Boolean Functions for Cryptography and Error Correcting Codes

Claude Carlet*

*LAGA, University of Paris 8, France; e-mail: claude.carlet@univ-paris8.fr.

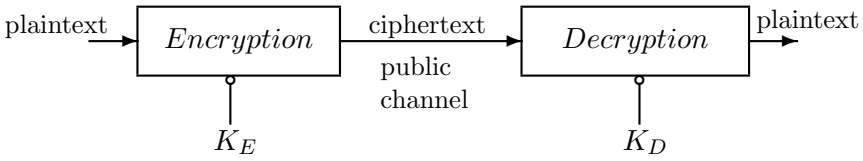
Contents

1 Introduction

A fundamental objective of *cryptography* is to enable two persons to communicate over an insecure channel (a public channel such as internet) in such a way that any other person is unable to recover their message (called the *plaintext*) from what is sent in its place over the channel (the *ciphertext*). The transformation of the plaintext into the ciphertext is called *encryption*, or enciphering. Encryption-decryption is the most ancient cryptographic activity (ciphers already existed four centuries B. C.) but its nature has deeply changed with the invention of computers, because the *cryptanalysis* (the activity of the third person, the eavesdropper, who aims at recovering the message) can use their power.

The encryption algorithm takes as input the plaintext and an encryption key K_E, and it outputs the ciphertext. If the encryption key is secret, then we speak of *conventional cryptography*, of *private key cryptography* or of *symmetric cryptography*. In practice, the principle of conventional cryptography relies on the sharing of a private key between the sender of a message (often called Alice in cryptography) and its receiver (often called Bob). If, on the contrary, the encryption key is public, then we speak of *public key cryptography*. Public key cryptography appeared in the literature in the late seventies.

The *decryption* (or deciphering) algorithm takes as input the ciphertext and a secret[1] decryption key K_D. It outputs the plaintext.

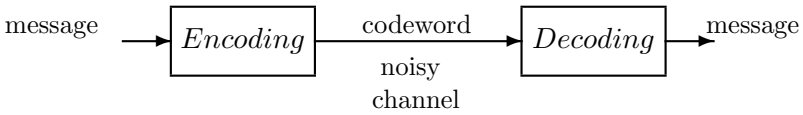


Public key cryptography is preferable to conventional cryptography, since it allows to securely communicate without having previously shared keys in a secure way: every person who wants to receive secret messages can keep secret a decryption key and publish an encryption key; if n persons want to secretly communicate pairwise using a public key cryptosystem, they need n

[1] According to principles already stated in 1883 by A. Kerckhoffs [212], who cited a still more ancient manuscript by R. du Carlet [50], only the secret keys must be kept secret – the confidentiality should not rely on the secrecy of the encryption method – and a cipher cannot be considered secure if it can be decrypted by the designer himself.

encryption keys and n decryption keys, when conventional cryptosystems will need $\binom{n}{2} = \frac{n(n-1)}{2}$ keys. But all known public key cryptosystems are much less efficient than conventional cryptosystems (they allow a much lower data throughput) and they also need much longer keys to ensure the same level of security. This is why conventional cryptography is still widely used and studied nowadays. Thanks to public key cryptosystems, the share-out of the necessary secret keys can be done without using a secure channel (the secret keys for conventional cryptosystems are strings of a few hundreds of bits only and can then be encrypted by public key cryptosystems). Protocols specially devoted to key-exchange can also be used.

The objective of *error correcting codes* is to enable digital communication over a noisy channel in such a way that the errors in the transmission of bits can be detected[2] and corrected by the receiver. This aim is achieved by using an encoding algorithm which transforms the information before sending it over the channel. In the case of block coding[3], the original message is treated as a list of binary words (vectors) of the same length – say k – which are encoded into *codewords* of a larger length – say n. Thanks to this extension of the length, called *redundancy*, the decoding algorithm can correct the errors of transmission (if their number is, for each sent word, smaller than or equal to the so-called correction capacity of the code) and recover the correct message. The set of all possible codewords is called the *code*. Sending words of length n over the channel instead of words of length k slows down the transmission of information in the ratio of $\frac{k}{n}$. This ratio, called the *transmission rate*, must be as high as possible, to allow fast communication.



In both cryptographic and error correcting coding activities, *Boolean functions* (that is, functions from the vectorspace $\mathbb{F}_2^n$ of all binary vectors of

[2] If the code is used only to detect errors, then when an error is detected, the information must be requested and sent again in a so-called "automatic request" procedure.

[3] We shall not address convolutional coding here.

length n, to the finite field with two elements[4] $\mathbb{F}_2$) play roles:
- every code of length 2^n, for some positive integer n, can be interpreted as a set of Boolean functions, since every n-variable Boolean function can be represented by its truth table (an ordering of the set of binary vectors of length n being first chosen) and thus associated with a binary word of length 2^n, and *vice versa*; important codes (Reed-Muller, Kerdock codes) can be defined this way as sets of Boolean functions;
- the role of Boolean functions in conventional cryptography is even more important: cryptographic transformations (pseudo-random generators in stream ciphers, S-boxes in block ciphers) can be designed by appropriate composition of nonlinear Boolean functions.

In both frameworks, n is rarely large, in practice. The error correcting codes derived from n-variable Boolean functions have length 2^n; so, taking $n = 10$ already gives codes of length 1024. For reason of efficiency, the S-boxes used in most block ciphers are concatenations of sub S-boxes on at most 8 variables. In the case of stream ciphers, n was in general at most equal to 10 until recently. This has changed with the algebraic attacks (see [113, 117, 150] and see below) but the number of variables is now most often limited to 20.

Despite the fact that Boolean functions are currently used in cryptography and coding with low numbers of variables, determining and studying those Boolean functions satisfying the desired conditions (see Subection 4.1 below) is not feasible through an exhaustive computer investigation: the number $|\mathcal{BF}_n| = 2^{2^n}$ of n-variable Boolean functions is too large when $n \geq 6$. We give in table 1 below the values of this number for n ranging between 4 and 8.

n	4	5	6	7	8
$\lvert\mathcal{BF}_n\rvert$	2^{16}	2^{32}	2^{64}	2^{128}	2^{256}
$\approx$	$6 \cdot 10^4$	$4 \cdot 10^9$	10^{19}	10^{38}	10^{77}

Table 1: Number of n-variable Boolean functions

Assume that visiting an n-variable Boolean function, and determining whether it has the desired properties, needs one nano-second (10^{-9} seconds), then it would need millions of hours to visit all functions in 6 variables, and about one hundred billions times the age of the universe to visit all those in 7 variables. The number of 8-variable Boolean functions approximately equals the number of atoms in the whole universe! We see that trying to find

[4]Denoted by $\mathcal{B}$ is some chapters of the present collection.

functions satisfying the desired conditions by simply picking up functions at random is also impossible for these values of n, since visiting a non-negligible part of all Boolean functions in 7 or more variables is not feasible, even when parallelizing. The study of Boolean functions for constructing or studying codes or ciphers is essentially mathematical. But clever computer investigation is very useful to imagine or to test conjectures, and sometimes to generate interesting functions.

2 Generalities on Boolean functions

In this chapter and in the chapter "Vectorial Boolean Functions for Cryptography" which follows, the set $\{0,1\}$ will be most often endowed with the structure of field (and denoted by $\mathbb{F}_2$), and the set $\mathbb{F}_2^n$ of all binary vectors[5] of length n will be viewed as an $\mathbb{F}_2$-vectorspace. We shall denote simply by 0 the null vector in $\mathbb{F}_2^n$. The vectorspace $\mathbb{F}_2^n$ will sometimes be also endowed with the structure of field – the field $\mathbb{F}_{2^n}$ (also denoted by $GF(2^n)$); indeed, this field being an n-dimensional vectorspace over $\mathbb{F}_2$, each of its elements can be identified with the binary vector of length n of its coordinates relative to a fixed basis. The set of all Boolean functions $f : \mathbb{F}_2^n \to \mathbb{F}_2$ will be denoted as usual by $\mathcal{BF}_n$. The *Hamming weight* $w_H(x)$ of a binary vector $x \in \mathbb{F}_2^n$ being the number of its nonzero coordinates (*i.*e. the size of $\{i \in N/\ x_i \neq 0\}$ where N denotes the set $\{1, \cdots, n\}$, called the *support of the codeword*), the Hamming weight $w_H(f)$ of a Boolean function f on $\mathbb{F}_2^n$ is (also) the size of the *support of the function* , *i.e.* the set $\{x \in \mathbb{F}_2^n/\ f(x) \neq 0\}$. The *Hamming distance* $d_H(f,g)$ between two functions f and g is the size of the set $\{x \in \mathbb{F}_2^n/\ f(x) \neq g(x)\}$. Thus it equals $w_H(f \oplus g)$.

Note. Some additions of bits will be considered in $\mathbb{Z}$ (in characteristic 0) and denoted then by $+$, and some will be computed modulo 2 and denoted by $\oplus$. These two different notations will be necessary because some representations of Boolean functions will live in characteristic 2 and some representations of the same functions will live in characteristic 0. But the additions of elements of the finite field $\mathbb{F}_{2^n}$ will be denoted by $+$, as it is usual in mathematics. So, for simplicity (since $\mathbb{F}_2^n$ will often be identified with $\mathbb{F}_{2^n}$) and because there will be no ambiguity, we shall also denote by $+$ the addition of vectors of $\mathbb{F}_2^n$ when $n > 1$.

[5] Coders say "words"

2.1 Representation of Boolean functions

Among the classical representations of Boolean functions, the one which is most usually used in cryptography and coding is the n-variable polynomial representation over $\mathbb{F}_2$, of the form

$$f(x) = \bigoplus_{I \in \mathcal{P}(N)} a_I \left(\prod_{i \in I} x_i \right) = \bigoplus_{I \in \mathcal{P}(N)} a_I \, x^I, \tag{1}$$

where $\mathcal{P}(N)$ denotes the power set of $N = \{1, \cdots, n\}$. Every coordinate x_i appears in this polynomial with exponents at most 1, because every bit in $\mathbb{F}_2$ equals its own square. This representation belongs to $\mathbb{F}_2[x_1, \cdots, x_n]/(x_1^2 \oplus x_1, \cdots, x_n^2 \oplus x_n)$. It is called the *Algebraic Normal Form* (in brief the ANF).

Example: let us consider the function f whose truth-table is

x_1	x_2	x_3	$f(x)$
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

It is the sum (modulo 2 or not, no matter) of the *atomic functions* f_1, f_2 and f_3 whose truth-tables are

x_1	x_2	x_3	$f_1(x)$	$f_2(x)$	$f_3(x)$
0	0	0	0	0	0
0	0	1	1	0	0
0	1	0	0	0	0
0	1	1	0	0	0
1	0	0	0	0	0
1	0	1	0	1	0
1	1	0	0	0	0
1	1	1	0	0	1

The function $f_1(x)$ takes value 1 if and only if $1 \oplus x_1 = 1$, $1 \oplus x_2 = 1$ and $x_3 = 1$, that is if and only if $(1 \oplus x_1)(1 \oplus x_2)\, x_3 = 1$. Thus the ANF of f_1

can be obtained by expanding the product $(1 \oplus x_1)(1 \oplus x_2)\, x_3$. After similar observations on f_2 and f_3, we see that the ANF of f equals $(1 \oplus x_1)(1 \oplus x_2)\, x_3 \oplus x_1(1 \oplus x_2)\, x_3 \oplus x_1x_2x_3 = x_1x_2x_3 \oplus x_2x_3 \oplus x_3$. □

Another possible representation of this same ANF uses an indexation by means of vectors of $\mathbb{F}_2^n$ instead of subsets of N; if, for any such vector u, we denote by a_u what is denoted by $a_{supp(u)}$ in Relation (1) (where $supp(u)$ denotes the support of u), we have the equivalent representation:

$$f(x) = \bigoplus_{u \in \mathbb{F}_2^n} a_u \left(\prod_{j=1}^{n} {x_j}^{u_j} \right).$$

The monomial $\prod_{j=1}^{n} {x_j}^{u_j}$ is often denoted by x^u.

Existence and uniqueness of the ANF By applying the Lagrange interpolation method described in the example above, it is a simple matter to show the existence of the ANF of every Boolean function. This implies that the mapping, from every polynomial $P \in \mathbb{F}_2[x_1, \cdots, x_n]/(x_1^2 \oplus x_1, \cdots, x_n^2 \oplus x_n)$ to the corresponding function $x \in \mathbb{F}_2^n \mapsto P(x)$, is onto $\mathcal{BF}_n$. Since the size of $\mathcal{BF}_n$ equals the size of $\mathbb{F}_2[x_1, \cdots, x_n]/(x_1^2 \oplus x_1, \cdots, x_n^2 \oplus x_n)$, this correspondence is one to one[6]. But more can be said.

Relationship between a Boolean function and its ANF The product $x^I = \prod_{i \in I} x_i$ is nonzero if and only if x_i is nonzero (*i*.e. equals 1) for every $i \in I$, that is, if I is included in the support of x; hence, the Boolean function $f(x) = \bigoplus_{I \in \mathcal{P}(N)} a_I\, x^I$ takes value

$$f(x) = \bigoplus_{I \subseteq supp(x)} a_I, \tag{2}$$

where $supp(x)$ denotes the support of x. If we use the notation $f(x) = \bigoplus_{u \in \mathbb{F}_2^n} a_u x^u$, we obtain the relation $f(x) = \bigoplus_{u \preceq x} a_u$, where $u \preceq x$ means that $supp(u) \subseteq supp(x)$ (we say that u is *covered* by x). A Boolean function f° can be associated to the ANF of f: for every $x \in \mathbb{F}_2^n$, we set $f^\circ(x) = a_{supp(x)}$, that is, with the notation $f(x) = \bigoplus_{u \in \mathbb{F}_2^n} a_u x^u$: $f^\circ(u) = a_u$. Relation (2) shows that f is the image of f° by the so-called *binary Möbius*

[6] Another argument is that this mapping is a linear mapping from a vectorspace over $\mathbb{F}_2$ of dimension 2^n to a vectorspace of the same dimension.

transform.
The converse is also true:

Proposition 1 *Let f be a Boolean function on $\mathbb{F}_2^n$ and let $\bigoplus_{I\in\mathcal{P}(N)} a_I\, x^I$ be its ANF. We have:*

$$\forall I \in \mathcal{P}(N),\ a_I = \bigoplus_{x\in\mathbb{F}_2^n/\ supp(x)\subseteq I} f(x). \tag{3}$$

Proof. Let us denote $\bigoplus_{x\in\mathbb{F}_2^n/\ supp(x)\subseteq I} f(x)$ by b_I and consider the function $g(x) = \bigoplus_{I\in\mathcal{P}(N)} b_I\, x^I$. We have

$$g(x) = \bigoplus_{I\subseteq supp(x)} b_I = \bigoplus_{I\subseteq supp(x)} \left(\bigoplus_{y\in\mathbb{F}_2^n/\ supp(y)\subseteq I} f(y) \right)$$

and thus

$$g(x) = \bigoplus_{y\in\mathbb{F}_2^n} f(y) \left(\bigoplus_{I\in\mathcal{P}(N)/\ supp(y)\subseteq I\subseteq supp(x)} 1 \right).$$

The sum $\bigoplus_{I\in\mathcal{P}(N)/\ supp(y)\subseteq I\subseteq supp(x)} 1$ is null if $y \neq x$, since the set $\{I \in \mathcal{P}(N)/\ supp(y) \subseteq I \subseteq supp(x)\}$ contains $2^{w_H(x)-w_H(y)}$ elements if $supp(y) \subseteq supp(x)$, and none otherwise. Hence, $g = f$ and, by uniqueness of the ANF, $b_I = a_I$ for every I. □

Algorithm There exists a simple divide-and-conquer butterfly algorithm to compute the ANF from the truth-table (or *vice-versa*), that we can call the *Fast Möbius Transform*. For every $u = (u_1, \cdots, u_n) \in \mathbb{F}_2^n$, the coefficient a_u of x^u in the ANF of f equals

$$\bigoplus_{(x_1,\cdots,x_{n-1})\preceq(u_1,\cdots,u_{n-1})} [f(x_1, \cdots, x_{n-1}, 0)] \quad \text{if } u_n = 0 \text{ and}$$
$$\bigoplus_{(x_1,\cdots,x_{n-1})\preceq(u_1,\cdots,u_{n-1})} [f(x_1, \cdots, x_{n-1}, 0) \oplus f(x_1, \cdots, x_{n-1}, 1)] \text{ if } u_n = 1.$$

Hence if, in the truth-table of f, the binary vectors are ordered in lexicographic order, with the bit of higher weight on the right, the table of the ANF equals the concatenation of the ANFs of the $(n-1)$-variable functions $f(x_1, \cdots, x_{n-1}, 0)$ and $f(x_1, \cdots, x_{n-1}, 0) \oplus f(x_1, \cdots, x_{n-1}, 1)$. We deduce the following algorithm:

1. write the truth-table of f, in which the binary vectors of length n are in lexicographic order as decribed above;

2. let f_0 and f_1 be the restrictions of f to $\mathbb{F}_2^{n-1} \times \{0\}$ and $\mathbb{F}_2^{n-1} \times \{1\}$, respectively[7]; replace the values of f_1 by those of $f_0 \oplus f_1$;

3. apply recursively step 2, separately to the functions now obtained in the places of f_0 and f_1.

When the algorithm ends (*i*.e. when it arrives to functions in one variable each), the global table gives the values of the ANF of f. The complexity of this algorithm is of $n\,2^n$ XORs.

Remark.
The algorithm works the same if the vectors are ordered in standard lexicographic order, with the bit of higher weight on the left (indeed, this corresponds to applying it to $f(x_n, x_{n-1}, \cdots, x_1)$).

The degree of the ANF is denoted by $d^\circ f$ and is called the *algebraic degree* of the function (this makes sense thanks to the existence and uniqueness of the ANF): $d^\circ f = \max\{|I| / \, a_I \neq 0\}$, where $|I|$ denotes the size of I. Some authors also call it the nonlinear order of f. According to Relation (3), we have:

Proposition 2 *The algebraic degree $d^\circ f$ of any n-variable Boolean function f equals the maximum dimension of the subspaces $\{x \in \mathbb{F}_2^n / \, supp(x) \subseteq I\}$ on which f takes value 1 an odd number of times.*

The algebraic degree is an *affine invariant* (it is invariant under the action of the general affine group): for every affine isomorphism $L : \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{F}_2^n \mapsto M \times \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \oplus \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{F}_2^n$ (where M is a nonsingular $n \times n$ matrix

[7]The truth-table of f_0 (resp. f_1) corresponds to the upper (resp. lower) half of the table of f.

over $\mathbb{F}_2$), we have $d^\circ(f \circ L) = d^\circ f$. Indeed, the composition by L clearly cannot increase the algebraic degree, since the coordinates of $L(x)$ have degree 1. Hence we have $d^\circ(f \circ L) \leq d^\circ f$ (this inequality is more generally valid for every affine homomorphism). And applying this inequality to $f \circ L$ in the place of f and to L^{-1} in the place of L shows the inverse inequality.
Two functions f and $f \circ L$ where L is an $\mathbb{F}_2$-linear automorphism of $\mathbb{F}_2^n$ (in the case case $a_1 = a_2 = \cdots = a_n = 0$ above) will be called *linearly equivalent* and two functions f and $f \circ L$, where L is an affine automorphism of $\mathbb{F}_2^n$, will be called *affinely equivalent*.
The algebraic degree being an affine invariant, Proposition 2 implies that it also equals the maximum dimension of all the affine subspaces of $\mathbb{F}_2^n$ on which f takes value 1 an odd number of times.
It is shown in [297] that, for every nonzero n-variable Boolean function f, denoting by g the binary Möbius transform of f, we have $d^\circ f + d^\circ g \geq n$. This same paper deduces characterizations and constructions of the functions which are equal to their binary Möbius transform, called *coincident functions*.

Remarks.
1. Every atomic function has algebraic degree n, since its ANF equals $(x_1 \oplus \epsilon_1)(x_2 \oplus \epsilon_2) \cdots (x_n \oplus \epsilon_n)$, where $\epsilon_i \in \mathbb{F}_2$. Thus, a Boolean function f has algebraic degree n if and only if, in its decomposition as a sum of atomic functions, the number of these atomic functions is odd, that is, if and only if $w_H(f)$ is odd. This property will have an important consequence on the Reed-Muller codes and it will be also useful in Section 3.
2. If we know that the algebraic degree of an n-variable Boolean function f is bounded above by $d < n$, then the whole function can be recovered from some of its restrictions (*i.e.*, a unique function corresponds to this *partially defined* Boolean function). Precisely, according to the existence and uniqueness of the ANF, the knowledge of the restriction $f_{|E}$ of the Boolean function f (of algebraic degree at most d) to a set E implies the knowledge of the whole function if and only if the system of the equations $f(x) = \bigoplus_{I \in \mathcal{P}(N)/\,|I| \leq d} a_I\, x^I$, with indeterminates $a_I \in \mathbb{F}_2$, and where x ranges over E (this makes $|E|$ equations), has a unique solution[8]. This happens with the set E_d of all words of Hamming weights smaller than or equal to d, since Relation (3) gives the value of a_I (when $I \in \mathcal{P}(N)$ has size

[8]Taking $f_{|E}$ null leads to determining the so-called annihilators of the *indicator* of E (the function 1_E, also called characteristic function of E, defined by $1_E(x) = 1$ if $x \in E$ and $1_E(x) = 0$ otherwise); this is the core analysis of Boolean functions from the viewpoint of algebraic attacks, see Subsection 4.1.

$|I| \leq d$). Notice that Relation (2) allows then to express the value of $f(x)$ for every $x \in \mathbb{F}_2^n$ by means of the values taken by f at all words of Hamming weights smaller than or equal to d. We have (using the notation a_u instead of a_I, see above):

$$\begin{aligned} f(x) &= \bigoplus_{u \preceq x} a_u = \bigoplus_{\substack{u \preceq x \\ u \in E_d}} a_u = \bigoplus_{\substack{y \preceq x \\ y \in E_d}} f(y) \, |\{u \in E_d \,/\, y \preceq u \preceq x| \\ &= \bigoplus_{\substack{y \preceq x \\ y \in E_d}} f(y) \left[\left[\sum_{i=0}^{d-w_H(y)} \binom{w_H(x) - w_H(y)}{i} \right] \pmod 2 \right]. \end{aligned}$$

More generally, the whole function f can be recovered from $f_{|E}$ for every set E affinely equivalent to E_d, according to the affine invariance of the algebraic degree. This also generalizes to "pseudo-Boolean" (that is, real-valued) functions, if we consider the numerical degree (see below) instead of the the algebraic degree, *cf.* [350]. □

The simplest functions, from the viewpoint of the ANF, are those Boolean functions of algebraic degrees at most 1, called *affine functions*:

$$f(x) = a_1 \, x_1 \oplus \cdots \oplus a_n \, x_n \oplus a_0.$$

They are the sums of linear and constant functions. Denoting by $a \cdot x$ the usual *inner product* $a \cdot x = a_1 \, x_1 \oplus \cdots \oplus a_n \, x_n$ in $\mathbb{F}_2^n$, or any other inner product (symmetric and such that, for every $a \neq 0$, the function $x \to a \cdot x$ is a nonzero linear form on $\mathbb{F}_2^n$), the general form of an n-variable affine function is $a \cdot x \oplus a_0$ (with $a \in \mathbb{F}_2^n$; $a_0 \in \mathbb{F}_2$).
Affine functions play an important role in coding (they are involved in the definition of the Reed-Muller code of order 1, see Subsection 3.1) and in cryptography (the Boolean functions used as "nonlinear functions" in cryptosystems must behave as differently as possible from affine functions, see Subsection 4.1).

Trace representation(s) A second kind of representation plays an important role in sequence theory, and is also used for defining and studying Boolean functions. It leads to the construction of the Kerdock codes (see Subsection 6.10). Recall that, for every n, there exists a (unique up to isomorphism) field $\mathbb{F}_{2^n}$ (also denoted by $GF(2^n)$) of order 2^n (see [248]). The vectorspace $\mathbb{F}_2^n$ can be endowed with the structure of this field $\mathbb{F}_{2^n}$. Indeed, we know that $\mathbb{F}_{2^n}$ has the structure of an n-dimensional $\mathbb{F}_2$-vectorspace; if

we choose an $\mathbb{F}_2$-basis $(\alpha_1, \cdots, \alpha_n)$ of this vectorspace, then every element $x \in \mathbb{F}_2^n$ can be identified with $x_1\,\alpha_1 + \cdots + x_n\,\alpha_n \in \mathbb{F}_{2^n}$. We shall still denote by x this element of the field.

1. It is shown in the chapter "Vectorial Boolean Functions for Cryptography" (see another proof below) that every mapping from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$ admits a (unique) representation as a polynomial

$$f(x) = \sum_{i=0}^{2^n-1} \delta_i x^i \tag{4}$$

over $\mathbb{F}_{2^n}$ in one variable and of (univariate) degree at most $2^n - 1$. Any Boolean function on $\mathbb{F}_{2^n}$ is a particular case of a vectorial function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^n}$ (since $\mathbb{F}_2$ is a subfield of $\mathbb{F}_{2^n}$) and admits therefore such a unique representation, that we shall call the *univariate representation* of f. For every $u, v \in \mathbb{F}_{2^n}$ we have $(u+v)^2 = u^2 + v^2$ and $u^{2^n} = u$. A univariate polynomial $\sum_{i=0}^{2^n-1} \delta_i x^i$, $\delta_i \in \mathbb{F}_{2^n}$, is then the univariate representation of a Boolean function if and only if $\left(\sum_{i=0}^{2^n-1} \delta_i x^i\right)^2 = \sum_{i=0}^{2^n-1} \delta_i^2 x^{2i} = \sum_{i=0}^{2^n-1} \delta_i x^i$ [mod $x^{2^n}+x$], that is, δ_0, $\delta_{2^n-1} \in \mathbb{F}_2$ and, for every $i = 1, \cdots, 2^n-2$, $\delta_{2i} = \delta_i^2$, where the index $2i$ is taken mod $2^n - 1$.

2. The function defined on $\mathbb{F}_{2^n}$ by $tr_n(u) = u + u^2 + u^{2^2} + \cdots + u^{2^{n-1}}$ is $\mathbb{F}_2$-linear and satisfies $(tr_n(u))^2 = tr_n(u^2) = tr_n(u)$; it is therefore valued in $\mathbb{F}_2$. This function is called the *trace function* from $\mathbb{F}_{2^n}$ to its prime field $\mathbb{F}_2$ or the absolute trace function on $\mathbb{F}_{2^n}$. The function $(u, v) \mapsto tr_n(u\,v)$ is an inner product in $\mathbb{F}_{2^n}$ (that is, it is symmetric and for every $v \neq 0$, the function $u \to tr_n(u\,v)$ is a nonzero linear form on $\mathbb{F}_{2^n}$). Every Boolean function can be written in the form $f(x) = tr_n(F(x))$ where F is a mapping from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$ (an example of such mapping F is defined by $F(x) = \lambda\, f(x)$ where $tr_n(\lambda) = 1$ and $f(x)$ is the univariate representation). Thus, every Boolean function can be also represented in the form

$$tr_n\left(\sum_{i=0}^{2^n-1} \beta_i\, x^i\right), \tag{5}$$

where $\beta_i \in \mathbb{F}_{2^n}$. Such a representation is not unique. Now, thanks to the fact that $tr_n(u^2) = tr_n(u)$ for every $u \in \mathbb{F}_{2^n}$, we can restrict the exponents i with nonzero coefficients β_i so that there is at most one such exponent in each *cyclotomic class* $\{i \times 2^j\,[\bmod (2^n-1)]\,;\, j \in \mathbb{N}\}$ of 2 modulo $2^n - 1$ (but this still does not make the representation unique). We shall call this expression the *absolute trace representation* of f.

3. We come back to the univariate representation. Let us see how it can be obtained from the truth table of the function and represented in a convenient way by using the notation tr_n. Denoting by α a *primitive element* of the field $\mathbb{F}_{2^n}$ (that is, an element such that $\mathbb{F}_{2^n} = \{0, 1, \alpha, \alpha^2, \cdots, \alpha^{2^n-2}\}$, which always exists [248]), the *Mattson-Solomon polynomial* of the vector $(f(1), f(\alpha), f(\alpha^2), \cdots, f(\alpha^{2^n-2}))$ is the polynomial [258]

$$A(x) = \sum_{j=1}^{2^n-1} A_j x^{2^n-1-j} = \sum_{j=0}^{2^n-2} A_{-j} x^j$$

with:

$$A_j = \sum_{k=0}^{2^n-2} f(\alpha^k)\alpha^{kj}.$$

Note that the Mattson Solomon transform is a discrete Fourier transform. We have, for every $0 \leq i \leq 2^n - 2$:

$$A(\alpha^i) = \sum_{j=1}^{2^n-1} A_j \alpha^{-ij} = \sum_{j=1}^{2^n-1} \sum_{k=0}^{2^n-2} f(\alpha^k)\alpha^{(k-i)j} = f(\alpha^i)$$

(since $\sum_{j=1}^{2^n-1} \alpha^{(k-i)j} = \sum_{j=0}^{2^n-2} \alpha^{(k-i)j} = \frac{\alpha^{(k-i)(2^n-1)}+1}{\alpha^{k-i}+1}$ equals 0 if $1 \leq k \neq i \leq 2^n - 2$), and A is therefore the univariate representation of f, if $f(0) = A_0 = \sum_{i=0}^{2^n-2} f(\alpha^i)$ (note that this works also for functions from $\mathbb{F}_{2^n}$ to $\mathbb{F}_{2^n}$) that is, if f has even weight, *i.e.* has algebraic degree strictly less than n. Otherwise, we have $f(x) = A(x) + 1 + x^{2^n-1}$, since $1 + x^{2^n-1}$ takes value 1 at 0 and 0 at every nonzero element of $\mathbb{F}_{2^n}$.
Note that $A_{2j} = A_j^2$. Denoting by $\Gamma(n)$ the set obtained by choosing one element in each cyclotomic class of 2 modulo $2^n - 1$ (the most usual choice for k is the smallest element in its cyclotomic class, called the *coset leader* of the class), this allows representing $f(x)$ in the form

$$\sum_{j \in \Gamma(n)} tr_{n_j}(A_{-j}x^j) + \epsilon(1 + x^{2^n-1}), \tag{6}$$

where $\epsilon = w_H(f)$ [mod 2] and where n_j is the size of the cyclotomic class containing j. Note that, for every $j \in \Gamma(n)$ and every $x \in \mathbb{F}_{2^n}$, we have $A_j \in \mathbb{F}_{2^{n_j}}$ (since $A_j^{2^{n_j}} = A_j$) and $x^j \in \mathbb{F}_{2^{n_j}}$ as well. We shall call this expression the *trace representation* of f. Obviously, it is nothing more than an alternate expression for the univariate representation. For this reason, it is unique (if we restrict the coefficient of x^j to live in $\mathbb{F}_{2^{n_j}}$). But it is

useful to distinguish the different expressions by different names. We shall call globally "trace representations" the three expressions (4), (5) and (6). Trace representations and the algebraic normal form are closely related. Let us see how the ANF can be obtained from the univariate representation: we express x in the form $\sum_{i=1}^{n} x_i\alpha_i$, where $(\alpha_1, \cdots, \alpha_n)$ is a basis of the $\mathbb{F}_2$-vectorspace $\mathbb{F}_{2^n}$. Recall that, for every $j \in \mathbb{Z}/(2^n-1)\mathbb{Z}$, the *binary expansion* of j has the form $\sum_{s\in E} 2^s$, where $E \subseteq \{0, 1, \cdots, n-1\}$. The size of E is often called the *2-weight* of j and written $w_2(j)$. We write more conveniently the binary expansion of j in the form: $\sum_{s=0}^{n-1} j_s 2^s$, $j_s \in \{0, 1\}$. We have then:

$$\begin{aligned} f(x) &= \sum_{j=0}^{2^n-1} \delta_j \left(\sum_{i=1}^{n} x_i\alpha_i\right)^j \\ &= \sum_{j=0}^{2^n-1} \delta_j \left(\sum_{i=1}^{n} x_i\alpha_i\right)^{\sum_{s=0}^{n-1} j_s 2^s} \\ &= \sum_{j=0}^{2^n-1} \delta_j \prod_{s=0}^{n-1} \left(\sum_{i=1}^{n} x_i\alpha_i^{2^s}\right)^{j_s}. \end{aligned}$$

Expanding these last products and simplifying gives the ANF of f.
Function f has then algebraic degree $\max_{j=0,\cdots,2^n-1/\,\delta_j\neq 0} w_2(j)$. Indeed, according to the above equalities, its algebraic degree is clearly bounded above by this number, and it can not be strictly smaller, because the number of Boolean n-variable functions of algebraic degrees at most d equals the number of the polynomials $\sum_{j=0}^{2^n-1} \delta_j x^j$ such that δ_0, $\delta_{2^n-1} \in \mathbb{F}_2$ and $\delta_{2j} = \delta_j^2 \in \mathbb{F}_{2^n}$ for every $j = 1, \cdots, 2^n-2$ and $\max_{j=0,\cdots,2^n-1/\,\delta_j\neq 0} w_2(j) \leq d$.
We have also:

Proposition 3 *[51] Let a be any element of $\mathbb{F}_{2^n}$ and k any integer [mod 2^n-1]. If $f(x) = tr_n(ax^k)$ is not the null function, then it has algebraic degree $w_2(k)$.*

Proof. Let n_k be again the size of the cyclotomic class containing k. Then the univariate representation of $f(x)$ equals

$$\left(a + a^{2^{n_k}} + a^{2^{2n_k}} + \cdots + a^{2^{n-n_k}}\right) x^k + \left(a + a^{2^{n_k}} + a^{2^{2n_k}} + \cdots + a^{2^{n-n_k}}\right)^2 x^{2k}$$

$$+ \cdots + \left(a + a^{2^{n_k}} + a^{2^{2n_k}} + \cdots + a^{2^{n-n_k}}\right)^{2^{n_k-1}} x^{2^{n_k-1}k}.$$

All the exponents of x have 2-weight $w_2(k)$ and their coefficients are nonzero if and only if f is not null. □

Remark. Another (more complex) way of showing Proposition 3 is used in [51] as follows: let $r = w_2(k)$; we consider the r-linear function ϕ over the field $\mathbb{F}_{2^n}$ whose value at $(x_1, ..., x_r)$ equals the sum of the images by f of all the 2^r possible linear combinations of the x_j's. Then $\phi(x_1, ..., x_r)$ equals the sum, for all bijective mappings σ from $\{1, \cdots, r\}$ onto E (where $k = \sum_{s \in E} 2^s$) of $tr_n(a \prod_{j=1}^{r} x_j^{2^{\sigma(j)}})$. Proving that f has degree r is equivalent to proving that ϕ is not null, and it can be shown that if ϕ is null, then f is null.

