On structural properties of the class of bent functions

Natalia Tokareva

Sobolev Institute of Mathematics,
Novosibirsk State University
Russia
tokareva@math.nsc.ru

Maximally nonlinear Boolean functions in n variables, where n is even, are called **bent functions**.

Bent functions form the special mysterious class, $\mathcal{B}_n$, studied from the early sixties in connection with cryptographic applications.

Too many problems related to this class are still open.
Constructions cover only separate parts of $\mathcal{B}_n$ while the core of it is still hidden from one's eyes.

In this talk let us try to deal not with separate constructions of bent functions, but with the set of bent functions $\mathcal{B}_n$ at whole.

Definitions

$\mathbb{F}_2^n$ — the vector space over $\mathbb{F}_2$;

$f, g : \mathbb{F}_2^n \to \mathbb{F}_2$ — Boolean functions;

$dist(f, g)$ — Hamming distance between f and g, i. e. the number of coordinates in which their vectors of values differ;

$x = (x_1, \ldots, x_n)$ — a binary vector;

$\langle x, y \rangle = x_1 y_1 + \ldots + x_n y_n$ — the standard inner product modulo 2;

$\langle a, x \rangle + b$ is an ***affine function*** in variables $x_1, \ldots, x_n$;

Bent function — a Boolean function in n variables (n is even) that is on the maximal possible distance from the set of all affine functions. This distance is $2^{n-1} - 2^{(n/2)-1}$.

$\mathcal{A}_n$ — the set of all affine functions in n variables.
$\mathcal{B}_n$ — the set of all bent functions in n variables.

A bit of history

Oscar Rothaus (1927-2003) was the recognized authority in this area. Bent functions were introduced by him in 1966 (declassified in 1976).

He graduated from Princeton University; served in the US Army Signal Corps during the Korean War, and then as a mathematician at the National Security Agency. From 1960 to 1966, he worked at the Defense Department's Institute for Defense Analyses.

«He was one of the most important teachers of cryptology to mathematicians and mathematics to cryptologists»

(a top of the Institute for Defense Analysis about O. Rothaus)

By O. Rothaus the main properties of bent functions were obtained, simple constructions of bent functions were given, and several steps for the classification of bent functions in six variables were made. In 1966, he joined Cornell University as a professor and worked there until 2003.

Oscar Rothaus

A bit of history

In the USSR, bent functions were also studied in the 1960s.

The names of the first Soviet researchers of bent functions are not too public. Also, their papers in this area have still not been declassified.

It is known that Yu. A. Vasiliev, B.M. Kloss, V.A.Eliseev, and O.P.Stepchenkov studied properties of the Walsh-Hadamard transform of a Boolean function at that time. In 1960, they studied the statistical structure of a Boolean function—that is, values

$$\Delta_a^f = 2^{n-1} - dist(f, \ell_{a,0}) = W_f(a)/2, \text{ where } a \text{ runs through } \mathbb{F}_2^n.$$

The notion of a minimal function was introduced in the USSR by **V.A. Eliseev** and **O.P. Stepchenkov** (1962). A Boolean function is **minimal** if the parameter $\Delta^f = max_a|\Delta_a^f|$ takes the minimal possible value $2^{(n/2)-1}$. Such functions exist only if n is even. Obviously, "minimal function" is just another name for "bent function."

An analog of the McFarland construction of bent functions was proposed by V.A. Eliseev in 1962. At the same year they proved that the degree of a minimal function is not more than $n/2$.

