3. for linear functions $2m = 0$ and the maximal number of nonzero linear structures is $L = 2^{n-0} - 1 = 2^n - 1$.

The property of functions that their correlations with the linear functions take only the value 0 and $\pm C$ is not influenced by a linear transformation of the input variables.

Theorem 2. *Let $f : \{0,1\}^n \to \{0,1\}$ be a function whose correlation with the linear functions takes only the values 0 and $\pm C$, and the number of the linear functions having nonzero correlation with f is 2^{2m}. Let A be a nonsingular $n \times n$ matrix over GF(2). Then for any $x, b \in \{0,1\}^n$ the*

correlation of the function $g(x) = f(Ax + b)$ *with the linear functions takes only the values 0 and* $\pm C$ *and the number of linear functions having nonzero correlation with* g *is* 2^{2m}.

Proof. Let $y = Ax \oplus b$ and l_a be the a-th row of the matrix H_n, which in other words is the $\{-1, 1\}$ sequence generated by the function ax

$$\overline{g}(x)l_a = \sum_{x\in\{0,1\}^n} (-1)^{g(x)\oplus ax} = \sum_{x\in\{0,1\}^n} (-1)^{f(Ax\oplus b)\oplus ax}.$$

Since A is nonsingular

$$x = A^{-1}(y \oplus b) \quad \text{is true.}$$

The function $aA^{-1}(y \oplus b)$ is affine, so there is such a $c \in \{0, 1\}^n$, that

$$aA^{-1}(y \oplus b) = cy \oplus d$$

is true. So

$$\sum_{x\in\{0,1\}^n} (-1)^{f(Ax\oplus b)\oplus ax} = \sum_{y\in\{0,1\}^n} (-1)^{f(y)\oplus aA^{-1}(y\oplus b)} =$$

$$= \sum_{y\in\{0,1\}^n} (-1)^{f(y)\oplus cy\oplus d} = (-1)^d c(f, ax).$$

Considering the functions fulfilling the conditions of Lemma 1 one can find that the maximal value of the nonlinearity is reached in the case when $2m = n-1$. In order to construct such functions let us examine their characteristics.

Theorem 3. *Let* n *be odd,* g_1 *and* $g_2 : \{0, 1\}^n \to \{0, 1\}$ *Boolean functions whose correlations with the linear functions take only 0 and* $\pm C$ *and the number of the linear functions having nonzero correlation with* g_1 *and* g_2 *is* 2^{2m}. *Then the set of linear functions having nonzero correlation with* g_1 *is disjoint from the set of linear functions having nonzero correlation with* g_2 *if and only if*

$$(\overline{g}_1(x)l_i)^2 + (\overline{g}_2(x)l_i)^2 = 2^{2n-2m} \qquad \text{for each } i = 0, 1, \ldots, 2^n - 1.$$

Remark. The two sets of functions are disjoint when from $(\overline{g}_1(x)l_i)^2 = 0$ follows that $(\overline{g}_2(x)l_i)^2 \neq 0$, and conversely.

Proof. The proof consists of two steps.

1. Necessity. When both functions fulfill the condition that their correlations with the linear functions take only the values 0 and $\pm C$ and there are 2^{2m} linear functions having nonzero correlation with them, then both functions satisfy

$$(\overline{g}_1(x)l_i)^2 = \begin{cases} 2^{2n-2m} \\ 0 \end{cases} \qquad \text{or} \qquad (\overline{g}_2(x)l_i)^2 = \begin{cases} 2^{2n-2m} \\ 0 \end{cases}.$$

Since it is true that $(\overline{g}_2(x)l_i)^2 = 2^{2n-2m}$ in every case when $(\overline{g}_1(x)l_i)^2 = 0$, so

$$(\overline{g}_1(x)l_i)^2 + (\overline{g}_2(x)l_i)^2 = 2^{2n-2m} \qquad \text{for each} \quad i = 0, 1, \ldots, 2^n - 1.$$

2. Sufficiency. If

$$(\overline{g}_1(x)l_i)^2 + (\overline{g}_2(x)l_i)^2 = 2^{2n-2m} = 2^{2(n-m)}$$

is met for each $i = 0, 1, \ldots, 2^n - 1$, we apply Lemma 9 of [8]. The value $2(n-m)$ is even in every case, so the above equality is met if $(\overline{g}_1(x)l_i)^2 = 0$ and $(\overline{g}_2(x)l_i)^2 = 2^{2n-2m}$, or vice versa. This, on the other hand, means that the correlations of the two functions with the same linear function can never take the same absolute value.

