S©MR ISSN 1813-3304

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ

Siberian Electronic Mathematical Reports
http://semr.math.nsc.ru

Том 11, стр. 745–751 (2014) УДК 512.5
MSC 13A99

ON DECOMPOSITION OF A BOOLEAN FUNCTION INTO SUM OF BENT FUNCTIONS

N.N. TOKAREVA

Abstract. It is proved that every Boolean function in n variables of a constant degree d, where $d \leq n/2$, n is even, can be represented as the sum of constant number of bent functions in n variables. It is shown that any cubic Boolean function in 8 variables is the sum of not more than 4 bent functions in 8 variables.

Keywords: Boolean function; bent function; affine classification; bent decomposition.

1. Introduction

Nonlinear Boolean functions and their generalizations are widely used in cryptographic applications, e. g. in filter and combining models of pseudorandom stream generators, for constructing of highly nonlinear S-boxes, see [1], [10].

Boolean functions with extremal nonlinear properties are called *bent functions*. They are exactly those functions that have the maximal possible Hamming distance to the class of all affine Boolean functions in n variables [7]. Note that degree of a bent function is not more than $n/2$. One of the most important problem in bent functions is to find the number of them. In [8] we introduced a new approach to this problem and formulated the following hypothesis.

Hypothesis 1. *Any Boolean function in n variables of degree not more than $n/2$ can be represented as the sum of two bent functions in n variables (n is even, $n \geq 2$).*

This hypothesis is an analog of the Goldbach's conjecture in number theory unsolved since 1742: any even number $n > 4$ can be represented as the sum of two

Tokareva, N.N., On decomposition of a Boolean function into sum of bent functions.

The work is supported by RFFI (grant 14-01-00507).
Received August, 14, 2014, published September, 21, 2014.

prime numbers. If one can prove the mentioned hypothesis on bent functions then the asymptotic value of the number of all bent functions will be found and hence an answer for the main question in bent functions will be given.

In [8] we checked the hypothesis for Boolean functions in n variables for all possible small cases: for $n = 2, 4, 6$. To check the case $n = 8$ is too hard now, since there is no complete affine classification of Boolean functions of degree 4 in 8 variables. Directly it requires to find bent decompositions for about 2^{163} Boolean functions in 8 variables. Recall that the number of all bent functions in 8 variables has been found only a few years ago [3] and is about $2^{106,29}$.

L. Qu and C. Li [6] continued the study of [8]. They confirmed the hypothesis in some particular cases. Namely, they proved that all quadratic Boolean functions, Maiorana-McFarland bent functions and partial spread functions can be represented as the sums of two bent functions.

In this paper a weakened variant of Hypothesis 1 is studied. It is proved that every Boolean function in n variables of a constant degree d, where $d \leq n/2$, n is even, can be represented as the sum of constant number of bent functions in n variables (Theorem 1). Using affine classification of homogeneous cubic Boolean functions in 8 variables we show that any cubic Boolean function in 8 variables is the sum of not more than 4 bent functions in 8 variables (Theorem 2).

2. Preliminaries

Recall that a Boolean function f in n variables can be uniquely represented by its *algebraic normal form* (ANF)

$$f(x) = \left(\sum_{k=1}^{n} \sum_{i_1,\dots,i_k} a_{i_1,\dots,i_k}\, x_{i_1} \cdot \ldots \cdot x_{i_k} \right) + a_0,$$

where for each k indices $i_1, \dots, i_k$ are pairwise distinct and all together run trough all k-element subsets of the set $\{1, \dots, n\}$. Here $+$ denotes sum modulo 2, coefficients $a_{i_1,\dots,i_k}$, a_0 belong to $\mathbb{Z}_2$. *Algebraic degree* (briefly *degree*) of a Boolean function f is the number of variables in the longest item of its ANF. A Boolean function is called *affine*, *quadratic*, *cubic*, etc. if its degree is ≤ 1 or equals to 2, 3 and so on. If all items of ANF of a Boolean function contain exactly k variables then such a function is called *homogeneous of degree k*.

A Boolean function f in n variables is called *bent* if for every nonzero vector y it holds $\sum_{x \in \mathbb{Z}_2^n} (-1)^{f(x)+f(x+y)} = 0$, where n is even. Equivalently, bent function is on the maximal possible Hamming distance from the class of all affine Boolean functions in n variables [7]. Note that degree of a bent function is between 2 and $n/2$.

Recall the known McFarland construction [5] of bent functions.

