Tight bounds between algebraic immunity and nonlinearities of high orders

Mikhail Lobanov

Mech. & Math. Department
Moscow State University
119992 Moscow, Russia
emails: misha_msu@mail.ru

Abstract

Among cryptographically significant characteristics of Boolean functions used in symmetric ciphers the algebraic immunity and the nonlinearities of high orders play the important role. Some bounds on the nonlinearities of high orders of Boolean functions via its algebraic immunity were obtained in recent papers. In this paper we improve these results and obtain new tight bounds. We prove new universal tight lower bound that reduces the problem of an estimation of high order nonlinearities to the problem of the finding of dimensions of some linear spaces of Boolean functions. As simple consequences we obtain all previously known bounds in this field. For polynomials with disjoint terms we reduce the finding of dimensions of linear spaces of Boolean functions mentioned above to a simple combinatorial analysis. Finally, we prove the tight lower bound on the nonlinearity of the second order via its algebraic immunity.

Keywords: stream cipher, nonlinear filter, algebraic attack, Boolean function, algebraic immunity, algebraic degree, nonlinearity, higher order nonlinearity, annihilator.

1 Introduction

Boolean functions have wide applications in cryptography, in particular, in symmetric cryptography. Stream ciphers use Boolean functions as nonlinear filters or nonlinear combiners, block ciphers use Boolean functions in S-boxes. Boolean functions have many cryptographically important characteristics. Good characteristics provide the resistance at least against known attacks. Among these characteristics the algebraic immunity and the nonlinearity of high orders play the important role. Recently, the significance of the algebraic immunity and the nonlinearities of high orders and their mutual relations were described in papers

[2, 4, 5, 6, 8]. Boolean functions with good algebraic immunity and nonlinearities of high orders allow to resist against many types of known cryptographical attacks including algebraic, correlation and linear attacks.

Some bounds between algebraic immunity and nonlinearities were obtained in [2, 4, 5, 6, 8]. It appeared that the good algebraic immunity provides also some guaranteed nonlinearity of rth order. These results are important since recently it was proposed some algorithms for the calculation of algebraic immunity whereas the effective calculation or estimation of high order nonlinearities is not an easy problem.

In [5] the tight lower bound on the nonlinearity (of the first order) of a Boolean function via the value of its algebraic immunity was obtained. The lower bounds on the nonlinearity of the rth order via its algebraic immunity were obtained in [2, 4, 6, 8], the strongest among them is the bound (4).

In this paper we propose the new approach that reduces the problem of an estimation of high order nonlinearities to the problem of the finding of dimensions of some linear spaces of Boolean functions. This result is given in our Theorem 1. This Theorem gives the new universal lower bound on the rth order nonlinearity of a Boolean function via its algebraic immunity. This bound is tight, i. e. for any possible set of parameters there exists a function that achieves this bound. We obtain all previously known bounds in this field as simple consequences of our Theorem 1.

Next, for the functions of the special form — for the polynomials with disjoint terms — we prove the Theorem 2 that allows to reduce the finding of dimensions of some linear spaces of Boolean functions mentioned above to a simple combinatorial analysis. Using Theorem 2 we obtain in Theorem 3 the tight lower bound on the nonlinearity of the second order via its algebraic immunity. The bound is tight, i. e. for all possible pairs of algebraic immunity and the number of variables there exists a function that achieves this bound.

The rest of the paper is organized as follows. In Section 2 we give the necessary definitions and some previously known results. In Section 3 we prove Theorem 1 with new universal tight lower bound that reduces the problem of an estimation of high order nonlinearities to the problem of the finding of dimensions of some linear spaces of Boolean functions. As simple consequences we obtain all previously known bounds in this field. In Section 4 for polynomials with disjoint terms we prove the Theorem 2 that allows to reduce the finding of dimensions of linear spaces of Boolean functions mentioned above to a simple combinatorial analysis. In Section 5 we prove in Theorem 3 the tight lower bound on the nonlinearity of the second order via its algebraic immunity.