Theorem 4. *Let n be odd, g_1 and $g_2 : \{0,1\}^n \to \{0,1\}$ Boolean functions whose correlations with the linear functions take only the values 0 and $\pm C$ and the number of the linear functions having nonzero correlation with g_1 and g_2 is 2^{2m}. Then the set of the linear functions having nonzero correlation with g_1 is disjoint from the set of linear functions having nonzero correlation with g_2 if and only if*

$$\overline{g}_1(x)\overline{g}_1(x \oplus b_j) + \overline{g}_2(x)\overline{g}_2(x \oplus b_j) = 0$$

for each $0 \neq b_j \in \{0,1\}^n$.

Proof. Let us use the Wiener-Khintcine relation for both functions

$$\big((\overline{g}_1(x)\overline{g}_1(x \oplus b_0)), \ldots, (\overline{g}_1(x)\overline{g}_1(x \oplus b_{2^n-1}))\big)H_n =$$

$$= \big((\overline{g}_1(x)l_0)^2, \ldots, (\overline{g}_1(x)l_{2^n-1})^2\big),$$

$$\big((\overline{g}_2(x)\overline{g}_2(x \oplus b_0)), \ldots, (\overline{g}_2(x)\overline{g}_2(x \oplus b_{2^n-1}))\big)H_n =$$

$$= \big((\overline{g}_2(x)l_0)^2, \ldots, (\overline{g}_2(x)l_{2^n-1})^2\big).$$

If we add up the two equations, multiply the sum H_n and then divide by 2^n, we receive

$$\big((\overline{g}_1(x)\overline{g}_1(x \oplus b_0) + \overline{g}_2(x)\overline{g}_2(x \oplus b_0)), \ldots$$

$$\ldots, (\overline{g}_1(x)\overline{g}_1(x \oplus b_{2^n-1}) + \overline{g}_2(x)\overline{g}_2(x \oplus b_{2^n-1})) =$$

$$= \Big(\big((\overline{g}_1(x)l_0^2 + (\overline{g}_2(x)l_0)^2)\big), \ldots, \big((\overline{g}_1(x)l_{2^n-1})^2 + (\overline{g}_2(x)l_{2^n-1})^2\big) \Big) H_n 2^{-n}.$$

1. Necessity

If it is true for both functions that $\overline{g}_2(x)l_i)^2 = 0$ follows from $(\overline{g}_1(x)l_i)^2 \neq 0$, and vice versa, then each element of the vector on the right hand side of the above equation, according to Theorem 3, has the value 2^{2n-2m} and so

$$\big(\ (\overline{g}_1(x)\overline{g}_1(x \oplus b_0) + \overline{g}_2(x)\overline{g}_2(x \oplus b_0)), \ldots$$

$$\ldots, (\overline{g}_1(x)\overline{g}_1(x \oplus b_{2^n-1}) + \overline{g}_2(x)\overline{g}_2(x \oplus b_{2^n-1})) \ \big) =$$

$$= 2^{2n-2m} \cdot (1, 1, \ldots, 1) \cdot H_n \cdot 2^{-n}.$$

In this case, however,

$$(\overline{g}_1(x)\overline{g}_1(x \oplus b_j) + \overline{g}_2(x)\overline{g}_2(x \oplus b_j)) = 2^n \cdot 2^{2n-2m} 2^{-n} = 2^{2n-2m} \qquad \text{if} \quad j = 0,$$

$$(\overline{g}_1(x)\overline{g}_1(x \oplus b_j) + \overline{g}_2(x)\overline{g}_2(x \oplus b_j)) = 0 \qquad \text{if} \quad j \neq 0.$$

2. Sufficiency

Let us assume that $(\overline{g}_1(x)\overline{g}_1(x \oplus b_j) + \overline{g}_2(x)\overline{g}_2(x \oplus b_j)) = 0$ is true for each $j \neq 0$. Having the Wiener-Khintchine theorem for both functions let us sum them and multiply the result by H_n:

$$2^n \ \big(\ (\overline{g}_1(x)\overline{g}_1(x \oplus b_0) + \overline{g}_2(x)\overline{g}_2(x \oplus b_0)), \ldots$$

$$\ldots, (\overline{g}_1(x)\overline{g}_1(x \oplus b_{2^n-1}) + \overline{g}_2(x)\overline{g}_2(x \oplus b_{2^n-1})) \ \big) \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} =$$

$$= \Big(\big(\big(\ \overline{g}_1(x)l_0 \ \big)^2 + \big(\ \overline{g}_2(x)l_0 \ \big)^2 \big), \ldots$$

$$\ldots, \big(\big(\ \overline{g}_1(x)l_{2^n-1} \ \big)^2 + \big(\ \overline{g}_2(x)l_{2^n-1} \ \big)^2 \big) \Big) \ H_n.$$