V.A.Eliseev

O.P.Stepchenkov

Robert McFarland; John Dillon

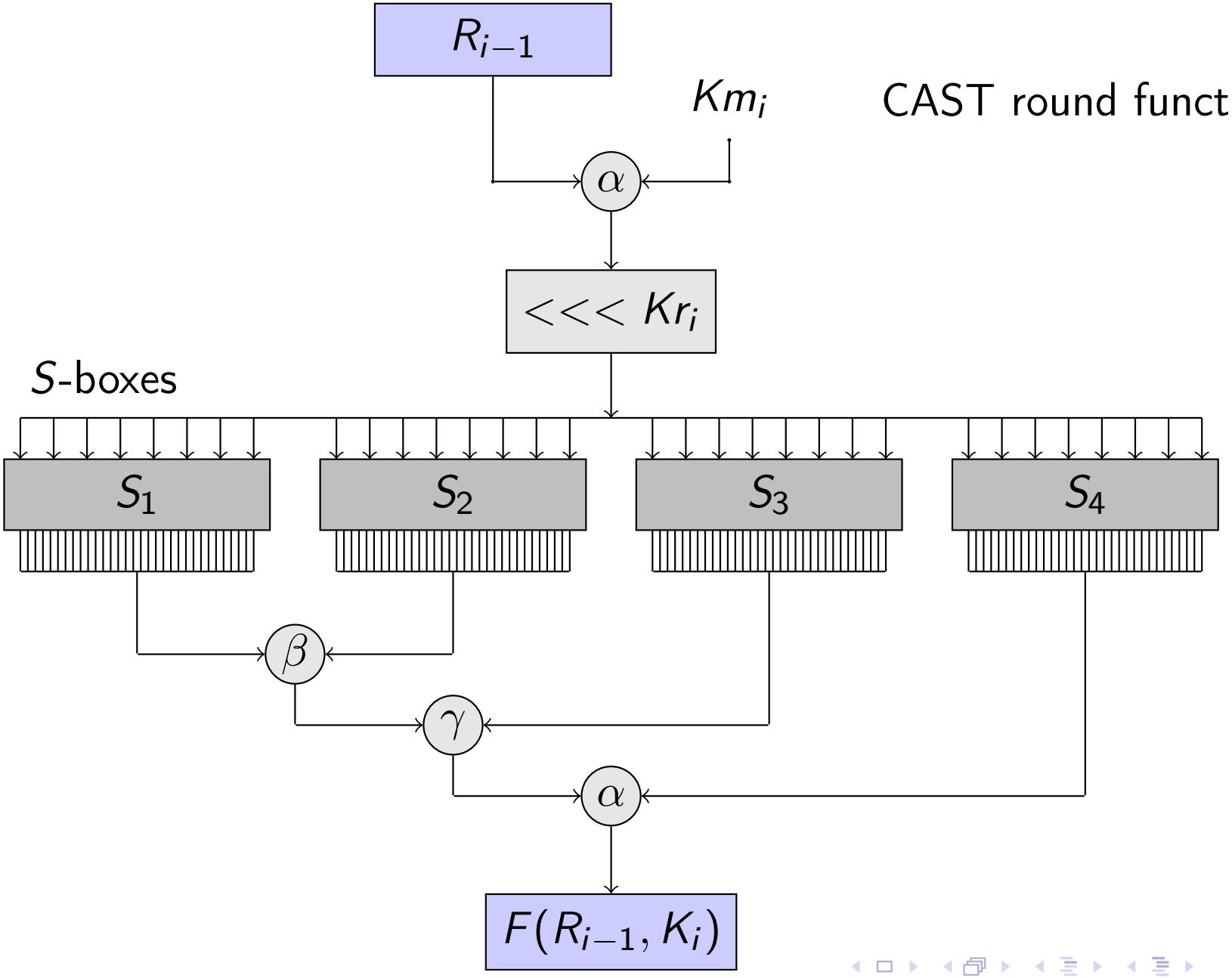
J.F. Dillon (1972) Bent functions in connection to differential sets;
R.L. McFarland (1973) Large class of bent functions.

Applications of bent functions

Now bent functions are studied very widely since they have numerous applications in computer science.

Hadamard matrices (combinatorics);
Classification problems for H. m. and bent functions are equivalent.
Differential sets (group theory);
Orthogonal spreads (finite geometries);
Codes of the constant amplitude in **CDMA systems** — the 3d generation mobile systems (communication theory);
Kerdock codes (coding theory);
S-boxes in block and stream ciphers resistant to linear cryptanalyses. E. g. CAST, Grain, etc. (cryptography);
Authentication schemes, hash functions; pseudo-random generators (cryptography)

CAST round function



An example

Each S-box of CAST is a vectorial Boolean function, $S_j : \mathbb{Z}_2^8 \to \mathbb{Z}_2^{32}$. One can express it with the set of 32 Boolean functions $f_k^{(j)}$, i. e.

$$S_j(x_1, \ldots, x_8) = (y_1, \ldots, y_{32}),\ j = 1, \ldots, 4$$

where

$$y_k = f_k^{(j)}(x_1, \ldots, x_8),\ k = 1, \ldots, 32.$$

In CAST all the functions $f_k^{(j)}$ are bent. Moreover any linear combination of component functions from one S-box has «good enough» nonlinear properties. It was done for making CAST secure to linear cryptanalysis.