Proposition 1. *(see* [5]*) Let π be a permutation on the set of all binary vectors of length $n/2$, let h be a Boolean function in $n/2$ variables. Then $f(x', x'') = \langle x', \pi(x'') \rangle + h(x'')$ is a bent function in n variables. Here x', x'' are of length $n/2$.*

Note that in McFarland construction n variables can be partitioned into halves in an arbitrary way.

Proposition 2. *(see* [6]*) Every McFarland bent function can be represented as the sum of two bent functions.*

It is easy to prove Proposition 2. Indeed [6], if one consider a Boolean function in n variables as a function from $GF(2^n)$ to $GF(2)$ then $f(x', x'') = \langle x', \pi(x'')\rangle + h(x'')$ can be represented as sum of two bent functions f_1 and f_2, where $f_1(x', x'') = \langle x', \beta\pi(x'')\rangle$ and $f_2(x', x'') = \langle x', (\beta+1)\pi(x'')\rangle + h(x'')$, where β is an arbitrary element from $GF(2^n) \setminus \{0, 1\}$.

Recall that a Boolean function f in n variables *depends on variable* x_i if its ANF contains x_i.

Proposition 3. *(see* [6]*) Every Boolean function in n variables that depends on $n/2$ variables (or less) is the sum of two bent functions.*

Indeed, if a function f depends only on variables $x_1, \ldots, x_d$, $d \le n/2$, then it is the sum of two McFarland bent functions $g(x', x'') = \langle x', x''\rangle$ and $h(x', x'') = \langle x', x''\rangle + f'(x')$, where $f'(x') = f(x', x'')$ and variables $x_1, \ldots, x_d$ are covered by vector x'.

3. Boolean function as the sum of constant number of bent functions

In this section we prove a weakened variant of Hypothesis 1.

Note that according to [6] an arbitrary affine or quadratic Boolean function can be represented as the sum of two bent functions with the same number of variables. Let us prove

Theorem 1. *Every Boolean function in n variables of degree d, where $3 \le d \le n/2$, n is even, can be represented as the sum of constant number N_d of bent functions in n variables. Moreover, $N_d \le 2\binom{2k}{d}$, where k is the least number, $k \ge d$, such that $n/2$ can be divided by k.*

Here $\binom{n}{m}$ is the binomial coefficient.

Proof. Let f be a Boolean function of degree d. Consider two cases.

Case 1. Let d divide $n/2$, i. e. $n = 2dm$ for some integer m. Then number k defined in the statement of the theorem is equal to d. Consider the partition of the set $\{1, 2, 3, \ldots, n\}$ into $2d$ subsets

$$A_1 = \{1, \ldots, m\},\ A_2 = \{m+1, \ldots, 2m\},$$
$$A_3 = \{2m+1, \ldots, 3m\},\ \ldots,\ A_{2d} = \{2m(d-1)+1, \ldots n\}.$$

For any monomial of degree up to d it is possible to choose d subsets A_i in such a way that union of them, say A, covers all variables of the monomial. Note that $|A| = n/2$. Every such a choice of d subsets A_i among the given $2d$ sets defines a partition of all variables into halves. Let x' be the vector of variables with numbers from A and x'' be the vector of variables with numbers from $\{1, 2, 3, \ldots, n\} \setminus A$. There are $\binom{2d}{d}$ such partitions (x', x'') of n variables. Here partitions (x', x'') and (x'', x') are distinct. We say that monomial is *associated* with a partition (x', x'') if all its variables are covered by x'. Divide all monomials of the Boolean function f into groups by the association with the same partition (x', x''). One can see that sum of monomials of one group is a Boolean function that depends on not more

that $n/2$ variables. Then by Proposition 3 it can be represented as the sum of two bent functions in n variables.

Thus, every Boolean function f can be expressed as the sum of not more than $2\left(\begin{array}{c}2d\\d\end{array}\right)$ bent functions.

Case 2. Number d does not divide $n/2$. Take the least k such that $k > d$ and $n = 2k\ell$, where $1 \le \ell \le n/8$. The similar operations as in Case 1 can be performed. Let us construct partitions (x', x'') using sets

$$B_1 = \{1, \dots, \ell\},\ B_2 = \{\ell + 1, \dots, 2\ell\},$$
$$B_3 = \{2\ell + 1, \dots, 3\ell\},\ \dots,\ B_{2k} = \{2\ell(k-1) + 1, \dots n\}.$$