The representation over the reals has proved itself to be useful for characterizing several cryptographic criteria [63, 87, 88] (see Sections 6 and 7). It represents Boolean functions, and more generally real-valued functions on $\mathbb{F}_2^n$ (that are called n-variable *pseudo-Boolean functions*) by elements of $\mathbb{R}[x_1, \cdots, x_n]/(x_1^2 - x_1, \cdots, x_n^2 - x_n)$ (or of $\mathbb{Z}[x_1, \cdots, x_n]/(x_1^2 - x_1, \cdots, x_n^2 - x_n)$ for integer-valued functions). We shall call it the *Numerical Normal Form* (NNF).
The existence of this representation for every pseudo-Boolean function is easy to show with the same arguments as for the ANFs of Boolean functions (writing $1 - x_i$ instead of $1 \oplus x_i$). The linear mapping from every element of the 2^n-th dimensional $\mathbb{R}$-vectorspace $\mathbb{R}[x_1, \cdots, x_n]/(x_1^2 - x_1, \cdots, x_n^2 - x_n)$ to the corresponding pseudo-Boolean function on $\mathbb{F}_2^n$ being onto, it is therefore one to one (the $\mathbb{R}$-vectorspace of pseudo-Boolean functions on $\mathbb{F}_2^n$ having also dimension 2^n). We deduce the uniqueness of the NNF.
We call the degree of the NNF of a function its *numerical degree*. Since the ANF is the mod 2 version of the NNF, the numerical degree is always bounded below by the algebraic degree. It is shown in [286] that, if a Boolean function f has no ineffective variable (*i.e.* if it actually depends on each of its variables), then the numerical degree of f is greater than or equal to $\log_2 n - \log_2 \log_2 n$.
The numerical degree is not an affine invariant. But the NNF leads to an affine invariant (see a proof of this fact in [88]; see also [191]) which is more discriminant than the algebraic degree:

Definition 1 *Let f be a Boolean function on $\mathbb{F}_2^n$. We call* generalized degree *of f the sequence $(d_i)_{i \geq 1}$ defined as follows:*
for every $i \geq 1$, d_i is the smallest integer $d > d_{i-1}$ (if $i > 1$) such that, for every multi-index I of size strictly greater than d, the coefficient λ_I of x^I in

the NNF of f is a multiple of 2^i.

Example: the generalized degree of any nonzero affine function is the sequence of all positive integers.

Similarly as for the ANF, a (pseudo-) Boolean function $f(x) = \sum_{I \in \mathcal{P}(N)} \lambda_I \, x^I$ takes value:

$$f(x) = \sum_{I \subseteq supp(x)} \lambda_I. \tag{7}$$

But, contrary to what we observed for the ANF, the reverse formula is not identical to the direct formula:

Proposition 4 *Let f be a pseudo-Boolean function on $\mathbb{F}_2^n$ and let its NNF be $\sum_{I \in \mathcal{P}(N)} \lambda_I \, x^I$. Then:*

$$\forall I \in \mathcal{P}(N), \, \lambda_I = (-1)^{|I|} \sum_{x \in \mathbb{F}_2^n \,|\, supp(x) \subseteq I} (-1)^{w_H(x)} f(x). \tag{8}$$

Thus, function f and its NNF are related through the *Möbius transform over integers*.

Proof. Let us denote the number $(-1)^{|I|} \sum_{x \in \mathbb{F}_2^n \,|\, supp(x) \subseteq I} (-1)^{w_H(x)} f(x)$ by μ_I and consider the function $g(x) = \sum_{I \in \mathcal{P}(N)} \mu_I \, x^I$. We have

$$g(x) = \sum_{I \subseteq supp(x)} \mu_I = \sum_{I \subseteq supp(x)} \left((-1)^{|I|} \sum_{y \in \mathbb{F}_2^n \,|\, supp(y) \subseteq I} (-1)^{w_H(y)} f(y) \right)$$

and thus

$$g(x) = \sum_{y \in \mathbb{F}_2^n} (-1)^{w_H(y)} f(y) \left(\sum_{I \in \mathcal{P}(N) / \; supp(y) \subseteq I \subseteq supp(x)} (-1)^{|I|} \right).$$

The sum $\sum_{I \in \mathcal{P}(N) / \; supp(y) \subseteq I \subseteq supp(x)} (-1)^{|I|}$ is null if $supp(y) \not\subseteq supp(x)$. It is also null if $supp(y)$ is included in $supp(x)$, but different. Indeed, denoting $|I| - w_H(y)$ by i, it equals $\pm \sum_{i=0}^{w_H(x)-w_H(y)} \binom{w_H(x)-w_H(y)}{i} (-1)^i = \pm(1-1)^{w_H(x)-w_H(y)} = 0$. Hence, $g = f$ and, by uniqueness of the NNF, we have $\mu_I = \lambda_I$ for every I. □

We have seen that the ANF of any Boolean function can be deduced from its NNF by reducing it modulo 2. Conversely, the NNF can be deduced from the ANF since we have

$$f(x) = \bigoplus_{I \in \mathcal{P}(N)} a_I \, x^I \iff (-1)^{f(x)} = \prod_{I \in \mathcal{P}(N)} (-1)^{a_I \, x^I}$$
$$\iff 1 - 2 \, f(x) = \prod_{I \in \mathcal{P}(N)} (1 - 2 \, a_I \, x^I).$$

Expanding this last equality gives the NNF of $f(x)$ and we have [87]:

$$\lambda_I = \sum_{k=1}^{2^n} (-2)^{k-1} \sum_{\substack{\{I_1, \ldots, I_k\} \mid \\ I_1 \cup \cdots \cup I_k = I}} a_{I_1} \cdots a_{I_k}, \tag{9}$$

where "$\{I_1, \ldots, I_k\} \mid I_1 \cup \cdots \cup I_k = I$" means that the multi-indices $I_1, \ldots, I_k$ are all distinct, in indefinite order, and that their union equals I.
A polynomial $P(x) = \sum_{J \in \mathcal{P}(N)} \lambda_J \, x^J$, with real coefficients, is the NNF of some Boolean function if and only if we have $P^2(x) = P(x)$, for every $x \in \mathbb{F}_2^n$ (which is equivalent to $P = P^2$ in $\mathbb{R}[x_1, \cdots, x_n]/(x_1^2 - x_1, \cdots, x_n^2 - x_n)$), or equivalently, denoting $supp(x)$ by I:

$$\forall I \in \mathcal{P}(N), \ \left(\sum_{J \subseteq I} \lambda_J \right)^2 = \sum_{J \subseteq I} \lambda_J. \tag{10}$$

Remark.
Imagine that we want to generate a random Boolean function through its NNF (this can be useful, since we will see below that the main cryptographic criteria, on Boolean functions, can be characterized, in simple ways, through their NNFs). Assume that we have already chosen the values λ_J for every $J \subseteq I$ (where $I \in \mathcal{P}(N)$ is some multi-index) except for I itself. Let us denote the sum $\sum_{J \subseteq I \mid J \neq I} \lambda_J$ by μ. Relation (10) gives $(\lambda_I + \mu)^2 = \lambda_I + \mu$. This equation of degree 2 has two solutions (it has same discriminant as the equation ${\lambda_I}^2 = \lambda_I$, that is 1). One solution corresponds to the choice $P(x) = 0$ (where $I = supp(x)$) and the other one corresponds to the choice $P(x) = 1$.
□
Thus, verifying that a polynomial $P(x) = \sum_{I \in \mathcal{P}(N)} \lambda_I \, x^I$ with real coefficients represents a Boolean function can be done by checking 2^n relations. But it can also be done by verifying a simple condition on P and checking a single equation.

Proposition 5 *Any polynomial* $P \in \mathbb{R}[x_1, \cdots, x_n]/(x_1^2 - x_1, \cdots, x_n^2 - x_n)$ *is the NNF of an integer-valued function if and only if all of its coefficients are integers. Assuming that this condition is satisfied,* P *is the NNF of a Boolean function if and only if:* $\sum_{x \in \mathbb{F}_2^n} P^2(x) = \sum_{x \in \mathbb{F}_2^n} P(x)$.

Proof. The first assertion is a direct consequence of Relations (7) and (8). If all the coefficients of P are integers, then we have $P^2(x) \geq P(x)$ for every x; this implies that the 2^n equalities, expressing that the corresponding function is Boolean, can be reduced to the single one $\sum_{x \in \mathbb{F}_2^n} P^2(x) = \sum_{x \in \mathbb{F}_2^n} P(x)$. □

The translation of this characterization in terms of the coefficients of P is given in Relation (32) below.

2.2 The discrete Fourier transform on pseudo-Boolean and on Boolean functions

Almost all the characteristics needed for Boolean functions in cryptography and for sets of Boolean functions in coding can be expressed by means of the weights of some related Boolean functions (of the form $f \oplus \ell$, where ℓ is affine, or of the form $D_a f(x) = f(x) \oplus f(x + a)$). In this framework, the *discrete Fourier transform* is a very efficient tool: for a given Boolean function f, the knowledge of the discrete Fourier transform of f is equivalent with the knowledge of the weights of all the functions $f \oplus \ell$, where ℓ is linear (or affine). Also called Hadamard transform, the discrete Fourier transform is the linear mapping which maps any pseudo-Boolean function φ on $\mathbb{F}_2^n$ to the function $\widehat{\varphi}$ defined on $\mathbb{F}_2^n$ by

$$\widehat{\varphi}(u) = \sum_{x \in \mathbb{F}_2^n} \varphi(x) \, (-1)^{x \cdot u} \tag{11}$$

where $x \cdot u$ is some chosen inner product (for instance the usual inner product $x \cdot u = x_1 \, u_1 \oplus \cdots \oplus x_n \, u_n$).

Algorithm There exists a simple divide-and-conquer butterfly algorithm to compute $\widehat{\varphi}$, called the *Fast Fourier Transform* (FFT). For every $a = (a_1, \cdots, a_{n-1}) \in \mathbb{F}_2^{n-1}$ and every $a_n \in \mathbb{F}_2$, the number $\widehat{\varphi}(a_1, \cdots, a_n)$ equals

$$\sum_{x=(x_1, \cdots, x_{n-1}) \in \mathbb{F}_2^{n-1}} (-1)^{a \cdot x} \left[\varphi(x_1, \cdots, x_{n-1}, 0) + (-1)^{a_n} \varphi(x_1, \cdots, x_{n-1}, 1) \right].$$

Hence, if in the tables of values of the functions, the vectors are ordered in lexicographic order with the bit of highest weight on the right, the table of $\widehat{\varphi}$ equals the concatenation of those of the discrete Fourier transforms of the $(n-1)$-variable functions $\psi_0(x) = \varphi(x_1, \cdots, x_{n-1}, 0) + \varphi(x_1, \cdots, x_{n-1}, 1)$ and $\psi_1(x) = \varphi(x_1, \cdots, x_{n-1}, 0) - \varphi(x_1, \cdots, x_{n-1}, 1)$. We deduce the following algorithm:

1. write the table of the values of φ (its truth-table if φ is Boolean), in which the binary vectors of length n are in lexicographic order as decribed above;

2. let φ_0 be the restriction of φ to $\mathbb{F}_2^{n-1} \times \{0\}$ and φ_1 the restriction of φ to $\mathbb{F}_2^{n-1} \times \{1\}$[9]; replace the values of φ_0 by those of $\varphi_0 + \varphi_1$ and those of φ_1 by those of $\varphi_0 - \varphi_1$;

3. apply recursively step 2, separately to the functions now obtained in the places of φ_0 and φ_1.

When the algorithm ends (*i*.e. when it arrives to functions in one variable each), the global table gives the values of $\widehat{\varphi}$. The complexity of this algorithm is of $n\,2^n$ additions/substractions.

Application to Boolean functions For a given Boolean function f, the discrete Fourier transform can be applied to f itself, viewed as a function valued in $\{0, 1\} \subset \mathbb{Z}$. We denote by $\widehat{f}$ the corresponding discrete Fourier transform of f. Notice that $\widehat{f}(0)$ equals the Hamming weight of f. Thus, the Hamming distance $d_H(f, g) = |\{x \in \mathbb{F}_2^n / f(x) \neq g(x)\}| = w_H(f \oplus g)$ between two functions f and g equals $\widehat{f \oplus g}(0)$.
The discrete Fourier transform can also be applied to the pseudo-Boolean function $\mathrm{f}_\chi(x) = (-1)^{f(x)}$ (often called the *sign function*[10]) instead of f itself. We have

$$\widehat{\mathrm{f}_\chi}(u) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x) \oplus x \cdot u}.$$

[9]The table of values of φ_0 (resp. φ_1) corresponds to the upper (resp. lower) half of the table of φ.

[10]The symbol χ is used here because the sign function is the image of f by the non-trivial character over $\mathbb{F}_2$ (usually denoted by χ); to be sure that the distinction between the discrete Fourier transforms of f and of its sign function is easily done, we change the font when we deal with the sign function; many other ways of denoting the discrete Fourier transform can be found in the literature.

x_1	x_2	x_3	x_4	$x_1x_2x_3$	x_1x_4	$f(x)$	$\mathrm{f}_\chi(\mathrm{x})$				$\widehat{\mathrm{f}_\chi}(x)$
0	0	0	0	0	0	0	1	2	4	0	0
1	0	0	0	0	0	0	1	0	0	0	0
0	1	0	0	0	0	1	-1	-2	-4	8	8
1	1	0	0	0	0	1	-1	0	0	0	8
0	0	1	0	0	0	0	1	2	0	0	0
1	0	1	0	0	0	0	1	0	0	0	0
0	1	1	0	0	0	1	-1	-2	0	0	0
1	1	1	0	1	0	0	1	0	0	0	0
0	0	0	1	0	0	0	1	0	0	0	4
1	0	0	1	0	1	1	-1	2	4	4	-4
0	1	0	1	0	0	1	-1	0	0	0	4
1	1	0	1	0	1	0	1	-2	0	4	-4
0	0	1	1	0	0	0	1	0	0	0	-4
1	0	1	1	0	1	1	-1	2	0	-4	4
0	1	1	1	0	0	1	-1	0	0	0	4
1	1	1	1	1	1	1	-1	2	-4	4	-4

Table 2: truth table and Walsh spectrum of $f(x) = x_1x_2x_3 \oplus x_1x_4 \oplus x_2$

We shall call *Walsh transform*[11] of f the Fourier transform of the sign function f_χ. We give in Table 2 an example of the computation of the Walsh transform, using the algorithm recalled above.

Notice that f_χ being equal to $1 - 2f$, we have

$$\widehat{\mathrm{f}_\chi} = 2^n \, \delta_0 - 2\widehat{f} \tag{12}$$

where δ_0 denotes the *Dirac symbol*, *i.e.* the indicator of the singleton $\{0\}$, defined by $\delta_0(u) = 1$ if u is the null vector and $\delta_0(u) = 0$ otherwise; see Proposition 7 for a proof of the relation $\widehat{1} = 2^n \, \delta_0$. Relation (12) gives conversely $\widehat{f} = 2^{n-1}\delta_0 - \frac{\widehat{\mathrm{f}_\chi}}{2}$ and in particular:

$$w_H(f) = 2^{n-1} - \frac{\widehat{\mathrm{f}_\chi}(0)}{2}. \tag{13}$$

[11] The terminology is not much more settled in the literature than is the notation; we take advantage here of the fact that many authors, when working on Boolean functions, use the term of Walsh transform instead of discrete Fourier transform: we call Fourier transform the discrete Fourier transform of the Boolean function itself and Walsh transform (some authors write "Walsh-Hadamard transform") the discrete Fourier transform of its sign function.

Relation (13) applied to $f \oplus \ell_a$, where $\ell_a(x) = a \cdot x$, gives:

$$d_H(f, \ell_a) = w_H(f \oplus \ell_a) = 2^{n-1} - \frac{\widehat{\mathrm{f}_\chi}(a)}{2}. \tag{14}$$

The mapping $f \mapsto \widehat{\mathrm{f}_\chi}(0)$ playing an important role, and being applied in the sequel to various functions deduced from f, we shall also use the specific notation

$$\mathcal{F}(f) = \widehat{\mathrm{f}_\chi}(0) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)}. \tag{15}$$

Properties of the Fourier transform The discrete Fourier transform, as any other Fourier transform, has very nice and useful properties. The number of these properties and the richness of their mutual relationship are impressive. All of these properties are very useful in practice for studying Boolean functions (we shall often refer to the relations below in the rest of the chapter). Almost all properties can be deduced from the next lemma and from the next two propositions.

Lemma 1 *Let E be any vectorspace over $\mathbb{F}_2$ and ℓ any nonzero linear form on E. Then $\sum_{x \in E} (-1)^{\ell(x)}$ is null.*

Proof. The linear form ℓ being not null, its support is an affine hyperplane of E and has $2^{dimE-1} = \frac{|E|}{2}$ elements[12]. Thus, $\sum_{x \in E} (-1)^{\ell(x)}$ being the sum of 1's and -1's in equal numbers, it is null. □

Proposition 6 *For every pseudo-Boolean function φ on $\mathbb{F}_2^n$ and every elements a, b and u of $\mathbb{F}_2^n$, the value at u of the Fourier transform of the function $(-1)^{a \cdot x} \varphi(x + b)$ equals $(-1)^{b \cdot (a+u)} \widehat{\varphi}(a + u)$.*

Proof. The value at u of the Fourier transform of the function $(-1)^{a \cdot x} \varphi(x + b)$ equals $\sum_{x \in \mathbb{F}_2^n} (-1)^{(a+u) \cdot x} \varphi(x + b) = \sum_{x \in \mathbb{F}_2^n} (-1)^{(a+u) \cdot (x+b)} \varphi(x)$ and thus equals $(-1)^{b \cdot (a+u)} \widehat{\varphi}(a + u)$. □

Proposition 7 *Let E be any vector subspace of $\mathbb{F}_2^n$. Denote by 1_E its indicator (recall that it is the Boolean function defined by $1_E(x) = 1$ if $x \in E$ and $1_E(x) = 0$ otherwise). Then:*

$$\widehat{1_E} = |E| \, 1_{E^\perp}, \tag{16}$$

where $E^\perp = \{x \in \mathbb{F}_2^n / \ \forall y \in E, \ x \cdot y = 0\}$ is the orthogonal *of E.*
In particular, for $E = \mathbb{F}_2^n$, we have $\widehat{1} = 2^n \, \delta_0$.

[12] Another way of seeing this is as follows: choose $a \in E$ such that $\ell(a) = 1$; then the mapping $x \mapsto x + a$ is one to one between $\ell^{-1}(0)$ and $\ell^{-1}(1)$.

Proof. For every $u \in \mathbb{F}_2^n$, we have $\widehat{1_E}(u) = \sum_{x \in E}(-1)^{u \cdot x}$. If the linear form $x \in E \mapsto u \cdot x$ is not null on E (*i.*e. if $u \notin E^\perp$) then $\widehat{1_E}(u)$ is null, according to Lemma 1. And if $u \in E^\perp$, then it clearly equals $|E|$. □

We deduce from Proposition 7 the *Poisson summation formula*, which has been used to prove many cryptographic properties in [243], [253], [54] and later in [41, 42], and whose most general statement is:

Corollary 1 *For every pseudo-Boolean function φ on $\mathbb{F}_2^n$, for every vector subspace E of $\mathbb{F}_2^n$, and for every elements a and b of $\mathbb{F}_2^n$, we have:*

$$\sum_{u \in a+E} (-1)^{b \cdot u}\, \widehat{\varphi}(u) = |E|\, (-1)^{a \cdot b} \sum_{x \in b+E^\perp} (-1)^{a \cdot x}\, \varphi(x). \tag{17}$$

Proof. Let us first assume that $a = b = 0$. The sum $\sum_{u \in E} \widehat{\varphi}(u)$, by definition, equals $\sum_{u \in E} \sum_{x \in \mathbb{F}_2^n} \varphi(x)(-1)^{u \cdot x} = \sum_{x \in \mathbb{F}_2^n} \varphi(x)\, \widehat{1_E}(x)$. Hence, according to Proposition 7:

$$\sum_{u \in E} \widehat{\varphi}(u) = |E| \sum_{x \in E^\perp} \varphi(x). \tag{18}$$

We apply this last equality to the function $(-1)^{a \cdot x}\, \varphi(x + b)$, whose Fourier transform's value at u equals $(-1)^{b \cdot (a+u)}\, \widehat{\varphi}(a + u)$, according to Proposition 6. We deduce $\sum_{u \in E}(-1)^{b \cdot (a+u)}\, \widehat{\varphi}(a+u) = |E| \sum_{x \in E^\perp}(-1)^{a \cdot x}\, \varphi(x+b)$, which is equivalent to Equality (17). □

Relation (17) with $a = 0$ and $E = \mathbb{F}_2^n$ gives:

Corollary 2 *For every pseudo-Boolean function φ on $\mathbb{F}_2^n$:*

$$\widehat{\widehat{\varphi}} = 2^n\, \varphi. \tag{19}$$

Thus, the Fourier transform is a permutation on the set of pseudo-Boolean functions on $\mathbb{F}_2^n$ and is its own inverse, up to division by a constant. In order to avoid this division, the Fourier transform is often normalized, that is, divided by $\sqrt{2^n} = 2^{n/2}$ so that it becomes its own inverse. We do not use this normalized transform here because the functions we consider are integer-valued, and we want their Fourier transforms to be also integer-valued.
Corollary 2 allows to show easily that some properties, valid for the Fourier transform of any function φ having some specificities, are in fact necessary and sufficient conditions for φ having these specificities. For instance, according to Proposition 7, the Fourier transform of any constant function φ

takes null value at every nonzero vector; since the Fourier transform of a function null at every nonzero vector is constant, Corollary 2 implies that a function is constant if and only if its Fourier transform is null at every nonzero vector. Similarly, φ is constant on $\mathbb{F}_2^n \setminus \{0\}$ if and only if $\widehat{\varphi}$ is constant on $\mathbb{F}_2^n \setminus \{0\}$.

A classical property of the Fourier transform is to be an isomorphism from the set of pseudo-Boolean functions on $\mathbb{F}_2^n$, endowed with the so-called convolutional product (denoted by $\otimes$), into this same set, endowed with the usual (Hadamard) product (denoted by $\times$). We recall the definition of the convolutional product between two functions φ and ψ:

$$(\varphi \otimes \psi)(x) = \sum_{y \in \mathbb{F}_2^n} \varphi(y)\psi(x+y)$$

(adding here is equivalent to substracting since the operations take place in $\mathbb{F}_2^n$).

Proposition 8 *Let φ and ψ be any pseudo-Boolean functions on $\mathbb{F}_2^n$. We have:*

$$\widehat{\varphi \otimes \psi} = \widehat{\varphi} \times \widehat{\psi}. \tag{20}$$

Consequently:

$$\widehat{\varphi} \otimes \widehat{\psi} = 2^n \, \widehat{\varphi \times \psi}. \tag{21}$$

Proof. We have

$$\widehat{\varphi \otimes \psi}(u) = \sum_{x \in \mathbb{F}_2^n} (\varphi \otimes \psi)(x) \, (-1)^{u \cdot x} = \sum_{x \in \mathbb{F}_2^n} \sum_{y \in \mathbb{F}_2^n} \varphi(y)\psi(x+y) \, (-1)^{u \cdot x}$$

$$= \sum_{x \in \mathbb{F}_2^n} \sum_{y \in \mathbb{F}_2^n} \varphi(y)\psi(x+y) \, (-1)^{u \cdot y \oplus u \cdot (x+y)}.$$

Thus

$$\widehat{\varphi \otimes \psi}(u) = \sum_{y \in \mathbb{F}_2^n} \varphi(y)(-1)^{u \cdot y} \left(\sum_{x \in \mathbb{F}_2^n} \psi(x+y) \, (-1)^{u \cdot (x+y)} \right)$$

$$= \left(\sum_{y \in \mathbb{F}_2^n} \varphi(y)(-1)^{u \cdot y} \right) \left(\sum_{x \in \mathbb{F}_2^n} \psi(x) \, (-1)^{u \cdot x} \right) = \widehat{\varphi}(u) \, \widehat{\psi}(u).$$

This proves the first equality. Applying it to $\widehat{\varphi}$ and $\widehat{\psi}$ in the places of φ and ψ, we obtain $\widehat{\widehat{\varphi} \otimes \widehat{\psi}} = 2^{2n} \, \varphi \times \psi$, according to Corollary 2. Using again this

same corollary, we deduce Relation (21). □

Relation (21) applied at 0 gives

$$\widehat{\varphi} \otimes \widehat{\psi}(0) = 2^n \, \widehat{\varphi \times \psi}(0) = 2^n \sum_{x \in \mathbb{F}_2^n} \varphi(x)\psi(x) = 2^n \, \varphi \otimes \psi(0). \qquad (22)$$

Taking $\psi = \varphi$ in (22), we obtain *Parseval's relation*:

Corollary 3 *For every pseudo-Boolean function φ, we have:*

$$\sum_{u \in \mathbb{F}_2^n} \widehat{\varphi}^{\,2}(u) = 2^n \sum_{x \in \mathbb{F}_2^n} \varphi^2(x).$$

If φ takes values ± 1 only, this becomes:

$$\sum_{u \in \mathbb{F}_2^n} \widehat{\varphi}^{\,2}(u) = 2^{2n}. \qquad (23)$$

This is why, when dealing with Boolean functions, we shall most often prefer using the Walsh transform of f (that is, the Fourier transform of the function $\mathrm{f}_\chi = (-1)^{f(x)}$) instead of the Fourier transform of f.

Relation (20) leads to another relation involving the derivatives of a Boolean function.

Definition 2 *Let f be an n-variable Boolean function and let b be any vector in $\mathbb{F}_2^n$. We call* derivative *of f in the direction of b the Boolean function $D_b f(x) = f(x) \oplus f(x+b)$.*

For instance, the derivative of a function of the form $g(x_1, \cdots, x_{n-1}) \oplus x_n \, h(x_1, \cdots, x_{n-1})$ in the direction of $(0, \cdots, 0, 1)$ equals $h(x_1, \cdots, x_{n-1})$.
Relation (20) applied with $\psi = \varphi = \mathrm{f}_\chi$ implies the so-called *Wiener-Khintchine Theorem*:

$$\widehat{\mathrm{f}_\chi \otimes \mathrm{f}_\chi} = \widehat{\mathrm{f}_\chi}^{\,2}. \qquad (24)$$

We have $(\mathrm{f}_\chi \otimes \mathrm{f}_\chi)(b) = \sum_{x \in \mathbb{F}_2^n} (-1)^{D_b f(x)} = \mathcal{F}(D_b f)$ (the notation $\mathcal{F}$ was defined at Relation (15)). Thus Relation (24) shows that $\widehat{\mathrm{f}_\chi}^{\,2}$ is the Fourier transform of the so-called *auto-correlation function* $b \mapsto \Delta_f(b) = \mathcal{F}(D_b f)$ (this property was first used in the domain of cryptography in [53]):

$$\forall u \in \mathbb{F}_2^n, \; \sum_{b \in \mathbb{F}_2^n} \mathcal{F}(D_b f)(-1)^{u \cdot b} = \widehat{\mathrm{f}_\chi}^{\,2}(u). \qquad (25)$$

Applied at vector 0, this gives

$$\sum_{b\in\mathbb{F}_2^n} \mathcal{F}(D_b f) = \mathcal{F}^2(f). \tag{26}$$

Corollary 1 and Relation (25) imply that, for every vector subspace E of $\mathbb{F}_2^n$ and every vectors a and b (cf. [42]):

$$\sum_{u\in a+E} (-1)^{b\cdot u}\, \widehat{\mathrm{f}_\chi}^2(u) = |E|(-1)^{a\cdot b} \sum_{e\in b+E^\perp} (-1)^{a\cdot e} \mathcal{F}(D_e f)\ . \tag{27}$$

Another interesting relation has been also shown in [42] (see also [250]):

Proposition 9 *Let E and E' be subspaces of $\mathbb{F}_2^n$ such that $E\cap E' = \{0\}$ and whose direct sum equals $\mathbb{F}_2^n$. For every $a\in E'$, let h_a be the restriction of f to the coset $a+E$ (h_a can be identified with a function on $\mathbb{F}_2^k$ where k is the dimension of E). Then*

$$\sum_{u\in E^\perp} \widehat{\mathrm{f}_\chi}^2(u) = |E^\perp| \sum_{a\in E'} \mathcal{F}^2(h_a)\ . \tag{28}$$

Proof. Every element of $\mathbb{F}_2^n$ can be written in a unique way in the form $x+a$ where $x\in E$ and $a\in E'$. For every $e\in E$, we have $\mathcal{F}(D_e f) = \sum_{x\in E; a\in E'} (-1)^{f(x+a)\oplus f(x+e+a)} = \sum_{a\in E'} \mathcal{F}(D_e h_a)$. We deduce from Relation (27), applied with $E^\perp$ instead of E, and with $a=b=0$, that

$$\begin{aligned}\sum_{u\in E^\perp} \widehat{\mathrm{f}_\chi}^2(u) &= |E^\perp| \sum_{e\in E} \mathcal{F}(D_e f) = |E^\perp| \sum_{e\in E} \left(\sum_{a\in E'} \mathcal{F}(D_e h_a)\right)\\ &= |E^\perp| \sum_{a\in E'} \left(\sum_{e\in E} \mathcal{F}(D_e h_a)\right).\end{aligned}$$

Thus, according to Relation (26) applied with E in the place of $\mathbb{F}_2^n$ (recall that E can be identified with $\mathbb{F}_2^k$ where k is the dimension of E): $\sum_{u\in E^\perp} \widehat{\mathrm{f}_\chi}^2(u) = |E^\perp| \sum_{a\in E'} \mathcal{F}^2(h_a)$. □

Fourier transform and linear isomorphisms A last relation that must be mentioned shows what the composition with a linear isomorphism implies on the Fourier transform of a pseudo-Boolean function:

Proposition 10 *Let φ be any pseudo-Boolean function on $\mathbb{F}_2^n$. Let M be a nonsingular $n\times n$ binary matrix and L the linear isomorphism $L : \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto M \times \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Let us denote by M' the transpose of M^{-1} and by L' the linear isomorphism $L' : \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto M' \times \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ (note that L' is the* adjoint operator *of L^{-1}, that is, satisfies $u \cdot L^{-1}(x) = L'(u) \cdot x$ for every x and u). Then*

$$\widehat{\varphi \circ L} = \widehat{\varphi} \circ L'. \tag{29}$$

Proof. For every $u \in \mathbb{F}_2^n$, we have $\widehat{\varphi \circ L}(u) = \sum_{x\in\mathbb{F}_2^n} \varphi(L(x))(-1)^{u\cdot x} = \sum_{x\in\mathbb{F}_2^n} \varphi(x)(-1)^{u\cdot L^{-1}(x)} = \sum_{x\in\mathbb{F}_2^n} \varphi(x)(-1)^{L'(u)\cdot x}$. □

A relationship between algebraic degree and Walsh transform was shown in [229] (see also [54]):

Proposition 11 *Let f be an n-variable Boolean function ($n \geq 2$), and let $1 \leq k \leq n$. Assume that the Walsh transform of f takes values divisible by 2^k (*i.e.*, according to Relation (12), that its Fourier transform takes values divisible by 2^{k-1}, or equivalently, according to Relation (14), that all the Hamming distances between f and affine functions are divisible by 2^{k-1}). Then f has algebraic degree at most $n - k + 1$.*

Proof. Let us suppose that f has algebraic degree $d > n - k + 1$ and, consider a term x^I of degree d in its algebraic normal form. The Poisson summation formula (18) applied to $\varphi = \mathrm{f}_\chi$ and to the vectorspace $E = \{u \in \mathbb{F}_2^n / \ \forall i \in I,\ u_i = 0\}$ gives $\sum_{u\in E} \widehat{\mathrm{f}_\chi}(u) = 2^{n-d} \sum_{x\in E^\perp} \mathrm{f}_\chi(x)$. The orthogonal $E^\perp$ of E equals $\{u \in \mathbb{F}_2^n / \ \forall i \notin I,\ u_i = 0\} = \{u \in \mathbb{F}_2^n / \ supp(u) \subseteq I\}$. According to Proposition 2, we have that $\sum_{x\in E^\perp} f(x)$ is not even and therefore $\sum_{x\in E^\perp} \mathrm{f}_\chi(x)$ is not divisible by 4. Hence, $\sum_{u\in E} \widehat{\mathrm{f}_\chi}(u)$ is not divisible by 2^{n-d+2} and it is therefore not divisible by 2^k. A contradiction. □

The converse of Proposition 11 is obviously valid if $k = 1$. It is also valid

if $k = 2$, since the n-variable Boolean functions of degrees at most $n-1$ are those Boolean functions of even Hamming weights, and $f(x) \oplus u \cdot x$ has degree at most $n-1$ too for every u, since $n \geq 2$. It is finally also valid for $k = n$, since the affine functions are characterized by the fact that their Walsh transforms take values $\pm 2^n$ and 0 only (more precisely, their Walsh transforms take value $\pm 2^n$ once, and all their other values are null, because of Parseval's relation). The converse is false for any other value of k. Indeed, we shall see below that it is false for $k = n-1$ ($n \geq 4$), since there exist quadratic functions f whose Walsh transforms take values $\pm 2^{n/2}$ for n even, ≥ 4, and $\pm 2^{(n+1)/2}$ for n odd, ≥ 5. It is then an easy task to deduce that the converse of Proposition 11 is also false for any value of k such that $3 \leq k \leq n-1$: we choose a quadratic function g in 4 variables, whose Walsh transform value at 0 equals 2^2, that is, whose weight equals $2^3 - 2 = 6$, and we take $f(x) = g(x_1, x_2, x_3, x_4)\, x_5 \cdots x_l$ ($5 \leq l \leq n$). Such function has algebraic degree $l-2$ and its weight equals 6; hence its Walsh transform value at 0 equals $2^n - 12$ and is therefore not divisible by 2^k with $k = n - (l-2) + 1 = n - l + 3 \geq 3$.
It is possible to characterize the functions whose Walsh transform values are divisible by 2^{n-1}: they are the affine functions and the sums of an indicator of a *flat* – an affine space – of co-dimension 2 and of an affine function (they have degree at most 2 according to Proposition 11 and the characterization follows from the results of subsection 5.2). Determining those Boolean functions whose Walsh transform is divisible by 2^k is an open problem for $3 \leq k \leq n-2$.
Note that it is possible to characterize the fact that a Boolean function has degree at most d by means of its Fourier or Walsh transform: since a Boolean function has algebraic degree at most d if and only if its restriction to any $(d+1)$-dimensional flat has an even weight, we can apply Poisson summation formula (17).

Characterizing the Fourier transforms of integer-valued pseudo-Boolean functions and of Boolean functions Obviously, according to the inverse Fourier transform property (19), the Fourier transforms of integer-valued functions (resp. the Walsh transforms of Boolean functions) are those integer-valued functions over $\mathbb{F}_2^n$ whose Fourier transforms take values divisible by 2^n (resp. equal to $\pm 2^n$). Also, the Walsh transforms of Boolean functions being those integer-valued functions φ over $\mathbb{F}_2^n$ such that $\widehat{\varphi}^2$ equals the constant function 2^{2n}, they are those integer-valued functions φ such that $\widehat{\varphi \otimes \varphi} = 2^{2n}$ (according to Relation (20) applied with $\psi = \varphi$),

that is $\varphi \otimes \varphi = 2^{2n}\ \delta_0$. But these characterizations are not easy to use mathematically and they are neither easily computable: they need to check 2^n divisibilities by 2^n for the Fourier transforms of integer-valued functions, and 2^n equalities for the Walsh transforms of Boolean functions.
Since the main cryptographic criteria on Boolean functions will be characterized below as properties of their Walsh transforms, it is important to have characterizations which are as simple as possible. We have seen that characterizing the NNFs of integer-valued (resp. Boolean) functions is easy (resp. easier than with Fourier transform). So it is useful to clarify the relationship between these two representations.

2.2.1 Fourier transform and NNF

There is a similarity between the Fourier transform and the NNF:
- the functions $(-1)^{u\cdot x}$, $u \in \mathbb{F}_2^n$, constitute an orthogonal basis of the space of pseudo-Boolean functions, and the Fourier transform corresponds, up to normalization, to a decomposition over this basis;
- the NNF is defined similarly with respect to the (non-orthogonal) basis of monomials.