Well-known open problems in bent functions

To find asymptotic value for the number of bent functions.
Now the exact number of bent functions is known only for $n \leqslant 8$.
It is very hard even to find good lower and upper bound for the number of bent functions.

Lower bound: $2^{2^{(n/2)+\log(n-2)-1}}$ (McFarland construction)

Upper bound: $2^{2^{n-1}+\frac{1}{2}\binom{n}{n/2}}$ (# of functions of degree $\leq n/2$)

To classify bent functions with respect to some (affine?) equivalence.

To find new constructions of bent functions.
There are known a few constructions that cover only the small part of all bent functions.

To reach a tradeoff between high nonlinearity and other cryptographic properties of a Boolean function.

Structural properties

Consider $\mathcal{B}_n$ as the subset of $\mathbb{F}_2^{2^n}$. What can we say about it? What problems we can formulate?

So, our object is the whole class of bent functions $\mathcal{B}_n$ and we are interested in its role in the set of all Boolean functions.

Automorphisms of the set of bent functions

Automorphisms of the set of bent functions

Let A be a binary nonsingular $n \times n$–matrix, b, c be any binary vectors of length n and d be a binary constant (0 or 1).

It is well known that $\mathcal{B}_n$ is closed under addition of affine functions and under affine transformations of variables, i. e. for any bent function g the function

$$g'(x) = g(Ax + b) + \langle c, x \rangle + d$$

is bent again. The functions g and g' are called ***EA-equivalent***.

In 2010 we have proven

Theorem. *For any non affine Boolean function f there exists a bent function g such that $f + g$ is not bent.*

By definition,

$$\mathcal{B}_n = \{f : dist(f, \mathcal{A}_n) \text{ is maximal, equal to } 2^{n-1} - 2^{(n/2)-1}\}.$$

Is it possible to invert this definition? In other words is it true that $\mathcal{A}_n$ is the set of all Boolean functions that are at the maximal distance from $\mathcal{B}_n$? What is this maximal distance?

We proved that YES,

$$\mathcal{A}_n = \{f : dist(f, \mathcal{B}_n) \text{ is maximal, equal to } 2^{n-1} - 2^{(n/2)-1}\}.$$

Thus, there is, so to say, a *duality* between definitions for bent and affine functions. Note that Theorem above is a key fact for it.

Mapping φ of the set of all Boolean functions in n variables into itself is ***isometric***, if it preserves Hamming distances, i. e.

$$dist(\varphi(f), \varphi(g)) = dist(f, g).$$

It is known that any such a mapping can be given as

$$g(x) \to g(s(x)) + f(x),$$

where $s : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ ia a substitution, f is a Boolean function.

Automorphism group of a subset of Boolean functions $\mathcal{M}$ is the group of all isometric mappings of the set of all Boolean functions into itself that transform $\mathcal{M}$ again to $\mathcal{M}$. Denote it by $Aut(\mathcal{M})$.

The automorphism group of all bent functions

Let $GA(n)$ be the **general affine group**,

$$GA(n) = GL(n) \ltimes \mathbb{Z}_2^n,$$

i. e. the group of all transforms $x \to Ax + b$, where A is a nonsingular matrix, b is any vector.

It is known that $Aut(\mathcal{A}_n)$ is a semidirect product of the general affine group $GA(n)$ and $\mathbb{Z}_2^{n+1}$. We proved the following fact (2010).

Theorem. *It is true* $Aut(\mathcal{B}_n) = Aut(\mathcal{A}_n) = GA(n) \ltimes \mathbb{Z}_2^{n+1}$.

Thus, any automorphism of $\mathcal{B}_n$ has the form $g \to g'$, where

$$g'(x) = g(Ax + b) + \langle c, x \rangle + d.$$

So, it is clear that definition of EA-equivalent bent functions is indeed very natural.

The set of bent functions as an extremal metrical regular set

The set of bent functions as an extremal metrical regular set

A.K. Oblaukhov continued and generalized the previous research.

Let $X \subseteq \mathbb{F}_2^n$ be an arbitrary set.
The **maximal distance** from a set X is $d(X) = \max\limits_{z \in \mathbb{F}_2^n} d(z, X)$.