2 Preliminaries

Let f be a Boolean function on $\mathbf{F}_2^n$. It is well known that f can be uniquely represented by a polynomial. *An algebraic degree* $\deg(f)$ of f is the length of the longest term in the polynomial of f. A Boolean function g is called *an annihilator* of f if $f \cdot g \equiv 0$. It is obvious that the set of all annihilators of

f forms the linear subspace in the space of all Boolean function on $\mathbf{F}_2^n$. *An algebraic immunity* $AI(f)$ of f is the minimum degree of a nonzero function g on $\mathbf{F}_2^n$ such that g is an annihilator of f or g is an annihilator of $f+1$. It is known [3, 7] that for any f on $\mathbf{F}_2^n$ the inequality $AI(f) \leq \lceil \frac{n}{2} \rceil$ holds.

The weight $wt(x)$ of a vector $x \in \mathbf{F}_2^n$ is the number of ones in x. *The distance* between two Boolean functions f_1 and f_2 is defined as $d(f_1, f_2) = \mid \{x \in \mathbf{F}_2^n \mid f_1(x) \neq f_2(x)\} \mid$. The nonlinearity of rth order $nl_r(f)$ of a Boolean function f over $\mathbf{F}_2^n$ is called the value $\min\limits_{l,\ deg(l) \leq r} d(f, l)$.

The nonlinearity $nl(f)$ of f is the distance between f and the set of affine functions, i. e. the nonlinearity of the first order.

In [4] it was proved the result equivalent to the next bound on the nonlinearity of the rth order:

$$nl_r(f) \geq \sum_{i=0}^{AI(f)-r-1} \binom{n}{i}. \tag{1}$$

Later in [6] it was proved the lower bound on the nonlinearity ($r = 1$) of a function via the value of its algebraic immunity:

$$nl(f) \geq 2 \sum_{i=0}^{AI(f)-2} \binom{n-1}{i}. \tag{2}$$

For all possible values of algebraic immunity in [5] it were constructed functions that achieve equality in this bound.

In [3] the bound (2) was generalized for the case of arbitrary r:

$$nl_r(f) \geq 2 \sum_{i=0}^{AI(f)-r-1} \binom{n-r}{i}. \tag{3}$$

Note that the bound (1) does not follow the bound (3) and visa versa.

In [8] and [6] it was proved the bound

$$nl_r(f) \geq \sum_{i=0}^{AI(f)-r-1} \binom{n}{i} + \sum_{i=AI(f)-2r}^{AI(f)-r-1} \binom{n-r}{i}. \tag{4}$$

that is better than both bounds (1) and (3).

3 The problem reduction to the estimation of linear subspaces dimensions

Definition 1 *Let h be a Boolean function on $\mathbf{F}_2^n$. Denote by $An_k(h)$ the linear subspace of all annihilators of degree at most k. Denote by $d_{k,h}$ the dimension of this subspace.*

Definition 2 *Let $C = \{\overline{x}_1, \ldots, \overline{x}_n\}$ be some set of vectors in $\mathbf{F}_2^n$. For any given k, $k \leq n$, and for any vector $x = (x_1, \ldots, x_n) \in \mathbf{F}_2^n$ we correspond to x the uniform linear equation with the left side generated by the substitution of components of the vector x into the expression*

$$a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j + \cdots + \sum_{1 \leq i_1 \leq \ldots \leq i_k \leq n} a_{i_1 \ldots i_k} x_{i_1} \ldots x_{i_k}.$$

The right side of the equation is 0. Then we call a k-rank of the set C the rank of the system of linear equations generated by such way from the vectors of the set C. Denote this rank by $r_k(C)$.

Next, for the function f we search the annihilators of degree at most k by the method of undefined coefficients:

$$g = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j + \cdots + \sum_{1 \leq i_1 \leq \ldots \leq i_k \leq n} a_{i_1 \ldots i_k} x_{i_1} \ldots x_{i_k}.$$

The function g is an annihilator of f if and only if $f(x) = 1$ follows $g(x) = 0$. Thus, we obtain the system of linear equations.

It is easy to see that $d_{k,f} = dim(An_k(f)) = \sum_{i=0}^{k} \binom{n}{i} - r_k(supp(f))$.

Proposition 1 *Let f and f_0 be Boolean functions on $\mathbf{F}_2^n$, $AI(f_0) \geq k$. Then $d(f, f_0) \geq dim(An_{k-1}(f)) + dim(An_{k-1}(f + 1))$.*

Proof. Since $AI(f_0) \geq k$, we have $r_{k-1}(supp(f_0)) = \sum_{i=0}^{k-1} \binom{n}{i}$.