Let us see now how each representation can be expressed by means of the other.
Let $\varphi(x)$ be any pseudo-Boolean function and let $\sum_{I\in\mathcal{P}(N)} \lambda_I x^I$ be its NNF. For every word $x \in \mathbb{F}_2^n$, we have: $\varphi(x) = \sum_{I\subseteq supp(x)} \lambda_I$. Setting $b = (1,\cdots,1)$, we have $\varphi(x+b) = \sum_{I\in\mathcal{P}(N)/\ supp(x)\cap I=\emptyset} \lambda_I$ (since the support of $x+b$ equals $\mathbb{F}_2^n \setminus supp(x)$).
For every $I \in \mathcal{P}(N)$, the set $\{x \in \mathbb{F}_2^n / \ supp(x) \cap I = \emptyset\}$ is an $(n-|I|)$-dimensional vector subspace of $\mathbb{F}_2^n$. Let us denote it by E_I. Its orthogonal equals $\{u \in \mathbb{F}_2^n / \ supp(u) \subseteq I\}$. We have $\varphi(x+b) = \sum_{I\in\mathcal{P}(N)} \lambda_I\, 1_{E_I}$. Applying Propositions 6 (with $a = 0$) and 7, we deduce:

$$\widehat{\varphi}(u) = (-1)^{w_H(u)} \sum_{I\in\mathcal{P}(N)\,|\,supp(u)\subseteq I} 2^{n-|I|}\lambda_I. \tag{30}$$

Using the same method as for computing λ_I by means of the values of f, it is an easy task to deduce:

$$\lambda_I = 2^{-n}(-2)^{|I|} \sum_{u\in\mathbb{F}_2^n\,|\,I\subseteq supp(u)} \widehat{\varphi}(u). \tag{31}$$

Note that if φ has numerical degree at most D, then, according to Relation (30), we have $\widehat{\varphi}(u) = 0$ for every vector u of weight strictly greater

than D and that the converse is true, according to Relation (31)

Applying Relation (30) to $\varphi(x) = P(x) = \sum_{I \in \mathcal{P}(N)} \lambda_I \, x^I$ and to $\varphi(x) = P^2(x) = \sum_{I \in \mathcal{P}(N)} \left(\sum_{J,J' \in \mathcal{P}(N) \,|\, I = J \cup J'} \lambda_J \, \lambda_{J'} \right) x^I$, with $u = 0$, we deduce from Proposition 5 that a polynomial $P(x) = \sum_{I \in \mathcal{P}(N)} \lambda_I \, x^I$, with integer coefficients, is the NNF of a Boolean function if and only if

$$\sum_{I \in \mathcal{P}(N)} 2^{n-|I|} \sum_{J,J' \in \mathcal{P}(N) \,|\, I = J \cup J'} \lambda_J \, \lambda_{J'} = \sum_{I \in \mathcal{P}(N)} 2^{n-|I|} \lambda_I. \qquad (32)$$

Remark. The NNF presents the interest of being a polynomial representation, but it can also be viewed as the transform which maps any pseudo-Boolean function $f(x) = \sum_{I \in \mathcal{P}(N)} \lambda_I \, x^I$ to the pseudo-Boolean function g defined by $g(x) = \lambda_{supp(x)}$. Let us denote this mapping by Φ. Three other transforms have also been used for studying Boolean functions:
- the mapping Φ^{-1} (the formulae relating this mapping and the Walsh transform are slightly simpler than for Φ; see [306]);
- a mapping defined by a formula similar to Relation (8), but in which $supp(x) \subseteq I$ is replaced by $I \subseteq supp(x)$; see [171];
- the inverse of this mapping. □

2.2.2 The size of the support of the Fourier transform and its relationship with Cayley graphs

Let f be a Boolean function and let G_f be the *Cayley graph* associated to f: the vertices of this graph are the elements of $\mathbb{F}_2^n$ and there is an edge between two vertices u and v if and only if the vector $u + v$ belongs to the support of f. Then (see [18]), if we multiply by 2^n the values $\widehat{f}(a)$, $a \in \mathbb{F}_2^n$, of the Fourier spectrum of f, we obtain the eigenvalues of the graph G_f (that is, by definition, the eigenvalues of the adjency matrix $(M_{u,v})_{u,v \in \mathbb{F}_2^n}$ of G_f, whose term $M_{u,v}$ equals 1 if $u + v$ belongs to the support of f, and equals 0 otherwise).
As a consequence, the cardinality $N_{\widehat{f}}$ of the support $\{a \in \mathbb{F}_2^n / \ \widehat{f}(a) \neq 0\}$ of the Fourier transform of any n-variable Boolean function f is greater than or equal to the cardinality $N_{\widehat{g}}$ of the support of the Fourier transform of any restriction g of f, obtained by keeping constant some of its input bits. Indeed, the adjency matrix M_g of the Cayley graph G_g is a submatrix of the adjency matrix M_f of the Cayley graph G_f; the number $N_{\widehat{g}}$ equals the rank of M_g, and is then smaller than or equal to the rank $N_{\widehat{f}}$ of M_f.
This property can be generalized to any pseudo-Boolean function φ. Moreover, a simpler proof is obtained by using the Poisson summation for-

mula (17): let I be any subset of $N = \{1, \cdots, n\}$; let E be the vector subspace of $\mathbb{F}_2^n$ equal to $\{x \in \mathbb{F}_2^n / \, x_i = 0, \, \forall i \in I\}$; we have $E^\perp = \{x \in \mathbb{F}_2^n / \, x_i = 0, \, \forall i \in N \setminus I\}$ and the sum of E and of $E^\perp$ is direct; then, for every $a \in E^\perp$ and every $b \in E$, the equality $\sum_{u \in a+E} (-1)^{b \cdot u} \, \widehat{\varphi}(u) = |E| \, (-1)^{a \cdot b} \, \widehat{\psi}(a)$, where ψ is the restriction of φ to $b + E^\perp$, implies that, if $N_{\widehat{f}} = k$, that is, if $\widehat{\varphi}(u)$ is nonzero for exactly k vectors $u \in \mathbb{F}_2^n$, then clearly $\widehat{\psi}(a)$ is nonzero for at most k vectors $a \in E^\perp$.
If φ is chosen to be a Boolean function of algebraic degree d and if we choose for I a multi-index of size d such that x^I is part of the ANF of φ, then the restriction ψ has odd weight and its Fourier transform takes therefore nonzero values only. We deduce (as proved in [18]) that $N_{\widehat{\varphi}} \geq 2^d$. Notice that $N_{\widehat{\varphi}}$ equals 2^d if and only if at most one element (that is, exactly one) satisfying $\widehat{\varphi}(u) \neq 0$ exists in each coset of E, that is, in each set obtained by keeping constant the coordinates x_i such that $i \in I$.
The number $N_{\widehat{\varphi}}$ is also bounded above by $\sum_{i=0}^{D} \binom{n}{i}$, where D is the numerical degree of φ. This is a direct consequence of Relation (30) and of the observation which follows Relation (31).
The graph viewpoint also gives insight on the Boolean functions whose Fourier spectra have at most three values (see [18]).
A hypergraph can also be related to the ANF of a Boolean function f. A related (weak) upper bound on the nonlinearity of Boolean functions (see definition in Subsection 4.1) has been pointed out in [364].

3 Boolean functions and coding

We explained in the introduction how, in error correcting coding, the message is divided into vectors of the same length k, which are transformed into codewords of length $N > k$, before being sent over a noisy channel, in order to enable the correction of the errors of transmission (or of storage, in the case of CD, CD-ROM and DVD) at their reception. A choice of the set of all possible codewords (called the code – let us denote it by C) allows to correct up to t errors (in the transmission of each codeword) if and only if the Hamming distance between any two different codewords is greater than or equal to $2t + 1$ (so, if d is the minimum distance between two codewords, the code can enable to correct up to $\left\lfloor \frac{d-1}{2} \right\rfloor$ errors, where "$\lfloor \, \rfloor$" denotes the integer part). Indeed, the only information the receiver has, concerning the sent word, is that it belongs to C. In order to be always able to recover the correct codeword, he needs that, for every word y at distance at most t from a codeword x, there does not exist another codeword x' at distance at most t

from y, and this is equivalent to saying that the Hamming distance between any two different codewords is greater than or equal to $2t+1$. This necessary condition is also sufficient[13]. Thus, the problem of generating a good code consists in finding a set C of binary words of the same length whose *minimum distance* $\min_{a \neq b \in C} d_H(a,b)$ (where $d_H(a,b) = |\{i / a_i \neq b_i\}|$) is high[14].

A code is called a *linear code* if it has the structure of a linear subspace of $\mathbb{F}_2^N$ where N is its length. The minimum distance of a linear code equals the minimum Hamming weight of all nonzero codewords, since the Hamming distance between two vectors equals the Hamming weight of their difference (*i.e.* their sum since we reduce ourselves here to binary vectors). We shall write that a linear code is an $[N,k,d]$-*code* if it has length N, dimension k and minimum distance d. It can then be described by a *generator matrix* G, obtained by choosing a basis of this vectorspace and writing its elements as the rows of this matrix. The code equals the set of all the vectors of the form $u \times G$, where u ranges over $\mathbb{F}_2^k$ (and $\times$ is the matrix product) and a possible encoding algorithm is therefore the mapping $u \in \mathbb{F}_2^k \mapsto u \times G \in \mathbb{F}_2^N$. The generator matrix is well suited for generating the codewords, but it is not for checking if a received word of length N is a codeword or not. A characterization of the codewords is obtained thanks to the generator matrix H of the *dual code* $C^\perp = \{x \in \mathbb{F}_2^N / \forall y \in C,\ x \cdot y = \bigoplus_{i=1}^N x_i\, y_i = 0\}$ (such a matrix is called a *parity-check matrix*): we have $x \in C$ if and only if $x \times H^t$ is the null vector. It is a simple matter to prove that the minimum distance of the code equals the minimum number of linearly dependent columns of H. For instance, the *Hamming code*, which has by definition for parity-check matrix the $n \times (2^n - 1)$ matrix whose columns are all the non-zero vectors of $\mathbb{F}_2^n$ in some order, has minimum distance 3. This code depends, *stricto sensu*, on the choice of the order, but we say that two binary codes are *equivalent codes* if they are equal, up to some permutation of the coordinates of their codewords.

We shall use in the sequel the notion of *covering radius* of a code: it is the smallest integer ρ such that the spheres of radius ρ centered at the

[13] In practice, we still need to have an efficient decoding algorithm to recover the sent codeword; the naive method consisting in visiting all codewords and keeping the nearest one from the received word is inefficient because the number 2^k of codewords is too large, in general.

[14] High with respect to some known bounds giving the necessary trade-offs between the length of the code, the minimum distance between codewords and the number of codewords, see [258, 298])

codewords cover the whole space, *i*.e. the minimum integer t such that every binary word of length N lies at Hamming distance at most t from at least one codeword, *i.e.* the maximum multiplicity of errors that have to be corrected when maximum likelihood decoding is used on a binary symmetric channel. The covering radius of a code is an important parameter [111], which can be used for analyzing and improving the decoding algorithms devoted to this code.

A linear code C is a *cyclic code* if it is invariant under cyclic shifts of the coordinates (see [258]). Cyclic codes have been extensively studied in coding theory. They have useful properties, that we briefly recall: representing each codeword $(c_0, \cdots, c_{N-1})$ by the polynomial $c_0 + c_1 X + \cdots + c_{N-1} X^{N-1}$, we obtain an ideal of the quotient algebra $\mathbb{F}_2[X]/(X^N + 1)$ (viewed as a set of polynomials of degrees at most $N-1$, each element of the algebra being identified to its minimum degree representent). This algebra is a principal domain, and any (linear) cyclic code has a unique element having minimal degree, called its *generator polynomial*. To simplify the presentation, we shall assume now that $N = 2^n - 1$ (which will be the case in the sequel). The generator polynomial being (as easily shown) a divisor of $X^{2^n-1} + 1$, its roots all belong to $\mathbb{F}_{2^n}^*$. The code equals the set of all those polynomials which include the roots of the generator polynomial among their own roots. The generator polynomial having all its coefficients in $\mathbb{F}_2$, its roots are of the form $\{\alpha^i,\, i \in I\}$ where $I \subseteq \mathbb{Z}/(2^n-1)\mathbb{Z}$ is a union of cyclotomic classes of 2 modulo $2^n - 1$. The set I is called the *defining set* of the code. The elements α^i, $i \in I$ are called the zeroes of the code, which has dimension $N - |I|$. The generator polynomial of $C^\perp$ is the reciprocal of the quotient of $X^{2^n-1} + 1$ by the generator polynomial of C, and its defining set therefore equals $\{2^n - 1 - i;\, i \in \mathbb{Z}/(2^n-1)\mathbb{Z} \setminus I\}$.
A very efficient bound on the minimum distance of cyclic codes is the *BCH bound* [258]: if I contains a string $\{l+1, \cdots, l+k\}$ of length k in $\mathbb{Z}/(2^n-1)\mathbb{Z}$, then the cyclic code has minimum distance greater than or equal to $k+1$. A proof of this bound (in the framework of Boolean functions) is given in the proof of Theorem 15. This bound is valid for cyclic codes over any finite field as well. When the length of such a cyclic code equals the order of the underlying field less 1, the set of zeros can be any set of nonzero elements of the field; when it is constituded of consecutive powers of a primitive element, the code is called a *Reed-Solomon code*.
A cyclic code C of length N being given, the extended code of C is the set of vectors $(c_{-\infty}, c_0, \cdots, c_{N-1})$, where $c_{-\infty} = c_0 \oplus \cdots \oplus c_{N-1}$. It is a linear code of length $N+1$ and of the same dimension as C.

Cyclic codes over $\mathbb{F}_2$ can also be considered in terms of the trace function and therefore viewed as sets of Boolean functions (when their length is 2^n-1, recall we assume this). Any codeword of a cyclic code with non-zeroes α^i for i in the cyclotomic classes containing $u_1, \cdots, u_l$ can be represented as $\sum_{i=1}^{l} tr_n(a_i x^{-u_i}), a_i \in \mathbb{F}_{2^n}$.

3.1 Reed-Muller codes

As explained in the introduction, every code whose length equals 2^n, for some positive integer n, can be interpreted as a set of Boolean functions. The existence of *Reed-Muller codes* comes from the following observation:

Theorem 1 *Any two distinct n-variable functions f and g of algebraic degrees at most r have mutual distances at least 2^{n-r}.*

Proof. In order to prove this property, it is necessary and sufficient to show that any nonzero Boolean function f of algebraic degree $d \leq r$ has weight at least 2^{n-r} (since the difference between two Boolean functions of algebraic degrees at most r has algebraic degree at most r). This can be proved by a double induction over r and n (see [258]), but there exists a simpler proof. Let $\prod_{i\in I} x_i$ be a monomial of degree d in the ANF of f; consider the 2^{n-d} restrictions of f obtained by keeping constant the $n-d$ coordinates of x whose indices lie outside I. Each of these restrictions, viewed as a function on $\mathbb{F}_2^d$, has an ANF of degree d because, when fixing these $n-d$ coordinates, the monomial $\prod_{i\in I} x_i$ is unchanged and all the monomials different from $\prod_{i\in I} x_i$ in the ANF of f give monomials of degrees strictly less than d. Thus any such restriction has an odd (and hence a nonzero) weight (see Subsection 2.1). The weight of f being equal to the sum of the weights of its restrictions, f has weight at least 2^{n-d}, which completes the proof. □

The functions of Hamming weight 2^{n-r} and degree r have been characterized, see a proof in [258]. We give below a proof which brings a little more insight on the reasons of this characterization.

Proposition 12 *The Boolean functions of algebraic degree r and of Hamming weight 2^{n-r} are the indicators of $(n-r)$-dimensional flats (*i.e. *the functions whose supports are $(n-r)$-dimensional affine subspaces of $\mathbb{F}_2^n$).*

Proof. The indicators of $(n-r)$-dimensional flats have clearly Hamming weight 2^{n-r} and they have degree r, since every $(n-r)$-dimensional flat equals $\{x \in \mathbb{F}_2^n \,/\, \ell_i(x) = 1, \forall i = 1, \cdots, r\}$ where the ℓ_i's are affine and

have linearly independent linear parts, and the ANF of its indicator equals $\prod_{i=1}^{r} \ell_i(x)$. Conversely, let f be a function of algebraic degree r and of Hamming weight 2^{n-r}. Let $\prod_{i\in I} x_i$ be a monomial of degree r in the ANF of f and let $J = \{1, \cdots, n\} \setminus I$. For every vector $\alpha \in \mathbb{F}_2^J$, let us denote by f_α the restriction of f to the flat $\{x \in \mathbb{F}_2^n;\ \forall j \in J,\ x_j = \alpha_j\}$. According to the proof of Theorem 1, and since f has Hamming weight 2^{n-r}, each function f_α is the indicator of a singleton $\{a_\alpha\}$. Let us prove that the mapping $a : \alpha \to a_\alpha$ is affine, *i.e.* that, for every $\alpha, \beta, \gamma \in \mathbb{F}_2^J$, we have $a_{\alpha+\beta+\gamma} = a_\alpha + a_\beta + a_\gamma$ (this will complete the proof of the proposition since, denoting by x_J the vector of $\mathbb{F}_2^J$ whose coordinates match the corresponding coordinates of x, the support of f equals the set $\{x \in \mathbb{F}_2^n \,/\, x_I = a_{x_J}\}$ and that the equality $x_I = a_{x_J}$ is equivalent to r linearly independent linear equations). Proving this is equivalent to proving that the function of Hamming weight at most 4 equal to $f_{\alpha+\beta+\gamma} \oplus f_\alpha \oplus f_\beta \oplus f_\gamma$ has algebraic degree at most $r-2$. But more generally, for every k-dimensional flat A of $\mathbb{F}_2^J$, the function $\bigoplus_{\alpha\in A} f_\alpha$ has degree at most $r-k$ (this can be easily proved by induction on k, using that f has degree r). □

Remark.
1. The proof of Theorem 1 shows in fact that, if a monomial $\prod_{i\in I} x_i$ has coefficient 1 in the ANF of f, and if every other monomial $\prod_{i\in J} x_i$ such that $I \subset J$ has coefficient 0, then the function has weight at least $2^{n-|I|}$. Applying this observation to the Möbius transform f° of f - whose definition has been given after Relation (2) - shows that, if there exists a vector $x \in \mathbb{F}_2^n$ such that $f(x) = 1$ and $f(y) = 0$ for every vector $y \neq x$ whose support contains $supp(x)$, then the ANF of f has at least $2^{n-w_H(x)}$ terms (this has been first observed in [364]). Indeed, the Möbius transform of f° is f.
2. The d-dimensional subspace $E = \{x \in \mathbb{F}_2^n \,/\, x_i = 0, \forall i \notin I\}$, in the proof of Theorem 1, is a *maximal odd weighting* subspace: the restriction of f to E has odd weight, and the restriction of f to any of its proper superspaces has even weight (*i.e.* the restriction of f to any coset of E has odd weight). Similarly as above, it can be proved, see [364], that any Boolean function admitting a d-dimensional maximal odd weighting subspace E has weight at least 2^{n-d}.

The Reed-Muller code of order r is by definition the set of all Boolean functions of algebraic degrees at most r (or more precisely the set of the binary words of length 2^n corresponding to (last columns of) the truth-tables of these functions). Denoted by $R(r, n)$, it is an $\mathbb{F}_2$-vectorspace of dimen-

sion $1+n+\binom{n}{2}+\cdots+\binom{n}{r}$ (since this is the number of monomials of degrees at most r, which constitute a basis of $R(r,n)$) and thus, it has $2^{1+n+\binom{n}{2}+\cdots+\binom{n}{r}}$ elements.
For $r=1$, it equals the set of all affine functions. Notice that the weight of any non-constant affine function being equal to the size of an affine hyperplane, it equals 2^{n-1}.

Historic note: the Reed-Muller code $R(1,5)$ was used in 1972 for transmitting the first black-and-white photographs of Mars. It has $2^6=64$ words of length $2^5=32$, with mutual distances at least $2^4=16$. Each codeword corresponded to a level of darkness (this made 64 different levels). Up to $\left\lfloor\frac{16-1}{2}\right\rfloor=7$ errors could be corrected in the transmission of each codeword. □

$R(r,n)$ is a linear code, *i*.e. an $\mathbb{F}_2$-vectorspace. Thus, it can be described by a generator matrix G. For instance, a generator matrix of the Reed-Muller code $R(1,4)$ is:

$$G=\begin{bmatrix}1&1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\0&1&0&0&0&1&1&1&0&0&0&1&1&1&0&1\\0&0&1&0&0&1&0&0&1&1&0&1&1&0&1&1\\0&0&0&1&0&0&1&0&1&0&1&1&0&1&1&1\\0&0&0&0&1&0&0&1&0&1&1&0&1&1&1&1\end{bmatrix}$$

(the first row corresponds to the constant function 1 and the other rows correspond to the coordinate functions $x_1,\cdots,x_4$)[15].
The duals of Reed-Muller codes are Reed-Muller codes:

Theorem 2 *The dual*

$$R(r,n)^\perp=\{f\in\mathcal{BF}_n/\,\forall g\in R(r,n),\ f\cdot g=\bigoplus_{x\in\mathbb{F}_2^n}f(x)\,g(x)=0\}$$

equals $R(n-r-1,n)$.

Proof. We have seen in Subsection 2.1 that the n-variable Boolean functions of even weights are the elements of $R(n-1,n)$. Thus, $R(r,n)^\perp$ is the set of those functions f such that, for every function g of algebraic degree at most

[15] We have chosen to order the words of length 4 by increasing weights; we could have chosen other orderings; this would have led to other codes, but equivalent ones, having the same parameters (a binary code C of length N is said to be equivalent to another binary code C' of the same length if there exists a permutation σ on $\{1,\cdots,N\}$ such that $C=\{(x_{\sigma(1)},\cdots,x_{\sigma(N)})/\,x\in C'\}$.

r, the product function fg (whose value at any $x \in \mathbb{F}_2^n$ equals $f(x)g(x)$) has algebraic degree at most $n-1$. This is clearly equivalent to the fact that f has algebraic degree at most $n-r-1$. □
If the vector-space $\mathbb{F}_2^n$ is identified with the field $\mathbb{F}_{2^n}$, the functions $tr_n(ax^j)$ such that $w_2(j) \leq n-r-1$ being a generating family of $R(n-r-1,n)$ (according to what we have seen on the trace representation of Boolean functions), we have that a Boolean function f belongs to $R(r,n)$ if and only if, for every j such that $w_2(j) \leq n-r-1$, we have $\sum_{x\in\mathbb{F}_{2^n}} f(x)\, tr_n(ax^j) = 0$ for every $a \in \mathbb{F}_{2^n}$, that is, $\sum_{x\in\mathbb{F}_{2^n}} f(x)\, x^j = 0$.

The Reed-Muller codes are invariant under the action of the general affine group. More precisely, it is a simple matter to show that, for any $1 \leq r \leq n-1$, the automorphism group of $R(r,n)$ (that is, the group of all permutations σ of $\mathbb{F}_2^n$ such that $f \circ \sigma \in R(r,n)$ for every $f \in R(r,n)$) equals the general affine group. The sets $R(r,n)$ or $R(r,n)/R(r',n)$ have been classified under this action for some values of r, of $r' < r$ and of n, see [183, 185, 30, 259, 341, 342].

The Reed-Muller code $R(r,n)$ is an extended cyclic code for every $r < n$ (see [258]): the zeroes of the corresponding cyclic code ($R^*(r,n)$, the *punctured Reed-Muller code of order* r) are the elements α^i such that $1 \leq i \leq 2^n-2$ and such that the 2-weight of i is at most equal to $n-r-1$. Indeed, the codewords of $R^*(r,n)$ are the vectors of the form $(g(1), g(\alpha), \cdots, g(\alpha^{2^n-2}))$ where g is a Boolean function of algebraic degree at most r; such function has univariate polynomial form $\sum_{\substack{0\leq j\leq 2^n-2\\ w_2(j)\leq r}} g_j x^j$ and we have

$$\sum_{0\leq l\leq 2^n-2} g(\alpha^l)\alpha^{li} = \sum_{\substack{0\leq j\leq 2^n-2\\ w_2(j)\leq r}} g_j \left(\sum_{0\leq l\leq 2^n-2} \alpha^{l(i+j)} \right)$$

and $\sum_{0\leq l\leq 2^n-2} \alpha^{l(i+j)}$ equals 0 when $w_2(i) \leq n-r-1$ and $w_2(j) \leq r$ since $i+j$ cannot be null and it cannot equal 2^n-1 either since $w_2(i+j) \leq w_2(i)+w_2(j)$. Hence, the α^i's such that $1 \leq i \leq 2^n-2$ and $w_2(i) \leq n-r-1$ are zeroes of the code. Since their number equals the co-dimension of the code, they are the only zeroes of the code.

The problem of determining the weight distributions of the Reed-Muller codes, the MacWilliams identity and the notion of dual distance: What are the possible distances between the words of $R(r,n)$,

or equivalently the possible weights in $R(r,n)$ (or better, the weight distribution of $R(r,n)$)? The answer, which is useful for improving the efficiency of the decoding algorithms and for evaluating their complexities, is known for every n if $r \leq 2$: see Subsection 5.2. For $r \geq n-3$, it can also be deduced from the very nice relationship, due to F. J. MacWilliams, existing between every linear code and its dual: let C be any binary linear code of length N; consider the polynomial $W_C(X,Y) = \sum_{i=0}^{N} A_i X^{N-i} Y^i$ where A_i is the number of codewords of weight i. This polynomial is called the *weight enumerator* of C and describes[16] the *weight distribution* $(A_i)_{0 \leq i \leq N}$ of C. Then (see [258, 298])

$$W_C(X+Y, X-Y) = |C|\, W_C(X,Y). \tag{33}$$

We give a sketch of proof of this *MacWilliams' identity*: we observe first that $W_C(X,Y) = \sum_{x \in C} \prod_{i=1}^{N} X^{1-x_i} Y^{x_i}$; we deduce $W_C(X+Y, X-Y) = \sum_{x \in C} \prod_{i=1}^{N} (X + (-1)^{x_i} Y)$; applying a classical method of expansion, we derive $W_C(X+Y, X-Y) = \sum_{x \in C} \sum_{b \in \mathbb{F}_2^N} \prod_{i=1}^{N} \left(X^{1-b_i} ((-1)^{x_i} Y)^{b_i} \right)$ (choosing X in the i-th factor $X + (-1)^{x_i} Y$ for $b_i = 0$ and $(-1)^{x_i} Y$ for $b_i = 1$; all the different possible choices are taken into account by considering all binary words b of length N). We obtain then $W_C(X+Y, X-Y) = \sum_{b \in \mathbb{F}_2^N} \left(X^{N-w_H(b)} Y^{w_H(b)} \sum_{x \in C} (-1)^{b \cdot x} \right)$ and we conclude by using Relation (16) with $E = C$.

The MacWilliams identity allows, theoretically, to deduce the weight distribution of $R(n-r-1,n)$ from the weight distribution of $R(r,n)$ (in fact, to actually determine this weight distribution, it is necessary to be able to explicitely expand the factors $(X+Y)^{N-i}(X-Y)^i$ and to simplify the obtained expression for $W_C(X+Y, X-Y)$; this is possible by running a computer up to some value of n). But this gives no information for the cases $3 \leq r \leq n-4$ which remain unsolved (except for small values of n, see [17], and for $n = 2r$, because the code is then self-dual, see [258, 298]). *McEliece's theorem* [272] (or *Ax's theorem* [12]; see also the *Stickelberger theorem*, e.g. in [232, 236]) shows that the weights (and thus the distances) in $R(r,n)$ are all divisible by $2^{\lceil \frac{n}{r} \rceil - 1} = 2^{\lfloor \frac{n-1}{r} \rfloor}$, where $\lceil u \rceil$ denotes the ceiling - the smallest integer greater than or equal to u - and $\lfloor u \rfloor$ denotes the integer part (this can also be shown by using the properties of the NNF, see [87]). Moreover, if f has degree d and g has degree $d' \leq d$, then $d_H(f,g) \equiv w_H(f) \left[\bmod 2^{\lceil \frac{n-d'}{d} \rceil} \right]$ [209] (see also [195]). In [36], A. Canteaut gives further properties of the weights

[16] W_C is a homogeneous version of the classical generating series for the weight distribution of C.

in $f \oplus R(1,n)$. Kasami and Tokura [207] have shown that the only weights in $R(r,n)$ occuring in the range $[2^{n-r}; 2^{n-r+1}[$ are of the form $2^{n-r+1}-2^i$ for some i; and they have completely characterized the codewords with these weights (and computed their number). The functions whose weights are between the minimum distance 2^{n-r} and 2.5 times the minimum distance have also been characterized, in [208].
The principle of MacWilliams' identity can also be applied to nonlinear codes. When C is not linear, the weight distribution of C has no great relevance. The distance distribution has more interest. We consider the *distance enumerator* of C: $D_C(X,Y) = \frac{1}{|C|}\sum_{i=0}^{N} B_i X^{N-i}Y^i$, where B_i is the size of the set $\{(x,y) \in C^2 / \, d_H(x,y) = i\}$. Note that, if C is linear, then $D_C = W_C$. Similarly as above, we see that $D_C(X,Y) = \frac{1}{|C|}\sum_{(x,y)\in C^2}\prod_{i=1}^{N} X^{1-(x_i\oplus y_i)}Y^{x_i\oplus y_i}$; we deduce that the polynomial $D_C(X+Y, X-Y)$ equals $\frac{1}{|C|}\sum_{(x,y)\in C^2}\prod_{i=1}^{N}(X+(-1)^{x_i\oplus y_i}Y)$. Expanding these products, we obtain $\frac{1}{|C|}\sum_{(x,y)\in C^2}\sum_{b\in\mathbb{F}_2^N}\prod_{i=1}^{N}\left(X^{1-b_i}((-1)^{x_i\oplus y_i}Y)^{b_i}\right)$, that is

$$D_C(X+Y,X-Y) = \frac{1}{|C|}\sum_{b\in\mathbb{F}_2^N} X^{N-w_H(b)}Y^{w_H(b)}\left(\sum_{x\in C}(-1)^{b\cdot x}\right)^2 \qquad (34)$$

Hence, $D_C(X+Y,X-Y)$ has non-negative coefficients.
The minimum exponent of Y with nonzero coefficient in the polynomial $D_C(X+Y,X-Y)$, that is, the number $\min\{w_H(b);\, b\neq 0, \sum_{x\in C}(-1)^{b\cdot x}\neq 0\}$, is usually denoted by $d^\perp$ and is called the *dual distance* of C. Note that the maximum number j such that the sum $\sum_{x\in C}(-1)^{b\cdot x}$ is null, for every nonzero vector b of weight at most j, equals $d^\perp - 1$ (see more in [129, 130]). This property will be useful in Subsection 4.1.

It is shown in [52] (see the remark of Subsection 5.2 in the present chapter) that for every Boolean function f on $\mathbb{F}_2^n$, there exists an integer m and a Boolean function g of algebraic degree at most 3 on $\mathbb{F}_2^{n+2m}$ whose Walsh transform satisfies: $\widehat{g_\chi}(0) = 2^m \widehat{f_\chi}(0)$. This means that the weight of f is related to the weight of a function of degree at most 3 (but in a number of variables which can be exponentially larger) in a simple way. This shows that the distances in $R(3,n)$ can be very diverse, contrary to those in $R(2,n)$.
□

4 Boolean functions and cryptography

Stream ciphers are based on the so-called *Vernam cipher* (see Figure 1) in which the plaintext (a binary string of some length) is bitwise added to a (binary) secret key of the same length, in order to produce the ciphertext. The Vernam cipher is also called the *one time pad* because a new random secret key must be used for every encryption. Indeed, the bitwise addition of two ciphertexts corresponding to the same key equals the addition of the corresponding plaintexts, which gives much information on these plaintexts (it is often enough to recover both plaintexts; some secret services and spies learned this at their own expenses!).

Figure 1: Vernam cipher

The Vernam cipher, which is the only known cipher offering unconditional security (see [332]) if the key is truly random and if it is changed for every new encryption, was used for the communication between the heads of USA and USSR during the cold war (the keys being carried by diplomats) and by some secret services.

In practice, since the length of the private key must be equal to the length of the plaintext, pseudo-random generators are most often used in order to minimize the size of the private key (but the unconditional security is then no longer ensured): a method is chosen for producing long *pseudo-random sequences* from short random secret keys (only the latter are actually shared; the method is supposed to be public; according to the Kerckhoffs principle, only the parameters which can be used by the attacker to break the system must be kept secret). The pseudo-random sequence is used in the place of the key in a Vernam cipher. For this reason, it is also called the *keystream*. If the keystream only depends on the key (and not on the plaintext), the cipher is called *synchronous*[17]. Stream ciphers, because they operate on data units as small as a bit or a few bits, are suitable for fast

[17]There also exist self-synchronous stream ciphers, in which each keystream bit depends

telecommunication applications. Having also a very simple construction, they are easily implemented both in hardware and software.
The first method for generating a pseudo-random sequence from a secret key has used *Linear Feedback Shift Registers* (*LFSR*). In such an LFSR

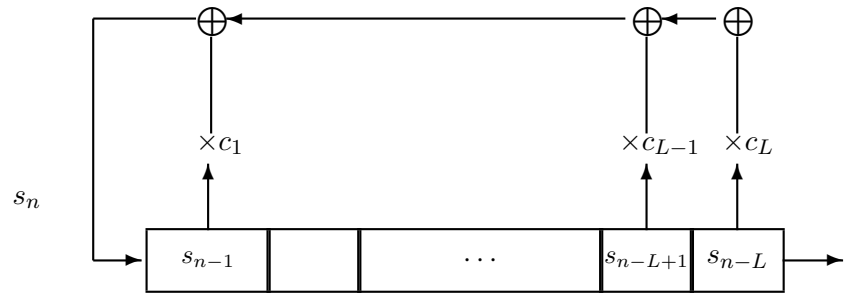


Figure 2: LFSR

(see Figure 2, where $\times$ means multiplication), at every clock-cycle, the bits $s_{n-1}, \cdots, s_{n-L}$ contained in the flip-flops of the LFSR move to the right. The right-most bit is the current output and the left-most flip-flop is feeded with the linear combination $\bigoplus_{i=1}^{L} c_i s_{n-i}$, where the c_i's are bits. Thus, such an LFSR outputs a recurrent sequence satisfying the relation

$$s_n = \bigoplus_{i=1}^{L} c_i s_{n-i}.$$

Such sequence is always ultimately periodic[18] (if $c_L = 1$, then it is periodic; we shall assume that $c_L = 1$ in the sequel, because otherwise, the same sequence can be output by an LFSR of a shorter length, except for its first bits, and this can be exploited in attacks) with period at most $2^L - 1$. The generating series $s(X) = \bigoplus_{i \geq 0} s_i X^i$ of the sequence can be expressed in a nice way (see the chapter by Helleseth and Kumar in [298]): $s(X) = \frac{G(X)}{F(X)}$, where $G(X) = \bigoplus_{i=0}^{L-1} X^i \left(\bigoplus_{j=0}^{i} c_{i-j} s_j \right)$ is a polynomial of degree smaller than L and $F(X) = 1 \oplus c_1 X \oplus \cdots \oplus c_L X^L$ is the *feedback polynomial*.
The short secret key contains the initialization $s_0, \cdots, s_{L-1}$ of the LFSR and

on the n preceding ciphertext bits, which allows re-synchronising after n bits if an error of transmission occurs between Alice and Bob

[18] Conversely, every ultimately periodic sequence can be generated by an LFSR.

the values of the *feedback coefficients* c_i (these must be kept secret; otherwise, the observation of L consecutive bits of the key would allow recovering all the subsequent sequence).