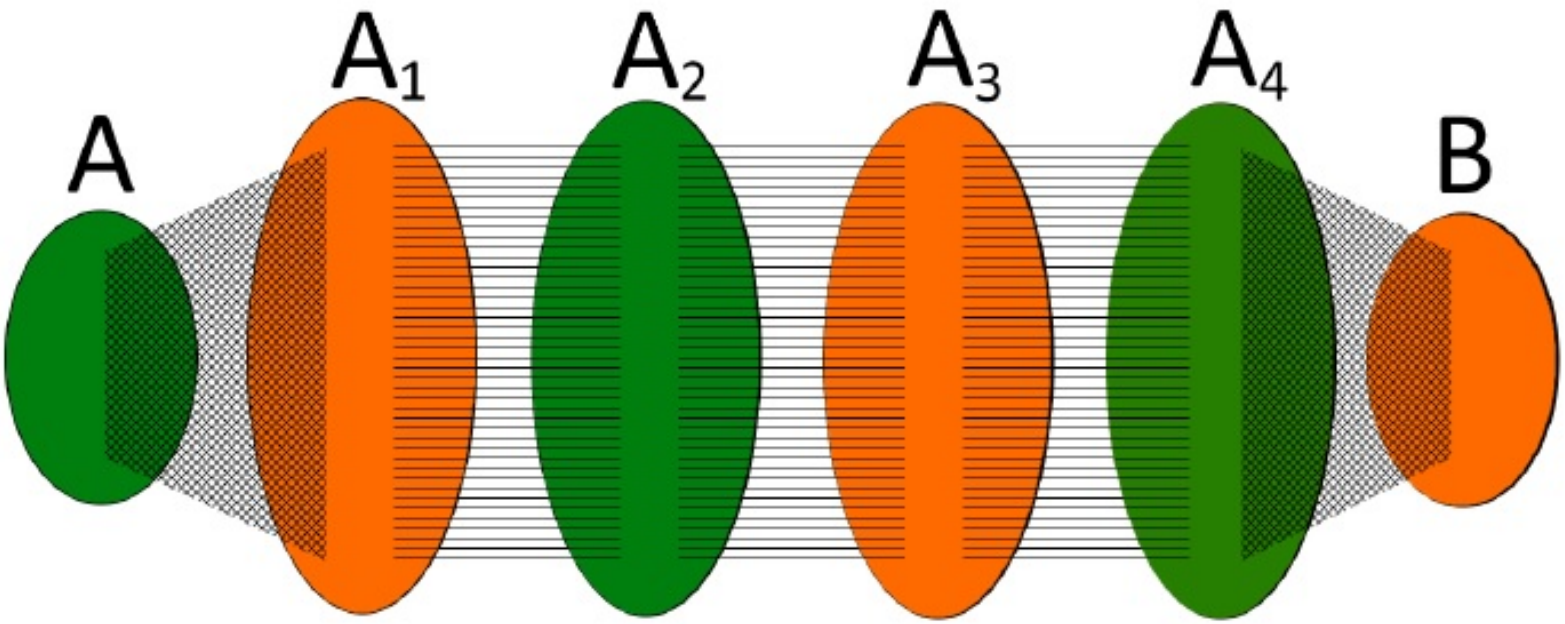
In coding th. this parameter is also known as the **covering radius** of a code.

Consider the set $\widehat{X}$ of vectors at maximal distance from X. This set is called the **metric complement** of X.

If $\widehat{\widehat{X}} = X$ then the set X is called **metrically regular**.

In 2016 A.K.Oblaukhov has proved

Theorem. Let A be an arbitrary subset of $\mathbb{F}_2^n$. Then, if we denote $A_0 := A$, $A_{k+1} = \widehat{A}_k$ for $k \geqslant 0$, there exists a number $m \leqslant n$ such that A_m is a metrically regular set, i. e. $A_m = A_{m+1}$.



A.K.Oblaukhov has proved (2017)

Theorem. Let A, B be a pair of metrically regular sets, i.e. $A = \widehat{B}, B = \widehat{A}$. Then there exists a pair of metrically regular sets A^*, B^* at distance 1 from each other such that $A \subseteq A^*, B \subseteq B^*$.

In other words, for every metrically regular set there exists a metrically regular superset with maximal distance 1. Therefore the largest metrically regular set has maximal distance 1, and it is the metric complement of the smallest metrically regular set with maximal distance 1.

Note that if A, B is a pair of metrically regular sets at distance 1 from each other, then $A \cup B = \mathbb{F}_2^n$.