After multiplication the j-th element of the resulted vector is

$$2^n \ \big(\ \overline{g}_1(x)\overline{g}_1(x \oplus b_j) + \overline{g}_2(x)\overline{g}_2(x \oplus b_j) \ \big) =$$

$$= \sum_{i=0}^{2^n-1} \left(\left(\overline{g}_1(x) l_i \right)^2 + \left(\overline{g}_2(x) l_i \right)^2 \right) (-1)^{a_i b_j}.$$

As for each $j \neq 0$ it is true that $(\overline{g}_1(x)\overline{g}_1(x \oplus b_j) + \overline{g}_2(x)\overline{g}_2(x \oplus b_j) = 0$ we get

a) if $j \neq 0$ then the components of the vector are 0;
b) if $j = 0$ $\overline{g}_1(x)\overline{g}_1(x + b_0) = 2^n$ and $\overline{g}_2(x)\overline{g}_2(x + b_0) = 2^n$. As we have

$$2^n \left(\overline{g}_1(x)\overline{g}_1(x \oplus b_j) + \overline{g}_2(x)\overline{g}_2(x \oplus b_j) \right) = \sum_{i=0}^{2^n-1} \left(\left(\overline{g}_1(x) l_i \right)^2 + \left(\overline{g}_2(x) l_i \right)^2 \right),$$

so we get

$$\sum_{i=0}^{2^n-1} \left(\left(\overline{g}_1(x) l_i \right)^2 + \left(\overline{g}_2(x) l_i \right)^2 \right) = 2^n 2^{n+1}.$$

Using the fact that $2m = n - 1$ and $2^{n+1} = 2^{2n-2m}$ the result is

$$\sum_{i=0}^{2^n-1} \left(\left(\overline{g}_1(x) l_i \right)^2 + \left(\overline{g}_2(x) l_i \right)^2 \right) = 2^n 2^{2n-2m}$$

and on the basis of Theorem 2 this proves the proposition.

As the highly nonlinear functions can be applied in many areas of cryptography, finding the conditions for the construction of such functions is important.

Lemma 2. *Let us assume*

1. *g_1 and $g_2 : \{0,1\}^n \to \{0,1\}$ are Boolean functions, whose correlations with linear functions take only the values 0 and $\pm C$ and n is odd;*
2. *the number of linear functions having nonzero correlation with g_i for $i = 1, 2$ is $P = 2^{2m}$, where $2m = n - 1$;*
3. *from $\left(\overline{g}_2(x) l_i \right)^2 = 0$ follows $\left(\overline{g}_1(x) l_i \right)^2 \neq 0$ and vice versa.*

Then the series

$$h_1(i) = \frac{(\overline{g}_1(x) l_i) + (\overline{g}_2(x) l_i)}{2^{n-m}}$$

and

$$h_2(i) = \frac{(\overline{g}_1(x) l_i) - (\overline{g}_2(x) l_i)}{2^{n-m}}$$

only take the values of 1 and -1 for each $i = 0, 1, \ldots, 2^n - 1$.

Proof. As both of g_1 and g_2 satisfy the conditions of Lemma 1 the correlations of both functions with the linear functions take the values of 0

and $\pm 2^{n-m}$. Considering a linear function, its correlation with g_1 differs from its correlation with g_2, so

$$h_1(i) = \frac{(\overline{g}_1(x)l_i) + (\overline{g}_2(x)l_i)}{2^{n-m}} = \begin{cases} +1 \\ -1 \end{cases}$$

and

$$h_2(i) = \frac{(\overline{g}_1(x)l_i) - (\overline{g}_2(x)l_i)}{2^{n-m}} = \begin{cases} +1 \\ -1 \end{cases}$$

for each $i = 0, \ldots, 2^n - 1$.