But these LFSRs are cryptographically weak because of the *Berlekamp-Massey algorithm* [269]: let $\mathcal{L}$ be the length of a minimum length LFSR producing the same sequence (this length, called the *linear complexity* of the sequence, is assumed to be unknown from the attacker; note that it equals L if and only if the polynomials F and G above are co-prime), then if we know at least $2\mathcal{L}$ consecutive bits, Berlekamp-Massey algorithm recovers the values of $\mathcal{L}$ and of the feedback coefficients of an LFSR of length $\mathcal{L}$ generating the sequence, and the initialization of this LFSR in $O(\mathcal{L}^2)$ elementary operations. A modern way of avoiding this attack is by using Boolean functions. The first model which appeared in the litterature for using Boolean functions is the *combiner model* (see Figure 3).

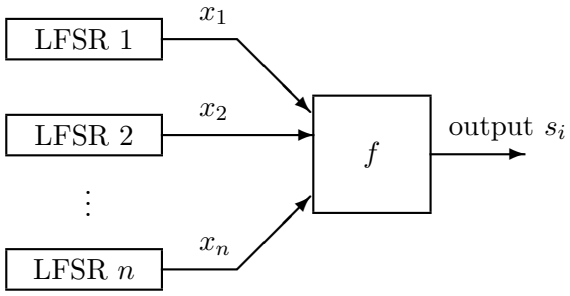


Figure 3: Combiner model

Notice that the feedback coefficients of the n LFSRs used in such a generator can be public. The Boolean function is also public, in general, and the short secret key gives only the initialization of the n LFSRs: if we want to use for instance a 128 bit long secret key, this allows using n LFSRs of lengths $L_1, \cdots, L_n$ such that $L_1 + \cdots + L_n = 128$.
Such system clearly outputs a periodic sequence whose period is at most the LCM of the periods of the sequences output by the n LFSRs (assuming that $c_L = 1$ in each LFSR; otherwise, the sequence is ultimately periodic). So, this sequence is also recurrent and can therefore be produced by a single LFSR. However, as we shall see, well-chosen Boolean functions allow the linear complexity of the sequence to be much larger than the sum of the lengths of the n LFSRs. Nevertheless, choosing LFSRs producing sequences

of large periods[19], choosing these periods pairwise co-prime so that to have then the largest possible global period, and choosing f such that the linear complexity is large enough too are not sufficient. As we shall see, the combining function should also not leak information about the individual LFSRs and behave as differently as possible from affine functions, in several different ways.
The combiner model is only a model, useful for studying attacks and related criteria. In practice, the systems are more complex (see for instance at URL http://www.ecrypt.eu.org/stream/ how are designed the stream ciphers of the *eSTREAM Project*).

An alternative model is the *filter model*, which uses a single LFSR (of a longer length). A filtered LFSR outputs $f(x_1, \cdots, x_n)$ where f is some n-variable Boolean function, called a filtering function, and where $x_1, \cdots, x_n$ are the bits contained in some flip-flops of the LFSR, see Figure 4.

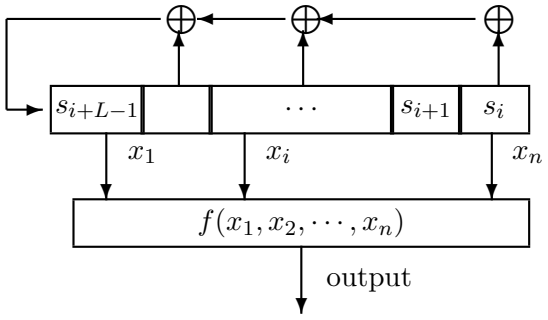


Figure 4: Filter model

Such system is equivalent to the combiner model using n copies of the LFSR. However, the attacks, even when they apply to both systems, do not work similarly (a first obvious difference is that the lengths of the LFSRs are different in the two models). Consequently, the criteria that the involved Boolean functions must satisfy because of these attacks may be different for

[19] *e.g. m-sequences* also called *maximum length sequences*, that is, sequences of period $2^{\mathcal{L}} - 1$ where $\mathcal{L}$ is the linear complexity – assuming that $L = \mathcal{L}$, this corresponds to taking a primitive feedback polynomial – which can be represented in the form $s_i = tr_n(a\alpha^i)$ where α is a primitive element of $\mathbb{F}_{2^n}$, and which have very strong properties; see the chapter by Helleseth and Kumar in [298].

the two models and we shall have to distinguish between the two models when describing the attacks and the related criteria.

Other pseudo-random generators exist. A *Feedback Shift Register* has the same structure as an LFSR, but the left-most flip-flop is feeded with $f(x_{i_1}, \cdots, x_{i_n})$ where $n \leq L$ and $x_{i_1}, \cdots, x_{i_n}$ are bits contained in the flip-flops of the FSR, and where f is some n-variable Boolean function. The linear complexity of the produced sequence can then be near 2^L, see [199] for general FSRs and [98] for FSRs with quadratic feedback function f. The linear complexity is difficult to study in general. Nice results similar to those on the m-sequences exist in the case of FCSRs (Feedback with Carry Shift-Registers), see [218, 167, 8, 168].

Boolean functions also play an important role in block ciphers. Every block cipher admits as input a binary vector $(x_1, \cdots, x_n)$ (a block of plaintext) and outputs a binary vector $(y_1, \cdots, y_m)$; the coordinates $y_1, \cdots, y_m$ are the outputs to Boolean functions (depending on the key) whose common input is $(x_1, \cdots, x_n)$, see Figure 5.

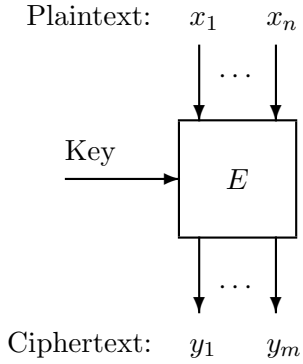


Figure 5: Block cipher

But the number n of variables of these Boolean functions being large (most often, more than a hundred), these functions can not be analyzed. Boolean functions on fewer variables are in fact involved in the ciphers. All known block ciphers are the iterations of a number of rounds.

We give in Figures 6 and 7 a description of the rounds of the DES and of the AES. The input to a DES round is a binary string of length 64, divided into two strings of 32 bits each (in the figure, they enter the round, from above, on the left and on the right); confusion (see below what this term

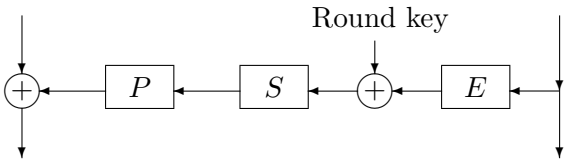


Figure 6: A DES ROUND

means) is achieved by the S-box, which is a nonlinear transformation of a binary string of 48 bits[20] into a 32 bit long one. So, 32 Boolean functions on 48 variables are involved. But, in fact, this nonlinear transformation is the concatenation of eight sub-S-boxes, which transform binary strings of 6 bits into 4 bit long ones. So, 32 (that is, 8×4) Boolean functions on 6 variables are involved.
In the (standard) AES round, the input is a 128 bit long string, divided into 16 strings of 8 bits each; the S-box is the concatenation of 16 sub-S-boxes corresponding to 16×8 Boolean functions on 8 variables.

A block cipher being considered, the individual properties of all the involved Boolean functions can be studied (see Subsection 4.1), but this is not sufficient. The whole sub-S-boxes must be globally studied (see the chapter "Vectorial Boolean Functions for Cryptography").

4.1 Cryptographic criteria for Boolean functions

The design of conventional cryptographic systems relies on two fundamental principles introduced by Shannon [332]: *confusion* and *diffusion*. Confusion aims at concealing any algebraic structure in the system. It is closely related to the complexity[21] of the involved Boolean functions. Diffusion consists in spreading out the influence of any minor modification of the input data or of the key over all outputs. These two principles were stated more than half a century ago. Since then, many attacks have been found against the diverse known cryptosystems, and the relevance of these two principles has always been confirmed. The known attacks on each cryptosystem lead to criteria [276, 300, 336] that the implemented cryptographic functions must satisfy. More precisely, the resistance of the cryptosystems to the known attacks can

[20] The E-box has expanded the 32 bit long string into a 48 bit long one.

[21] That is, the cryptographic complexity, which is different from circuit complexity, for instance.

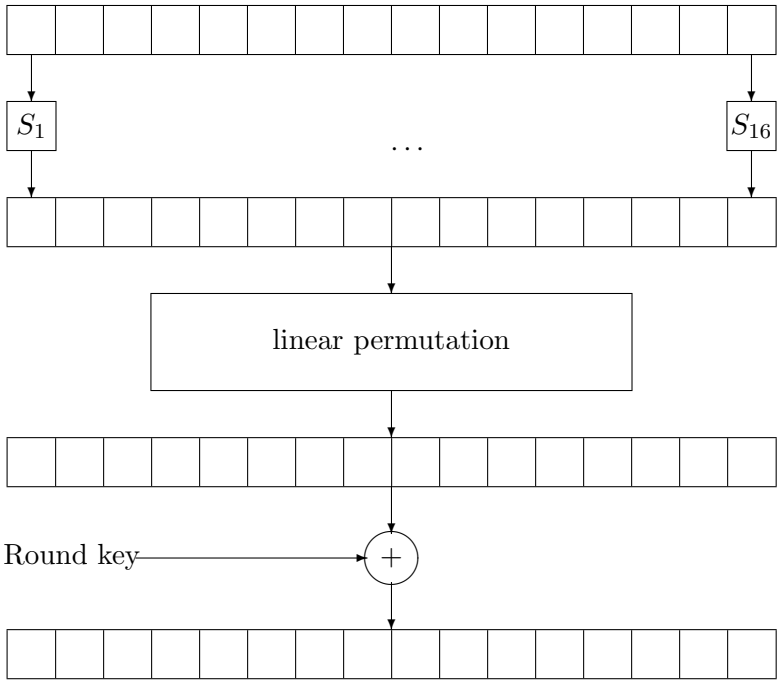


Figure 7: An AES round

be quantified through some fundamental characteristics (some, more related to confusion, and some, more related to diffusion) of the Boolean functions used in them; and the design of these cryptographic functions needs to consider various characteristics simultaneously. Some of these characteristics are affine invariants, *i*.e. are invariant under affine equivalence (recall that two functions f and g on $\mathbb{F}_2^n$ are called affinely equivalent if there exists a linear isomorphism L from $\mathbb{F}_2^n$ to $\mathbb{F}_2^n$ and a vector a such that $f(x) = g(L(x)+a)$ for every input $x \in \mathbb{F}_2^n$) and some are not. Of course, all characteristics cannot be optimum in the same time, and trade-offs must be considered (see below).

4.1.1 The algebraic degree

Cryptographic functions must have high algebraic degrees. Indeed, all cryptosystems using Boolean functions for confusion (combining or filtering functions in stream ciphers, functions involved in the S-boxes of block ciphers, ...) can be attacked if the functions have low degrees. For instance, in

the case of combining functions, if n LFSRs having lengths $L_1, \cdots, L_n$ are combined by the function

$$f(x) = \bigoplus_{I \in \mathcal{P}(N)} a_I \left(\prod_{i \in I} x_i \right),$$

where $\mathcal{P}(N)$ denotes the power set of $N = \{1, \cdots, n\}$, then (see [318]) the sequence produced by f has linear complexity

$$\mathcal{L} \leq \sum_{I \in \mathcal{P}(N)} a_I \left(\prod_{i \in I} L_i \right)$$

(and $\mathcal{L}$ equals this number under the sufficient condition that the sequences output by the LFSRs are m-sequences and the lengths of these LFSRs are pairwise co-prime), see [367]. In the case of the filter model, we have a less precise result [317]: if L is the length of the LFSR and if the feedback polynomial is primitive, then the linear complexity of the sequence satisfies:

$$\mathcal{L} \leq \sum_{i=0}^{d^\circ f} \binom{L}{i}.$$

Moreover, if L is a prime, then

$$\mathcal{L} \geq \binom{L}{d^\circ f},$$

and the fraction of functions f of given algebraic degree which output a sequence of linear complexity equal to $\sum_{i=0}^{d^\circ f} \binom{L}{i}$ is at least $e^{-1/L}$. In both models, the algebraic degree of f (recall that this is the largest size of I such that $a_I = 1$) has to be high so that $\mathcal{L}$ can have high value (the number of those nonzero coefficients a_I, in the ANF of f, such that I has large size, can also play a role, but clearly a less important one). In the case of block ciphers, using Boolean functions of low degrees makes the higher order differential attack [215, 227] effective.

When n tends to infinity, random Boolean functions have almost surely algebraic degrees at least $n-1$ since the number of Boolean functions of algebraic degrees at most $n-2$ equals $2^{\sum_{i=0}^{n-2} \binom{n}{i}} = 2^{2^n - n - 1}$ and is negligible with respect to the number 2^{2^n} of all Boolean functions. But we shall see that the functions of algebraic degrees $n-1$ or n do not allow achieving some other characteristics (balancedness, resiliency, ...).

We have seen in Subsection 2.1 that the algebraic degree is an affine invariant.

4.1.2 The nonlinearity

In order to provide confusion, cryptographic functions must lie at large Hamming distance to all affine functions. Let us explain why. We shall say that there is a correlation between a Boolean function f and a linear function ℓ if $d_H(f, \ell)$ is different from 2^{n-1}. Because of Parseval's Relation (23) applied to the sign function f_χ and of Relation (14), any Boolean function has correlation with some linear functions of its input. But this correlation should be small: the existence of affine approximations of the Boolean functions involved in a cryptosystem allows in various situations (block ciphers, stream ciphers) to build attacks on this system (see [173, 271]).

In the case of stream ciphers, these attacks are the so-called *fast correlation attacks* [47, 109, 156, 200, 201, 202, 275]: let g be a linear approximation of f (or $f \oplus 1$, then we change f into $f \oplus 1$) whose distance to f is smaller than 2^{n-1}. Then, denoting by $Pr[E]$ the probability of an event E:

$$p = Pr[f(x_1, \cdots, x_n) \neq g(x_1, \cdots, x_n)] = \frac{d_H(f, g)}{2^n} = \frac{1}{2} - \epsilon,$$

where $\epsilon > 0$. The pseudo-random sequence s corresponds then to the transmission with errors of the sequence σ which would be produced by the same model, but with g instead of f. Attacking the cipher can be done by correcting the errors as in the transmission of the sequence σ over a noisy channel. Assume that we have N bits $s_u, \cdots, s_{u+N-1}$ of the pseudo-random sequence s, then $Pr[s_i \neq \sigma_i] \approx p$. The set of possible sequences $\sigma_u, \cdots, \sigma_{u+N-1}$ is a vectorspace, that is, a linear code of length N and dimension at most L. We then use a decoding algorithm to recover $\sigma_u, \cdots, \sigma_{u+N-1}$ from $s_u, \cdots, s_{u+N-1}$ and since g is linear, the linear complexity of the sequence σ is small and we obtain by the Berlekamp-Massey algorithm the initialisation of the LFSR. We can then compute the whole sequence s.
There are several ways for performing the decoding. The method exposed in [275] and improved by [109] is as follows. We call a *parity check polynomial* any polynomial $a(x) = 1 + \sum_{j=1}^{r} a_j x^j$ ($a_r \neq 0$) which is a multiple of the feedback polynomial of an LFSR generating the sequence σ_i. Denoting by $\sigma(x)$ the generating function $\sum_{i \geq 0} \sigma_i x^i$, the product $a(x)\,\sigma(x)$ is a polynomial of degree less than r. We use for the decoding a set of parity check polynomials satisfying three conditions: their degrees are bounded by some integer m, the number of nonzero coefficients a_j in each of them is at most some number t (*i.e.*, each polynomial has Hamming weight at most $t+1$) and for every $j = 1, \cdots, m$, at most one polynomial has nonzero coefficient a_j. Each parity check polynomial $a(x) = 1 + \sum_{j=1}^{r} a_j x^j$ gives a linear relation

$\sigma_i = \sum_{j=1}^{r} a_j \sigma_{i-j} = \sum_{j=1,\cdots,r \,/\, a_j \neq 0} \sigma_{i-j}$ for every $i \geq m$ and the relations corresponding to different polynomials involve different indices $i-j$. If we replace the (unknown) σ_i's by the s_i's then some of these relations become false but it is possible by using the method of Gallager [160] to compute a sequence z_i such that $Pr(z_i = \sigma_i) > 1-p$. Then it can be proved that iterating this process converges to the sequence σ (with a speed which depends on m, t and p).

In the case of block ciphers, we shall see in the chapter "Vectorial Boolean Functions for Cryptography" that the Boolean functions involved in their S-boxes must also lie at large Hamming distances to affine functions, to allow resistance to the linear attacks.

The *nonlinearity* of f is the minimum Hamming distance between f and affine functions. The larger is the nonlinearity, the larger is p in the fast correlation attack and the less efficient is the attack. Hence, the nonlinearity must be high (in a sense that will be clarified below) and we shall see that this condition happens to be necessary against other attacks as well. A high nonlinearity is surely one of the most important cryptographic criteria.
The nonlinearity is an affine invariant, by definition, since $d_H(f \circ L, \ell \circ L) = d_H(f, \ell)$, for every functions f and ℓ, and for every affine automorphism L, and since $\ell \circ L$ ranges over the whole set of affine functions when ℓ does.
It can be computed through the Walsh transform: let $\ell_a(x) = a_1 x_1 \oplus \cdots \oplus a_n x_n = a \cdot x$ be any linear function; according to Relation (14), we have $d_H(f, \ell_a) = 2^{n-1} - \frac{1}{2}\widehat{\mathrm{f}_\chi}(a)$ and we deduce $d_H(f, \ell_a \oplus 1) = 2^{n-1} + \frac{1}{2}\widehat{\mathrm{f}_\chi}(a)$; the nonlinearity of f is therefore equal to:

$$nl(f) = 2^{n-1} - \frac{1}{2} \max_{a \in \mathbb{F}_2^n} |\widehat{\mathrm{f}_\chi}(a)|. \tag{35}$$

Hence a function has high nonlinearity if all of its Walsh values have low magnitudes.
Parseval's Relation (23) applied to f_χ gives $\sum_{a \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^{\,2}(a) = 2^{2n}$, and implies that the mean of $\widehat{\mathrm{f}_\chi}^{\,2}(a)$ equals 2^n. The maximum of $\widehat{\mathrm{f}_\chi}^{\,2}(a)$ being greater than or equal to its mean (equality occurs if and only if $\widehat{\mathrm{f}_\chi}^{\,2}(a)$ is constant), we deduce that $\max_{a \in \mathbb{F}_2^n} |\widehat{\mathrm{f}_\chi}(a)| \geq 2^{n/2}$. This implies

$$nl(f) \leq 2^{n-1} - 2^{n/2-1}. \tag{36}$$

This bound, valid for every Boolean function and tight for every even n, as we shall see, will be called the *covering radius bound* (since this is the value of the covering radius of the Reed-Muller code of order 1 if n is even;

indeed, in the case of the Reed-Muller code of order 1, the covering radius coincides with the maximum nonlinearity of Boolean functions). The covering radius bound can be improved when we restrict ourselves to sub-classes of functions (e.g. resilient and correlation-immune functions, see Section 7). A Boolean function will be considered as highly nonlinear if its nonlinearity lies near the upper bound corresponding to the class of functions to which it belongs. The meaning of "near" depends on the framework, see [203]. Olejár and Stanek [289] have shown that, when n tends to infinity, random Boolean functions on $\mathbb{F}_2^n$ have almost surely nonlinearity greater than $2^{n-1}-\sqrt{n}\,2^{\frac{n-1}{2}}$ (this is easy to prove by counting – or more precisely by upper bounding – the number of functions whose nonlinearities are bounded above by a given number, see [66]). Rodier [311] has shown later more precisely that, asymptotically, almost all Boolean functions have nonlinearity between $2^{n-1}-2^{n/2-1}\sqrt{n}\left(\sqrt{2\ln 2}+\frac{4\ln n}{n}\right)$ and $2^{n-1}-2^{n/2-1}\sqrt{n}\left(\sqrt{2\ln 2}-\frac{5\ln n}{n}\right)$ and therefore located in the neighbourhood of $2^{n-1}-2^{n/2-1}\sqrt{2n\ln 2}$.
Equality occurs in (36) if and only if $|\widehat{f_\chi}(a)|$ equals $2^{n/2}$ for every vector a. The corresponding functions are called *bent functions*. They exist only for even values of n, because $2^{n-1}-2^{n/2-1}$ must be an integer (in fact, they exist for every n even, see Section 6). The whole Section 6 is devoted to bent functions.
For n odd, Inequality (36) cannot be tight. The maximum nonlinearity of n-variable Boolean functions lies then between $2^{n-1}-2^{\frac{n-1}{2}}$ (which can always be achieved *e.g.* by quadratic functions, see Subsection 5.2) and $2^{n-1}-2^{n/2-1}$. It has been shown in [177, 284] that it equals $2^{n-1}-2^{\frac{n-1}{2}}$ when $n=1,\ 3,\ 5,\ 7$, and in [295, 296], by Patterson and Wiedemann[22] , that it is strictly greater than $2^{n-1}-2^{\frac{n-1}{2}}$ if $n\geq 15$ (a review on what was known in 1999 on the best nonlinearities of functions on odd numbers of variables was given in [154], see also [29, 237]). This value $2^{n-1}-2^{\frac{n-1}{2}}$ is called the *quadratic bound* because, as we already mentioned, such nonlinearity can be achieved by quadratic functions. It is also called the *bent-concatenation bound* since it can also be achieved by the concatenation of two bent functions in $n-1$ variables. Very recently it has been proved in [210] (see also [262]) that the best nonlinearity of Boolean functions in odd numbers of variables is strictly greater than the quadratic bound for any $n>7$.

[22]It has been later proved (see [328, 141] and [267, 216]) that balanced functions with nonlinearity strictly greater than $2^{n-1}-2^{\frac{n-1}{2}}$, and with algebraic degree $n-1$, or satisfying $PC(1)$, exist for every odd $n\geq 15$.

The nonlinearity of a Boolean function f equals the minimum distance of the linear code $R(1,n) \cup (f \oplus R(1,n))$. More generally, the minimum distance of a code defined as the union of cosets $f \oplus R(1,n)$ of the Reed-Muller code of order 1, where f ranges over a set $\mathcal{F}$, equals the minimum nonlinearity of the functions $f \oplus g$, where f and g are distinct and range over $\mathcal{F}$. This observation allows constructing good nonlinear codes such as Kerdock codes (see Subsection 6.10).

Bent functions being not balanced (*i*.e. their values being not uniformly distributed, see below), they are improper for use in cryptosystems[23] (see below). For this reason, even when they exist (for n even), it is also necessary to study those functions which have large but not optimal nonlinearities, say between $2^{n-1} - 2^{\frac{n-1}{2}}$ and $2^{n-1} - 2^{n/2-1}$, among which some balanced functions exist. The maximum nonlinearity of balanced functions is unknown for any $n \geq 8$.

Two relations have been first observed in [360, 363] between the nonlinearity and the derivatives of Boolean functions (we give here simpler proofs): applying Relation (27), relating the values of the Walsh transform of a function on a flat $a + E$ to the autocorrelation coefficients of the function on a flat $b + E^{\perp}$, to all linear hyperplanes $E = \{0, e\}^{\perp}$, $e \neq 0$, to all vectors a and to $b = 0$, and using that $\max_{u \in E} \widehat{f_\chi}^2(u) \geq \frac{1}{|E|} \sum_{u \in E} \widehat{f_\chi}^2(u)$, we deduce:

$$nl(f) \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + \max_{e \neq 0} |\mathcal{F}(D_e f)|}.$$

And the obvious relation $w_H(f) \geq \frac{1}{2} w_H(D_e f)$, valid for every $e \in \mathbb{F}_2^n$, leads when applied to the functions $f \oplus \ell$, where ℓ is affine, to the lower bound:

$$nl(f) \geq 2^{n-2} - \frac{1}{4} \min_{e \neq 0} |\mathcal{F}(D_e f)|. \tag{37}$$

Another lower bound on the nonlinearity is a consequence of Remark 2 after Theorem 1: if f admits a maximal odd weighting subspace E of dimension $d \geq 2$, then for every affine function ℓ, the function $f \oplus \ell$ also admits E as maximal odd weighting subspace (since the restriction of ℓ to E and to any of its superspaces has an even weight) and thus has nonlinearity at least

[23] As soon as n is large enough (say $n \geq 20$), the difference $2^{n/2-1}$ between their weights and the weight 2^{n-1} of balanced functions is very small with respect to this weight. However, according to [13, Theorem 6], 2^n bits of the pseudo-random sequence output by f are enough to distinguish it from a random sequence. Nevertheless, we shall see in Section 6 that highly nonlinear functions can be built from bent functions.

2^{n-d}.

The r-th order nonlinearity: changing one or a few bits in the output to a low degree Boolean function (that is, in its truth-table) gives a function with high degree and does not fundamentally modify the robustness of the system using this function (explicit attacks using approximations by low degree functions exist for block ciphers but not for all stream ciphers however, see e.g. [219]). A relevant criterion is the *nonlinearity profile*, that is, the sequence of the Hamming distances to the Reed-Muller code of order r, for all values of $r < n$. This distance is called the r-th order nonlinearity (and if r is not specified, the *higer order nonlinearity*) of f and denoted by $nl_r(f)$. This criterion is related to the maximum correlation of the Boolean function with respect to a subset of variables, or equivalently, to the minimal distance of the function to functions depending on a subset of variables (which plays a role with respect to the correlation attack, see below in Subsection 4.1.7) since a function depending on k variables has algebraic degree at most k. Hence the r-th order nonlinearity is a lower bound to the distance to functions depending of at most k variables. The former is much more difficult to study than the latter.
The best known asymptotic upper bound on $nl_r(f)$ is

$$2^{n-1} - \frac{\sqrt{15}}{2} \cdot (1+\sqrt{2})^{r-2} \cdot 2^{n/2} + O(n^{r-2})$$

(see [92], where a non-asymptotic - and more complex - bound is also given). Counting the number of functions whose r-th order nonlinearities are bounded above by a given number allows proving that, when n tends to infinity, there exist functions with r-th order nonlinearity greater than $2^{n-1} - \sqrt{\sum_{i=0}^{r} \binom{n}{i}} \quad 2^{\frac{n-1}{2}}$. But this does not help obtaining explicit functions with non-weak r-th order nonlinearity.
Computing the r-th order nonlinearity of a given function with algebraic degree strictly greater than r is a hard task for $r > 1$ (in the case of the first order, we have seen that much is known in theory and also algorithmically since the nonlinearity is related to the Walsh transform, which can be computed by the algorithm of the Fast Fourier Transform; but for $r > 1$, very little is known). Even the second order nonlinearity is known only for a few peculiar functions and for functions in small numbers of variables. A nice algorithm due to G. Kabatiansky and C. Tavernier and improved and implemented by Fourquet and Tavernier [157] works well for $r = 2$ and $n \leq 11$ (in some cases, $n \leq 13$), only. It can be applied for higher orders, but it is

then efficient only for very small numbers of variables. No better algorithm is known. Proving lower bounds on the r-th order nonlinearity of functions (and therefore proving their good behavior with respect to this criterion) is also a quite difficult task, even for the second order. Until recently, there had been only one attempt, by Iwata-Kurosawa [198], to construct functions with lower bounded r-th order nonlinearity. But the obtained value, $2^{n-r-3}(r+5)$, of the lower bound was small. Also, lower bounds on the r-th order nonlinearity by means of the algebraic immunity of Boolean functions have been derived (see Section 9) but they are small too. In [73] is introduced a method for efficiently bounding below the nonlinearity profile of a given function in the case lower bounds exist for the $(r-1)$-th order nonlinearities of the derivatives of f:

Proposition 13 *Let f be any n-variable function and r a positive integer smaller than n. We have:*

$$nl_r(f) \geq \frac{1}{2} \max_{a \in \mathbb{F}_2^n} nl_{r-1}(D_a f)$$

and

$$nl_r(f) \geq 2^{n-1} - \frac{1}{2}\sqrt{2^{2n} - 2\sum_{a \in \mathbb{F}_2^n} nl_{r-1}(D_a f)}.$$

The first bound is easy to prove and the second one comes from the equalities $nl_r(f) = 2^{n-1} - \frac{1}{2} \max_{h \in \mathcal{BF}_n \,/\, d^\circ f \leq r} \left| \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x) \oplus h(x)} \right|$ and:

$$\left(\sum_{x \in \mathbb{F}_2^n} (-1)^{f(x) \oplus h(x)} \right)^2 = \sum_{a \in \mathbb{F}_2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{D_a f(x) \oplus D_a h(x)}.$$

Lower bounds for the second order nonlinearities of some functions (known for being highly nonlinear) are deduced in [73], as well as bounds for the whole nonlinearity profile of the multiplicative inverse function $tr_n(x^{2^n-2})$ (used in the S-box of the AES with $n = 8$, see the chapter "Vectorial Boolean Functions for Cryptography"): the r-th order nonlinearity of this function is approximately bounded below by $2^{n-1} - 2^{(1-2^{-r})\,n}$ and therefore asymptotically equivalent to 2^{n-1}, for every fixed r. Note that the extension of the Weil bound recalled in Subsection 5.6 is efficient for lower bounding the r-th order nonlinearity of the inverse function only for $r = 1$. Indeed, already for $r = 2$, the univariate degree of a quadratic function in trace representation

form can be bounded above by $2^{\lfloor n/2 \rfloor} + 1$ only and this gives a bound in 2^n on the maximum magnitude of the Walsh transform and therefore no information on the nonlinearity.

4.1.3 Balancedness and resiliency

Cryptographic functions must be *balanced functions* (their output must be uniformly – that is, equally – distributed over $\{0,1\}$) for avoiding statistical dependence between the plaintext and the ciphertext. Notice that f is balanced if and only if $\widehat{\mathrm{f}_\chi}(0) = \mathcal{F}(f) = 0$.
A stronger condition is necessary in the filtering model of pseudo-random generators, in order to avoid so-called *distinguishing attacks*. These attacks are able to distinguish the pseudorandom sequence $(s_i)_{i\in\mathbb{N}}$ from a random sequence. A way of doing so is to observe that the distribution of the sequences $(s_{i+\gamma_1}, \cdots, s_{i+\gamma_n})$ is not uniform, where $\gamma_1, \cdots, \gamma_n$ are the positions where the input bits to the filtering function are chosen. J. Golić [163] has observed that if the feedback polynomial of the LFSR is primitive and if the filtering function has the form $x_1 \oplus g(x_2, \cdots, x_n)$ or $g(x_1, \cdots, x_{n-1}) \oplus x_n$, then the property of uniformity is satisfied. A. Canteaut [40] has proved that this condition on the function is also necessary for having uniformity. For choosing a filtering function, we shall have to choose a function g satisfying the cryptographic criteria listed in the present section, and use f defined by means of g in one of the two ways above.
There is an additional condition to balancedness in the case of the combiner model: any combining function $f(x)$ must stay balanced if we keep constant some number of coordinates x_i of x.

Definition 3 *Let n be a positive integer and $m < n$ a non-negative integer. An n-variable function f is called an* m-resilient function[24] *if any of its restrictions obtained by fixing at most*[25] *m of its input coordinates x_i is*

[24] More generally, a (non necessarily balanced) combining function whose output distribution probability is unaltered when any m (or, equivalently, at most m) of the inputs are kept constant is called an m-th order *correlation-immune function*. Similarly with resiliency, correlation immunity is characterized by the set of zero values in the Walsh spectrum of the function: f is m-th order correlation-immune if and only if $\widehat{\mathrm{f}_\chi}(u) = 0$, *i.*e. $\widehat{f}(u) = 0$, for all $u \in \mathbb{F}_2^n$ such that $1 \leq w_H(u) \leq m$. The notion of correlation-immune function is related to the notion of orthogonal array (see [35]). Only resilient functions are of interest as cryptographic functions (but Boolean correlation-immune functions play a role with respect to vectorial resilient functions, see the chapter "Vectorial Boolean Functions for Cryptography").

[25] Or exactly m, this is equivalent.

balanced.

This definition of resiliency was introduced by Siegenthaler[26] in [336]; it is related to an attack on the combiner model[27], called *correlation attack*: if f is not m-resilient, then there exists a correlation between the output to the function and (at most) m coordinates of its input; if m is small, a divide-and-conquer attack due to Siegenthaler [337] uses this weakness for attacking a system using f as combining function; in the original attack by Siegenthaler, all the possible initializations of the m LFSRs corresponding to these coordinates are tested (in other words, an exhaustive search of the initializations of these specific LFSRs is done); when we arrive to the correct initialization of these LFSRs, we observe a correlation (before that, the correlation is negligible, as for random pairs of sequences); now that the initializations of the m LFSRs are known, those of the remaining LFSRs can be found with an independent exhaustive search. The fast correlation attacks that we saw above can be more efficient if the Boolean combining function is not highly nonlinear. More precisely, Canteaut and Trabbia in [47] and Canteaut in [38] show that, to make the correlation attack on the combiner model with an m-resilient combining function as inefficient as possible, the coefficient $\widehat{f_\chi}(u)$ of the function has to be small for every vector u of Hamming weight higher than but close to m. This condition is satisfied if the function is highly nonlinear. Hence we see that nonlinearity plays a role with respect to all the main attacks.
Note that, when we say that a function f is m-resilient, we do not mean that m is the maximum value of k such that f is k-resilient. We will call this maximum value the *resiliency order* of f.
Resiliency has been characterized by Xiao and Massey through the Fourier and the Walsh transforms:

Theorem 3 *[174] Any n-variable Boolean function f is m-resilient if and only if $\widehat{f_\chi}(u) = 0$ for all $u \in \mathbb{F}_2^n$ such that $w_H(u) \leq m$. Equivalently, f is m-resilient if and only if it is balanced and $\widehat{f}(u) = 0$ for all $u \in \mathbb{F}_2^n$ such that $0 < w_H(u) \leq m$.*

We give here a first direct proof of this fact: we apply Relation (28) to $E = \{x \in \mathbb{F}_2^n / \ x_i = 0, \ \forall i \in I\}$ where I is any set of indices of size m; the sum of E and $E^\perp = \{x \in \mathbb{F}_2^n / \ x_i = 0, \ \forall i \notin I\}$ is direct and equals $\mathbb{F}_2^n$; hence

[26] The term of resiliency was, in fact, introduced in [110], in relationship with another cryptographic problem.

[27] This attack has no equivalent for the filter model, where first order resiliency seems sufficient; see more precisely in [170] the status of resiliency in the filter generator.

we can take $E' = E^\perp$ and we get $\sum_{u\in E^\perp} \widehat{\mathrm{f}_\chi}^2(u) = |E^\perp| \sum_{a\in E^\perp} \mathcal{F}^2(h_a)$, where h_a is the restriction of f to $a+E$, that is, the restriction obtained by fixing the coordinates of x whose indices belong to I to the corresponding coordinates of a. The number $\mathcal{F}(h_a)$ is null if and only if h_a is balanced and clearly, all the numbers $\mathcal{F}(h_a)$, $a \in E^\perp$ are null if and only if all the numbers $\widehat{\mathrm{f}_\chi}(u)$, $u \in E^\perp$ are null. Since this is valid for every muti-index I of size m, this completes the proof.