A.K.Oblaukhov continues (2017) the investigation.

Theorem. Let A, B be a pair of metrically regular sets at distance d from each other of sizes M and N respectively. Then

$$M + N \geqslant \frac{2^{n+1}(n-2)}{n(n-1)^{d-1} + n - 4}.$$

Hypothesis. The class of bent functions $\mathcal{B}_n$ forms the maximal possible metrically regular set with distance $d = 2^{n-1} - 2^{n/2-1}$.

Again! we meet another kind of **extremality property** of bent functions. But now this property is for the whole class of bent functions.

Properties of $\mathcal{B}_n$ as a binary code of length 2^n

Properties of $\mathcal{B}_n$ as a binary code of length 2^n

$\mathcal{B}_n$ can be considered as the binary code of length 2^n.

The minimum distance of it is $2^{n/2}$.

The weight spectrum of the code defined by $\mathcal{B}_n$ contains only two nonzero components $A_{2^{n-1}-2^{n/2-1}}$ and $A_{2^{n-1}+2^{n/2-1}}$.
But the values of them are unknown :)

* * *

N.A.Kolomeec studied such «coding» aspects of the class of bent functions and state new problems in this area.

Minimal distances in $\mathcal{B}_n$

In 1993 C. Carlet proposed a very important construction of bent functions based on affine properties of Boolean functions.

Theorem. Let f be a bent function in n variables. Let L be an affine subspace of $\mathbb{F}_2^n$ of dimension $n/2$. Let f be affine on L. Then a Boolean function $f \oplus Ind_L$ is also a bent function in n variables.

In 2009 N. A. Kolomeec and A. V. Pavlov proved that if two bent functions are on the minimal possible distance $2^{n/2}$ then one has to be obtained from the other via Carlet's construction.

Theorem. Let f, g be Boolean functions in n variables. Let f be a bent function. Suppose that $dist(f, g) = 2^{n/2}$. Then g is bent if and only if the set $supp(f + g)$ is an affine subspace and f is affine on it.

But not for every bent function in n variables there exists a bent function on distance $2^{n/2}$, since not every bent function is normal and weakly normal (A.Canteaut, Daum M., Dobbertin H., Leander G., 2006).

Graph of minimal distances of bent functions

Let GB_n be the special graph:

- vertices — bent functions;
- there is an edge between vertices if they are on distance $2^{n/2}$.

N. Kolomeec studies such a minimal graph. He proved that

- degree of a vertex is not more than $2^{n/2}\prod_{i=1}^{n/2}(2^i+1)$;
- this bound is achieved for and only for quadratic bent functions;

Since for every even $n \geqslant 14$ there are found non weakly normal bent functions (A.Canteaut, et al. 2006), graph GB_n is not connected if $n \geqslant 14$. It is proven (N.Kolomeec, 2014) that GB_n is connected for $n = 2, 4, 6$.

Is the graph GB_n connected / disconnected if $8 \leqslant n \leqslant 12$?

Let GB'_n be the graph obtained from GB_n after elimination of all pendant vertices (corresp. to non weakly normal bent func.s).
Is GB'_n connected for all even $n \geqslant 2$?

Duality as the magic transformation of $\mathcal{B}_n$

Dual function

Recall that f is bent iff $W_f(y) = \sum_{y \in \mathbb{F}_2^n} (-1)^{f(x)+\langle x,y \rangle} = \pm 2^{n/2}$.
For every bent function there is its dual function.

A Boolean function $\tilde{f}$ is said to be **dual** of f, if

$$W_f(x) = (-1)^{\tilde{f}(x)} 2^{n/2} \text{ for any } x \in \mathbb{F}_2^n.$$

Some properties of dual functions:

- Every dual function is a bent function;
- If $\tilde{f}$ is dual to f and $\tilde{\tilde{f}}$ is dual to $\tilde{f}$, then $\tilde{\tilde{f}} = f$.

Isometry of the set of all Boolean functions

A mapping φ of the set of all Boolean functions in n variables into itself is **isometric** if it preserves Hamming distances between functions, i.e.

$$\mathrm{dist}(\varphi(f), \varphi(g)) = \mathrm{dist}(f, g).$$

It is known (A. A. Markov, 1956) that every such a map has the unique representation of the form

$$f(x) \longrightarrow f(s(x)) \oplus h(x),$$

where $s : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^n$ is a permutation and h is a Boolean function in n variables.