At the same time $r_{k-1}(supp(f)) = \sum_{i=0}^{k-1} \binom{n}{i} - d_{k-1,f}$. Hence, there exist at least $d_{k-1,f}$ vectors where f_0 is equal to 1, and f is equal to 0.

Analogously, considering the pair of functions $f + 1$ and $f_0 + 1$ we obtain the lower bound on the number of vectors where f is 1 and f_0 is 0. □

Definition 3 *Let h be a Boolean function on $\mathbf{F}_2^n$. Denote by $B_k(h)$ the linear space of all such functions f on $\mathbf{F}_2^n$ that $\deg(f) \leq k$ and $\deg(fh) \leq k$.*

Proposition 2 *The sum of $dim(An_k(f))$ and $dim(An_k(f + 1))$ is equal to $dim(B_k(f))$.*

Proof. Consider the pair (g_1, g_2), where $g_1 \in An_k(f)$, $\quad g_2 \in An_k(f + 1)$. Then we have that $fg_1 + (f + 1)g_2 = 0$, it follows $f(g_1 + g_2) = g_2$. We obtain the correspondence between the pairs of functions, the first of which is from $An_k(f)$, the second is from $An_k(f+1)$, and the functions from $B_k(f)$. It is easy to check that the correspondence is one to one. □

Lemma 1 *Suppose $r_k(supp(f)) = wt(f)$ where $k < \lceil \frac{n}{2} \rceil$. Then $dim(An_k(f + 1)) = 0$.*

Proof. The condition of the Lemma implies that for any vector x such that $f(x) = 1$, there exists the function g of degree at most k such that the product fg is 1 only at the vector x. In the opposite case there exists the function that differs from f only at the vector x, with k-rank $r_k(supp(f)) = wt(f)$ and weight $wt(f) - 1$ that is impossible.

Suppose that there exists the function f', $deg(f') \leq k$ and $f \neq 0$ such that $(f + 1)f' = 0$. Choose the vector x such that $f'(x) = 1$. The relation $supp(f') \subseteq supp(f)$ follows that there exists the function g' of degree at most k such that the product $f'g'$ is 1 only at one vector x. At the same time the degree of the product of two Boolean functions does not exceed the sum of its degrees, it follows $deg(f'g') < n$ that contradicts to the fact that $f'g'$ is 1 only at one vector. □

Corollary 1 *Suppose* $dim(An_k(f)) = \sum_{i=0}^{k} \binom{n}{i} - wt(f)$ *where* $k < \lceil \frac{n}{2} \rceil$. *Then* $dim(An_{\lceil \frac{n}{2} \rceil - 1}(f + 1)) = 0$.

Corollary 2 *Suppose* $n = 2k + 1$ *and* $An_k(f) = 0$ *then* $AI(f) = k + 1$.

The Corollary 2 was proved in [1].

Proposition 3 *Suppose that* $deg(f) \leq \lceil \frac{n}{2} \rceil$, $k \leq \lceil \frac{n}{2} \rceil$. *Then there exists the function* g *such that* $AI(g) \geq k$ *and* $d(f, g) = dim(B_{k-1}(f))$.

Proof. In the set of vectors where f is 1 there exist $r_{k-1}(supp(f))$ vectors such that their $(k-1)$-rank is also equal to $r_{k-1}(supp(f))$, denote this set of vectors by C_1. Analogously, considering the function $f + 1$ we obtain the set C_0 with $r_{k-1}(supp(f + 1))$ vectors. Lemma 1 follows that we can supplement the set C_1 by $\sum_{i=0}^{k-1} \binom{n}{i} - r_{k-1}(supp(f))$ vectors which do not belong to C_0 and f is 0 at these vectors, such that the k-rank of a new set is exactly $\sum_{i=0}^{k-1} \binom{n}{i}$. Analogously, we can supplement the set C_0 by $\sum_{i=0}^{k-1} \binom{n}{i} - r_{k-1}(supp(f + 1))$ vectors which do not belong to C_1 and f is 1 at these vectors such that the k-rank of a new set is exactly $\sum_{i=0}^{k-1} \binom{n}{i}$.