An alternate proof of this same result is obtained by applying the Poisson summation formula (17) to $\varphi = \mathrm{f}_\chi$, $a = 0$ and $E = \{x \in \mathbb{F}_2^n /\ x_i = 0,\ \forall i \notin I\}$, b ranging over $\mathbb{F}_2^n$. We obtain that f is m-resilient if and only if, for every b and every I of size m, we have $\sum_{u\in\mathbb{F}_2^n /\ u_i=0,\ \forall i\notin I} (-1)^{b\cdot u} \widehat{\mathrm{f}_\chi}(u) = 0$ and it can easily be shown that this is equivalent to $\widehat{\mathrm{f}_\chi}(u) = 0$ for every u of weight at most m.

Theorem 3 shows that f is m-resilient if and only if its support has size 2^{n-1} and dual distance at least $m+1$. Indeed, if C denotes the support of f, the dual distance of C equals the number $\min\{w_H(b);\ b \neq 0, \sum_{x\in C}(-1)^{b\cdot x} \neq 0\}$, according to Relation (34) and to the observation which follows it; we have, for every vector b: $\sum_{x\in C}(-1)^{b\cdot x} = \widehat{f}(b)$ and therefore, for every $b \neq 0$: $\sum_{x\in C}(-1)^{b\cdot x} = -\frac{1}{2}\widehat{\mathrm{f}_\chi}(b)$. More generally, f is m-th order correlation immune if and only if its support has dual distance at least $m+1$. This had been observed by Delsarte in [129, 130] (see also in a paper by J. Massey [270] a generalization of this result to arrays over finite fields and other related nice results).

An easily provable related property is that, if G is the generator matrix of an $[n, k, d]$ linear code, then for every k-variable balanced function g, the n-variable function $f(x) = g(x \times G^t)$ is $(d-1)$-resilient [128] (but such function has nonzero linear structures, see below).

Contrary to the algebraic degree, to the nonlinearity and to the balancedness, the resiliency order is not an affine invariant, except for the null order (and for the order n, but the set of n-resilient functions is empty, because of Parseval's relation). It is invariant under any translation $x \mapsto x + b$, according to Proposition 6 and Theorem 3. The symmetry group of the set of m-resilient functions and the orbits under its action have been studied in [194]).

The whole Section 7 is devoted to resilient functions.

4.1.4 Strict avalanche criterion and propagation criterion

The *Strict Avalanche Criterion* (SAC) was introduced by Webster and Tavares [352] and this concept was generalized into the *Propagation Criterion* (PC) by Preneel et al. [300] (see also [301]). The SAC, and its generalizations, are based on the properties of the derivatives of Boolean functions. These properties describe the behavior of a function whenever some coordinates of the input are complemented. Thus, they are related to the property of diffusion of the cryptosystems using the function. They concern the Boolean functions involved in block ciphers. Let f be a Boolean function on $\mathbb{F}_2^n$ and $E \subset \mathbb{F}_2^n$. The function f satisfies the *propagation criterion PC with respect to E* if, for all $a \in E$, the derivative $D_a f(x) = f(x) \oplus f(a+x)$ (see Definition 2) is balanced. It satisfies $PC(l)$ if it satisfies PC with respect to the set of all nonzero vectors of weights at most l. In other words, f satisfies $PC(l)$ if the auto-correlation coefficient $\mathcal{F}(D_a f)$ is null for every $a \in \mathbb{F}_2^n$ such that $1 \leq w_H(a) \leq l$. Criterion SAC corresponds to $PC(1)$.
It is needed, for some cryptographic applications, to have Boolean functions which still satisfy $PC(l)$ when a certain number k of coordinates of the input x are kept constant (whatever are these coordinates and whatever are the constant values chosen for them). We say that such functions satisfy the *propagation criterion $PC(l)$ of order k*. This notion, introduced in [300], is a generalization of the strict avalanche criterion of order k, $SAC(k)$ (which is equivalent to $PC(1)$ of order k), introduced in [155]. Obviously, if a function f satisfies $PC(l)$ of order $k \leq n-l$, then it satisfies $PC(l)$ of order k' for any $k' \leq k$.
There exists another notion, which is similar to $PC(l)$ of order k, but stronger [300, 302] (see also [61]): a Boolean function satisfies the *extended propagation criterion $EPC(l)$* of order k if every derivative $D_a f$, with $a \neq 0$ of weight at most l, is k-resilient.
All of these criteria are not affine invariants, in general.
A weakened version of the PC criterion has been studied in [222].

4.1.5 Non-existence of nonzero linear structure

We shall call the *linear kernel* of f the set of those vectors e such that $D_e f$ is a constant function. The linear kernel of any Boolean function is an $\mathbb{F}_2$-subspace of $\mathbb{F}_2^n$. Any element e of the linear kernel of f is said to be a *linear structure* of f. Nonlinear cryptographic functions used in block ciphers should have no nonzero linear structure (see [148]). The existence

of nonzero (involuntary) linear structures, for the functions implemented in stream ciphers, is a potential risk that should also be avoided, despite the fact that such existence could not be used in attacks, so far.

Proposition 14 *An n-variable Boolean function admits a nonzero linear structure if and only if it is linearly equivalent to a function of the form $f(x_1, \cdots, x_n) = g(x_1, \cdots, x_{n-1}) \oplus \varepsilon\, x_n$ where $\varepsilon \in \mathbb{F}_2$. More generally, its linear kernel has dimension at least k if and only if it is linearly equivalent to a function of the form $f(x_1, \cdots, x_n) = g(x_1, \cdots, x_{n-k}) \oplus \varepsilon_{n-k+1}\, x_{n-k+1} \oplus \cdots \oplus \varepsilon_n\, x_n$ where $\varepsilon_{n-k+1}, \cdots, \varepsilon_n \in \mathbb{F}_2$.*

Indeed, if we compose f on the right with a linear automorphism L such that $L(0, \cdots, 0, 1) = e$ is a nonzero linear structure, we have then $D_{(0,\cdots,0,1)}(f \circ L)(x) = f \circ L(x) \oplus f \circ L(x + (0, \cdots, 0, 1)) = f \circ L(x) \oplus f(L(x) + e) = D_e f(L(x))$. The case of dimension k is similar.
Note that, according to Proposition 14, if f admits a nonzero linear structure, then the nonlinearity of f is bounded above by $2^{n-1} - 2^{\frac{n-1}{2}}$ (this implies that the functions obtained by Patterson and Wiedemann cannot have nonzero linear structure), since it equals twice that of g and since, g being an $(n-1)$-variable function, it has nonlinearity bounded above by $2^{n-2} - 2^{\frac{n-1}{2}-1}$. Similarly, if k is the dimension of the linear kernel of f, we have straighforwardly $nl(f) \leq 2^{n-1} - 2^{\frac{n+k-2}{2}}$ [41].
Another characterization of linear structures [228, 146] (see also [43]) is a direct consequence of Relation (27), relating the values of the Walsh transform of a function on a flat $a + E$ to the autocorrelation coefficients of the function on a flat $b + E^\perp$, with $b = 0$ and $E = \{0, e\}^\perp$, that is $\sum_{u \in a+E} \widehat{\mathrm{f}_\chi}^2(u) = 2^{n-1}(2^n + (-1)^{a \cdot e} \mathcal{F}(D_e f))$.

Proposition 15 *Let f be any n-variable Boolean function. The derivative $D_e f$ equals the null function (resp. the function 1) if and only if the support $S_{\widehat{\mathrm{f}_\chi}} = \{u \in \mathbb{F}_2^n / \widehat{\mathrm{f}_\chi}(u) \neq 0\}$ of $\widehat{\mathrm{f}_\chi}$ is included in $\{0, e\}^\perp$ (resp. its complement).*

This is a direct consequence of the relation above deduced from (27), with $a \cdot e = 1$ if $D_e f$ is null and $a \cdot e = 0$ if $D_e f = 1$. Notice that, if $D_e f$ is the constant function 1 for some $e \in \mathbb{F}_2^n$, then f is balanced (indeed, the relation $f(x + e) = f(x) \oplus 1$ implies that f takes the values 0 and 1 equally often). Thus, a non-balanced function f has no nonzero linear structure if and only if there is no nonzero vector e such that $D_e f$ is null. According to Proposition 15, this is equivalent to saying that the support of its Walsh

transform has rank n. A similar characterization exists for balanced functions by replacing the function $f(x)$ by a non-balanced function $f(x) \oplus b \cdot x$. It is deduced in [105] (see more in [347]) that resilient functions of high orders must have linear structures.
The existence/non-existence of nonzero linear structures is clearly an affine invariant. But, contrary to the other criteria, it is an all-or-nothing criterion. Meier and Staffelbach introduced in [276] a related criterion, leading to a characteristic (that is, a criterion which can be satisfied at levels quantified by numbers): a Boolean function on $\mathbb{F}_2^n$ being given, its *distance to linear structures* is its distance to the set of all Boolean functions admitting nonzero linear structures (among which we have all affine functions – hence, this distance is bounded above by the nonlinearity – but also other functions, such as all non bent quadratic functions). This distance is always bounded above by 2^{n-2}. More precisely, it equals[28]: $2^{n-2} - \frac{1}{4}\max_{e\in\mathbb{F}_2^{n*}} |\mathcal{F}(D_e f)|$, since a function g, which admits some nonzero vector e as a linear structure, and which lies at minimum distance from f among all such functions, can be obtained by choosing an affine hyperplane H such that $\mathbb{F}_2^n = H \cup (e+H)$, and defining $g(x) = f(x)$ for every $x \in H$ and $g(x) = f(x+e) \oplus \epsilon$ for every $x \in (e+H)$, where ϵ is chosen in $\mathbb{F}_2$; the Hamming distance between f and this function g equals $|\{x \in e+H/\ D_e f(x) = \epsilon \oplus 1\}| = \frac{1}{2}|\{x \in \mathbb{F}_2^n/\ D_e f(x) = \epsilon \oplus 1\}| = \frac{1}{2}\left(2^{n-1} - \frac{(-1)^\epsilon}{2}\mathcal{F}(D_e f)\right)$. Recall that $\Delta_f(e) = \mathcal{F}(D_e f)$ is the auto-correlation function of f. We see (according to Theorem 8) that the distance of f to linear structures equals 2^{n-2} if and only if f is bent.

4.1.6 Algebraic immunity

A new kind of attacks, called *algebraic attacks*, has been introduced recently (see [117, 150, 113]). Algebraic attacks recover the secret key, or at least the initialization of the system, by solving a system of multivariate algebraic equations. The idea that the key bits can be characterized as the solutions of a system of multivariate equations comes from C. Shannon [332]. In practice, for cryptosystems which are robust against the usual attacks such as the Berlekamp-Massey attack, this system is too complex to be solved (its equations being highly nonlinear). However, in the case of stream ciphers, we can get a very overdefined system (i.e. a system with a number of linearly independent equations much greater than the number of unknowns). Let us consider the combiner or the filter model, with a linear part (the n LFSRs

[28] Note that this proves again Relation (37).

in the case of the combiner model, the single LFSR in the case of the filter model) of size N and with an n-variable Boolean function f as combining or filtering function; then there exists a linear permutation $L : \mathbb{F}_2^N \mapsto \mathbb{F}_2^N$ and a linear mapping $L' : \mathbb{F}_2^N \mapsto \mathbb{F}_2^n$ such that, denoting by $u_1, \cdots, u_N$ the initialisation of the LFSR and by $(s_i)_{i \geq 0}$ the pseudo-random sequence output by the generator, we have, for every $i \geq 0$:

$$s_i = f(L' \circ L^i(u_1, \cdots, u_N)).$$

The number of equations can then be much larger than the number of unknowns. This makes less complex the resolution of the system by using Groebner basis (see [150]), and even allows linearizing the system (i.e. obtaining a system of linear equations by replacing every monomial of degree greater than 1 by a new unknown); the resulting linear system has however too many unkwnowns and cannot be solved. Nevertheless, Courtois and Meier have had a simple but very efficient idea. Assume that there exist functions $g \neq 0$ and h of low degrees (say, of degrees at most d) such that $f \times g = h$ (where $f \times g$ denotes the Hadamard product of f and g, whose support is the intersection of the supports of f and g, we shall omit writing $\times$ in the sequel). We have then, for every $i \geq 0$:

$$s_i \, g(L' \circ L^i(u_1, \cdots, u_N)) = h(L' \circ L^i(u_1, \cdots, u_N)).$$

This equation in $u_1, \cdots, u_N$ has degree at most d, since L and L' are linear, and the system of equations obtained after linearization can then be solved by Gaussian elimination.

Low degree relations have been shown to exist for several well known constructions of stream ciphers, which were immune to all previously known attacks.

Note that if we only know the existence of a nonzero low degree multiple h of f, then the support of h being included in that of f, we have $(f \oplus 1)h = 0$, and taking $g = h$, we have the desired relation $fg = h$ (the paper [117] mentioned the existence of low degree multiples of f for making the attack feasible). It is a simple matter to see also that the existence of functions $g \neq 0$ and h, of degrees at most d, such that $fg = h$ is equivalent to the existence of a function $g \neq 0$ of degree at most d such that $fg = 0$ or $(f \oplus 1)g = 0$. Indeed, $fg = h$ implies $f^2 g = fh$, that is (since $f^2 = f$), $f(g \oplus h) = 0$, which gives the desired equality if $g \neq h$ by replacing $g \oplus h$ by g, and if $g = h$ then $fg = h$ is equivalent to $(f \oplus 1)g = 0$. A function g such that $fg = 0$ is called an *annihilator* of f. The set of all annihilators is equal to the ideal of all the multiples of $f \oplus 1$.

Let g be a function of degree d. Let the ANF of g equal $a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{1\leq i<j\leq n} a_{i,j} x_i x_j + ... + \sum_{1\leq i_1\leq ...\leq i_d\leq n} a_{i_1,...i_d} x_{i_1}...x_{i_d}$. Note that g is an annihilator of f if and only if $f(x) = 1$ implies $g(x) = 0$. Hence, g is an annihilator of f if and only if the coefficients in its ANF satisfy the system of homogeneous linear equations which translates this fact. In this system, we have $\sum_{i=0}^{d} \binom{n}{i}$ number of variables (the coefficients of the monomials of degrees at most d) and $w_H(f)$ many equations.
The minimum degree of $g \neq 0$ such that $fg = 0$ (i.e. such that g is an annihilator of f) or $(f \oplus 1)g = 0$ (i.e. such that g is a multiple of f) is called the (standard) *algebraic immunity* of f and denoted by $AI(f)$. This important characteristic is an affine invariant. More precisely, its automorphism group (that is, the group of all permutations σ of $\mathbb{F}_2^n$ such that $AI(f \circ \sigma) = AI(f)$ for every Boolean function f) equals the general affine group. Indeed, denoting by $AN(f)$ the set of annihilators of f, we have $AN(f \circ \sigma) = AN(f) \circ \sigma$. Hence this automorphism group equals the automorphism group of the Reed-Muller codes.

As shown in [117], the algebraic immunity of any n-variable function is bounded above[29] by $\lceil n/2 \rceil$. Indeed, the sum of the number of monomials of degrees at most $\lceil n/2 \rceil$ and of the (equal) number of the products between f and these monomials being greater than 2^n, these functions are necessarily linearly dependent elements of the 2^n-dimensional vectorspace of all Boolean functions. This linear dependence gives two functions g and h of degrees at most $\lceil n/2 \rceil$ such that $fg = h$ and $(g, h) \neq (0, 0)$, i.e. $g \neq 0$.
Let us see now what are the consequences of the existence of this attack on the design of stream ciphers: let an n-variable function f, with algebraic immunity $\lceil n/2 \rceil$ be used for instance as a filtering function on an LFSR) of length $N \geq 2k$, where k is the length of the key (otherwise, it is known that the system is not robust against an attack called time-memory-data trade-off attack). Then the complexity of an algebraic attack using one annihilator of degree $\lceil n/2 \rceil$ is roughly $7\left(\binom{N}{0} + \cdots + \binom{N}{\lceil n/2 \rceil}\right)^{\log_2 7} \approx 7\left(\binom{N}{0} + \cdots + \binom{N}{\lceil n/2 \rceil}\right)^{2.8}$ (see [117]). Let us choose $k = 128$ (which is usual) and $N = 256$, then the complexity of the algebraic attack is at least 2^{80} (which is considered nowadays as a just sufficient complexity) for $n \geq 13$; and it is greater than the complexity of an exhaustive search, that is 2^{128},

[29] Consequently, it is bounded above by $\lceil k/2 \rceil$ if, up to affine equivalence, it depends only on k variables, and by $\lceil k/2 + 1 \rceil$ if it has a linear kernel of dimension $n - k$, since it is then equivalent, according to Proposition 14, to a function in k variables plus an affine function.

for $n \geq 15$. If the attacker knows several linearly independent annihilators of degree $\lceil n/2 \rceil$, then the number of variables must be enhanced! In practice, the number of variables will have to be near 20 (but this poses then a problem of efficiency of the stream cipher).

A high value of $AI(f)$ is not a sufficient property for a resistance to algebraic attacks, because of fast algebraic attacks, which work if one can find g of low degree and $h \neq 0$ of reasonable degree such that $fg = h$, see [113, 176] (note however that fast algebraic attacks need more data than standard ones). This has been exploited in [115] to present an attack on a stream cipher called SFINKS. Similarly as above, when the number of monomials of degrees at most e, plus the number of monomials of degrees at most d, is strictly greater than 2^n – that is, when $e + d \geq n$ – there exist $g \neq 0$ of degree at most e and h of degree at most d such that $fg = h$. An n-variable function f is then optimal with respect to fast algebraic attacks if there do not exist two functions $g \neq 0$ and h such that $fg = h$, $d^\circ g < \lceil n/2 \rceil$ and $d^\circ g + d^\circ h < n$. Since $fg = h$ implies $fh = ffg = fg = h$, we see that h is then an annihilator of $f \oplus 1$, and if $h \neq 0$, its degree is then at least equal to the algebraic immunity of f. This means that having a high algebraic immunity is not only a necessary condition for a resistance to standard algebraic attacks but also for a resistance to fast algebraic attacks.

The pseudo-random generator must also resist algebraic attacks on the augmented function [153], that is (considering now f as a function in N variables, to simplify description), the vectorial function $F(x)$ whose output equals the vector $(f(x), f(L(x)), \cdots, f(L^{m-1}(x)))$, where L is the (linear) update function of the linear part of the generator. Algebraic attacks can be more efficient when applied to the augmented function rather than to the function f itself. The efficiency of the attack depends not only on the function f, but also on the update function (and naturally also on the choice of m), since for two different update functions L and L', the vectorial functions $F(x)$ and $F'(x) = (f(x), f(L'(x)), ..., f(L'^{m-1}(x))$ are not linearly equivalent (neither equivalent in the more general sense called CCZ-equivalence, that is, affine equivalence of the graphs of the functions, see the chapter "Vectorial Boolean Functions for Cryptography"). Testing the behavior of a function with respect to this attack is therefore a long term work (all possible update functions have to be investigated).

Finally, a powerful attack on the filter generator has been introduced by S. Rønjom and T. Helleseth in [313], which also adapts the idea of algebraic attacks due to Shannon, but in a different way. The complexity of the attack is in about $\sum_{i=0}^{d} \binom{N}{i}$ operations, where d is the algebraic degree of the

filter function and N is the length of the LFSR. It needs about $\sum_{i=0}^{d}\binom{N}{i}$ consecutive bits of the keystream output by the pseudo-random generator. Since d is supposed to be close to the number n of variables of the filter function, the number $\sum_{i=0}^{d}\binom{N}{i}$ is comparable to $\binom{N}{n}$, while in the case of a standard algebraic attack with the method due to Courtois and Meier, the complexity of the attack is in $O\left(\left(\sum_{i=0}^{AI(f)}\binom{N}{i}\right)^{\omega}\right)$ operations, where $\omega \approx 3$ is the exponent of the Gaussian reduction[30] and $AI(f)$ is the algebraic immunity of the filter function, and it needs about $\sum_{i=0}^{AI(f)}\binom{N}{i}$ consecutive bits of the keystream. Since $AI(f)$ is supposed to be close to $\lceil n/2 \rceil$, we can see that denoting by C the complexity of the Courtois-Meier attack and by C' the amount of data it needs, the complexity of the Rønjom-Helleseth attack roughly equals $C^{2/3}$ and the amount of data it needs is roughly C'^2. From the viewpoint of complexity, it is more efficient and from the viewpoint of data it is less efficient.
The whole Section 9 is devoted to the algebraic immunity of Boolean functions.

4.1.7 Other criteria

- The second moment of the auto-correlation coefficients:

$$\mathcal{V}(f) = \sum_{e \in \mathbb{F}_2^n} \mathcal{F}^2(D_e f) \tag{38}$$

has been introduced by Zhang and Zheng [359] for measuring the *global avalanche criterion* (GAC), and also called the *sum-of-squares indicator*. The *absolute indicator* is by definition $\max_{e \in \mathbb{F}_2^n,\ e \neq 0} \mid \mathcal{F}(D_e f) \mid$. Both indicators are clearly affine invariants. In order to achieve good diffusion, cryptographic functions should have low sum-of-squares indicators and absolute indicators. Obviously, we have $\mathcal{V}(f) \geq 2^{2n}$, since $\mathcal{F}^2(D_0 f) = 2^{2n}$. Note that every lower bound of the form $\mathcal{V}(f) \geq V$ straightforwardly implies that the absolute indicator is bounded below by $\sqrt{\frac{V-2^{2n}}{2^n-1}}$. The functions achieving $\mathcal{V}(f) = 2^{2n}$ are those functions whose derivatives $D_e f(x)$, $e \neq 0$, are all balanced. We shall see in Section 6 that these are the bent functions. If f has a k-dimensional linear kernel, then $\mathcal{V}(f) \geq 2^{2n+k}$ (with equality if and only if f

[30] As already seen, it can be taken equal to $\log_2 7 \approx 2.8$ and the coefficient in the O can be taken equal to 7, according to Strassen [340]; a still better exponent is due to Coppersmith and Winograd but the multiplicative constant is then inefficiently high for our framework.

is partially bent, see below).
Note that, according to Relation (26) applied to $D_e f$ for every e, we have

$$\mathcal{V}(f) = \sum_{a,e \in \mathbb{F}_2^n} \mathcal{F}(D_a D_e f),$$

where $D_a D_e f(x) = f(x) \oplus f(x+a) \oplus f(x+e) \oplus f(x+a+e)$ is the *second order derivative* of f.
Note also that, according to Relation (21) applied to $\varphi(e) = \psi(e) = \mathcal{F}(D_e f)$, we have, for any n-variable Boolean function f:

$$\forall a \in \mathbb{F}_2^n,\ \sum_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(e) \widehat{\mathrm{f}_\chi}^2(a+e) = 2^n \sum_{e \in \mathbb{F}_2^n} \mathcal{F}^2(D_e f)(-1)^{e \cdot a}\ ,$$

as shown in [42] (indeed, the Fourier transform of φ equals $\widehat{\mathrm{f}_\chi}^2$, according to Relation (25)), and thus, for $a = 0$:

$$\sum_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^4(e) = 2^n\, \mathcal{V}(f). \tag{39}$$

We have: $\sum_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^4(e) \leq \left(\sum_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(e) \right) \left(\max_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(e) \right) \leq 2^n \max_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^4(e)$.
According to Parseval's relation $\sum_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(e) = 2^{2n}$, we deduce, using Relation (39): $\max_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(e) \geq \frac{\mathcal{V}(f)}{2^n} \geq \sqrt{\mathcal{V}(f)}$; thus, according to Relation (35) relating the nonlinearity to the Walsh transform, we have (as first shown in [360, 363]):

$$nl(f) \leq 2^{n-1} - 2^{-n/2-1} \sqrt{\mathcal{V}(f)} \leq 2^{n-1} - \frac{1}{2} \sqrt[4]{\mathcal{V}(f)}.$$

Denoting again by $N_{\widehat{\mathrm{f}_\chi}}$ the cardinality of the support $\{a \in \mathbb{F}_2^n / \ \widehat{\mathrm{f}_\chi}(a) \neq 0\}$ of the Walsh transform of f, Relation (39) also implies the following relation, first observed in [363]: $\mathcal{V}(f) \times N_{\widehat{\mathrm{f}_\chi}} \geq 2^{3n}$. Indeed, using for instance the Cauchy-Schwartz inequality, we see that $\left(\sum_{a \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(a) \right)^2 \leq \left(\sum_{a \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^4(a) \right) \times N_{\widehat{\mathrm{f}_\chi}}$ and we have $\sum_{a \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(a) = 2^{2n}$, according to Parseval's Relation (23). Clearly, the functions satisfying $nl(f) = 2^{n-1} - 2^{-n/2-1} \sqrt{\mathcal{V}(f)}$ (resp. $\mathcal{V}(f) \times N_{\widehat{\mathrm{f}_\chi}} = 2^{3n}$) are the functions whose Walsh transforms take at most one nonzero magnitude. These functions are called

plateaued functions (see Subsection 6.8 for further properties of plateaued functions). The functions satisfying $nl(f) = 2^{n-1} - \frac{1}{2}\sqrt[4]{\mathcal{V}(f)}$ are (also clearly) the bent functions.
Constructions of balanced Boolean functions with low absolute indicators and high nonlinearities have been studied in [260].

- The *maximum correlation* of an n-variable Boolean function f with respect to a subset I of $N = \{1, \cdots, n\}$ equals by definition (see [358]) $C_f(I) = \max_{g \in \mathcal{BF}_{I,n}} \frac{\mathcal{F}(f \oplus g)}{2^n}$, where $\mathcal{BF}_{I,n}$ is the set of n-variable Boolean functions depending on $\{x_i, i \in I\}$ only. According to Relation (13), the distance from f to $\mathcal{BF}_{I,n}$ equals $2^{n-1}(1 - C_f(I))$. As we saw above, denoting the size of I by r, this distance is bounded below by the r-th order nonlinearity.
The maximum correlation of any combining function with respect to any subset I of small size should be small (*i.e.* its distance to $\mathcal{BF}_{I,n}$ should be high). It is straightforward to prove, by decomposing the sum $\mathcal{F}(f \oplus g)$, that $C_f(I)$ equals $\sum_{j=1}^{2^{|I|}} \frac{|\mathcal{F}(h_j)|}{2^n}$, where $h_1, \cdots, h_{2^{|I|}}$ are the restrictions of f obtained by keeping constant the x_i's for $i \in I$, and to see that the distance from f to $\mathcal{BF}_{I,n}$ is achieved by the functions g taking value 0 (resp. 1) when the corresponding value of $\mathcal{F}(h_j)$ is positive (resp. negative), and that we have $C_f(I) = 0$ if and only if all h_j's are balanced (thus, f is m-resilient if and only if $C_f(I) = 0$ for every set I of size at most m). Also, according to the Cauchy-Schwartz inequality, we have $\left(\sum_{j=1}^{2^{|I|}} |\mathcal{F}(h_j)|\right)^2 \leq 2^{|I|} \sum_{j=1}^{2^{|I|}} \mathcal{F}^2(h_j)$, and Relation (28) directly implies the following inequality observed in [38]:

$$C_f(I) \leq 2^{-n} \left(\sum_{u \in \mathbb{F}_2^n / \, u_i = 0, \forall i \notin I} \widehat{f_\chi}^2(u) \right)^{\frac{1}{2}} \leq 2^{-n + \frac{|I|}{2}} \left(2^n - 2nl(f)\right) \qquad (40)$$

or equivalently:

$$d_H\left(f, \mathcal{BF}_{I,n}\right) \geq 2^{n-1} - \frac{1}{2} \left(\sum_{\substack{u \in \mathbb{F}_2^n / \\ supp(u) \subseteq I}} \widehat{f_\chi}^2(u) \right)^{\frac{1}{2}} \geq 2^{n-1} - 2^{\frac{|I|}{2} - 1} \max_{u \in \mathbb{F}_2^n} |\widehat{f_\chi}(u)|.$$

This inequality shows that, contrary to the case of approximations by functions of algebraic degrees at most r (higher order nonlinearity), it is sufficient that the first-order nonlinearity of a combining function be high for avoiding close approximations of f by functions of $\mathcal{BF}_{I,n}$ (when I has small size).

An affine invariant criterion related to the maximum correlation and also related to the "distance to linear structures" is the following: the distance to the Boolean functions g such that the space $\{e \in \mathbb{F}_2^n / D_e g = 0\}$ has dimension at least k (the functions of $\mathcal{BF}_{I,n}$ can be viewed as n-variable functions g such that the set $\{e \in \mathbb{F}_2^n / D_e g = 0\}$ contains $\mathbb{F}_2^{N \setminus I}$). The results on the maximum correlation above generalize to this criterion [38].

- The main cryptographic *complexity criteria* for a Boolean function are the algebraic degree and the nonlinearity, but other criteria have also been studied: the minimum number of terms in the algebraic normal forms of all affinely equivalent functions, called the *algebraic thickness* (studied in [66] and first evoked in [276]), the maximum dimension k of those flats E such that the restriction of f to E is constant (f is then called a *k-normal function*) or is affine (f is called a *k-weakly-normal function*) [66] (see Subsection 5.4), the number of nonzero coefficients of the Walsh transform [302, 316]. It has been shown in [66, 289, 316] that (asymptotically) almost all Boolean functions have high complexities with respect to all these criteria.
For every even integer k such that $4 \leq k \leq 2^n$, the kth-order *nonhomomorphicity* [362] of a Boolean function equals the number of k-tuples $(u_1, \cdots, u_k)$ of vectors of $\mathbb{F}_2^n$ such that $u_1 + \cdots + u_k = 0$ and $f(u_1) \oplus \cdots \oplus f(u_k) = 0$. It is a simple matter to show (more directly than in [362]) that it equals $2^{(k-1)n-1} + 2^{-n-1} \sum_{u \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^k(u)$. This parameter should be small (but no related attack exists on stream ciphers). It is maximum and equals $2^{(k-1)n}$ if and only if the function is affine. It is minimum and equals $2^{(k-1)n-1} + 2^{\frac{nk}{2}-1}$ if and only if the function is bent, and some relationship obviouly exists between nonhomomorphicity and nonlinearity.

Conclusion of this subsection: As we can see, there are numerous cryptographic criteria for Boolean functions. The ones which must be necessarily satisfied are balancedness, a high algebraic degree, a high nonlinearity, a high algebraic immunity and a good resistance to fast algebraic attacks. It is difficult but not impossible to find functions satisfying good trade-offs between all these criteria (see Section 9). It is not clear whether it is possible to achieve additionally resiliency of a sufficient order, which is necessary for the combiner model. Hence, the filter model may be more appropriate (future research will determine this). Once we know the criteria above are satisfied by some function f (except resiliency), it is a simple matter to render f 1-resilient by composing it with a linear automorphism (we just

need for this that there exist n linearly independent vectors at which the Walsh transform vanishes). First-order resiliency is useful for resisting some distinguishing (less dreadful) attacks.

5 Classes of functions for which restrictions on the possible values of the weights, Walsh spectra and nonlinearities can be proved

5.1 Affine functions

The weights and the Walsh spectra of affine functions are peculiar: the Walsh transform of the function $\ell(x) = a \cdot x \oplus \varepsilon$ takes null value at every vector $u \neq a$ and takes value $2^n (-1)^\varepsilon$ at a.
Concatenating affine functions gives the so-called Maiorana-McFarland functions: for every n-variable function f, if we order all the binary words of length n in lexicographic order, with the bit of higher weight on the right (for instance), then the truth-table of f is the concatenation of the restrictions of f obtained by setting the values of the (say) s last bits of the input and letting the others freely range over $\mathbb{F}_2$. If all these restrictions are affine then f is called a Maiorana-McFarland function. These Maiorana-McFarland functions will be studied in Section 6 (Subsection 6.4, for bent functions) and Section 7 (Subsection 7.5, for resilient functions). The computation of their weights, Walsh spectra and nonlinearities are easier than for general Boolean functions, and in some cases can be completely determined.

5.2 Quadratic functions

The behavior of the functions of $R(2, n)$, called *quadratic functions*, is also peculiar. Recall that Relation (26) states that, for every Boolean function f:

$$\mathcal{F}^2(f) = \sum_{b \in \mathbb{F}_2^n} \mathcal{F}(D_b f).$$

If f is quadratic, then $D_b f$ is affine for every $b \in \mathbb{F}_2^n$, and is therefore either balanced or constant. Since $\mathcal{F}(g) = 0$ for every balanced function g, we deduce:

$$\mathcal{F}^2(f) = 2^n \sum_{b \in \mathcal{E}_f} (-1)^{D_b f(0)}, \tag{41}$$

where $\mathcal{E}_f$ is the set of all $b \in \mathbb{F}_2^n$ such that $D_b f$ is constant. The set $\mathcal{E}_f$ is the linear kernel of f (see Subsection 4.1). In the case of quadratic functions,

it also equals the kernel $\{x \in \mathbb{F}_2^n / \ \forall y \in \mathbb{F}_2^n,\ \varphi_f(x,y) = 0\}$ of the symplectic (*i*.e. bilinear, symmetric, and null over the diagonal) form associated to f: $\varphi_f(x,y) = f(0) \oplus f(x) \oplus f(y) \oplus f(x+y)$. The restriction of the function $b \mapsto D_b f(0) = f(b) \oplus f(0)$ to this vectorspace is linear, as can be easily checked; we deduce that $\mathcal{F}^2(f)$ equals $2^n |\mathcal{E}_f|$ if this linear form on $\mathcal{E}_f$ is null, that is, if f is constant on $\mathcal{E}_f$, and is null otherwise. According to Relation (13), this proves the following:

Theorem 4 *Any quadratic function f is balanced if and only if its restriction to its linear kernel $\mathcal{E}_f$ (*i.e. *the kernel of its associated symplectic form) is not constant. If it is not balanced, then its weight equals $2^{n-1} \pm 2^{\frac{n+k}{2}-1}$ where k is the dimension of $\mathcal{E}_f$.*

Note that Theorem 4 implies that f is balanced if and only if there exists $b \in \mathbb{F}_2^n$ such that the derivative $D_b f(x) = f(x) \oplus f(x+b)$ equals the constant function 1 (take b in $\mathcal{E}_f$ such that $f(b) \neq f(0)$). For general Boolean functions, this condition is sufficient for f being balanced, but it is not necessary. Theorem 4 applied to $f \oplus \ell$, where ℓ is a linear function such that $f \oplus \ell$ is not balanced (such function ℓ always exists, according to Parseval's relation) shows that the co-dimension of $\mathcal{E}_f$ must be even (this co-dimension is the *rank of φ_f*).

The weight of a quadratic function can be any element of the set $\{2^{n-1}\} \cup \{2^{n-1} \pm 2^i;\ i = \lceil \frac{n}{2} \rceil - 1, \cdots, n-1\}$. Its nonlinearity can be any element of the set $\{2^{n-1} - 2^i;\ i = \frac{n}{2} - 1, \cdots, n-1\}$, and if f has weight $2^{n-1} \pm 2^i$, then for every affine function l, the weight of the function $f \oplus l$ belongs to the set $\{2^{n-1} - 2^i, 2^{n-1}, 2^{n-1} + 2^i\}$.