Properties of isometries of bent functions

It is known that the map $f \longrightarrow \tilde{f}$ which acts on the set of bent functions, preserves Hamming distance.

Answer to the following natural question was obtained by A.V. Kutsenko (2016).

Proposition. The map $f(x) \longrightarrow \tilde{f}(x)$ defined on the set of all bent function in n variables can not be extended to the isometric mapping of the set of all Boolean functions in n variables.

Isometries of bent functions

Recall that there are other isometries that keeps the set $\mathcal{B}_n$ without changes. Yes, we mean EA-transformations of $\mathcal{B}_n$.

A subset M of all Boolean functions in n variables is **transitive** with respect to the set K of some special mappings, defined on it, if for any two distinct functions $f, g \in M$ function f can be transformed to the function g by using a map from K.
A.V.Kutsenko (2017) has proposed the statements.

Proposition. The set of bent functions in n variables for $n \geqslant 6$ is not transitive with a respect to compositions of duality mapping and EA-transformations of class $\mathcal{B}_n$.

Proposition. Any bent function in at most 6 variables is affinely equivalent to its dual.

Questions for the future research in this area.

Do there exist other isometrical transformations of the set of all bent functions in n variables into itself?

Is it possible to construct such isometric mappings of $\mathcal{B}_n$ that act transitively on it?

Study the duality function in more details.

Bent sum decomposition problem

Bent sum decomposition problem

It is known that for any bent function f in n variables it holds

$$2 \leqslant deg(f) \leqslant n/2.$$

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, $n \geqslant 2$).*

We call this open question (to prove or disprove Hypothesis 1) the **Bent sum decomposition problem**. Hypothesis 1 is closely connected to the problem of asymptotic of the number of all bent functions.

This question appeared (2011) from the following considerations.

Bent sum decomposition problem

Define the set $X_n = \{\, h + g \;:\; h, g \in \mathcal{B}_n \,\}$ and consider the sets $C_g = \mathcal{B}_n + g$ for all $g \in \mathcal{B}_n$. So,

$$X_n = \bigcup_{g \in \mathcal{B}_n} C_g.$$

Let f be an element of X_n. The number of subsets C_g that cover f we call **multiplicity** of f and denote it by $m(f)$. It is clear that

$$\sum_{f \in X_n} m(f) = |\mathcal{B}_n|^2.$$

There is a low bound on the number of bent functions.

Theorem. (2011) $|\mathcal{B}_{n+2}| \geqslant \sum_{f \in X_n} m(f)^2$.

Bent sum decomposition problem

What Boolean functions can be represented as the sum of two bent functions in n variables?

Recall that for any bent function $f \in \mathcal{B}_n$ it holds

$$2 \leq deg(f) \leq n/2.$$

So, $|X_n| \leq 2^{2^{n-1}+\frac{1}{2}\binom{n}{n/2}}$. ***But what is the exact value of $|X_n|$?***

How many such representations does a Boolean function admit?

Analogy with Goldbach's conjecture

Hypothesis (unproved since 1742). *Any even number $n \geq 4$ can be represented as the sum of two prime numbers (binary variant).*

Partial results.

• Schnirelmann has proved a weak form of the Goldbach conjecture: he has shown (1931) that every number is the sum of not more than 20 primes.

• It is possible (1937) to represent big numbers as the sums of three prime numbers.

• It is known (1996), that any even number is the sum of not more than 6 prime numbers.

• Correctness of the Goldbach's conjecture is checked for numbers $\leqslant 1.2 \cdot 10^{18}$ (2008).

Christian Goldbach

Mathematician Ch. Goldbach (1690–1764) was the first professional cryptanalyst in Russia. He worked on the «special position» in the College of Foreign Affaires since 1742. He had decrypted several diplomatic ciphers.

Progress in decomposition problem

For $n = 2, 4, 6$ the set X_n contains all Boolean functions of degree less or equal to $n/2$.

$|X_2| = 2^3$, $|X_4| = 2^{11}$, $|X_6| = 2^{42}$.

The hypothesis 1 in the case $n = 6$ was checked first via exhaustive search (2011) end then (2014, L. Qu, C. Li) it was proved analytically.

L. Qu and C. Li (2014) continued the study. They confirmed the hypothesis in some particular cases. Namely, they proved that

- quadratic Boolean functions,
- Maiorana—McFarland bent functions,
- partial spread functions

can be represented as the sums of two bent functions.