It follows that it is possible to change the values of f at $\sum_{i=0}^{k-1} \binom{n}{i} - r_{k-1}(supp(f)) + \sum_{i=0}^{k-1} \binom{n}{i} - r_{k-1}(supp(f + 1)) = dim(An_{k-1}(f)) + dim(An_{k-1}(f + 1)) = dim(B_{k-1}(f))$ vectors and obtain the function g such that $dim(An_{k-1}(g)) = dim(An_{k-1}(g + 1)) = 0$, hence, $AI(g) \geq k$.□

Thus, in Propositions 1–3 we have proved that the problem of the finding of the most strong bound for the nonlinearity of the rth order via the value of its algebraic immunity k is completely reduced to the finding of the value $min_{deg(g) \leq r} dim(B_{k-1}(g))$. We formulate this statement as a theorem:

Theorem 1 *Suppose that* $f(x_1, \ldots, x_n)$ *has* $AI(f) = k \leq \lceil \frac{n}{2} \rceil$. *Then*

$$nl_r(f) \geq min_{deg(g) \leq r} dim(B_{k-1}(g)).$$

Moreover, there exists the function f_0, $AI(f_0) = k$, *such that*

$$nl_r(f_0) = min_{deg(g) \leq r} dim(B_{k-1}(g)).$$

A few bounds are deduced from Theorem 1.

Proposition 4 *Suppose that $deg(f) = r$. Then $dim(B_{k-1}(f)) \geq \sum_{i=0}^{k-r-1} \binom{n}{i}$.*

Proof. We take all functions of degree at most $(k - r - 1)$. □

Corollary 3 *Suppose that $AI(g) = k$. Then*

$$nl_r(g) \geq \sum_{i=0}^{k-r-1} \binom{n}{i}.$$

We obtain the bound (1) from [4].

Proposition 5 *Suppose that $deg(f) = r$. Then*

$$dim(B_{k-1}(f)) \geq 2 \sum_{i=0}^{k-r-1} \binom{n-r}{i}.$$

Proof. Without loss of generality we can assume that the polynomial of f contains the term $x_1x_2 \dots x_r$. Consider the functions of the form $fg_1 + (f+1)g_2$ where g_1 and g_2 are any functions of variables $x_{r+1}, \dots, x_n$, whose degree at most $(k - r - 1)$. It is easy to check that all such functions are different and belong to $B_{k-1}(f)$. □

Corollary 4 *Suppose that $AI(g) = k$. Then*

$$nl_r(g) \geq 2 \sum_{i=0}^{k-r-1} \binom{n-r}{i}.$$

We obtain the bound (3) from [2].

Proposition 6 *Suppose that $deg(f) = r$. Then $dim(B_{k-1}(f))$ is not less than*

$$\sum_{i=0}^{k-r-1} \binom{n}{i} + \sum_{i=k-2r}^{k-r-1} \binom{n-r}{i}.$$

Proof. Without loss of generality we can assume that the polynomial of f contains the term $x_1x_2 \dots x_r$. Consider the functions of the form $g_1 + fg_2$ where g_1 is an arbitrary function of degree at most $(k - r - 1)$, and g_2 is an arbitrary function of variables $x_{r+1}, \dots, x_n$, whose degree at most $(k - r - 1)$, that contains only terms of the length at least $k - 2r$.

It is easy to check that all such functions belong to $B_{k-1}(f)$. The checking of the fact that all such functions are different is reduced to the checking that $g_1 + fg_2 = 0$ follows $g_1 = 0$ and $g_2 = 0$. The equality $g_2 = 0$ follows from the fact that in the opposite case the function fg_2 contains the term of length at least $(k-r)$ which was also in the polynomial of f (since $\deg(f_1) \leq (k-r-1)$). The equality $g_1 = 0$ follows straightforwardly from $g_1 + fg_2 = 0$ and $g_2 = 0$.□

Corollary 5 *Suppose $AI(g) = k$. Then*

$$nl_r(g) \geq \sum_{i=0}^{k-r-1} \binom{n}{i} + \sum_{i=k-2r}^{k-r-1} \binom{n-r}{i}.$$

We deduce the bound (4) from [8, 6].

Theorem 1 can give more strong corollaries. Some of them we obtain in the following sections.