Namely for every monomial of degree not more than d one can choose d subsets B_i in such a way that union of them (denote it by B) covers all variables of the monomial. Note that $|B| = d\ell < n/2$. If we add to the set B any other $k-d$ subsets B_i (say, with the smallest possible indices) then we obtain the set of size $n/2$ (it is considered to be the set of coordinates of x') and hence get a partition (x', x'') of all variables into halves. Thus, using not more than $\left(\begin{array}{c}2k\\d\end{array}\right)$ of such partitions (x', x'') we can represent a Boolean function f as the sum of at most $2\left(\begin{array}{c}2k\\d\end{array}\right)$ bent functions in the same way as in Case 1. □

It is easy to derive

Corollary 1. *Let even n be multiple of 3. Then every cubic Boolean function in n variables is the sum of not more than 40 bent functions in n variables.*

Corollary 2. *Let $n/2$ be multiple of 4. Then an arbitrary Boolean function in n variables of degree 4 is the sum of not more than 140 bent functions in n variables.*

Corollary 3. *Let even n be multiple of 5. Then every Boolean function in n variables of degree 5 is the sum of not more than 504 bent functions in n variables.*

The next proposition is devoted to decompositions of bent functions. It shows the speciality of case $k = 2$ in decomposition of a bent function into the sum of k bent functions.

Proposition 4. *Let k be a positive integer, $k \ne 2$. An arbitrary bent function in n variables can be represented as the sum of k distinct bent functions in n variables.*

Proof. Let f be a bent function in n variables. Let us show how to get the required decompositions. Case $k = 1$ is out of our interest. Consider $k = 3$. Take an arbitrary bent function g in n variables. Let g_ℓ be a bent function obtained from g by adding an affine (nonzero) Boolean function ℓ. Then $f = f_\ell + g + g_\ell$. It is easy to see that by considering another bent function h and its shift $h_{\ell'}$ (that can not be obtained from f, g by adding an affine functions) one can get a decomposition of f in 5 distinct bent functions, $f = f_{\ell+\ell'} + g + g_\ell + h + h_{\ell'}$, and so on.

If k is even, consider a bent function g from the McFarland class. By Proposition 2 there exist bent functions m and m' such that $g = m + m'$. Then we get a bent decomposition for the function f in 4 bent functions, namely $f = f_\ell + m + m' + g_\ell$. It is clear now how to get decompositions of f into 6 and more bent functions. □

4. Any cubic Boolean function in 8 variables is the sum of at most 4 bent functions

In this section we find bent decompositions of Boolean functions in 8 variables of degree up to 3.

Recall that Boolean functions f and g in n variables are *affine equivalent*, if there exist nonsingular binary $n \times n$ matrix A, vectors u, v of length n and constant $\lambda \in \mathbb{Z}_2$, such that $g(x) = f(Ax + u) + \langle v, x \rangle + \lambda$.

We can study bent decompositions only for affine nonequivalent Boolean functions due to the following facts.

Proposition 5. *A Boolean function affine equivalent to a bent function is bent too.*

Proposition 6. *Let a Boolean function f in n variables be represented as the sum of k bent functions. Then every Boolean function affine equivalent to f also can be represented as the sum of k bent functions.*

Proposition 5 is well known [7]. Note that in [9] it is proved that there is no other isometric transformation of bent functions that save the property to be bent.

Let us recall the following result.

Proposition 7. *(see* [6]*) Every quadratic Boolean function in n variables (n is even) is the sum of two bent functions in n variables.*

The proof of this fact was based on the known affine classification of all quadratic Boolean functions in n variables (due to the Dickson's theorem). Thus, let us consider Boolean functions of degree 3.

Theorem 2. *Every cubic Boolean function in* 8 *variables is the sum of not more than* 4 *bent functions.*

Proof. Consider all affine nonequivalent bent functions in 8 variables of degree not more than 3. We list them bellow according to classification obtained in [2]. To be short we write monomial $x_1x_2x_3$ as 123 and so on. Let $f(x) = f_3(x) + f_2(x)$ be an arbitrary cubic Boolean function in 8 variables, where $f_3(x)$ is a homogeneous part of degree 3 and $f_2(x)$ has degree ≤ 2. W.l.o.g. assume that f_3 is from the table bellow (otherwise consider a function affine equivalent to f).

It is not hard to get decompositions of the Boolean function f up to the quadratic part. It is enough to use only following nonequivalent bent functions:

$a = 123 + 14 + 25 + 36 + 78$;
$b = 123 + 145 + 34 + 16 + 27 + 58$;
$c = 123 + 145 + 346 + 35 + 16 + 15 + 27 + 48$;
$d = 123 + 347 + 356 + 14 + 76 + 25 + 45 + 38$;
$e = 123 + 145 + 247 + 346 + 35 + 17 + 25 + 26 + 48$.