The series $h_1(i)$, $h_2(i)$, generated in Lemma 2, where $i = 0, 1, \ldots, 2^{n-1}$, can be regarded as $\{-1, 1\}$ series, generated by the functions $h_1 = \overline{f}_1$ and $h_2 = \overline{f}_2$, both of them belonging to a Boolean function of n variables, to f_1 and f_2. These functions have the following properties.

Theorem 5. *Let n be odd and let f_1 and f_2 be such Boolean functions that the series $h_1(i) = \overline{f}_1(i)$ and $h_2(i) = \overline{f}_2(i)$, where $i = 0, 1, \ldots, 2^{n-1}$, are equal to the sequences constructed in Lemma 2. Then the correlations of f_1 and f_2 with the linear functions take only the values 0 and $\pm C$ and the number of the linear functions having nonzero correlation with them is 2^{2m} and $(\overline{f}_2(x)l_i)^2 + +(\overline{f}_2(x)l_i)^2 = 2^{2n-2m}$.*

Proof. The construction of the series h_1 and h_2 is given in Lemma 2. We can calculate the autocorrelation value for the two functions f_1 and f_2 with an $a \in \{0, 1\}^n$ as follows:

$$\overline{f}_1(i) = \frac{(\overline{g}_1(x)l_i) + (\overline{g}_2(x)l_i)}{2^{n-m}}, \qquad \overline{f}_2(i) = \frac{(\overline{g}_1(x)l_i) - (\overline{g}_2(x)l_i)}{2^{n-m}}$$

and

$$\overline{f}_1(i \oplus a) = \frac{(\overline{g}_1(x)l_{i\oplus a}) + (\overline{g}_2(x)l_{i\oplus a})}{2^{n-m}}, \qquad \overline{f}_2(i \oplus a) = \frac{(\overline{g}_1(x)l_{i\oplus a}) - (\overline{g}_2(x)l_{i\oplus a})}{2^{n-m}},$$

$$\overline{f}_1(i)\overline{f}_1(i \oplus a) = \sum_{i=0}^{2^n-1} \left(\frac{(\overline{g}_1(x)l_i) + (\overline{g}_2(x)l_i)}{2^{n-m}} \cdot \frac{(\overline{g}_1(x)l_{i\oplus a}) + (\overline{g}_2(x)l_{i\oplus a})}{2^{n-m}} \right)$$

and

$$\overline{f}_2(i)\overline{f}_2(i \oplus a) = \sum_{i=0}^{2^n-1} \left(\frac{(\overline{g}_1(x)l_i) - (\overline{g}_2(x)l_i)}{2^{n-m}} \cdot \frac{(\overline{g}_1(x)l_{i\oplus a}) - (\overline{g}_2(x)l_{i\oplus a})}{2^{n-m}} \right).$$

The sum of the autocorrelations is

$$\overline{f}_1(i)\overline{f}_1(i\oplus a)+\overline{f}_2(i)\overline{f}_2(i\oplus a)=$$

$$=\sum_{i=0}^{2^n-1}\frac{(\overline{g}_1(x)l_i)+(\overline{g}_2(x)l_i)}{2^{n-m}}\cdot\frac{(\overline{g}_1(x)l_{i\oplus a})+(\overline{g}_2(x)l_{i\oplus a})}{2^{n-m}}+$$

$$+\sum_{i=0}^{2^n-1}\frac{(\overline{g}_1(x)l_i)-(\overline{g}_2(x)l_i)}{2^{n-m}}\cdot\frac{(\overline{g}_1(x)l_{i\oplus a})-(\overline{g}_2(x)l_{i\oplus a})}{2^{n-m}}.$$

Upon having done the operations we get

$$\overline{f}_1(i)\overline{f}_1(i\oplus a)+\overline{f}_2(i)\overline{f}_2(i\oplus a)=$$

$$=\frac{2}{2^{2n-2m}}\sum_{i=0}^{2^n-1}\left((\overline{g}_1(x)l_i)(\overline{g}_1(x)l_{i\oplus a})+(\overline{g}_2(x)l_i)(\overline{g}_2(x)l_{i\oplus a})\right).$$

The Wiener-Khintchine theorem can be used for both functions

$$\big(\overline{f}_1(i)\overline{f}_1(i\oplus a_0),\ldots,\overline{f}_1(i)\overline{f}_1(i\oplus a_{2^n-1})\big)H_n=\big((\overline{f}_1(i)l_0)^2,\ldots,(\overline{f}_1(i)l_{2^n-1})^2\big),$$

$$\big(\overline{f}_2(i)\overline{f}_2(i\oplus a_0),\ldots,\overline{f}_2(i)\overline{f}_2(i\oplus a_{2^n-1})\big)H_n=\big((\overline{f}_2(i)l_0)^2,\ldots,(\overline{f}_2(i)l_{2^n-1})^2\big).$$