4 Exact value of $dim(B_k(f))$ for polynomials with disjoint terms

Definition 4 *Let $a_1 \geq a_2 \geq \ldots \geq a_q > 0$ be the set of integer numbers such that $\sum_{i=1}^{q} a_i \leq n$. Then it is possible to map to any vector $x = (x_1, \ldots, x_n) \in \mathbf{F}_2^n$ the set $s_{a_1,\ldots,a_q}(x)$ of integer numbers by the next way: $(s_1(x), \ldots, s_q(x)) = (\sum_{i=1}^{a_1} x_i, \sum_{i=a_1+1}^{a_1+a_2} x_i, \ldots, \sum_{i=a_1+\ldots+a_{q-1}+1}^{a_1+\ldots+a_q} x_i)$. Denote by $S_{a_1,\ldots,a_q}(k)$ the set of all vectors $x \in \mathbf{F}_2^n$ such that $s_t(x) = 0$ for some $t \leq q$, $0 < s_i(x) < a_i$ for all $i < t$ and also $k - a_t < wt(x) \leq k$.*

Proposition 7 *Suppose that any two terms in the polynomial of the function $f(x_1, \ldots, x_n)$ do not contain joint variables. Let q be the number of terms in the polynomial of f, and $a_1 \geq a_2 \geq \ldots \geq a_q$ are the lengths of these terms. Then $dim(B_k(f)) \leq \sum_{i=0}^{k} \binom{n}{i} - |S_{a_1,\ldots,a_q}(k)|$.*

Proof. Without loss of generality we can assume that the function has the form $f = x_1x_2 \cdots x_{a_1} + x_{a_1+1} \cdots x_{a_1+a_2} + \cdots + x_{a_1+\ldots+a_{q-1}+1} \cdots x_{a_1+\ldots+a_q}$.

It is possible to map any vector x from $\mathbf{F}_2^n$ to the term that contains all such variables from $x_1, \ldots, x_n$ that correspond to ones in the vector x.

Consider the linear subspace $C_{f,k}$ in the space of Boolean functions on $\mathbf{F}_2^n$ stretched on the monomials corresponded to the vectors from $S_{a_1,\ldots,a_q}(k)$. This space is also the subspace in the space of functions of degree at most k. Any nonzero function from $C_{f,k}$ does not belong to $B_k(f)$. Indeed, suppose that $g \in C_{f,k}$, then for any monomial from g and corresponding vector $x = (x_1, \ldots, x_n)$ it is possible to assign its own $t \leq q$ according Definition 4 such that $s_t(x) = 0$ and $0 < s_i(x) < a_i$ for $i < t$. Choose from the terms of the function g the terms with the maximal length, and among them take some term $x_{i_1} \ldots x_{i_h}$ with the minimal assigned t. Then the polynomial of the function gf contains the term $x_{i_1} \ldots x_{i_h} x_{a_1+\ldots+a_{t-1}+1} \cdots x_{a_1+\ldots+a_t}$ that cannot be cancelled. Thus, $deg(fg) > k$ and g does not belong to $B_k(f)$.

The dimension of $C_{f,k}$ is equal to $|S_{a_1,\ldots,a_q}(k)|$ that follows the conclusion of this Proposition. □

Now we prove the converse inequality.

Proposition 8 *Suppose that any two terms in the polynomial form of $f(x_1, \ldots, x_n)$ do not contain joint variables. Let q be the number of terms in the* *polynomial of f, and $a_1 \geq a_2 \geq \ldots \geq a_q$ are the lengths of these terms. Then $dim(B_k(f)) \geq \sum_{i=0}^{k} \binom{n}{i} - |S_{a_1,\ldots,a_q}(k)|$.*

Proof. Without loss of generality it is possible to assume that the function has the form $f = x_1 x_2 \cdots x_{a_1} + x_{a_1+1} \cdots x_{a_1+a_2} + \cdots + x_{a_1+\ldots+a_{q-1}+1} \cdots x_{a_1+\ldots+a_q}$.

Denote by $\overline{S_{a_1,\ldots,a_q}(k)}$ the set of vectors $x = (x_1, \ldots, x_n)$, $wt(x) \leq k$ and $x \notin S_{a_1,\ldots,a_q}(k)$.