Determining whether the weight is $2^{n-1} - 2^i$ or $2^{n-1} + 2^i$ (when the function is not balanced), and more generally studying the sign of the Walsh transform is in general much more difficult than determining the value of i, or equivalently the magnitude of the Walsh transform. In [226] is studied the sign of the values of the Walsh transform of Gold and Kasami functions. The former are quadratic (the latter are not but they are related to quadratic functions, see the chapter "Vectorial Boolean Functions for Cryptography"). In [164], the result of [226] is generalized: for every AB power function x^d over $\mathbb{F}_{2^n}$ (see definition in the chapter "Vectorial Boolean Functions for Cryptography") whose restriction to any subfield of $\mathbb{F}_{2^n}$ is also AB, the value $\sum_{x \in \mathbb{F}_{2^n}} (-1)^{tr_n(x^d + x)}$ equals $2^{\frac{n+1}{2}}$ if $n \equiv \pm 1$ [mod 8] and $-2^{\frac{n+1}{2}}$ if $n \equiv \pm 3$ [mod 8].

Any quadratic non-affine function f having a monomial of degree 2 in its ANF, we can assume without loss of generality that, up to a non-singular linear transformation, this monomial is $x_1 x_2$. The function has

then the form $x_1x_2 \oplus x_1f_1(x_3,\cdots,x_n) \oplus x_2f_2(x_3,\cdots,x_n) \oplus f_3(x_3,\cdots,x_n)$ where f_1, f_2 are affine functions and f_3 is quadratic. Then, $f(x)$ equals $(x_1 \oplus f_2(x_3,\cdots,x_n))(x_2 \oplus f_1(x_3,\cdots,x_n)) \oplus f_1(x_3,\cdots,x_n)f_2(x_3,\cdots,x_n) \oplus f_3(x_3,\cdots,x_n)$ and is therefore affinely equivalent to the function $x_1x_2 \oplus f_1(x_3,\cdots,x_n)f_2(x_3,\cdots,x_n) \oplus f_3(x_3,\cdots,x_n)$. Applying this method recursively shows:

Theorem 5 *Every quadratic non-affine function is affinely equivalent to* $x_1x_2 \oplus \cdots \oplus x_{2l-1}x_{2l} \oplus x_{2l+1}$ *(where* $l \leq \frac{n-1}{2}$*) if it is balanced, to* $x_1x_2 \oplus \cdots \oplus x_{2l-1}x_{2l}$ *(where* $l \leq n/2$*) if it has weight smaller than* 2^{n-1} *and to* $x_1x_2 \oplus \cdots \oplus x_{2l-1}x_{2l} \oplus 1$ *(where* $l \leq n/2$*) if it has weight greater than* 2^{n-1}.

This allows describing precisely the weight distribution of $R(2,n)$ [258].

Remark. Let f_1, f_2 and f_3 be any Boolean functions on $\mathbb{F}_2^n$. Define the function on $\mathbb{F}_2^{n+2}$: $f(x,y_1,y_2) = y_1y_2 \oplus y_1f_1(x) \oplus y_2f_2(x) \oplus f_3(x)$. Then we have

$$\mathcal{F}(f) = \sum_{x\in\mathbb{F}_2^n/\ y_1,y_2\in\mathbb{F}_2} (-1)^{(y_1\oplus f_2(x))(y_2\oplus f_1(x))\oplus f_1(x)f_2(x)\oplus f_3(x)}$$

$$= \sum_{x\in\mathbb{F}_2^n/\ y_1,y_2\in\mathbb{F}_2} (-1)^{y_1y_2\oplus f_1(x)f_2(x)\oplus f_3(x)} = 2\sum_{x\in\mathbb{F}_2^n} (-1)^{f_1(x)f_2(x)\oplus f_3(x)}.$$

So, starting with a function $g = f_1f_2 \oplus f_3$, we can relate $\mathcal{F}(g)$ to $\mathcal{F}(f)$, on two more variables, in which the term f_1f_2 has been replaced by $y_1y_2 \oplus y_1f_1(x) \oplus y_2f_2(x)$. Applying this repeatedly ("breaking" this way all the monomials of degrees at least 4), this allows showing easily (see [52]) that, for every Boolean function g on $\mathbb{F}_2^n$, there exists an integer m and a Boolean function f of algebraic degree at most 3 on $\mathbb{F}_2^{n+2m}$ whose Walsh transform takes value $\widehat{f_\chi}(0) = 2^m\,\widehat{g_\chi}(0)$ at 0. As we already mentioned, this proves that the functions of algebraic degree 3 can have weights much more diverse than functions of degrees at most 2.

The trace representation of quadratic functions is $tr_n\left(\beta_\emptyset + \sum_{i=0}^{\frac{n-1}{2}} \beta_i\, x^{2^i+1}\right)$ for n odd and $tr_n\left(\beta_\emptyset + \sum_{i=0}^{\frac{n}{2}-1} \beta_i\, x^{2^i+1}\right) + tr_{\frac{n}{2}}(\gamma x^{2^{n/2}+1})$ for n even, where the β_i's belong to $\mathbb{F}_{2^n}$ and γ belongs to $\mathbb{F}_{2^{n/2}}$. For n odd, the quadratic functions of nonlinearity $2^{n-1} - 2^{\frac{n-1}{2}}$ (called *semi-bent functions* because their extended Walsh spectra only contain the values 0 and $2^{\frac{n+1}{2}}$) of the form $tr_n(\sum_{i=1}^{(n-1)/2} c_i x^{2^i+1})$ have been studied in [106] and cubic bent functions have been deduced by concatenation of these semi-bent functions. Further

functions of this kind have been given and studied in [197, 217].
Concatenating quadratic functions gives a super-class of the class of Maiorana-McFarland functions, studied in [64], and presented in Section 7 (Subsection 7.5.2) below.

5.3 Indicators of flats

As we have already seen, a Boolean function f is the indicator of a flat A of co-dimension r if and only if it has the form $f(x) = \prod_{i=1}^{r}(a_i \cdot x \oplus \varepsilon_i)$ where $a_1, \cdots, a_r \in \mathbb{F}_2^n$ are linearly independent and $\varepsilon_1, \cdots, \varepsilon_r \in \mathbb{F}_2$. Then f has weight 2^{n-r}. Moreover, set $a \in \mathbb{F}_2^n$. If a is linearly independent of $a_1, \cdots, a_r$, then the function $f(x) \oplus a \cdot x$ is balanced (and hence $\widehat{\mathrm{f}_\chi}(a) = 0$), since it is linearly equivalent to a function of the form $g(x_1, \cdots, x_r) \oplus x_{r+1}$. If a is linearly dependent of $a_1, \cdots, a_r$, say $a = \sum_{i=1}^{r} \eta_i \, a_i$, then $a \cdot x$ takes constant value $\bigoplus_{i=1}^{r} \eta_i \, (a_i \cdot x) = \bigoplus_{i=1}^{r} \eta_i \, (\epsilon_i \oplus 1)$ on the flat; hence, $\widehat{f}(a) = \sum_{x \in A}(-1)^{a \cdot x}$ equals then $2^{n-r}(-1)^{\bigoplus_{i=1}^{r} \eta_i \, (\epsilon_i \oplus 1)}$. Thus, if $a = \sum_{i=1}^{r} \eta_i \, a_i \neq 0$, then we have $\widehat{\mathrm{f}_\chi}(a) = -2^{n-r+1}(-1)^{\bigoplus_{i=1}^{r} \eta_i \, (\epsilon_i \oplus 1)}$; and we have $\widehat{\mathrm{f}_\chi}(0) = 2^n - 2|A| = 2^n - 2^{n-r+1}$.
Note that the nonlinearity of f equals 2^{n-r} and is bad as soon as $r \geq 2$. But indicators of flats can be used to design Boolean functions with good nonlinearities: concatenating sums of indicators of flats and of affine functions gives another super-class of the Maiorana-McFarland functions, studied in [68] and presented in Section 7 (Subsection 7.5.2) below.

Note. As recalled in Section 3.1, the functions of $R(r, n)$ whose weights occur in the range $[2^{n-r}; 2^{n-r+1}[$ have been characterized by Kasami and Tokura [207]; any such function is the product of the indicator of a flat and of a quadratic function or is the sum (modulo 2) of two indicators of flats. The Walsh spectra of such functions can also be precisely computed.

5.4 Normal functions

Let E and E' be subspaces of $\mathbb{F}_2^n$ such that $E \cap E' = \{0\}$ and whose direct sum equals $\mathbb{F}_2^n$. Denote by k the dimension of E. For every $a \in E'$, let h_a be the restriction of f to the coset $a + E$. Then, Relation (28) in Proposition 9 implies

$$\max_{u \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(u) \geq \sum_{a \in E'} \mathcal{F}^2(h_a)$$

(indeed, the maximum of $\widehat{\mathrm{f}_\chi}^2(u)$ is greater than or equal to its mean). Hence we have: $\max_{u\in\mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(u) \geq \mathcal{F}^2(h_a)$ for every a. Applying this property to $f \oplus \ell$, where ℓ is any linear function, and using Relation (35) relating the nonlinearity of a function to the maximum magnitude of its Walsh transform, we deduce:

$$\forall a \in E',\ nl(f) \leq 2^{n-1} - 2^{k-1} + nl(h_a). \tag{42}$$

This bound was first proved (in a different way) by Zheng et al. in [364]. The present proof is from [42]. Relation (42) can also be deduced from the Poisson summation formula (17) applied to the sign function of f, and in which the roles of E and $E^\perp$ are exchanged: let us choose $b \in \mathbb{F}_2^n$ such that $\left|\sum_{x\in a\oplus E}(-1)^{f(x)\oplus b\cdot x}\right|$ is maximum, that is, equals $\left(2^k - 2nl(h_a)\right)$. Then

$$\left|\sum_{u\in b\oplus E^\perp} (-1)^{a\cdot u}\, \widehat{\mathrm{f}_\chi}(u)\right| = |E^\perp| \left(2^k - 2nl(h_a)\right).$$

Then the mean of $(-1)^{a\cdot u}\widehat{\mathrm{f}_\chi}(u)$, when u ranges over $b \oplus E^\perp$, is equal to $\pm\left(2^k - 2nl(h_a)\right)$. Thus, the maximum magnitude of $\widehat{\mathrm{f}_\chi}(u)$ is greater than or equal to $2^k - 2nl(h_a)$. This implies Relation (42). These two methods, for proving (42), lead to two different necessary conditions for the case of equality (see [66]).
Relation (42) implies in particular that, if the restriction of f to a k-dimensional flat of $\mathbb{F}_2^n$ is affine (say equals ℓ), then $nl(f) \leq 2^{n-1} - 2^{k-1}$, and that, if equality occurs, then $f \oplus \ell$ is balanced on every other coset of this flat.

Definition 4 *A function is called k-weakly-normal (resp. k-normal) if its restriction to some k-dimensional flat is affine (resp. constant).*

H. Dobbertin introduced this terminology by calling normal the functions that we call $n/2$-normal here (we shall also call normal the $n/2$-normal functions, in the sequel). He used this notion for constructing balanced functions with high nonlinearities (see Subsection 7.5.1). It is proved in [66] that, for every $\alpha > 1$, when n tends to infinity, random Boolean functions are almost surely $[\alpha \log_2 n]$-non-normal. This means that almost all Boolean functions have high complexity with respect to this criterion. As usual, the proof of existence of non-normal functions does not give examples of such functions. Alon, Goldreich, Hastad and Peralta give in [2] several constructions of functions which are nonconstant on flats of dimension $n/2$. This is not explicitly mentioned in the paper. What is shown is that the functions are

not constant on flats defined by equations $x_{i_1} = a_1, ..., x_{i_{n/2}} = a_{n/2}$. As the proof still works when composing the function by an affine automorphism, it implies the result.
There are also explicit constructions which work for dimensions $(1/2 - \epsilon)\, n$, for some small $\epsilon > 0$ very recently found by Jean Bourgain [24].
Functions which are nonconstant on flats of dimensions n^δ for every $\delta > 0$ are also given in [14]. These constructions are very good asymptotically (but may not be usable to obtain functions in explicit numbers of variables).
As far as we know, no construction is known below n^δ.

5.5 Functions admitting partial covering sequences

The notion of covering sequence of a Boolean function has been introduced in [95].

Definition 5 *Let f be an n-variable Boolean function. An integer-valued*[31] *sequence $(\lambda_a)_{a\in\mathbb{F}_2^n}$ is called a* covering sequence *of f if the integer-valued function $\sum_{a\in\mathbb{F}_2^n} \lambda_a D_a f(x)$ takes a constant value. This constant value is called the* level of a covering sequence. *If the level is nonzero, we say that the covering sequence is a* non-trivial covering sequence.

Note that the sum $\sum_{a\in\mathbb{F}_2^n} \lambda_a D_a f(x)$ involves both kinds of additions: the addition $\sum$ in $\mathbb{Z}$ and the addition $\oplus$ in $\mathbb{F}_2$ (which is concealed inside $D_a f$).
It was shown in [95] that any function admitting a non-trivial covering sequence is balanced (see Theorem 6 below for a proof) and that any balanced function admits the constant sequence 1 as covering sequence (the level of this sequence is 2^{n-1}).
A characterization of covering sequences by means of the Walsh transform was also given in [95]: denote again by $S_{\widehat{f_\chi}}$ the support $\{u \in \mathbb{F}_2^n \mid \widehat{f_\chi}(u) \neq 0\}$ of $\widehat{f_\chi}$; then f admits an integer-valued sequence $\lambda = (\lambda_a)_{a\in\mathbb{F}_2^n}$ as covering sequence if and only if the Fourier transform $\widehat{\lambda}$ of the function $a \mapsto \lambda_a$ takes a constant value on $S_{\widehat{f_\chi}}$. Indeed, replacing $D_a f(x)$ by $\frac{1}{2} - \frac{1}{2}(-1)^{D_a f(x)} = \frac{1}{2} - \frac{1}{2}(-1)^{f(x)}(-1)^{f(x+a)}$ in the equality $\sum_{a\in\mathbb{F}_2^n} \lambda_a D_a f(x) = \rho$, we see that f admits the covering sequence λ with level ρ if and only if, for every $x \in \mathbb{F}_2^n$, we have $\sum_{a\in\mathbb{F}_2^n} \lambda_a (-1)^{f(x+a)} = \left(\sum_{a\in\mathbb{F}_2^n} \lambda_a - 2\rho\right)(-1)^{f(x)}$; since two integer-valued functions are equal if and only if their Fourier transforms

[31] or real-valued, or complex-valued; but taking real or complex sequences instead of integer-valued ones has no practical sense.

are equal, the characterization follows, thanks to the straightforward relation $\sum_{a,x\in\mathbb{F}_2^n}\lambda_a(-1)^{f(x+a)+x\cdot b}=\left(\sum_{a\in\mathbb{F}_2^n}\lambda_a(-1)^{a\cdot b}\right)\widehat{\mathrm{f}_\chi}(b)=\widehat{\lambda}(b)\,\widehat{\mathrm{f}_\chi}(b)$.
Knowing a covering sequence (trivial or not) of a function f allows knowing that all the vectors a such that $f(x)\oplus a\cdot x$ is non-balanced belong to the set $\widehat{\lambda}^{-1}(\mu)$, where $\mu=\widehat{\lambda}(0)-2\rho$ is the constant value of $\widehat{\lambda}$ on $S_{\widehat{\mathrm{f}_\chi}}$; hence, if f admits a covering sequence $\lambda=(\lambda_a)_{a\in\mathbb{F}_2^n}$ with level ρ (resp. with level $\rho\neq 0$), then f is k-th order correlation-immune (resp. k-resilient) where $k+1$ is the minimum Hamming weight of nonzero $b\in\mathbb{F}_2^n$ such that $\widehat{\lambda}(b)=\mu$. Conversely, if f is k-th order correlation-immune (resp. k-resilient) and if it is not $(k+1)$-th order correlation-immune (resp. $(k+1)$-resilient), then there exists at least one (non-trivial) covering sequence $\lambda=(\lambda_a)_{a\in\mathbb{F}_2^n}$ with level ρ such that $k+1$ is the minimum Hamming weight of $b\in\mathbb{F}_2^n$ satisfying $\widehat{\lambda}(b)=\widehat{\lambda}(0)-2\rho$.
A particularly simple covering sequence is the indicator of the set of vectors of weight one. The functions which admit this covering sequence are called regular; they are $(\rho-1)$-resilient (where ρ is the level); more generally, any function, admitting as covering sequence the indicator of a set of vectors whose supports are disjoint, has this same property. See further properties in [95].
But knowing a covering sequence for f gives no information on the nonlinearity of f, since it gives only information on the support of the Walsh transform, not on the nonzero values it takes. In [69] is weakened the definition of covering sequence, so that it can help computing the (nonzero) values of the Walsh transform.

Definition 6 *Let f be a Boolean function on $\mathbb{F}_2^n$. A partial covering sequence for f is a sequence $(\lambda_a)_{a\in\mathbb{F}_2^n}$ such that $\sum_{a\in\mathbb{F}_2^n}\lambda_a D_a f(x)$ takes two values ρ and ρ' (distinct or not) called the* levels *of the sequence. The partial covering sequence is called* non-trivial *if one of the constants is nonzero.*

A simple example of non-trivial partial covering sequence is as follows: let $\mathcal{E}$ be any set of derivatives of f. Assume that $\mathcal{E}$ contains a nonzero function and is stable under addition (*i*.e. is a non-trivial $\mathbb{F}_2$-vectorspace). Then $\sum_{g\in\mathcal{E}}g$ takes on values 0 and $\frac{|\mathcal{E}|}{2}$. Thus, if $\mathcal{E}=\{D_af/\ a\in E\}$ (where we choose E minimum, so that any two different vectors of the set E give different functions of $\mathcal{E}$), then 1_E is a non-trivial partial covering sequence.

The interest of non-trivial partial covering sequences is that they allow simplifying the computation of the weight and of the Walsh transform of f.

Theorem 6 *Let $(\lambda_a)_{a\in\mathbb{F}_2^n}$ be a partial covering sequence of a Boolean function f, of levels ρ and ρ'.*
Let $A = \{x \in \mathbb{F}_2^n \,/\, \sum_{a\in\mathbb{F}_2^n} \lambda_a D_a f(x) = \rho'\}$ (assuming that $\rho' \neq \rho$; otherwise, when λ is in fact a covering sequence of level ρ, we set $A = \emptyset$).
Then, for every vector $b \in \mathbb{F}_2^n$, we have:

$$\left(\widehat{\lambda}(b) - \widehat{\lambda}(0) + 2\,\rho\right)\ \widehat{\mathrm{f}_\chi}(b) = 2\,(\rho - \rho') \sum_{x\in A} (-1)^{f(x)\oplus b\cdot x}.$$

Proof. By definition, we have, for every $x \in \mathbb{F}_2^n$:

$$\sum_{a\in\mathbb{F}_2^n} \lambda_a D_a f(x) = \rho'\, 1_A(x) + \rho\, 1_{A^c}(x)$$

and therefore:

$$\begin{aligned}\sum_{a\in\mathbb{F}_2^n} \lambda_a (-1)^{D_a f(x)} &= \sum_{a\in\mathbb{F}_2^n} \lambda_a (1 - 2\, D_a f(x)) \\ &= \sum_{a\in\mathbb{F}_2^n} \lambda_a - 2\,\rho'\, 1_A(x) - 2\,\rho\, 1_{A^c}(x).\end{aligned}$$

We deduce:

$$\sum_{a\in\mathbb{F}_2^n} \lambda_a (-1)^{f(x+a)} = (-1)^{f(x)} \left(\sum_{a\in\mathbb{F}_2^n} \lambda_a - 2\,\rho'\, 1_A(x) - 2\,\rho\, 1_{A^c}(x)\right). \quad (43)$$

The Fourier transform of the function $(-1)^{f(x+a)}$ maps every vector $b \in \mathbb{F}_2^n$ to the value $\sum_{x\in\mathbb{F}_2^n} (-1)^{f(x+a)\oplus x\cdot b} = \sum_{x\in\mathbb{F}_2^n} (-1)^{f(x)\oplus(x+a)\cdot b} = (-1)^{a\cdot b}\, \widehat{\mathrm{f}_\chi}(b)$.
Hence, taking the Fourier transform of both terms of equality (43), we get:

$$\left(\sum_{a\in\mathbb{F}_2^n} \lambda_a (-1)^{a\cdot b}\right) \widehat{\mathrm{f}_\chi}(b) =$$

$$\left(\sum_{a\in\mathbb{F}_2^n} \lambda_a\right) \widehat{\mathrm{f}_\chi}(b) - 2\,\rho' \sum_{x\in A} (-1)^{f(x)\oplus b\cdot x} - 2\,\rho \sum_{x\in A^c} (-1)^{f(x)\oplus b\cdot x},$$

that is

$$\widehat{\lambda}(b)\, \widehat{\mathrm{f}_\chi}(b) = \widehat{\lambda}(0)\, \widehat{\mathrm{f}_\chi}(b) - 2\,\rho\, \widehat{\mathrm{f}_\chi}(b) + 2\,(\rho - \rho') \sum_{x\in A} (-1)^{f(x)\oplus b\cdot x}.$$

Hence:

$$\left(\widehat{\lambda}(b) - \widehat{\lambda}(0) + 2\,\rho\right)\ \widehat{\mathrm{f}_\chi}(b) = 2\,(\rho - \rho')\ \sum_{x\in A}(-1)^{f(x)\oplus b\cdot x}. \qquad \square$$

Hence, if $\rho \neq 0$, we have in particular an information on the weight of f:

$$2^n - 2w_H(f) = \widehat{\mathrm{f}_\chi}(0) = \left(1 - \frac{\rho'}{\rho}\right) \sum_{x\in A}(-1)^{f(x)}.$$

Examples are given in [69] of computations of the weights or Walsh spectra of some Boolean functions (quadratic functions, Maiorana-McFarland's functions and their extensions, and other examples of functions), using Theorem 6.

5.6 Functions with low univariate degree

The following Weil's Theorem is very well-known in finite field theory (*cf.* [248, Theorem 5.38]):

Theorem 7 *Let q be a prime power and $f \in \mathbb{F}_q[x]$ a univariate polynomial of degree $d \geq 1$ with $gcd(d, q) = 1$. Let χ be a non-trivial character of $\mathbb{F}_q$. Then*

$$\left|\sum_{x\in\mathbb{F}_q} \chi(f(x))\right| \leq (d-1)\,q^{1/2}.$$

For $q = 2^n$, this *Weil's bound* means that, for every nonzero $a \in \mathbb{F}_{2^n}$: $\left|\sum_{x\in\mathbb{F}_{2^n}}(-1)^{tr_n(af(x))}\right| \leq (d-1)\,2^{n/2}$. And since adding a linear function $tr_n(bx)$ to the function $tr_n(af(x))$ corresponds to adding $(b/a)\,x$ to $f(x)$ and does not change its univariate degree, we deduce that, if $d > 1$ is odd and $a \neq 0$, then:

$$nl(tr_n(af)) \geq 2^{n-1} - (d-1)\,2^{n/2-1}.$$

An extension of the Weil bound to the character sums of functions of the form $f(x) + g(1/x)$ (where $1/x = x^{2^n-2}$ takes value 0 at 0), among which are the so-called *Kloosterman sums* $\sum_{x\in\mathbb{F}_{2^n}}(-1)^{tr_n(1/x+ax)}$, has been first obtained by Carlitz and Uchiyama [97] and extended by Shanbhag, Kumar and Helleseth [330]: if f and g have odd univariate degrees, then

$$\sum_{x\in\mathbb{F}_{2^n}} (-1)^{tr_n(f(1/x)+g(x))} \leq (d^\circ f + d^\circ g)2^{n/2}.$$

6 Bent functions

We recall the definition of bent functions:

Definition 7 *A Boolean function f on $\mathbb{F}_2^n$ (n even) is called bent if its Hamming distance to the set $R(1,n)$ of all n-variable affine functions (the nonlinearity of f) equals $2^{n-1} - 2^{n/2-1}$ (the covering radius of the Reed-Muller code of order 1).*

Equivalently, as seen in Subsection 4.1, f is bent if and only if $\widehat{f_\chi}$ takes on values $\pm 2^{n/2}$ only (this characterization is independent of the choice of the inner product on $\mathbb{F}_2^n$, since any other inner product has the form $\langle x, s\rangle = x \cdot L(s)$, where L is an auto-adjoint linear automorphism, *i*.e. an automorphism whose associated matrix is symmetric). Hence, f is bent if and only if its distance to any affine function equals $2^{n-1} \pm 2^{n/2-1}$. Note that, for any bent function f, half of the elements of the Reed-Muller code of order 1 lie at distance $2^{n-1} + 2^{n/2-1}$ from f and half lie at distance $2^{n-1} - 2^{n/2-1}$ (indeed, if ℓ lies at distance $2^{n-1} + 2^{n/2-1}$ from f, then $\ell \oplus 1$ lies at distance $2^{n-1} - 2^{n/2-1}$ and *vice versa*). In fact, the condition on $\widehat{f_\chi}$ can be weakened, without losing the property of being necessary and sufficient:

Lemma 2 *Any n-variable (n even ≥ 2) Boolean function f is bent if and only if, for every $a \in \mathbb{F}_2^n$, $\widehat{f_\chi}(a) \equiv 2^{n/2} \left[\bmod 2^{n/2+1} \right]$, or equivalently $\widehat{f}(a) \equiv 2^{n/2-1} \left[\bmod 2^{n/2} \right]$.*

Proof. This necessary condition is also sufficient, since, if it is satisfied, then writing $\widehat{f_\chi}(a) = 2^{n/2}\lambda_a$, where λ_a is odd for every a, Parseval's Relation (23) implies $\sum_{a\in\mathbb{F}_2^n} \lambda_a^2 = 2^n$, which implies that $\lambda_a^2 = 1$ for every a. □

A slightly different viewpoint is that of bent sequences[32] but we shall not adopt it here because it most often gives no extra insight on the problems. The nonlinearity being an affine invariant, so is the notion of bent function. Clearly, if f is bent and ℓ is affine, then $f \oplus \ell$ is bent. A class of bent functions is called a *complete class of functions* if it is globally invariant under the

[32] For each vector X in $\{-1,1\}^{2^n}$, define: $\hat{X} = \frac{1}{\sqrt{2^n}} H_n X$, where H_n is the Walsh-Hadamard matrix, recursively defined by:

$$H_n = \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}, H_0 = [1].$$

The vectors X such that $\hat{X}$ belongs to $\{-1,1\}^{2^n}$ are called bent sequences. They are the images by the character $\chi = (-1)^{\cdot}$ of the bent functions on $\mathbb{F}_2^n$.

action of the general affine group and the addition of affine functions.
The automorphism group of the set of bent functions is the general affine group. This is a direct consequence of the property that, given a Boolean function g, if for every bent function f, function $f \oplus g$ is also bent, then g has degree at most 1 (which is easily proved).
Thanks to Relation (25) and to the fact that the Fourier transform of a function is constant if and only if the function equals δ_0 times some constant, we see that any function f is bent if and only if, for any nonzero vector a, the Boolean function $D_a f(x) = f(x) \oplus f(x+a)$ is balanced. In other words:

Theorem 8 *Any n-variable Boolean function (n even[33]) is bent if and only if it satisfies $PC(n)$.*

For this reason, bent functions are also called *perfect nonlinear functions*[34]. Equivalently, f is bent if and only if the $2^n \times 2^n$ matrix $H = [(-1)^{f(x+y)}]_{x,y \in \mathbb{F}_2^n}$ is a Hadamard matrix (*i*.e. satisfies $H \times H^t = 2^n I$, where I is the identity matrix), and if and only if the support of f is a *difference set*[35] of the elementary Abelian 2-group $\mathbb{F}_2^n$ [136, 204] (other types of difference sets exist, see *e.g.* [139]). This implies that the Cayley graph G_f (see Subsection 2.2.2) is strongly regular (see [18] for more precision).
The functions whose derivatives $D_a f$, $a \in H$, $a \neq 0$ are all balanced, where H is a linear hyperplane of $\mathbb{F}_2^n$, are characterized in [41, 42] for every n; they are all bent if n is even. The functions whose derivatives $D_a f$, $a \in E$, $a \neq 0$ are all balanced, where E is a vector subspace of $\mathbb{F}_2^n$ of dimension $n-2$, are also characterized in these two papers.

Bent functions have the property that, for every even positive integer w, the sum $\sum_{a \in \mathbb{F}_2^n} \widehat{f_\chi}^{\,w}(a)$ is minimum. Such sums (for even or odd w) play a role with respect to fast correlation attacks [47, 40] (when these sums have small magnitude for low values of w, this contributes to a good resistance to fast correlation attacks).

[33] In fact, according to the observations above, "n even" is implied by "f satisfies $PC(n)$"; functions satisfying $PC(n)$ do not exist for odd n.

[34] The characterization of Theorem 8 leads to a generalization of the notion of bent function to non-binary functions. In fact, several generalizations exist [3, 220, 257] (see [78] for a survey); the equivalence between being bent and being perfect nonlinear is no more valid if we consider functions defined over residue class rings (see [80]).

[35] Thus, bent functions are also related to designs, since any difference set can be used to construct a symmetric design, see [11], pages 274-278. The notion of difference set is anterior to that of bent function, but it had not been much studied in the case of elementary 2-groups before the introduction of bent functions.

A last way of looking at bent functions deals with linear codes: let f be any n-variable Boolean function (n even). Denote its support $\{x \in \mathbb{F}_2^n \mid f(x) = 1\}$ by S_f and write $S_f = \{u_1, \cdots, u_{w_H(f)}\}$. Consider a matrix G whose columns are all the vectors of S_f, without repetition, and let C be the linear code generated by the rows of this matrix. Thus, C is the set of all the vectors $U_v = (v \cdot u_1, \cdots, v \cdot u_{w_H(f)})$, where v ranges over $\mathbb{F}_2^n$. Then:

Proposition 16 *Let n be any even positive integer. Any n-variable Boolean function f is bent if and only if the linear code C defined above has dimension n (*i.e. *G is a generator matrix of C) and has exactly two nonzero Hamming weights: 2^{n-2} and $w_H(f) - 2^{n-2}$.*

Indeed, $w_H(U_v)$ equals $\sum_{x \in \mathbb{F}_2^n} f(x) \times v \cdot x = \sum_{x \in \mathbb{F}_2^n} f(x) \frac{1-(-1)^{v \cdot x}}{2} = \frac{\widehat{f}(0) - \widehat{f}(v)}{2}$. Hence, according to Relation (12), $w_H(U_v)$ equals $2^{n-2} + \frac{\widehat{\mathrm{f}_\chi}(v) - \widehat{\mathrm{f}_\chi}(0)}{4}$, for every nonzero vector v. Thus, C has dimension n and has the two nonzero Hamming weights 2^{n-2} and $w_H(f) - 2^{n-2}$ if and only if, for every $v \neq 0$, U_v is nonzero and $\widehat{\mathrm{f}_\chi}(v) = \widehat{\mathrm{f}_\chi}(0)$ or $\widehat{\mathrm{f}_\chi}(v) = \widehat{\mathrm{f}_\chi}(0) + 4w_H(f) - 2^{n+1} = \widehat{\mathrm{f}_\chi}(0) - 2\widehat{\mathrm{f}_\chi}(0) = -\widehat{\mathrm{f}_\chi}(0)$. If f is bent, then this condition is clearly satisfied. Conversely, according to Parseval's Relation (23), if this condition is satisfied, then $\widehat{\mathrm{f}_\chi}(v)$ equals $\pm 2^{n/2}$ for every v, *i.e.* f is bent.

There exist two other characterizations [353] dealing with C:
1. C has dimension n and C has exactly two weights, whose sum equals $w_H(f)$;
2. The length $w_H(f)$ of C is even, C has exactly two weights, and one of these weights is 2^{n-2}.

6.1 The dual

If f is bent, then the *dual function* $\widetilde{f}$ of f, defined on $\mathbb{F}_2^n$ by:

$$\widehat{\mathrm{f}_\chi}(u) = 2^{n/2}(-1)^{\widetilde{f}(u)}$$

is also bent and its own dual is f itself. Indeed, the inverse Fourier transform property (19) applied to $\varphi = \mathrm{f}_\chi$ (the sign function of f) gives, for every vector a: $\sum_{u \in \mathbb{F}_2^n} (-1)^{\widetilde{f}(u) \oplus a \cdot u} = 2^{n/2} \mathrm{f}_\chi(a) = 2^{n/2}(-1)^{f(a)}$.
Let f and g be two bent functions, then Relation (22) applied with $\varphi = \mathrm{f}_\chi$ and $\psi = \mathrm{g}_\chi$ shows that

$$\mathcal{F}(\widetilde{f} \oplus \widetilde{g}) = \mathcal{F}(f \oplus g). \tag{44}$$

Thus, $f \oplus g$ and $\widetilde{f} \oplus \widetilde{g}$ have the same weight and the mapping $f \mapsto \widetilde{f}$ is an isometry.
According to Proposition 6, for every $a, b \in \mathbb{F}_2^n$ and for every bent function f, the dual of the function $f(x+b) \oplus a \cdot x$ equals $\widetilde{f}(x+a) \oplus b \cdot (x+a) = \widetilde{f}(x+a) \oplus b \cdot x \oplus a \cdot b$. Denoting $b \cdot x$ by $\ell_b(x)$, Relation (44), applied with $g(x) = f(x+b) \oplus a \cdot x$, gives $\mathcal{F}(D_a\widetilde{f} \oplus \ell_b) = (-1)^{a \cdot b}\mathcal{F}(D_b f \oplus \ell_a)$, and applied with $g(x) = f(x) \oplus \ell_a(x)$ and with $f(x+b)$ in the place of $f(x)$, it gives the following property, first observed in [61] (and rediscovered in [43]):

$$\mathcal{F}(D_a\widetilde{f} \oplus \ell_b) = \mathcal{F}(D_b f \oplus \ell_a) \tag{45}$$

(from these two relations, we deduce that, if $a \cdot b = 1$, then $\mathcal{F}(D_a\widetilde{f} \oplus \ell_b) = \mathcal{F}(D_b f \oplus \ell_a) = 0$). Notice that, for every a and b, we have $D_b f = \ell_a \oplus \epsilon$ if and only if $D_a\widetilde{f} = \ell_b \oplus \epsilon$).
Moreover, if a pair of Boolean functions f and f' satisfies the relation $\mathcal{F}(D_a f' \oplus \ell_b) = \mathcal{F}(D_b f \oplus \ell_a)$, then these functions are bent (indeed, taking $a = 0$ shows that $D_b f$ is balanced for every $b \neq 0$ and taking $b = 0$ shows that $D_a f'$ is balanced for every $a \neq 0$), and are then the duals of each other up to the addition of a constant. Indeed, summing up the relation $\mathcal{F}(D_a f' \oplus \ell_b) = \mathcal{F}(D_b f \oplus \ell_a)$ for b ranging over $\mathbb{F}_2^n$ shows that $f'(0) \oplus f'(a) = \widetilde{f}(0) \oplus \widetilde{f}(a)$ for every a, since we have $\sum_{x,b \in \mathbb{F}_2^n}(-1)^{f'(x) \oplus f'(x+a) \oplus b \cdot x} = 2^n(-1)^{f'(0) \oplus f'(a)}$, and $\sum_{x,b \in \mathbb{F}_2^n}(-1)^{f(x) \oplus f(x+b) \oplus a \cdot x} = \widehat{\mathrm{f}_\chi}(0) \times \widehat{\mathrm{f}_\chi}(a)$.