If we add the two equations and then multiply the sum by H_n, we get

$$\left(\sum_{j=0}^{2^n-1}\left(\big(\overline{f}_1(i)\overline{f}_1(i+a_j)+\overline{f}_2(i)\overline{f}_2(i+a_j)\big)(-1)^{a_ja_0}\right),\ldots\right.$$

$$\left.\ldots,\sum_{j=0}^{2^n-1}\left(\big(\overline{f}_1(i)\overline{f}_1(i+a_j)+\overline{f}_2(i)\overline{f}_2(i+a_j)\big)(-1)^{a_ja_{2^n-1}}\right)\right)=$$

$$=\left(\Big((\overline{f}_1(i)l_0)^2+(\overline{f}_2(i)l_0)^2\Big),\ldots,\Big((\overline{f}_1(i)l_{2^n-1})^2+(\overline{f}_2(i)l_{2^n-1})^2\Big)\right).$$

On both sides of the equation there is a vector, the equation means that these two vectors are equal in terms of the elements contained. Let us consider the k-th element

$$\sum_{j=0}^{2^n-1}\left(\Big(\overline{f}_1(i)\overline{f}_1(i+a_j)+\overline{f}_2(i)\overline{f}_2(i+a_j)\Big)(-1)^{a_ja_i}\right)=\big(\overline{f}_1(i)l_k\big)^2+\big(\overline{f}_2(i)l_k\big)^2.$$

Substituting the value of the autocorrelation and doing the possible transformations we receive

$$(\overline{f}_1(i)l_k)^2 + (\overline{f}_2(i)l_k)^2 = \frac{2^{2n+2}}{2^{2n-2m}}$$

for any k. As $2m = n - 1$ is true, the above equation can be written in the form

$$(\overline{f}_1(i)l_k)^2 + (\overline{f}_2(i)l_k)^2 = 2^{2n-2m}$$

and on the basis of Theorem 3 our proposition is proved.

5. Summary

The results described in this article offer a methodology for constructing highly nonlinear functions. Theorem 1 and Theorem 2 show the characteristics of the highly nonlinear functions, Theorem 3 and Theorem 4 give sufficient and necessary condition for the existence of special pairs of highly nonlinear functions. Theorem 5 gives the basis of an algorithm to generate highly nonlinear Boolean functions of odd dimension.

References

[1] **Licskó I.,** On highly nonlinear functions, *Annales Univ. Sci. Budapest. Sect. Comp.,* **21** (2002), 165-175.

[2] **Licskó I.,** Characteristics of highly nonlinear functions, *Hajducrypt'02, Debrecen,* 2002.

[3] **Pásztor-Varga K.,** Boole függvények Boole algebrájának strukturális tulajdonságait felhasználó Boole függvény optimizációs módszer, *Alk. Mat. Lapok,* **13** (1987-88), 69-76.

[4] **Rothaus O.S.,** On "bent" functions, *J. Combinatorial Theory, Ser. A,* **20** (1976), 300-305.

[5] **Siegenthaler T.,** Correlation-immunity of nonlinear combining functions for cryptographic applications, *IEEE Transactions on Information Theory,* **30** (5) (1984), 776-779.

[6] **Seberry J., Zhang X.M. and Zheng Y.,** Nonlinearity and propogation characteristics of balanced Boolean functions, *Information and Computation,* **119** (1) (1995), 1-13.

[7] **Vajda I., Buttyán L. and Szekeres B.,** *Az algoritmikus adatvédelem módszerei* (Methods of algorithmic data protection), Technical University of Budapest (nonpublished manuscript)

[8] **Zhang X.M. and Zheng Y.,** New lower bounds on nonlinearity and class of highly nonlinear functions, *Information Security and Privacy. Second Australian Conf., Sidney, Australia, July 1997,* Lecture Notes in Computer Science **1270**, Springer, 1997, 147-158.

(Received March 18, 2002)

I. Licskó
Budapest Gazdasági Főiskola
Kereskedelmi, Vendéglátóipari és Idegenforgalmi Kar
Informatikai Intézet
V. Alkotmány u. 9-11.
H-1054 Budapest, Hungary
licsko@mailbox.hu