Suppose that $x = (x_1, \ldots, x_n) \in \overline{S_{a_1,\ldots,a_q}(k)}$. Then we map the vector x to the function f_x by the next rules:

1. If $deg(x_{i_1} x_{i_2} \cdots x_{i_{wt(x)}} f) \leq k$ where $i_1, \ldots, i_{wt(x)}$ are the indexes of positions of ones in the vector $x = (x_1, \ldots, x_n)$ then $f_x = x_{i_1} x_{i_2} \cdots x_{i_{wt(x)}}$.

2. If x does not satisfy the first item and for any $t \leq q$ it holds $0 < s_t(x) \leq a_i$ then $f_x = (x_{i_1} \cdots x_{i_{s_1(k)}} + 1) \ldots (x_{i_{s_1(k)+\ldots+s_{q-1}(k)+1}} \cdots x_{i_{s_1(k)+\ldots+s_q(k)}} + 1)$, where $i_1, \ldots, i_{s_1(k)+\ldots+s_q(k)}$ are the indexes of positions of ones in the vector $x = (x_1, \ldots, x_n)$.

3. If x does not satisfy any of two previous items and $s_t(x) = 0$, $0 < s_i(x) < a_i$ for all $i < t$, then $f_x = (x_{i_1} \cdots x_{i_{s_1(k)}} + 1) \ldots (x_{i_{s_1(k)+\ldots+s_{q-1}(k)+1}} \cdots x_{i_{s_1(k)+\ldots+s_q(k)}} + 1)$, where $i_1, \ldots, i_{s_1(k)+\ldots+s_q(k)}$ are the indexes of positions of ones in the vector $x = (x_1, \ldots, x_n)$.

4. If x does not satisfy any of three previous items then $s_t(x) = 1$ for some $t \leq q$ and $s_i(k) = 0$ for $i = b_1, \ldots, b_u$, where $b_h > t$ ($0 < s_i(x) < a_i$ for $i < t$ and $0 < s_i(x)$ for $i \neq t, b_1, \ldots, b_u$) then
$$f_x = (x_{i_1} \ldots x_{i_{s_1(k)}} + 1)\,(x_{i_{s_1(k)+1}} \ldots x_{i_{s_1(k)+s_2(k)}} + 1)\ \ldots$$
$$\ldots (x_{i_{s_1(k)+\ldots+s_{t-2}(k)+1}} \ldots x_{i_{s_1(k)+\ldots+s_{t-1}(k)}} + 1)$$
$$(x_{i_{s_1(k)+\ldots+s_t(k)+1}} \ldots x_{i_{s_1(k)+\ldots+s_{t+1}(k)}} + 1) \ldots$$
$$\ldots (x_{i_{s_1(k)+\ldots+s_{q-1}(k)+1}} \ldots x_{i_{s_1(k)+\ldots+s_q(k)}} + 1)$$
$$(x_{a_1+\ldots+a_{t-1}+1} \ldots x_{a_1+\ldots+a_t} + x_{a_1+\ldots+a_{(b_1-1)}+1} \ldots x_{a_1+\ldots+a_{b_1}} + \ldots$$
$$\ldots + x_{a_1+\ldots+a_{b_{u-1}}+1} \ldots x_{a_1+\ldots+a_{b_u}}),$$
where $i_1, \ldots, i_{s_1(k)+\ldots+s_q(k)}$ are the indexes of positions of ones in the vector $x = (x_1, \ldots, x_n)$.

It is possible to check that the rule given above maps any vector $x = (x_1, \ldots, x_n) \in \overline{S_{a_1,\ldots,a_q}(k)}$ to the unique function f_x.

The polynomial of f_x for any vector x described above contains the term that contains all variables which correspond to ones in the vector x; all other terms have smaller length or lexicographically greater. It follows the linear independence of all f_x corresponded to vectors $x = (x_1, \ldots, x_n) \in \overline{S_{a_1,\ldots,a_q}(k)}$.

Indeed, suppose that we have the vectors $x^1, \ldots, x^h \in \overline{S_{a_1,\ldots,a_q}(k)}$, choose from them the vectors with the maximal weight , and among them the first vector in the lexicographical order. The term corresponded to this vector enters into the polynomial of $f_{x^1} + \ldots + f_{x^h}$, all other terms have smaller length or lexicographically greater. Therefore $f_{x^1} + \ldots + f_{x^h}$ is not identically zero.