The NNF of $\widetilde{f}$ can be deduced from the NNF of f. Indeed, using equality $\widetilde{f} = \frac{1-(-1)^{\widetilde{f}}}{2}$, we have $\widetilde{f} = \frac{1}{2} - 2^{-n/2-1}\widehat{\mathrm{f}_\chi} = \frac{1}{2} - 2^{n/2-1}\delta_0 + 2^{-n/2}\widehat{f}$ (according to Relation (12)). Applying now Relation (30) (expressing the value of the Fourier transform by means of the coefficients of the NNF) to $\varphi = f$, we deduce that if $\sum_{I \in \mathcal{P}(N)} \lambda_I x^I$ is the NNF of f then:

$$\widetilde{f}(x) = \frac{1}{2} - 2^{n/2-1}\delta_0(x) + (-1)^{w_H(x)} \sum_{I \in \mathcal{P}(N) \,|\, supp(x) \subseteq I} 2^{n/2-|I|}\lambda_I.$$

Changing I into $N \setminus I$ in this relation, and observing that $supp(x)$ is included in $N \setminus I$ if and only if $x_i = 0, \forall i \in I$, we obtain the NNF of $\widetilde{f}$ by expanding the following relation:

$$\widetilde{f}(x) = \frac{1}{2} - 2^{n/2-1}\prod_{i=1}^{n}(1-x_i) + (-1)^{w_H(x)} \sum_{I \in \mathcal{P}(N)} 2^{|I|-n/2}\lambda_{N \setminus I}\prod_{i \in I}(1-x_i).$$

We deduce (as shown in [87]):

Proposition 17 *Let f be any n-variable bent function (n even). For every $I \neq N$ such that $|I| > n/2$, the coefficient of x^I in the NNF of $\widetilde{f}$ (resp. of f) is divisible by $2^{|I|-n/2}$.*

Reducing this equality modulo 2 proves Rothaus' bound (see below) and that, for $n \geq 4$ and $|I| = n/2$, the coefficient of x^I in the ANF of $\tilde{f}$ equals the coefficient of $x^{N\setminus I}$ in the ANF of f. Using Relation (9), the equality above can be related to the main result of [191] (but this result by Hou was stated in a complex way).

The Poisson summation formula (17) applied to $\varphi = \mathrm{f}_\chi$ gives (see [54]) that for every vector subspace E of $\mathbb{F}_2^n$, and for every elements a and b of $\mathbb{F}_2^n$, we have:

$$\sum_{x\in a+E} (-1)^{\widetilde{f}(x)\oplus b\cdot x} = 2^{-n/2}|E|\,(-1)^{a\cdot b} \sum_{x\in b+E^\perp} (-1)^{f(x)\oplus a\cdot x}. \qquad (46)$$

Self-dual bent functions are studied in [77].

6.2 Bent functions of low algebraic degrees

Obviously, no affine function can be bent. All the quadratic bent functions are known: according to the properties recalled in Subsection 5.2, any such function

$$f(x) = \bigoplus_{1\leq i<j\leq n} a_{i,j}\, x_i\, x_j \oplus h(x) \ (h \text{ affine}, a_{i,j} \in \mathbb{F}_2)$$

is bent if and only if one of the following equivalent properties is satisfied:

1. its Hamming weight is equal to $2^{n-1} \pm 2^{n/2-1}$;
2. its associated symplectic form: $\varphi_f : (x, y) \mapsto f(0) \oplus f(x) \oplus f(y) \oplus f(x + y)$ is non-degenerate (*i*.e. has kernel $\{0\}$);
3. the skew-symmetric matrix $M = (m_{i,j})_{i,j\in\{1,\cdots,n\}}$ over $\mathbb{F}_2$, defined by: $m_{i,j} = a_{i,j}$ if $i < j$, $m_{i,j} = 0$ if $i = j$, and $m_{i,j} = a_{j,i}$ if $i > j$, is regular (*i*.e. has determinant 1); indeed, M is the matrix of the bilinear form φ_f;
4. $f(x)$ is equivalent, up to an affine nonsingular transformation, to the function: $x_1x_2 \oplus x_3x_4 \oplus \cdots \oplus x_{n-1}x_n \oplus \varepsilon$ ($\varepsilon \in \mathbb{F}_2$).

It is interesting to charaterize quadratic bent functions in the trace representation. This leads for instance to the Kerdock code; see Subsection 6.10 where the bent functions leading to this code are given.
Let us study for example the case of the *Gold function* $tr_n(vx^{2^i+1})$, where $gcd(i,n)=1$. It is bent if and only if there is no nonzero $x \in \mathbb{F}_{2^n}$ such that $tr_n(vx^{2^i}y+vxy^{2^i})=0$ for every $y \in \mathbb{F}_{2^n}$, *i.e.*, the equation $vx^{2^i}+(vx)^{2^{n-i}}=0$ has no non-zero solution. Raising this equation to the 2^i-th power gives $v^{2^i}x^{2^{2i}}+vx=0$ and 2^i-1 being co-prime with 2^n-1, it is equivalent, after dividing by vx (when $x \neq 0$) and taking the (2^i-1)th root, to $vx^{2^i+1} \in \mathbb{F}_2$. Hence, the function $tr_n(vx^{2^i+1})$ is bent if and only if v is not the (2^i+1)-th power of an element of $\mathbb{F}_{2^n}$, that is (since $gcd(2^i+1,2^n-1)=3$), v is not the third power of an element of $\mathbb{F}_{2^n}$. The same result exists with the *Kasami function* $tr_n(x^{2^{2i}-2^i+1})$, $gcd(i,n)=1$ (this is proved in [139, Theorem 11] for n not divisible by 3 and true also for n divisible by 3 as seen by Leander [240]).
Another example of quadratic bent function in the trace representation uses two trace functions, the trace function tr_n on the whole field $\mathbb{F}_{2^n}$ and the trace function $tr_{\frac{n}{2}}$ on the subfield $\mathbb{F}_{2^{n/2}}$, is: $f(x)=tr_n(\sum_{i=1}^{\frac{n}{2}-1}x^{2^i+1}) \oplus tr_{\frac{n}{2}}(x^{2^{n/2}+1})$.
A third example did not appear yet in the literature (as far as we know): let n be coprime with 3 and i be coprime with n, then the function $f(x,y)=tr_{\frac{n}{2}}(x^{2^i+1}+y^{2^i+1}+xy)$, $x,y \in \mathbb{F}_{2^{n/2}}$ is bent. Indeed, its associated symplectic form equals the function $((x,y),(x',y')) \to f(0,0) \oplus f(x,y) \oplus f(x',y') \oplus f(x+x',y+y')=tr_{\frac{n}{2}}(x^{2^i}x'+xx'^{2^i}+y^{2^i}y'+yy'^{2^i}+xy'+x'y)$. The kernel of this symplectic form equals $\left\{(x,y) \in \mathbb{F}_{2^{n/2}}^2 \,/\, \left\{ \begin{array}{l} x^{2^i}+x^{2^{n-i}}+y=0 \\ y^{2^i}+y^{2^{n-i}}+x=0 \end{array} \right. \right\}$; this set is reduced to $\{(0,0)\}$, since denoting $z=x+y$ we have $z^{2^i}+z^{2^{n-i}}+z=0$ which implies $z^{2^{2i}}=z^{2^i}+z$ and therefore $z^{2^{3i}}=z$, that is $z \in \mathbb{F}_{2^{3i}}$, and therefore $z \in \mathbb{F}_2$ and since 1 is not solution $z=0$. Then x and y must be null.

Open problem: characterize the bent functions of algebraic degrees at least 3 (that is, classify them under the action of the general affine group). This has been done for $n \leq 6$ in [315] (see also [302] where the number of bent functions is computed for these values of n). For $n=8$, it has been done in [190], for functions of algebraic degrees at most 3 only; all of these functions have at least one affine derivative D_af, $a \neq 0$ (it has been proved in [43] that this happens for $n \leq 8$ only). The determination of all bent 8-variable functions has been completed very recently, see [233].

Hans Dobbertin (with G. Leander) has presented in the posthumous paper [143] a nice approach for generating new bent functions by recursively gluing so-called $\mathbb{Z}$-bent functions.

6.3 Bound on algebraic degree

The algebraic degree of any Boolean function f being equal to the maximum size of the multi-index I such that x^I has an odd coefficient in the NNF of f, Proposition 17 gives:

Proposition 18 *Let n be any even integer greater than or equal to 4. The algebraic degree of any bent function on $\mathbb{F}_2^n$ is at most $n/2$.*

In the case that $n = 2$, the bent functions have degree 2, since they have odd weight (in fact, they are the functions of odd weights).
The bound of Proposition 18 (which is obviously also true for $\widetilde{f}$) was first proved in [315] and will be called *Rothaus' bound* in the sequel. It can also be proved (see below) by using a similar method as in the proof of Proposition 11. This same method also allows obtaining a bound, shown in [192], relating the gaps between $n/2$ and the algebraic degrees of f and $\widetilde{f}$:

Proposition 19 *The algebraic degrees of any n-variable bent function and of its dual satisfy:*

$$n/2 - d^\circ f \geq \frac{n/2 - d^\circ \widetilde{f}}{d^\circ \widetilde{f} - 1}. \tag{47}$$

A proof of Proposition 19 and a second proof of Proposition 18. Let us denote by d (resp. by $\widetilde{d}$) the algebraic degree of f (resp. of $\widetilde{f}$) and consider a term x^I of degree d in the ANF of f. The Poisson summation formula (18) applied to $\varphi = \mathrm{f}_\chi$ (or Relation (46) with $a = b = 0$) and to the vectorspace $E = \{u \in \mathbb{F}_2^n / \ \forall i \in I,\, u_i = 0\}$ gives $\sum_{u\in E}(-1)^{\widetilde{f}(u)} = 2^{n/2-d}\sum_{x\in E^\perp} \mathrm{f}_\chi(x)$. The orthogonal $E^\perp$ of E equals $\{u \in \mathbb{F}_2^n / \ \forall i \notin I,\, u_i = 0\}$. According to Relation (3), the restriction of f to $E^\perp$ has odd weight w, thus $\sum_{x\in E^\perp} \mathrm{f}_\chi(x) = 2^d - 2w$ is not divisible by 4. Hence, $\sum_{u\in E}(-1)^{\widetilde{f}(u)}$ is not divisible by $2^{n/2-d+2}$. We deduce the proof of Proposition 18: suppose that $d > n/2$, then $\sum_{u\in E}(-1)^{\widetilde{f}(u)}$ is not even, a contradiction with the fact that E has an even size. We prove now Proposition 19: according to McEliece's theorem (or Ax's theorem), $\sum_{u\in E}(-1)^{\widetilde{f}(u)}$ is divisible by $2^{\left\lceil \frac{n-d}{\widetilde{d}} \right\rceil}$. We deduce the inequality $n/2 - d + 2 > \left\lceil \frac{n-d}{\widetilde{d}} \right\rceil$, that is, $n/2 - d + 1 \geq \frac{n-d}{\widetilde{d}}$, which is

equivalent to (47). □

Using Relation (7) instead of Relation (3) gives a more precise result than Proposition 18, first shown in [87], which will be given in Subsection 6.6.

Proposition 19 can also be deduced from Proposition 17 and from some divisibility properties, shown in [87], of the coefficients of the NNFs of Boolean functions of algebraic degree d.

More on the algebraic degree of bent functions can be said for homogeneous functions (whose ANF contain monomials of fixed degree), see [279].

6.4 Constructions

There does not exist for $n \geq 10$ a classification of bent functions under the action of the general affine group. In order to understand better the structure of bent functions, we can try to design constructions of bent functions. It is useful also to deduce constructions of highly nonlinear balanced functions. Some of the known constructions of bent functions are direct, that is, do not use as building blocks previously constructed bent functions. We will call *primary constructions* these direct constructions. The others, sometimes leading to recursive constructions, will be called *secondary constructions*.

6.4.1 Primary constructions

1. The *Maiorana-McFarland original class* $\mathcal{M}$ (see [136, 273]) is the set of all the Boolean functions on $\mathbb{F}_2^n = \{(x, y); x, y \in \mathbb{F}_2^{n/2}\}$, of the form:

$$f(x, y) = x \cdot \pi(y) \oplus g(y) \tag{48}$$

where π is any permutation on $\mathbb{F}_2^{n/2}$ and g any Boolean function on $\mathbb{F}_2^{n/2}$ ("$\cdot$" denotes here an inner product in $\mathbb{F}_2^{n/2}$). Any such function is bent. More precisely, the bijectivity of π is a necessary and sufficient condition[36] for f being bent, according to Relation (49) below, applied with $r = n/2$. Note that for every function $h(y)$, the function $f(x, y) \oplus h(y)$ is bent. This property is characteristic of the functions of the form (48); indeed, taking $h = \delta_a$, the indicator of the singleton $\{a\}$, we have for every $a, u, v \in \mathbb{F}_2^{n/2}$ that $\pm 2^{n/2} = \sum_{x,y \in \mathbb{F}_2^{n/2}} (-1)^{f(x,y) \oplus u \cdot x \oplus v \cdot y \oplus \delta_a(y)} = \sum_{x,y \in \mathbb{F}_2^{n/2}} (-1)^{f(x,y) \oplus u \cdot x \oplus v \cdot y} - 2 \sum_{x \in \mathbb{F}_2^{n/2}} (-1)^{f(x,a) \oplus u \cdot x \oplus v \cdot a} = \pm 2^{n/2} \pm 2 \sum_{x \in \mathbb{F}_2^{n/2}} (-1)^{f(x,a) \oplus u \cdot x}$. Hence for

[36] It is, because the input has been cut in two pieces x and y of the same length; it is also possible to cut them in pieces of different lengths, see Proposition 20 below, and bentness is then obviously not characterized by the bijectivity of π.

every $a, u \in \mathbb{F}_2^{n/2}$, we have $\sum_{x\in\mathbb{F}_2^{n/2}}(-1)^{f(x,a)\oplus u\cdot x} \in \{0, \pm 2^{n/2}\}$. Clearly, having "$\sum_{x\in\mathbb{F}_2^{n/2}}(-1)^{f(x,a)\oplus u\cdot x} = 0$ for every u" for some a is impossible because of Parseval's relation. Then, for every $a \in \mathbb{F}_2^{n/2}$, there exists $u \in \mathbb{F}_2^{n/2}$ such that $\sum_{x\in\mathbb{F}_2^{n/2}}(-1)^{f(x,a)\oplus u\cdot x} = \pm 2^{n/2}$ that is $f(x,a) = u \cdot x$ or $f(x,a) = u \cdot x \oplus 1$.
The dual function $\widetilde{f}(x,y)$ equals: $y \cdot \pi^{-1}(x) \oplus g(\pi^{-1}(x))$, where π^{-1} is the inverse permutation of π. The completed class of $\mathcal{M}$ (that is, the smallest possible complete class including $\mathcal{M}$) contains all the quadratic bent functions (according to Alinea 4 of the characterization of quadratic bent functions given in Subsection 6.2; take $\pi = id$ and g constant in (48)) and all bent functions in at most 6 variables [135].
As we saw already in Subsection 5.1, the fundamental idea of Maiorana-McFarland's construction consists in *concatenating affine functions*. If we order all the binary words of length n in lexicographic order, with the bit of higher weight on the right, then the truth-table of f is the concatenation of the restrictions of f obtained by setting the value of y and letting x freely range over $\mathbb{F}_2^{n/2}$. These restrictions are affine. In fact, Maiorana-McFarland's construction is a particular case of a more general construction of bent functions [65] (see the next proposition), which is properly speaking a secondary construction for $r < n/2$ and which is the original Maiorana-McFarland construction for $r = n/2$ (this is why we give it in this subsection).

Proposition 20 *Let $n = r + s$ ($r \leq s$) be even. Let ϕ be any mapping from $\mathbb{F}_2^s$ to $\mathbb{F}_2^r$ such that, for every $a \in \mathbb{F}_2^r$, the set $\phi^{-1}(a)$ is an $(n-2r)$-dimensional affine subspace of $\mathbb{F}_2^s$. Let g be any Boolean function on $\mathbb{F}_2^s$ whose restriction to $\phi^{-1}(a)$ (viewed as a Boolean function on $\mathbb{F}_2^{n-2r}$ via an affine isomorphism between $\phi^{-1}(a)$ and this vectorspace) is bent for every $a \in \mathbb{F}_2^r$, if $n > 2r$ (no condition on g being imposed if $n = 2r$). Then the function $f_{\phi,g} = x \cdot \phi(y) \oplus g(y)$ is bent on $\mathbb{F}_2^n$.*

Proof. This is a direct consequence of the equality (valid for every ϕ and every g):

$$\widehat{(-1)^{f_{\phi,g}}}(a,b) = 2^r \sum_{y\in\phi^{-1}(a)} (-1)^{g(y)\oplus b\cdot y}, \tag{49}$$

which comes from the fact that every function $x \mapsto f_{\phi,g}(x,y) \oplus a \cdot x \oplus b \cdot y$ being affine, and thus constant or balanced, it contributes for a nonzero value in the sum $\sum_{x\in\mathbb{F}_2^r, y\in\mathbb{F}_2^s}(-1)^{f_{\phi,g}(x,y)\oplus x\cdot a\oplus y\cdot b}$ only if $\phi(y) = a$. According to Relation (49), the function $f_{\phi,g}$ is bent if and only if $r \leq n/2$ and

$\sum_{y\in\phi^{-1}(a)}(-1)^{g(y)\oplus b\cdot y} = \pm 2^{n/2-r}$ for every $a \in \mathbb{F}_2^r$ and every $b \in \mathbb{F}_2^s$. The hypothesis in Proposition 20 is a sufficient condition for that (but it is not a necessary one). □

This construction is pretty general: the choice of any partition of $\mathbb{F}_2^s$ in 2^r flats of dimension $(n-2r)$ and of an $(n-2r)$-variable bent function on each of these flats leads to an n-variable bent function.

Obviously, every Boolean function can be represented (in several ways) in the form $f_{\phi,g}$ for some values of $r \geq 1$ and s and for some mapping ϕ from $\mathbb{F}_2^s$ to $\mathbb{F}_2^r$ and Boolean function g on $\mathbb{F}_2^s$. It has been observed in [257] that, if a bent function has this form, then ϕ is balanced (*i.e.* is uniformly distributed over $\mathbb{F}_2^r$). This is a direct consequence of the fact that, for every nonzero $a \in \mathbb{F}_2^r$, the Boolean function $a \cdot \phi$ is balanced, since it equals the derivative $D_{(a,0)}f_{\phi,g}$, and of the characterization of balanced vectorial functions given in the chapter "Vectorial Boolean Functions for Cryptography".

It is shown in [25] that every bent function in 6 variables is affinely equivalent to a function of the Maiorana-McFarland class.

Remark: There exist $n/2$-dimensional vector spaces of n-variable Boolean functions whose non-zero elements are all bent. The Maiorana-McFarland construction easily allows constructing such vector spaces. A result by K. Nyberg (see the chapter "Vectorial Boolean Functions for Cryptography") shows that k-dimensional vector spaces of n-variable Boolean functions whose non-zero elements are all bent cannot exist for $k > n/2$.

2. The *Partial Spreads class* $\mathcal{PS}$, introduced in [136] by J. Dillon, is the set of all the sums (modulo 2) of the indicators of $2^{n/2-1}$ or $2^{n/2-1}+1$ "disjoint" $n/2$-dimensional subspaces of $\mathbb{F}_2^n$ ("disjoint" meaning that any two of these spaces intersect in 0 only, and therefore that their sum is direct and equals $\mathbb{F}_2^n$). The bentness of such function is a direct consequence of Theorem 12 below. This is why we omit the proof of this fact here. According to this same theorem, the dual of such a function has the same form, all the $n/2$-dimensional spaces E being replaced by their orthogonals. Note that the Boolean functions equal to the sums of the indicators of "disjoint" $n/2$-dimensional subspaces of $\mathbb{F}_2^n$ share with quadratic functions the property of being bent if and only if they have the weight of a bent function (which is $2^{n-1} \pm 2^{n/2-1}$). J. Dillon denotes by $\mathcal{PS}^-$ (resp. $\mathcal{PS}^+$) the class of those bent functions for which the number of $n/2$-dimensional subspaces is $2^{n/2-1}$ (resp. $2^{n/2-1}+1$). All the elements of $\mathcal{PS}^-$ have algebraic degree $n/2$ exactly (indeed, by applying a linear isomorphism of $\mathbb{F}_2^n$, we may assume

that $\mathbb{F}_2^{n/2} \times \{0\}$ is among the $2^{n/2-1}$ "disjoint" spaces defining the function, and since the function vanishes at 0, Relation (3) shows that the monomial $x_1 \cdots x_{n/2}$ appears in its ANF), but not all those of $\mathcal{PS}^+$ (which contains for instance all the quadratic functions, if $n/2$ is even, see below). It is an open problem to characterize the algebraic normal forms of the elements of class $\mathcal{PS}$, and it is not a simple matter to construct, practically, elements of this class. J. Dillon exhibits in [136] a subclass of $\mathcal{PS}^-$, denoted by $\mathcal{PS}_{ap}$, whose elements (that we shall call *Dillon's functions*) are defined in an explicit form: $\mathbb{F}_2^{n/2}$ is identified to the Galois field $\mathbb{F}_{2^{n/2}}$ (an inner product in this field being defined as $x \cdot y = tr_{\frac{n}{2}}(xy)$, where $tr_{\frac{n}{2}}$ is the trace function from $\mathbb{F}_{2^{n/2}}$ to $\mathbb{F}_2$; we know that the notion of bent function is independent of the choice of the inner product); the space $\mathbb{F}_2^n \approx \mathbb{F}_{2^{n/2}} \times \mathbb{F}_{2^{n/2}}$, viewed[37] as a 2-dimensional $\mathbb{F}_{2^{n/2}}$-vectorspace, is equal to the "disjoint" union of its $2^{n/2}+1$ lines through the origin; these lines are $n/2$-dimensional $\mathbb{F}_2$-subspaces of $\mathbb{F}_2^n$. Choosing any $2^{n/2-1}$ of the lines, and taking them different from those of equations $x = 0$ and $y = 0$, leads, by definition, to an element of $\mathcal{PS}_{ap}$, that is, to a function of the form $f(x, y) = g\left(x\, y^{2^{n/2}-2}\right)$, *i*.e. $g\left(\frac{x}{y}\right)$ with $\frac{x}{y} = 0$ if $y = 0$, where g is a balanced Boolean function on $\mathbb{F}_2^{n/2}$ which vanishes at 0. The complements $g\left(\frac{x}{y}\right) \oplus 1$ of these functions are the functions $g(\frac{x}{y})$ where g is balanced and does not vanish at 0; they belong to the class $\mathcal{PS}^+$. In both cases, the dual of $g(\frac{x}{y})$ is $g(\frac{y}{x})$ (this is a direct consequence of Theorem 12). Any function $f(x, y) = g\left(x\, y^{2^{n/2}-2}\right)$ can be represented as a function of a single variable X belonging to $\mathbb{F}_{2^n}$: we have $x = aX + (aX)^{2^{n/2}}$ and $y = bX + (bX)^{2^{n/2}}$ for some elements $a, b \in \mathbb{F}_{2^n}^*$ linearly independent over $\mathbb{F}_2^{n/2}$, and we have then $f(X) = g\left(\left(a + a^{2^{n/2}} X^{2^{n/2}-1}\right)\left(b + b^{2^{n/2}} X^{2^{n/2}-1}\right)^{2^{n/2}-2}\right)$, for every $X \neq 0$.

Given a primitive element α of $\mathbb{F}_{2^n}$, we have then for $i = 0, \cdots, 2^{n/2}$ and $j = 0, \cdots, 2^{n/2} - 2$:

$$f\left(\alpha^{i+j(2^{n/2}+1)}\right) = g\left((a + a^{2^{n/2}}\beta^i)\,(b + b^{2^{n/2}}\beta^i)^{2^{n/2}-2}\right),$$

where $\beta = \alpha^{2^{n/2}-1}$. The elements of the class $\mathcal{PS}_{ap}^{\#}$, of those Boolean functions over $\mathbb{F}_{2^n}$ which can be obtained from those of $\mathcal{PS}_{ap}$ by composition by the transformations $x \in \mathbb{F}_{2^n} \mapsto \delta x$, $\delta \neq 0$, and by addition of a constant[38]

[37] Let ω be an element of $\mathbb{F}_{2^n} \setminus \mathbb{F}_{2^{n/2}}$; the pair $(1, \omega)$ is a basis of the $\mathbb{F}_{2^{n/2}}$-vectorspace $\mathbb{F}_{2^n}$; hence, we have $\mathbb{F}_{2^n} = \mathbb{F}_{2^{n/2}} + \omega\mathbb{F}_{2^{n/2}}$.

[38] The functions of $\mathcal{PS}_{ap}$ are among them those satisfying $f(0) = f(1) = 0$.

are those Boolean functions f of weight $2^{n-1} \pm 2^{n/2-1}$ on $\mathbb{F}_{2^n}$ such that, denoting by α a primitive element of this field, $f(\alpha^{2^{n/2}+1}x) = f(x)$ for every $x \in \mathbb{F}_{2^n}$. It is proved in [136, 82] that these functions are the functions of weight $2^{n-1} \pm 2^{n/2-1}$ which can be written as $\sum_{i=1}^{r} tr_n(a_i x^{j_i})$ for $a_i \in \mathbb{F}_{2^n}$ and j_i a multiple of $2^{n/2}-1$ with $j_i \leq 2^n - 1$.
Dillon [136] shows that, when $n/2$ is even, all quadratic bent functions are equal to $\mathcal{PS}^+$ functions or to their complements (while they cannot be affinely equivalent to $\mathcal{PS}^{\#}_{ap}$ functions because their degree does not equal $n/2$): by affine equivalence we can restrict ourselves to the function $(x, \epsilon, y, \eta) \in \mathbb{F}_{2^{n/2-1}} \times \mathbb{F}_2 \times \mathbb{F}_{2^{n/2-1}} \times \mathbb{F}_2 \to tr(xy) + \epsilon\eta + 1$, where tr is the trace function from $\mathbb{F}_{2^{n/2-1}}$ to $\mathbb{F}_2$; the support of this function equals the union of the $2^{n/2-1} + 1$ $n/2$-dimensional vector spaces (very much related to the Kerdock code) $S_\infty = \{0\} \times \{0\} \times \mathbb{F}_{2^{n/2-1}} \times \mathbb{F}_2$ and $S_a = \{(x, \epsilon, a^2x + atr(ax) + a\epsilon, tr(ax)); (x, \epsilon) \in \mathbb{F}_{2^{n/2-1}} \times \mathbb{F}_2\}$ for $a \in \mathbb{F}_{2^{n/2-1}}$.
3. Dobbertin gives in [141] the construction of a class of bent functions which contains both $\mathcal{PS}_{ap}$ and $\mathcal{M}$. The elements of this class are the functions f defined by $f(x, \phi(y)) = g\left(\frac{x+\psi(y)}{y}\right)$, where g is a balanced Boolean function on $\mathbb{F}_{2^{n/2}}$ and ϕ, ψ are two mappings from $\mathbb{F}_{2^{n/2}}$ to itself such that, if T denotes the affine subspace of $\mathbb{F}_{2^{n/2}}$ spanned by the support of the function $\widehat{\mathrm{g}_\chi}$ (where $\mathrm{g}_\chi = (-1)^g$), then, for any a in $\mathbb{F}_{2^{n/2}}$, the functions ϕ and ψ are affine on $aT = \{ax, x \in T\}$. The mapping ϕ must additionally be one to one. The elements of this class do not have an explicit form, but Dobbertin gives two explicit examples of bent functions constructed this way. In both, ϕ is a power function (see below).

4. If $n/2$ is odd, then it is possible to deduce a bent Boolean function on $\mathbb{F}_2^n$ from any almost bent function from $\mathbb{F}_2^{n/2}$ to $\mathbb{F}_2^{n/2}$. A vectorial Boolean function $F : \mathbb{F}_2^m \to \mathbb{F}_2^m$ is called almost bent if all of the component functions $v \cdot F$, $v \neq 0$ in $\mathbb{F}_2^m$, are plateaued with amplitude $2^{\frac{m+1}{2}}$ (see in Subsection 6.8 the definition of these terms). The function $\gamma_F(a, b)$, $a, b \in \mathbb{F}_2^m$, equal to 1 if the equation $F(x) + F(x + a) = b$ admits solutions, with $a \neq 0$ in $\mathbb{F}_2^m$, and equal to 0 otherwise is then bent (see the proof of this result in the chapter "Vectorial Boolean Functions for Cryptography"). This gives new bent functions related to the almost bent functions listed in this same chapter. However, determining the ANF or the univariate representation of γ_F is an open problem when F is a Kasami, Welch or Niho almost bent function.

5. Some infinite classes of bent functions have also been obtained, thanks to the identification between the vectorspace $\mathbb{F}_2^n$ and the field $\mathbb{F}_{2^n}$, as *power*

functions (which can also be called *monomial functions*), that is, functions of the form $tr_n(ax^i)$, where tr_n is the trace function on $\mathbb{F}_{2^n}$ and where $a \neq 0$ and x belong to this same field. Obviously, a power function $tr_n(ax^i)$ can be bent only if the mapping $x \to x^i$ is not one to one (otherwise, the function would be balanced, a contradiction), that is, if i is not co-prime with $2^n - 1$. It has been proved in [240] that i must be co-prime either with $2^{n/2} - 1$ or with $2^{n/2} + 1$: it is a simple matter to show that $\widehat{f_\chi}(0)$ equals 1 modulo $gcd(i, 2^n - 1)$, and this implies that $\widehat{f_\chi}(0) = 2^{n/2}$ if and only if $gcd(i, 2^{n/2} + 1) = 1$ and $\widehat{f_\chi}(0) = -2^{n/2}$ if and only if $gcd(i, 2^{n/2} - 1) = 1$ (this is easy to show by using that $2^{n/2} - 1$ and $2^{n/2} + 1$ are co-prime). The known values of i for which there exists at least one a such that $tr_n(ax^i)$ is bent are (up to conjugacy $i \to 2i \pmod{2^n - 1}$):

- the Gold exponents $i = 2^j + 1$, where $\frac{n}{gcd(j,n)}$ is even (the corresponding function $tr_n(ax^i)$ is bent if and only if $a \notin \{x^i,\ x \in \mathbb{F}_{2^n}\}$; the condition" $\frac{n}{gcd(j,n)}$ even" is for allowing existence of such a; the function belongs to the Maiorana-McFarland class);
- the Dillon exponents [135] of the form $j \cdot (2^{n/2} - 1)$, where $gcd(j, 2^{n/2} + 1) = 1$ (the function $tr_n(ax^i)$, where $a \in \mathbb{F}_{2^{n/2}}$ without loss of generality, and $i = j\,(2^{n/2} - 1)$ is then bent if and only if the Kloosterman sum $\sum_{x \in \mathbb{F}_{2^{n/2}}} (-1)^{tr_{\frac{n}{2}}(1/x + ax)}$ is null[39], where $1/0 = 0$ and where $tr_{\frac{n}{2}}$ is the trace function on the field $\mathbb{F}_{2^{n/2}}$; this equivalence has been first proved by Dillon [136]; more recently, Leander [240] has found another proof which gives more insight; a small error in his proof has been corrected in [102]; the function $tr_n(ax^i)$ belongs then to the PS_{ap} class);
- the Kasami exponents $i = 2^{2j} - 2^j + 1$, where $gcd(j, n) = 1$ (the corresponding function $tr_n(ax^i)$ is bent if and only if $a \notin \{x^3,\ x \in \mathbb{F}_{2^n}\}$, see [139] and [240]);
- and two exponents more recently found: $i = (2^{n/4} + 1)^2$ where n is divisible by 4 but not by 8 (see [240], where the Gold and Dillon exponents are also revisited, see also [104] where (at page 2) the set of all a's such that the corresponding function $tr_n(ax^i)$ is bent is determined: $a = a'b^i$, $a' \in w\mathbb{F}_{2^{n/4}}$, $w \in \mathbb{F}_4 \setminus \mathbb{F}_2$, $b \in \mathbb{F}_{2^n}$; the function belongs to the Maiorana-McFarland class) and $i = 2^{n/3} + 2^{n/6} + 1$, where n is divisible by 6 [44] (the corresponding function $tr_n(ax^i)$ is bent if and only if $a = a'b^i$, $a' \in \mathbb{F}_{2^{n/2}}$

[39] The existence of a such that the Kloosterman sum is null had been conjectured by Dillon. It has been proved by Lachaud and Wolfmann [225] who proved that the values of such Kloosterman sums are all the numbers divisible by 4 in the range $[-2^{n/4+1} + 1; 2^{n/4+1} + 1]$.

such that $tr_{n/2}^{n/6}(a') := a' + a'^{2^{n/6}} + a'^{2^{2n/6}} = 0$, $b \in \mathbb{F}_{2^n}$; it belongs to the Maiorana-McFarland class).
Note that a still simpler bent function (but which is not expressed by means of the function tr_n itself) is $f(x) = tr_{\frac{n}{2}}(x^{2^{n/2}+1})$, that is, $f(x) = x^{2^{n/2}+1} + \left(x^{2^{n/2}+1}\right)^2 + \left(x^{2^{n/2}+1}\right)^{2^2} + \cdots + \left(x^{2^{n/2}+1}\right)^{2^{n/2-1}}$. The symplectic form $\varphi_f(x,y)$ associated to f equals $tr_n(y^{2^{n/2}}x)$; its kernel is therefore trivial and f is bent.
Some other functions are defined as the sums of a few power functions, see [136, 139, 144, 145, 102, 197, 239, 242, 357].
Note that power functions and sums of power functions represent for the designer of the cryptosystem using them the interest of being more easily computable than general functions (which allows using them with more variables while keeping a good efficiency). Power functions have the peculiarity that, denoting the set $\{x^i;\ x \in \mathbb{F}_{2^n}^*\}$ by U, two functions $tr_n(ax^i)$ and $tr_n(bx^i)$ such that $a/b \in U$ are linearly equivalent. It is not clear whether this is more an advantage for the designer or for the attacker of a system using a nonlinear balanced function derived from such bent function.
Finally, bent functions have been also obtained by Dillon and McGuire [140] as the restrictions of functions on $\mathbb{F}_{2^{n+1}}$, with $n+1$ odd, to a hyperplane of this field: these functions are the Kasami functions $tr_n\left(x^{2^{2k}-2^k+1}\right)$ and the hyperplane has equation $tr_n(x) = 0$. The restriction is bent under the condition that $n+1 = 3k \pm 1$.

Remark. The bent sequences given in [355] are particular cases of the constructions given above (using also some of the secondary constructions given below).
In [99] are constructed homogeneous bent functions (*i*.e. bent functions whose ANFs are the sums of monomials of the same degree) on 12 (and less) variables by using the invariant theory (which makes feasible the computer searchs).

6.4.2 Secondary constructions

We have already seen in Proposition 20 a secondary construction based on the Maiorana-McFarland construction. We describe now the others (which have been found so far).

1. The first secondary construction given by J. Dillon and O. Rothaus

in [136, 315] is very simple: let f be a bent function on $\mathbb{F}_2^n$ (n even) and g a bent function on $\mathbb{F}_2^m$ (m even) then the function h defined on $\mathbb{F}_2^{n+m}$ by $h(x,y) = f(x) \oplus g(y)$ is bent. Indeed, we have clearly $\widehat{h_\chi}(a,b) = \widehat{f_\chi}(a) \times \widehat{g_\chi}(b)$. This construction, called the *direct sum* has unfortunately no great interest from a cryptographic point of view, since it produces *decomposable functions* (a Boolean function is called decomposable if it is equivalent to the sum of two functions that depend on two disjoint subsets of coordinates; such peculiarity is easy to detect and can be used for designing divide-and-conquer attacks, as pointed out by J. Dillon in [137]).