Progress

We proved (2014) a some weakened variant of the hypothesis.

Theorem. *Any Boolean function in n variables of degree d, where $d \leq n/2$, n is even, can be represented as the sum of constant number A_d of bent functions in n variables. Moreover, $A_d \leq 2\binom{2b}{b}$, where b is the least number, $b \geq d$, such that $n/2$ can be divided by b.*

E. g. any Boolean function of degree 3 in $n = 6m$ variables can be represented as the sum of not more than 40 bent functions.

But the number of bent functions in decomposition *depends* on degree of the function.

Progress

In 2014 the following results were obtained.

Proposition. *Every cubic bent function in* 8 *variables can be presented as the sum of not more than* 4 *bent functions.*

In fact, to construct decompositions of these functions into sum of exactly two bent functions requires a more complicated technique while working with quadratic parts.

Proposition. *A bent function in* n *variables,* $n \geqslant 4$*, is decomposable into the sum of two bent functions in* n *variables if and only if the dual bent function is decomposable.*

Progress. A new approach

A new approach to the problem is under the work now (2017).

Let z be the vector of length n in alphabet $\{0, 1, *\}$ such that
1) exactly $n/2$ of its coordinates are equal to $*$;
2) all ones in the vector z stay before all $*$.
We call such a vector **admissible**.

For instance, for $n = 4$ there are 11 admissible vectors:

$(00**)$	$(01**)$
$(0*0*)$	$(10**)$
$(0**0)$	$(1*0*)$
$(*00*)$	$(1**0)$
$(*0*0)$	$(11**)$
$(**00)$	

Progress. A new approach

In general there are always $2^{1+\binom{n}{1}+\binom{n}{2}+\ldots+\binom{n}{n/2}} = 2^{2^{n-1}+\frac{1}{2}\binom{n}{n/2}}$ admissible vectors z of length n.
Let L_z be the facet of $\mathbb{F}_2^n$ generated by z.
For example, if $z = (1 * 0*)$ then
$L_z = \{(\mathbf{10}0), (\mathbf{10}1), (\mathbf{11}0), (\mathbf{11}1)\}$.

Theorem. *Any Boolean function in n variables of degree not more than $n/2$ can be uniquely represented as the sum modulo 2 of indicators of facets in $\mathbb{F}_2^n$ corresponding to some admissible vectors.*

So, $f(x) = \bigoplus_{z \text{ is addm.}} a_z \cdot Ind_{L_z}(x)$, where $a_z \in \mathbb{F}_2$.

Let us call this representation as $n/2$-**facet form** of a Boolean function of degree $\leqslant n/2$.

Progress. A new approach

Now we are looking for restrictions on the $n/2$-facet forms of bent functions. And the next step will be to represent any $n/2$-facet form (of an arbitrary Boolean function) as the some of two ones satisfied the restrictions. This work is still in progress.

Note that according to papers of C. Carlet and Ph. Guillot (1995–1996) the following representation for bent functions has a place:

Theorem. *Let f be a Boolean function in n variables. Then f is bent if and only if there exist linear $n/2$-dimensional subspaces $E_1, \ldots, E_k$ of $\mathbb{F}_2^n$ and integers $m_1, \ldots, m_k$ (positive or negative) such that for any element x of $\mathbb{F}_2^n$:*

$$\sum_{i=1}^{k} m_i Ind_{E_i}(x) = 2^{n/2-1}\delta_0(x) + f(x) \,[\mathrm{mod} n].$$

Here $\delta_0(x) = 1$ iff $x = 0$. But there is a some problem to get from here the characterization of bent functions in «usual» form (i. e. over $\mathbb{F}_2$).

The derivatives of bent functions

The derivatives of bent functions

Remember that f is bent iff every its derivative $D_f(x) = f(x) \oplus f(x \oplus y)$ is a balanced function, where y is a nonzero vector of length n.

Concerning the question: are there «many» or «not too many» bent functions in n variables? the following conjecture arises.

Hypothesis 2. *Every balanced Boolean function g in n variables of degree not more than $n/2 - 1$, such that for all x and for some y it holds $g(x) = g(x + y)$, is a derivative of a bent function in n variables.*

It means that any possible balanced function in n variables we can meet as the derivative of a some bent function in n variables.

We proved (2016) this hypothesis in cases $n = 4, 6$ and look now on the case $n = 8$. What about the general case?