Now show that for any $x = (x_1, \ldots, x_n) \in \overline{S_{a_1,\ldots,a_q}(k)}$ the corresponding function f_x belongs to $B_k(f)$. If the vector x satisfies the item 1, the desired fact follows from the definition f_x for such vectors. If x satisfies the item 2, then the product of f and f_x is identically zero. If x satisfies the item 3, the product f and f_x has the degree at most k, since in the opposite case $x \in S_{a_1,\ldots,a_q}(k)$. Suppose that x satisfies the item 4. Then we represent f as the sum of two functions $f = f_1 + f_2$ where f_1 contains the terms of f with ordinal numbers $t, b_1, \ldots, b_u$ and f_2 contains all remained terms. It is easy to check that the product of f_2 and f_x is identically zero, and the product of f_1 and f_x is equal to f_x since f_1 enters as the last factor in f_x. Taking into account that $deg(f_x) = wt(x)$, we deduce that for any $x = (x_1, \ldots, x_n) \in \overline{S_{a_1,\ldots,a_q}(k)}$ the corresponding f_x belongs to $B_k(f)$.

Thus, $dim(B_k(f)) \geq |\overline{S_{a_1,\ldots,a_q}(k)}| = \sum_{i=0}^{k} \binom{n}{i} - |S_{a_1,\ldots,a_q}(k)|$. □

We combine the propositions 7 and 8 into the next theorem.

Theorem 2 *Suppose that any two terms in the polynomial of the function $f(x_1, \ldots, x_n)$ do not contain joint variables. Let q be the number of terms in the polynomial of the function f, and $a_1 \geq a_2 \geq \ldots \geq a_q$ are the lenghts of these terms. Then $dim(B_k(f)) = \sum_{i=0}^{k} \binom{n}{i} - |S_{a_1,\ldots,a_q}(k)|$.*

Thus, for the quite wide class of functions we have reduced the problem of the calculation of $\dim(B_k(f))$ to a simple combinatorial analysis.

5 Tight bound between algebraic immunity and nonlinearity of the second order

Remark 1 *Below we assume that the binomial coefficient $\binom{n}{m}$ is equal to 0 if n or m are less than 0.*

Proposition 9 *Suppose $f(x_1, \ldots, x_n) = x_1x_2 + x_3x_4 + \cdots + x_{2q-1}x_{2q}$; Then $dim(B_k(f)) = \sum_{i=0}^{k} \binom{n}{i} - \sum_{i=0}^{q-1} 2^i \binom{n-2i-1}{k-i}$.*

Proof. The set $S_{a_1,\ldots,a_q}(k)$ contains $\binom{n-2}{k}$ vectors of the weight k and $\binom{n-2}{k-1}$ vectors of the weight $k-1$ equal to zero in first two components. Summing, we obtain $\binom{n-1}{k}$ vectors.

The set $S_{a_1,\ldots,a_q}(k)$ contains $2\binom{n-4}{k-1}$ vectors of the weight k and $2\binom{n-4}{k-2}$ vectors of the weight $k-1$ equal to zero in the second pair of components and equal to 1 in exactly one of the first two components Summing, we obtain $2\binom{n-3}{k-1}$ vectors.

The set $S_{a_1,\dots,a_q}(k)$ contains $2^{t-1}\binom{n-2t}{k-t+1}$ vectors of the weight k and $2^{t-1}\binom{n-2t}{k-t}$ vectors of the weight $k-1$ equal to zero in the tth pair of components and equal to 1 in exactly one of two components for all previous pairs of variables. Summing, we obtain $2^{t-1}\binom{n-2t+1}{k-t+1}$ vectors.

Thus, we exhaust all vectors from $S_{a_1,\dots,a_q}(k)$ and obtain that their number is equal to $\binom{n-1}{k}+2\binom{n-3}{k-1}+4\binom{n-5}{k-2}+\ldots+2^{q-1}\binom{n-2q+1}{k-q+1}=\sum_{i=0}^{q-1}2^i\binom{n-2i-1}{k-i}$. Using the Theorem 2 we obtain the conclusion of this Proposition. □

The next proposition is analogous.