We give the required decomposition in the form $f(x) = g(\pi(x)) + h(\sigma(x)) + q(x)$, where g and h are bent functions from the set $\{a, b, c, d, e\}$, substitutions π, σ are nonsingular affine transformations of variables (permutations in most cases), function q is a certain Boolean function of degree ≤ 2 (we do not concretize it). According to [6] any quadratic function q is the sum of two bent functions. Thus, f can be represented as the sum of not more than 4 bent functions in 8 variables.

For example, function $f(x) = x_1x_2x_3 + x_2x_4x_6 + x_3x_5x_7 + x_1x_2x_8 + x_1x_3x_8$ (number 15 in the table) is the sum $b(x_2 + x_3, x_1, x_8, x_4, x_6, x_3, x_5, x_7) + d(x_1 + x_2, x_2, x_3, x_4, x_5, x_7, x_6, x_8) + q(x)$, where q is a quadratic function.

No	Affine nonequivalent homogeneous Boolean functions of degree 3	g	h	π	σ
1	123	a	b	$[1, 4, 5, 2, 3, 6, 7, 8]$	id
2	$123 + 145$	a	a	id	$[1, 4, 5, 2, 3, 6, 7, 8]$
3	$123 + 456$	a	a	id	$[4, 5, 6, 1, 2, 3, 7, 8]$
4	$123 + 135 + 236$	a	b	id	$[3, 1, 5, 2, 6, 4, 7, 8]$
5	$123 + 124 + 135 + 236 + 456$	c	c	$[1 + 6, 2, 3, 4, 5, 6, 7, 8]$	$[3 + 4, 5, 1, 4, 6, 2, 7, 8]$
6	$123 + 145 + 167$	a	b	id	[1,4,5,6,7,2,3,8]
7	$123 + 246 + 357$	b	d	$[4, 2, 6, 3, 8, 1, 7, 5]$	$[1, 2, 3, 4, 5, 7, 8, 6]$
8	$123 + 145 + 167 + 246$	a	c	id	$[1, 5, 4, 6, 7, 2, 3, 8]$
9	$123 + 145 + 246 + 357$	d	d	$[1, 2, 3, 4, 5, 7, 8, 6]$	$[1, 5, 4, 2, 3, 8, 6, 7]$
10	$123 + 124 + 135 + 236 + 456 + 167$	b	d	$[1 + 6, 2, 3, 4, 5, 6, 7, 8]$	$[2 + 5, 4, 1, 3, 6, 7, 5, 8]$
11	$123 + 145 + 167 + 246 + 357$	b	c	$[6, 1, 7, 2, 4, 3, 5, 8]$	$[1, 2, 3, 5, 4, 7, 6, 8]$
12	$123 + 478 + 568$	a	b	id	$[8, 4, 7, 5, 6, 1, 2, 3]$
13	$123 + 145 + 167 + 568$	a	c	id	$[1, 4, 5, 6, 7, 8, 2, 3]$
14	$123 + 246 + 357 + 568$	c	d	[4,2,6,8,3,5,1,7]	[1,2,3,4,5,7,8,6]
15	$123 + 246 + 357 + 128 + 138$	b	d	$[2 + 3, 1, 8, 4, 6, 3, 5, 7]$	$[1 + 2, 2, 3, 4, 5, 7, 6, 8]$
16	$123 + 145 + 167 + 357 + 568$	a	e	id	$[1, 6, 7, 5, 4, 3, 8, 2]$
17	$123 + 145 + 478 + 568$	a	c	id	$[4, 1, 5, 8, 7, 6, 2, 3]$
18	$123 + 124 + 135 + 236 + 456 + 167 + 258$	e	e	$[1, 2 + 5, 3, 5, 4, 6, 8, 7]$	$[1, 2 + 5, 4, 6, 7, 5, 3, 8]$
19	$123 + 124 + 135 + 236 + 456 + 178$	b	d	$[1 + 6, 2, 3, 4, 5, 6, 7, 8]$	$[2 + 5, 4, 1, 3, 7, 8, 5, 6]$
20	$123 + 145 + 246 + 357 + 568$	d	e	$[1, 2, 3, 4, 5, 7, 8, 6]$	$[5, 6, 8, 4, 1, 3, 2, 7]$
21	$123 + 145 + 246 + 467 + 578$	c	e	$[4, 3, 8, 7, 6, 5, 1, 2]$	$[1, 2, 3, 4, 5, 8, 6, 7]$
22	$123 + 145 + 357 + 478 + 568$	a	e	id	$[4, 7, 8, 5, 1, 6, 3, 2]$
23	$123 + 246 + 357 + 478 + 568$	c	e	$[1, 2, 3, 5, 4, 7, 6, 8]$	$[5, 6, 8, 4, 1, 7, 2, 3]$
24	$123 + 246 + 357 + 148 + 178 + 258$	c	c	$[1, 2, 3, 7, 8, 5, 4, 6]$	$[2, 5, 8, 4, 6, 1, 3, 7]$
25	$123 + 145 + 167 + 246 + 357 + 568$	c	d	$[1, 2, 3, 5, 4, 7, 6, 8]$	$[1, 7, 6, 2, 5, 8, 4, 3]$
26	$123 + 145 + 167 + 246 + 238 + 258 + 348$	c	e	$[1, 7 + 8, 6, 4, 5, 2, 3, 8]$	$[2, 1 + 8, 3, 8, 5, 4, 6, 7]$
27	$123 + 145 + 167 + 258 + 268 + 378 + 468$	c	e	$[1, 3 + 8, 2, 5, 4, 8, 6, 7]$	$[6, 1 + 6, 7, 8, 4, 3, 2, 5]$
28	$123 + 145 + 246 + 357 + 238 + 678$	c	c	$[1, 2, 3, 5, 4, 7, 6, 8]$	$[2, 3, 8, 6, 4, 7, 1, 5]$
29	$123 + 145 + 246 + 357 + 478 + 568$	c	c	$[1, 2, 3, 5, 4, 7, 6, 8]$	$[4, 2, 6, 8, 7, 5, 1, 3]$
30	$123 + 124 + 135 + 236 + 456 + 167 + 258 + 378$	c	e	$[1, 2, 3 + 4, 6, 7, 5, 4, 8]$	$[5, 8, 2 + 5, 3, 1, 6, 7, 4]$
31	$123 + 156 + 246 + 256 + 147 + 157 + +357 + 348 + 258 + 458$	c	e	$[5, 2 + 4, 8, 3, 7, 4, 1, 6]$	$[2, 4 + 5, 6, 1, 3, 5, 7, 8]$