General ideas in classification

While studying to care about the possibility for bent functions to be... irrational.

* * *

To get the lower bound (or asymptotic) for the number of bent functions via a nonconstructive way.

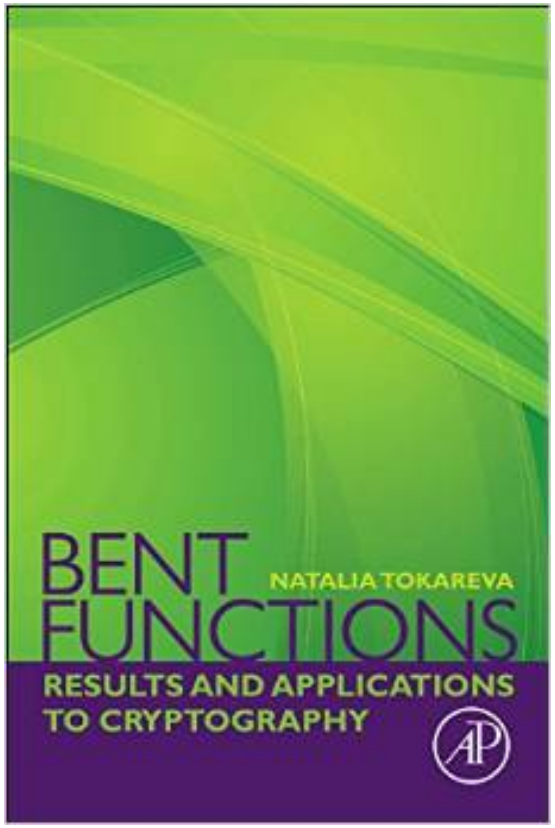
* * *

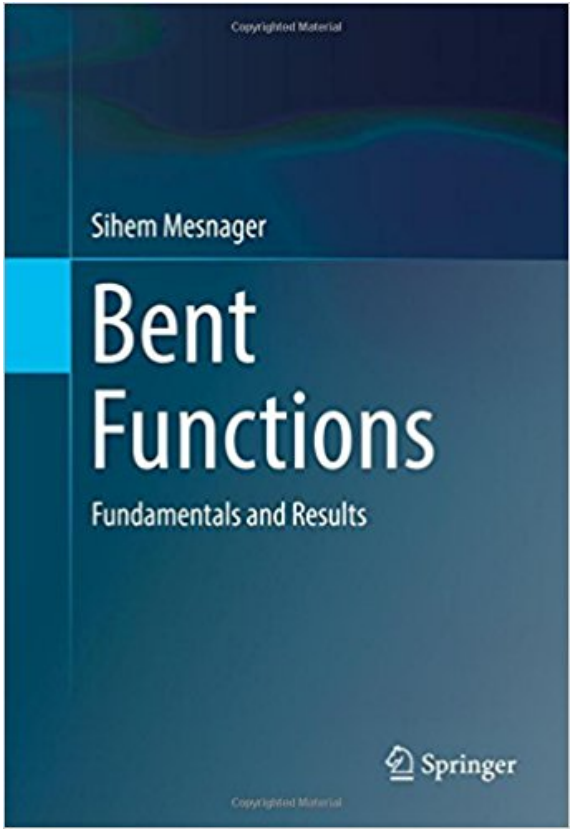
To consider problems in more general mathematical form, rely with the basic concepts. Very often partial cases and projections of objects are more complicated than the general ones.

To conclude: books in bent functions

N.T. «Bent functions: results and applications to cryptography» Elsevier, 2015).

S. Mesnager «Bent functions: Fundamentals and Results» (Springer, 2016).





NSUCRYPTO-2017: welcome!

October 22–30, 2017.

NSUCRYPTO-2017: October 22, 2017. Welcome!

International Students' Olympiad in Cryptography. It is organized by

Novosibirsk State University

Sobolev Institute of Mathematics (Novosibirsk)

University of Leuven (KU Leuven, Belgium)

Belarusian State University

Tomsk State University

NSUCRYPTO is the unique cryptographic Olympiad containing scientific mathematical problems for senior pupils, students and professionals from any country. The concept of the Olympiad is not to focus on solving only olympic tasks but on including hard and unsolved research problems at the intersection of mathematics and cryptography. It holds in two rounds via Internet. Welcome to participate!

www.nsucrypto.nsu.ru

Thank you for the attention!