Proposition 10 *Suppose $f(x_1,\dots,x_n)=x_1x_2+x_3x_4+\ldots+x_{2q-1}x_{2q}+x_{2q+1}$; then $dim(B_k(f))=\sum_{i=0}^{k}\binom{n}{i}-\sum_{i=0}^{q}2^i\binom{n-2i-1}{k-i}$.*

Theorem 3 *Suppose that the function $f(x_1,\dots,x_n)$ has the algebraic immunity $AI(f)=k\le\lceil\frac{n}{2}\rceil$. Then*

$$nl_2(f)\ge\sum_{i=0}^{k-1}\binom{n}{i}-\sum_{i=0}^{k-1}2^i\binom{n-2i-1}{k-1-i}.$$

Moreover, there exists the function f_0, $AI(f_0)=k$, such that

$$nl_2(f_0)=\sum_{i=0}^{k-1}\binom{n}{i}-\sum_{i=0}^{k-1}2^i\binom{n-2i-1}{k-1-i}.$$

Proof. It is well known (see, for example, [9]), that the function of degree at most 2 can be reduced by an affine transformation either to the form from Proposition 9 or to the form from Proposition 10. Then these propositions and Theorems 1 and 2 follow the conclusion of the Theorem 3. □

Table 1. The lower bounds on $nl_2(f)$ given by our Theorem 3 and by bound (4) [8, 6].

n	$AI(f)$	The bound of Theorem 3	The bound (4) [8, 6]
n>5	3	2	2
7	4	16	14
8	4	18	16
9	4	20	18
9	5	90	74
10	4	22	20
10	5	110	92
11	4	24	22
11	5	132	112
11	6	440	352
12	4	26	24
12	5	156	134
12	6	572	464
13	4	28	26
13	5	182	158
13	6	728	598
13	7	2004	1588

n	$AI(f)$	The bound of Theorem 3	The bound (4) [8, 6]
14	4	30	28
14	5	210	184
14	6	910	756
14	7	2732	2186
15	4	32	30
15	5	240	212
15	6	1120	940
15	7	3642	2942
15	8	8768	6946
16	4	34	32
16	5	272	242
16	6	1360	1152
16	7	4762	3882
16	8	12410	9888
17	4	36	32
17	5	306	274
17	6	1632	1394
17	7	6122	5034
17	8	17172	13770
17	9	37434	29786
18	4	38	36
18	5	342	308
18	6	1938	1668
18	7	7754	6428
18	8	23294	18804
18	9	54606	43556

References

[1] A. Canteaut. Open problems related to algebraic attacks on stream ciphers. In Proceedings of the 2005 International Workshop on Coding and Cryptography (WCC 2005), Bergen (Norway), pages 1-11, March 2005.

[2] C.Carlet. On the higher order nonlinearities of algebraic immune functions. CRYPTO 2006, Lecture Notes in Computer Science, vol.4117, pp. 584-601.

[3] N.Courtois and W.Meier. Algebraic attacks on stream ciphers with linear feedback. In Anvances in Cryptology — EUROCRYPT 2003, number 2656 in Lecture Notes in Computer Science, pages 345-359. Springer Verlag, 2003.

[4] D.K.Dalai, K.C.Gupta and S.Maitra. Results on Algebraic Immunity for Cryptographically Significant Boolean Functions. Indocrypt 2004, Chennai, India, December 20-22, pages 92-106, number 3348 in Lecture Notes in Computer Science, Springer Verl ag, 2004.

[5] M.Lobanov. Exact relation between nonlinearity and algebraic immunity. Discrete Mathematics and Applications, Vol. 16, Issue 5, pp. 453–460, 2006.

[6] M.Lobanov. The bound on the nonlinearity of high orders of Boolean function via the value of its algebraic immunity. Proceedings of 6th school of young researchers in discrete mathematics and its applications, Moscow, April 2007, Part 2, pp. 11–16 (in Russian).

[7] W.Meier, E.Pasalic and C.Carlet. Algebraic attacks and decomposition of Boolean functions. In Advances in Cryptology — EUROCRYPT 2004, number 3027 in Lecture Notes in Computer Science, pages 474-491. Springer Verlag, 2004.

[8] S.Mesnager. Improving the lower bound on the higher order nonlinearity of Boolean functions with prescribed algebraic immunity. Cryptology ePrint archive(http://eprint.iacr.org/), Report 2007/117.

[9] F. J. McWilliams and N. J. A. Sloane, The Theory of Error Correcring Codes. New York: North-Holland, 1977.