□

5. Conclusion

The weakened variant of hypothesis about bent decompositions of Boolean functions is proved. For further research it is interesting to decrease the number of bent functions in decomposition proposed in Theorem 1 and use for these decompositions not only McFarland bent functions. For small number of variables $n = 2, 4, 6$ and 8 (if degree ≤ 3) decomposition of a Boolean function in 2 or 4 bent functions is now provided. The author is supported by the Russian Foundation for Basic Research (grant 14-01-00507).

References

[1] Cusick T. W., Stănică P., *Cryptographic Boolean Functions and Applications.* Acad. Press. Elsevier. 2009. 245 pages. MR2530579.

[2] Hou X.-D., *Cubic bent functions* // Discrete Mathematics, **189** (1998), 149–161. MR1637733.

[3] Langevin P., Leander G., *Counting all bent functions in dimension eight 99270589265934370305785861242880* // Designs, Codes and Crypt., **59** (2011), 193–205. MR2781609.

[4] Logachev O. A., Sal'nikov A. A., Smyshlyaev S. V., Yashenko V. V., *Boolean functions in coding theory and cryptology* // Moscow center for the uninterrupted mathematical education, 2012. 584 pages.

[5] McFarland R. L., *A family of difference sets in non-cyclic groups* // J. of Combin. Theory, Ser. A. **15**:1 (1973), 1–10. MR0314647.

[6] Qu L. and Li C., *When a Boolean Function can be Expressed as the Sum of two Bent Functions* // Cryptology ePrint Archive. 2014/048.

[7] Rothaus O. *On bent functions* // J. Combin. Theory. Ser. A. **20**:3 (1976), 300–305. MR0403988.

[8] Tokareva N. N., *On the number of bent functions from iterative constructions: lower bounds and hypotheses* // Advances in Mathematics of Communications (AMC). **5**:4 (2011), 609–621. MR2855274.

[9] Tokareva N. N., *Duality between bent functions and affine functions* // Discrete Mathematics. **312** (2012), 666–670. MR2854814.

[10] Tokareva N., *Nonlinear Boolean functions: bent functions and their generalizations.* LAP LAMBERT Academic Publishing (Saarbrucken, Germany), 2011. ISBN: 978-3-8433-0904-2. 180 pages.

Natalia Nikolaevna Tokareva
Sobolev Institute of Mathematics,
pr. Koptyuga, 4,
630090, Novosibirsk, Russia
Novosibirsk State University,
Pirogova st., 2,
630090, Novosibirsk, Russia
E-mail address: tokareva@math.nsc.ru