2. A more interesting result, by the same authors, is the following: if g, h, k and $g \oplus h \oplus k$ are bent on $\mathbb{F}_2^n$ (n even), then the function defined at every element (x_1, x_2, x) of $\mathbb{F}_2^{n+2}$ ($x_1, x_2 \in \mathbb{F}_2$, $x \in \mathbb{F}_2^n$) by:

$$f(x_1, x_2, x) =$$

$$g(x)h(x) \oplus g(x)k(x) \oplus h(x)k(x) \oplus [g(x) \oplus h(x)]x_1 \oplus [g(x) \oplus k(x)]x_2 \oplus x_1x_2$$

is bent (this is a particular case of Theorem 10 below). No general class of bent functions has been deduced from this *Rothaus construction*.

3. Two classes of bent functions have been derived in [54] from Maiorana-McFarland's class, by adding to some functions of this class the indicators of some vector subspaces:

- the class $\mathcal{D}_0$ whose elements are the functions of the form $f(x,y) = x \cdot \pi(y) \oplus \delta_0(x)$ (recall that δ_0 is the Dirac symbol; the ANF of $\delta_0(x)$ is $\prod_{i=1}^{n/2}(x_i \oplus 1)$). The dual of such a function f is the function $y \cdot \pi^{-1}(x) \oplus \delta_0(y)$. It is proved in [54] that this class is not included[40] in the completed versions $\mathcal{M}^\#$ and $\mathcal{PS}^\#$ of classes $\mathcal{M}$ and $\mathcal{PS}$ (*i*.e. the smallest possible classes including them) and that every bent function in 6 variables is affinely equivalent to a function of this class, up to the addition of an affine function. Class $\mathcal{D}_0$ is a subclass of the class denoted by $\mathcal{D}$, whose elements are the functions of the form $f(x,y) = x \cdot \pi(y) \oplus 1_{E_1}(x)1_{E_2}(y)$, where π is any permutation on $\mathbb{F}_2^{n/2}$ and where E_1, E_2 are two linear subspaces of $\mathbb{F}_2^{n/2}$ such that $\pi(E_2) = {E_1}^\perp$ (1_{E_1} and 1_{E_2} denote their indicators). The dual of f belongs to the completed version of this same class;

- the class $\mathcal{C}$ of all the functions of the form $x \cdot \pi(y) \oplus 1_L(x)$, where L is any linear subspace of $\mathbb{F}_2^{n/2}$ and π any permutation on $\mathbb{F}_2^{n/2}$ such that, for

[40] It is easy to show that a function f does not belong to $\mathcal{M}^\#$ by showing that there does not exist an $n/2$-dimensional vector-subspace E of $\mathbb{F}_2^n$ such that D_aD_bf is null for every $a, b \in E$; it is much more difficult to show that it does not belong to $\mathcal{PS}^\#$.

any element a of $\mathbb{F}_2^{n/2}$, the set $\pi^{-1}(a+L^\perp)$ is a flat. It is a simple matter to see, as shown in [45], that, under the same hypothesis on π, if g is a Boolean function whose restriction to every flat $\pi^{-1}(a+L^\perp)$ is affine, then the function $x\cdot\pi(y)\oplus 1_L(x)\oplus g(y)$ is also bent.
The fact that any function in class $\mathcal{D}$ or class $\mathcal{C}$ is bent comes from the following theorem proved in [54], which has its own interest:

Theorem 9 *Let $b+E$ be any flat in $\mathbb{F}_2^n$ (E being a linear subspace of $\mathbb{F}_2^n$). Let f be any bent function on $\mathbb{F}_2^n$. The function $f^\star=f\oplus 1_{b+E}$ is bent if and only if one of the following equivalent conditions is satisfied:*

1. *For any a in $\mathbb{F}_2^n\setminus E$, the function D_af is balanced on $b+E$;*
2. *The restriction of the function $\widetilde{f}(x)\oplus b\cdot x$ to any coset of $E^\perp$ is either constant or balanced.*

If f and $f^\star$ are bent, then E has dimension greater than or equal to $n/2$ and the algebraic degree of the restriction of f to $b+E$ is at most $\dim(E)-n/2+1$.
If f is bent, if E has dimension $n/2$, and if the restriction of f to $b+E$ has algebraic degree at most $\dim(E)-n/2+1=1$, i.*e. is affine, then conversely $f^\star$ is bent too.*

Proof. Recall that a function is bent if and only if it satisfies $PC(n)$. The equivalence between Condition *1.* and the bentness of $f^\star$ comes then from the fact that $\mathcal{F}(D_af^\star)$ equals $\mathcal{F}(D_af)$ if $a\in E$, and equals $\mathcal{F}(D_af)-4\sum_{x\in b+E}(-1)^{D_af(x)}$ otherwise.
We have $\widehat{\mathrm{f}_\chi}(a)-\widehat{\mathrm{f}^\star_\chi}(a)=2\sum_{x\in b+E}(-1)^{f(x)\oplus a\cdot x}$. Using Relation (46), applied with $E^\perp$ in the place of E, we deduce that for every $a\in\mathbb{F}_2^n$:

$$\sum_{u\in a+E^\perp}(-1)^{\widetilde{f}(u)\oplus b\cdot u}=2^{\dim(E^\perp)-n/2-1}(-1)^{a\cdot b}\left(\widehat{\mathrm{f}_\chi}(a)-\widehat{\mathrm{f}^\star_\chi}(a)\right),$$

and $\widehat{\mathrm{f}_\chi}(a)-\widehat{\mathrm{f}_\chi}^\star(a)$ takes value 0 or $\pm 2^{n/2+1}$ for every a if and only if Condition *2.* is satisfied. So Condition *2.* is necessary and sufficient, according to Lemma 2 (at the beginning of Section 6).
Let us now assume that f and $f^\star$ are bent. Then $1_{b+E}=f^\star\oplus f$ has algebraic degree at most $n/2$, according to Rothaus' bound, and thus $\dim(E)\geq n/2$. The values of the Walsh transform of the restriction of f to $b+E$ being equal to those of $\frac{1}{2}\left(\widehat{\mathrm{f}_\chi}-\widehat{\mathrm{f}_\chi^\star}\right)$, they are divisible by $2^{n/2}$ and thus the restriction of f to $b+E$ has algebraic degree at most $\dim(E)-n/2+1$, according to

Proposition 11.
If f is bent, if E has dimension $n/2$, and if the restriction of f to $b+E$ is affine, then the relation $\widehat{f_\chi}(a) - \widehat{f_\chi}^\star(a) = 2\sum_{x\in b+E}(-1)^{f(x)\oplus a\cdot x}$ shows that $f^\star$ is bent too, according to Lemma 2. □

Remarks.
- Relation (46) applied to $E^\perp$ in the place of E, where E is some $n/2$-dimensional subspace, shows straightforwardly that, if f is a bent function on $\mathbb{F}_2^n$, then $f(x)\oplus a\cdot x$ is constant on $b+E$ if and only if $\widetilde{f}(x)\oplus b\cdot x$ is constant on $a+E^\perp$. The same relation shows that $f(x)\oplus a\cdot x$ is then balanced on every other coset of E and $\widetilde{f}(x)\oplus b\cdot x$ is balanced on every other coset of $E^\perp$. Notice that Relation (46) shows also that $f(x)\oplus a\cdot x$ cannot be constant on a flat of dimension strictly greater than $n/2$ (*i.e.* that f cannot be k-weakly-normal with $k > n/2$).
- Let f be bent on $\mathbb{F}_2^n$. Let a and a' be two linearly independent elements of $\mathbb{F}_2^n$. Let us denote by E the orthogonal of the subspace spanned by a and a'. According to condition *2.* of Theorem 9, the function $f\oplus 1_E$ is bent if and only if $D_aD_{a'}\widetilde{f}$ is null (indeed, a 2-variable function is constant or balanced if and only if it has even weight, and $\widetilde{f}$ has even weight on any coset of the vector subspace spanned by a and a' if and only if, for every vector x, we have $f(x)\oplus f(x+a)\oplus f(x+a')\oplus f(x+a+a') = 0$). This result has been restated in [43] and used in [45] to design (potentially) new bent functions.

4. Other classes of bent functions have been deduced from a construction given in [57], which generalizes the secondary constructions given in 1 and 2 above:

Theorem 10 *Let n and m be two even positive integers. Let f be a Boolean function on $\mathbb{F}_2^{n+m} = \mathbb{F}_2^n\times\mathbb{F}_2^m$ such that, for any element y of $\mathbb{F}_2^m$, the function on $\mathbb{F}_2^n$:*

$$f_y : x \mapsto f(x,y)$$

is bent. Then f is bent if and only if, for any element s of $\mathbb{F}_2^n$, the function

$$\varphi_s : y \mapsto \widetilde{f_y}(s)$$

is bent on $\mathbb{F}_2^m$. If this condition is satisfied, then the dual of f is the function $\widetilde{f}(s,t) = \widetilde{\varphi_s}(t)$ (taking as inner product in $\mathbb{F}_2^n\times\mathbb{F}_2^m$: $(x,y)\cdot(s,t) = x\cdot s\oplus y\cdot t$).

This very general result is easy to prove, using that, for every $s \in \mathbb{F}_2^n$,

$$\sum_{x\in\mathbb{F}_2^n} (-1)^{f(x,y)\oplus x\cdot s} = 2^{n/2}(-1)^{\widetilde{f_y}(s)} = 2^{n/2}(-1)^{\varphi_s(y)},$$

and thus that

$$\widehat{\mathrm{f}_\chi}(s,t) = 2^{n/2} \sum_{y\in\mathbb{F}_2^m} (-1)^{\varphi_s(y)\oplus y\cdot t}.$$

This construction has also been considered in a particular case by Adams and Tavares [1] under the name of bent-based functions, and later studied by J. Seberry and X.-M. Zhang in [326] in special cases too.
A case of application of this construction is nicely simple:

Corollary 4 *[67] Let f_1 and f_2 be two n-variable bent functions (n even) and let g_1 and g_2 be two m-variable bent functions (m even). Define*[41]

$$h(x,y) = f_1(x) \oplus g_1(y) \oplus (f_1 \oplus f_2)(x)\,(g_1 \oplus g_2)(y);\ x \in \mathbb{F}_2^n,\ y \in \mathbb{F}_2^m.$$

Then h is bent and its dual is obtained from $\widetilde{f_1}, \widetilde{f_2}, \widetilde{g_1}$ and $\widetilde{g_2}$ by the same formula as h is obtained from f_1, f_2, g_1 and g_2.

Proof. For every y, the function $h_y(x)$ of Theorem 10 equals $f_1(x)$ plus the constant $g_1(y)$ if $g_1(y) = g_2(y)$ and $f_2(x)$ plus the constant $g_1(y)$ if $g_1(y) \neq g_2(y)$; thus it is bent and function $\varphi_s(y)$ equals $\widetilde{f_1}(s) \oplus g_1(y)$ if $g_1(y) = g_2(y)$ and $\widetilde{f_2}(s) \oplus g_1(y)$ if $g_1(y) \neq g_2(y)$, that is, equals $\widetilde{f_1}(s) \oplus g_1(y) \oplus (\widetilde{f_1} \oplus \widetilde{f_2})(s)\,(g_1 \oplus g_2)(y)$. Hence, $\varphi_s(y)$ is bent too and according to Theorem 10, h is then bent and its dual equals:

$$\widetilde{h}(s,t) = \widetilde{f_1}(s) \oplus \widetilde{g_1}(t) \oplus (\widetilde{f_1} \oplus \widetilde{f_2})(s)(\widetilde{g_1} \oplus \widetilde{g_2})(t).$$

□

What is interesting in this particular case of Theorem 10 (sometimes called the *indirect sum of bent functions*) is that we only assume the bentness of f_1, f_2, g_1, and g_2 for deducing the bentness of h; no extra condition is needed, contrary to the general construction.
Another simple application of Theorem 10, called the *extension of Maiorana-McFarland type* is given in [79]: Let π be a permutation on $\mathbb{F}_2^{n/2}$ and $f_{\pi,g}(x,y) = x \cdot \pi(y) \oplus g(y)$ a related Maiorana-McFarland bent function.

[41] h is the concatenation of the four functions f_1, $f_1 \oplus 1$, f_2 and $f_2 \oplus 1$, in an order controled by $g_1(y)$ and $g_2(y)$. This construction $(f_1, f_2, g_1, g_2) \mapsto h$ will appear again below to construct resilient functions; see Theorem 14.

Let $(h_y)_{y\in\mathbb{F}_2^{n/2}}$ be a collection of bent functions on $\mathbb{F}_2^m$ for some even integer m. Then the function $(x,y,z)\in\mathbb{F}_2^{n/2}\times\mathbb{F}_2^{n/2}\times\mathbb{F}_2^m \to h_y(z)\oplus f_{\pi,g}(x,y)$ is bent.

Several classes have been deduced from Theorem 10 in [57], and later in [192].
- Let n and m be two even positive integers. The elements of $\mathbb{F}_2^{n+m}$ are written (x,y,z,τ), where x,y are elements of $\mathbb{F}_2^{n/2}$ and z,τ are elements of $\mathbb{F}_2^{m/2}$. Let π and π' be permutations on $\mathbb{F}_2^{n/2}$ and $\mathbb{F}_2^{m/2}$ (respectively) and h a Boolean function on $\mathbb{F}_2^{m/2}$. Then, the following Boolean function on $\mathbb{F}_2^{n+m}$ is bent:

$$f(x,y,z,\tau)=x\cdot\pi(y)\oplus z\cdot\pi'(\tau)\oplus\delta_0(x)h(\tau)$$

(recall that $\delta_0(x)$ equals 1 if $x=0$ and is null otherwise). It is possible to prove, see [57], that such a function does not belong, in general, to the completed version of class $\mathcal{M}$. It is also easy to prove that f does not belong, in general, to the completed version of class $\mathcal{D}_0$, since any element of $\mathcal{D}_0$ has algebraic degree $\frac{n+m}{2}$, and it is a simple matter to produce examples of functions f whose algebraic degree is smaller than $\frac{n+m}{2}$.
- Let n and m be two even positive integers. We identify $\mathbb{F}_2^{n/2}$ (resp. $\mathbb{F}_2^{m/2}$) with the Galois field $\mathbb{F}_{2^{n/2}}$ (resp. with $\mathbb{F}_{2^{m/2}}$). Let k be a Boolean function on $\mathbb{F}_{2^{n/2}}\times\mathbb{F}_{2^{m/2}}$ such that, for any element x of $\mathbb{F}_{2^{n/2}}$, the function $z\mapsto k(x,z)$ is balanced on $\mathbb{F}_{2^{m/2}}$, and for any element z of $\mathbb{F}_{2^{m/2}}$, the function $x\mapsto k(x,z)$ is balanced on $\mathbb{F}_{2^{n/2}}$. Then the function

$$f(x,y,z,\tau)=k\left(\frac{x}{y},\frac{z}{\tau}\right)$$

is bent on $\mathbb{F}_2^{n+m}$.
- Let r be a positive integer. We identify $\mathbb{F}_2^r$ with $\mathbb{F}_{2^r}$. Let π and π' be two permutations on $\mathbb{F}_{2^r}$ and g a balanced Boolean function on $\mathbb{F}_{2^r}$. The following Boolean function on $\mathbb{F}_2^{4r}=(\mathbb{F}_2^r)^4$:

$$f(x,y,z,\tau)=z\cdot\pi'\left[\tau+\pi\left(\frac{x}{y}\right)\right]\oplus\delta_0(z)g\left(\frac{x}{y}\right)$$

is a bent function.

5. X.-D. Hou and P. Langevin have made in [196] a very simple observation which leads to potentially new bent functions:

Proposition 21 *Let f be a Boolean function on $\mathbb{F}_2^n$, n even. Let σ be a permutation on $\mathbb{F}_2^n$. We denote its coordinate functions by $\sigma_1, \cdots, \sigma_n$ and we assume that, for every $a \in \mathbb{F}_2^n$:*

$$d_H(f, \bigoplus_{i=1}^{n} a_i \, \sigma_i) = 2^{n-1} \pm 2^{n/2-1}.$$

Then $f \circ \sigma^{-1}$ is bent.

Indeed, the Hamming distance between $f \circ \sigma^{-1}$ and the linear function $\ell_a(x) = a \cdot x$ equals $d_H(f, \bigoplus_{i=1}^{n} a_i \, \sigma_i)$.
Hou and Langevin deduced that, if h is an affine function on $\mathbb{F}_2^n$, if f_1, f_2 and g are Boolean functions on $\mathbb{F}_2^n$, and if the following function is bent:

$$f(x_1, x_2, x) = x_1 \, x_2 \, h(x) \oplus x_1 \, f_1(x) \oplus x_2 \, f_2(x) \oplus g(x) / \; x \in \mathbb{F}_2^n, \; x_1, \, x_2 \in \mathbb{F}_2,$$

then the function

$$f(x_1, x_2, x) \oplus (h(x) \oplus 1) \, f_1(x) f_2(x) \oplus f_1(x) \oplus (x_1 \oplus h(x) \oplus 1) \, f_2(x) \oplus x_2 \, h(x)$$

is bent.
They also deduced that, if f is a bent function on $\mathbb{F}_2^n$ whose algebraic degree is at most 3, and if σ is a permutation on $\mathbb{F}_2^n$ such that, for every $i = 1, \cdots, n$, there exists a subset U_i of $\mathbb{F}_2^n$ and an affine function h_i such that:

$$\sigma_i(x) = \bigoplus_{u \in U_i} (f(x) \oplus f(x+u)) \oplus h_i(x),$$

then $f \circ \sigma^{-1}$ is bent.
Finally, X.-D. Hou [192] deduced that if $f(x, y)$ $(x, y \in \mathbb{F}_2^{n/2})$ is a Maiorana-McFarland's function of the particular form $x \cdot y \oplus g(y)$ and if $\sigma_1, \cdots, \sigma_n$ are all of the form $\bigoplus_{1 \le i < j \le n/2} a_{i,j} x_i \, y_j \oplus b \cdot x \oplus c \cdot y \oplus h(y)$, then $f \circ \sigma^{-1}$ is bent. He gave several examples of application of this result.

6. Note that the construction of 5. does not increase the number of variables, contrary to most other secondary constructions. Another secondary construction without extension of the number of variables was introduced in [70]. It is based on the following result:

Proposition 22 *Let f_1, f_2 and f_3 be three Boolean functions on $\mathbb{F}_2^n$. Denote by s_1 the Boolean function equal to $f_1 \oplus f_2 \oplus f_3$ and by s_2 the Boolean function equal to $f_1 f_2 \oplus f_1 f_3 \oplus f_2 f_3$. Then we have $f_1 + f_2 + f_3 = s_1 + 2 s_2$. This implies*

the following equality between the Fourier transforms: $\widehat{f_1}+\widehat{f_2}+\widehat{f_3} = \widehat{s_1}+2\widehat{s_2}$ *and the similar equality between the Walsh transforms:*

$$\widehat{\mathrm{f}_{1_\chi}} + \widehat{\mathrm{f}_{2_\chi}} + \widehat{\mathrm{f}_{3_\chi}} = \widehat{\mathrm{s}_{1_\chi}} + 2\,\widehat{\mathrm{s}_{2_\chi}}. \tag{50}$$

Proof. The fact that $f_1 + f_2 + f_3 = s_1 + 2s_2$ (the sums being computed in $\mathbb{Z}$ and not modulo 2) can be checked easily. The linearity of the Fourier transform with respect to the addition in $\mathbb{Z}$ implies then $\widehat{f_1}+\widehat{f_2}+\widehat{f_3} = \widehat{s_1}+2\widehat{s_2}$. The equality $f_1+f_2+f_3 = s_1+2s_2$ also directly implies $\mathrm{f}_{1_\chi}+\mathrm{f}_{2_\chi}+\mathrm{f}_{3_\chi} = \mathrm{s}_{1_\chi}+2\mathrm{s}_{2_\chi}$, thanks to the equality $\mathrm{f}_\chi = 1 - 2f$ valid for every Boolean function, which implies Relation (50). □

Proposition 22 leads to the following double construction of bent functions:

Corollary 5 *Let f_1, f_2 and f_3 be three n-variable bent functions, n even. Denote by s_1 the function $f_1 \oplus f_2 \oplus f_3$ and by s_2 the function $f_1f_2 \oplus f_1f_3 \oplus f_2f_3$. Then:*
- if s_1 is bent and if $\tilde{s_1} = \tilde{f_1} \oplus \tilde{f_2} \oplus \tilde{f_3}$, then s_2 is bent, and $\tilde{s_2} = \tilde{f_1}\tilde{f_2} \oplus \tilde{f_1}\tilde{f_3} \oplus \tilde{f_2}\tilde{f_3}$;
*- if $\widehat{\mathrm{s}_{2_\chi}}(a)$ is divisible by $2^{n/2}$ for every a (*e.g. *if s_2 is bent, or if it is quadratic, or more generally if it is plateaued; see the definition in Subsection 6.8), then s_1 is bent.*

Proof. - If s_1 is bent and if $\tilde{s_1} = \tilde{f_1} \oplus \tilde{f_2} \oplus \tilde{f_3}$, then, for every a, Relation (50) implies:

$$\begin{aligned}\widehat{\mathrm{s}_{2_\chi}}(a) &= \left[(-1)^{\tilde{f_1}(a)} + (-1)^{\tilde{f_2}(a)} + (-1)^{\tilde{f_3}(a)} - (-1)^{\tilde{f_1}(a)\oplus\tilde{f_2}(a)\oplus\tilde{f_3}(a)}\right] 2^{\frac{n-2}{2}} \\ &= (-1)^{\tilde{f_1}(a)\tilde{f_2}(a)\oplus\tilde{f_1}(a)\tilde{f_3}(a)\oplus\tilde{f_2}(a)\tilde{f_3}(a)}\, 2^{n/2}.\end{aligned}$$

Indeed, as we already saw above with the relation $\mathrm{f}_{1_\chi}+\mathrm{f}_{2_\chi}+\mathrm{f}_{3_\chi} = \mathrm{s}_{1_\chi}+2\mathrm{s}_{2_\chi}$, for every bits ϵ, η and τ, we have $(-1)^\epsilon + (-1)^\eta + (-1)^\tau - (-1)^{\epsilon\oplus\eta\oplus\tau} = 2\,(-1)^{\epsilon\eta\oplus\epsilon\tau\oplus\eta\tau}$.
- If $\widehat{\mathrm{s}_{2_\chi}}(a)$ is divisible by $2^{n/2}$ for every a, then the number $\widehat{\mathrm{s}_{1_\chi}}(a)$, which is equal to $\left[(-1)^{\tilde{f_1}(a)} + (-1)^{\tilde{f_2}(a)} + (-1)^{\tilde{f_3}(a)}\right] 2^{n/2} - 2\,\widehat{\mathrm{s}_{2_\chi}}(a)$, according to Relation (50), is congruent with $2^{n/2}$ modulo $2^{n/2+1}$ for every a. This is sufficient to imply that s_1 is bent, according to Lemma 2 (at the beginning of Section 6). □

7. A construction related to the notion of normal extension of bent function can be found in Proposition 31.

6.4.3 Decompositions of bent functions

The following theorem, proved in [42], is a direct consequence of Relation (28), applied to $f \oplus \ell$ where ℓ is linear, and to a linear hyperplane E of $\mathbb{F}_2^n$, and of the well-known (easy to prove) fact that, for every even integer $n \geq 4$, the sum of the squares of two integers equals 2^n (resp. 2^{n+1}) if and only if one of these squares is null and the other one equals 2^n (resp. both squares equal 2^n):

Theorem 11 *Let n be an even integer, $n \geq 4$, and let f be an n-variable Boolean function. Then the following properties are equivalent.*

1. *f is bent.*
2. *For every (resp. for some) linear hyperplane E of $\mathbb{F}_2^n$, the Walsh transforms of the restrictions h_1, h_2 of f to E and to its complement (viewed as Boolean functions on $\mathbb{F}_2^{n-1}$) take values $\pm 2^{n/2}$ and 0 only, and the disjoint union of their supports equals the whole space $\mathbb{F}_2^{n-1}$.*

Hence, a simple way of obtaining a plateaued function in an odd number of variables and with optimal nonlinearity is to take the restriction of a bent function to an affine hyperplane. Note that we have also (see [42]) that, if a function in an odd number of variables is such that, for some nonzero $a \in \mathbb{F}_2^n$, every derivative $D_u f$, $u \neq 0$, $u \in a^\perp$, is balanced, then its restriction to the linear hyperplane $a^\perp$ or to its complement is bent.
It is also proved in [42] that the Walsh transforms of the four restrictions of a bent function to an $(n-2)$-dimensional vector subspace E of $\mathbb{F}_2^n$ and to its cosets have the same sets of magnitudes. It is a simple matter to see that, denoting by a and b two vectors such that $E^\perp$ is the linear space spanned by a and b, these four restrictions are bent if and only if $D_a D_b \widetilde{f}$ takes on constant value 1.
More on decomposing bent functions can be found in [42, 43, 101].

6.5 On the number of bent functions

The class of bent functions produced by the original Maiorana-McFarland's construction is far the widest class, compared to the classes obtained from the other primary constructions.
The number of bent functions of the form (48) equals $(2^{n/2})! \times 2^{2^{n/2}}$, which is asymptotically equivalent to $\left(\frac{2^{n/2+1}}{e}\right)^{2^{n/2}} \sqrt{2^{n/2+1}\pi}$ (according to Stirling's formula) while the only other important construction of bent functions,

$\mathcal{PS}_{ap}$, leads only to $\binom{2^{n/2}}{2^{n/2-1}} \approx \frac{2^{2^{n/2}+\frac{1}{2}}}{\sqrt{\pi 2^{n/2}}}$ functions. However, the number of provably bent Maiorana-McFarland's functions seems negligible with respect to the total number of bent functions. The number of (bent) functions which are affinely equivalent to Maiorana-McFarland's functions is unknown; it is at most equal to the number of Maiorana-McFarland's functions times the number of affine automorphisms, which equals $2^n(2^n-1)(2^n-2)\cdots(2^n-2^{n-1})$. It seems also negligible with respect to the total number of bent functions. The problem of determining an efficient lower bound on the number of n-variable bent functions is open.
Rothaus' inequality recalled in Subsection 6.3 (Proposition 18) states that any bent function has algebraic degree at most $n/2$. Thus, the number of bent functions is at most

$$2^{1+n+\ldots+\binom{n}{n/2}} = 2^{2^{n-1}+\frac{1}{2}\binom{n}{n/2}}.$$

We shall call this upper bound the *naive bound*. For $n=6$, the number of bent functions is known and is approximately equal to $2^{32.3}$ (see [302]), which is much less than what gives the naive bound: 2^{42}. For $n=8$, the number is also known: it has been first shown in [234] that it is inferior to $2^{129.2}$; it has been very recently calculated by Langevin, Leander et al. [233] and equals approximately $2^{106.3}$ (the naive bound gives 2^{163}). Hence picking at random an 8-variable Boolean function of algebraic degree bounded above by 4 does not allow obtaining bent functions (but more clever methods exist, see [127, 82]). An upper bound improving upon the naive bound has been found in [90]. It is exponentially better than the naive bound since it divides it by approximately $2^{2^{n/2}-n/2-1}$. But it seems to be still far from the exact number of bent functions: for $n=6$ it gives roughly 2^{38} (to be compared with $2^{32.3}$) and for $n=8$ it gives roughly 2^{152} (to be compared with $2^{106.3}$).

6.6 Characterizations of bent functions

6.6.1 characterization through the NNF

Proposition 23 *Let $f(x)=\sum_{I\in\mathcal{P}(N)}\lambda_I x^I$ be the NNF of a Boolean function f on $\mathbb{F}_2^n$. Then f is bent if and only if:*
1. for every I such that $n/2<|I|<n$, the coefficient λ_I is divisible by $2^{|I|-n/2}$;
2. λ_N (with $N=\{1,\cdots,n\}$) is congruent with $2^{n/2-1}$ modulo $2^{n/2}$.

Proof. According to Lemma 2, f is bent if and only if, for every $a\in\mathbb{F}_2^n$, $\widehat{f}(a)\equiv 2^{n/2-1}\left[\bmod 2^{n/2}\right]$. We deduce that, according to Relation (30) applied with $\varphi=f$, Conditions *1.* and *2.* imply that f is bent.

Conversely, Condition *1.* is necessary, according to Proposition 17. Condition *2.* is also necessary since $\widehat{f}(1,\cdots,1) = (-1)^n \lambda_N$ (from Relation (30)). □

Proposition 23 and Relation (9) imply some restrictions on the coefficients of the ANFs of bent functions, observed and used in [90] (and also partially observed by Hou and Langevin in [196]).

Proposition 23 can be seen as a (much) stronger version of Rothaus' bound, since the algebraic degree of a Boolean function whose NNF is $f(x) = \sum_{I \in \mathcal{P}(N)} \lambda_I \, x^I$ equals the maximum size of I, such that λ_I is odd.

6.6.2 Geometric characterization

Proposition 23 also allows proving the following characterization:

Theorem 12 *[85] Let f be a Boolean function on $\mathbb{F}_2^n$. Then f is bent if and only if there exist $n/2$-dimensional subspaces $E_1, \ldots, E_k$ of $\mathbb{F}_2^n$ (there is no constraint on the number k) and integers $m_1, \ldots, m_k$ (positive or negative) such that, for any element x of $\mathbb{F}_2^n$:*

$$f(x) \equiv \sum_{i=1}^{k} m_i 1_{E_i}(x) - 2^{n/2-1}\delta_0(x) \quad \left[\ mod \ 2^{n/2} \right]. \tag{51}$$

If we have $f(x) = \sum_{i=1}^{k} m_i 1_{E_i}(x) - 2^{n/2-1}\delta_0(x)$ then the dual of f equals $\widetilde{f}(x) = \sum_{i=1}^{k} m_i 1_{E_i^\perp}(x) - 2^{n/2-1}\delta_0(x)$.

Proof (sketch of). Relation (51) is a sufficient condition for f being bent, according to Lemma 2 (at the beginning of Section 6) and to Relation (16). This same Relation (16) also implies the last sentence of Theorem 12. Conversely, if f is bent, then Proposition 23 allows to deduce Relation (51), by expressing all the monomials x^I by means of the indicators of subspaces of dimension at least $n - |I|$ (indeed, the NNF of the indicator of the subspace $\{x \in \mathbb{F}_2^n / \, x_i = 0, \, \forall i \in I\}$ being equal to $\prod_{i \in I}(1 - x_i) = \sum_{J \subseteq I}(-1)^{|J|} x^J$, the monomial x^I can be expressed by means of this indicator and of the monomials x^J, where J is strictly included in I) and by using Lemma 3 below (note that $d \geq n - |I|$ implies $|I| - n/2 \geq n/2 - d$ and that $\prod_{i \in N}(1 - x_i) = \delta_0(x)$). □

Lemma 3 *Let F be any d-dimensional subspace of $\mathbb{F}_2^n$. There exist $n/2$-dimensional subspaces $E_1, \cdots, E_k$ of $\mathbb{F}_2^n$ and integers $m, m_1, \cdots, m_k$ such*

that, for any element x of $\mathbb{F}_2^n$:

$$2^{n/2-d}\ 1_F(x) \equiv m + \sum_{i=1}^{k} m_i 1_{E_i}(x) \left[\ mod\ 2^{n/2}\right]\ \ if\ d < n/2,\ \ and$$

$$1_F(x) \equiv \sum_{i=1}^{k} m_i 1_{E_i}(x) \left[\ mod\ 2^{n/2}\right]\ \ if\ d > n/2.$$

The class of those functions f which satisfy the relation obtained from (51) by withdrawing "[mod $2^{n/2}$]" is called *Generalized Partial Spread* class and denoted by $\mathcal{GPS}$ (it includes $\mathcal{PS}$), see [55]. The dual $\widetilde{f}$ of such function f of $\mathcal{GPS}$ equaling $\widetilde{f}(x) = \sum_{i=1}^{k} m_i 1_{E_i^{\perp}}(x) - 2^{n/2-1}\delta_0(x)$, it belongs to $\mathcal{GPS}$ too.

There is no uniqueness of the representation of a given bent function in the form (51). There exists another characterization, shown in [86], in the form $f(x) = \sum_{i=1}^{k} m_i 1_{E_i}(x) \pm 2^{n/2-1}\delta_0(x)$, where $E_1, \ldots, E_k$ are vector subspaces of $\mathbb{F}_2^n$ of dimensions $n/2$ or $n/2+1$ and where $m_1, \ldots, m_k$ are integers (positive or negative). There is not a unique way, either, to choose these spaces E_i. But it is possible to define some subclass of $n/2$-dimensional and $(n/2+1)$-dimensional spaces such that there is uniqueness, if the spaces E_i are chosen in this subclass.
P. Guillot has proved subsequently in [171] that, up to composition by a translation $x \mapsto x + a$, every bent function belongs to $\mathcal{GPS}$.

6.6.3 characterization by second-order covering sequences

Proposition 24 *[93] A Boolean function f defined on $\mathbb{F}_2^n$ is bent if and only if:*

$$\forall x \in \mathbb{F}_2^n,\ \sum_{a,b \in \mathbb{F}_2^n} (-1)^{D_a D_b f(x)} = 2^n. \tag{52}$$

Proof. If we multiply both terms of Relation (52) by $\mathrm{f}_\chi(x) = (-1)^{f(x)}$, we obtain the (equivalent) relation:

$$\forall x \in \mathbb{F}_2^n, \mathrm{f}_\chi \otimes \mathrm{f}_\chi \otimes \mathrm{f}_\chi(x) = 2^n\, \mathrm{f}_\chi(x);$$

indeed, we have $\mathrm{f}_\chi \otimes \mathrm{f}_\chi \otimes \mathrm{f}_\chi(x) = \sum_{b \in \mathbb{F}_2^n} \left(\sum_{a \in \mathbb{F}_2^n} (-1)^{f(a) \oplus f(a+b)} \right) (-1)^{f(b+x)} = \sum_{a,b \in \mathbb{F}_2^n} (-1)^{f(a+x) \oplus f(a+b+x) \oplus f(b+x)}$. According to the bijectivity of the Fourier transform and to Relation (20), this is equivalent to :

$$\forall u \in \mathbb{F}_2^n,\ \widehat{\mathrm{f}_\chi}^3(u) = 2^n \widehat{\mathrm{f}_\chi}(u)\,.$$

Thus, we have $\sum_{a,b\in\mathbb{F}_2^n}(-1)^{D_aD_bf(x)} = 2^n$ if and only if, for every $u \in \mathbb{F}_2^n$, $\widehat{\mathrm{f}_\chi}(u)$ equals $\pm\sqrt{2^n}$ or 0. According to Parseval's relation, the value 0 cannot be achieved by $\widehat{\mathrm{f}_\chi}$ and this is therefore equivalent to the bentness of f. □

Relation (52) is equivalent to the relation $\sum_{a,b\in\mathbb{F}_2^n}(1-2D_aD_bf(x)) = 2^n$, that is $\sum_{a,b\in\mathbb{F}_2^n} D_aD_bf(x) = 2^{2n-1} - 2^{n-1}$, and hence to the fact that f admits the *second order covering sequence* with all-1 coefficients and with level $2^{2n-1} - 2^{n-1}$.

It is shown similarly in [93] that the relation similar to (52) but with any integer in the place of 2^n characterizes the class of plateaued functions (see Subsection 6.8).

A characterization of bent functions through Cayley graphs also exists, see [18].

6.7 Subclasses: hyper-bent functions

In [356], A. Youssef and G. Gong study the Boolean functions f on the field $\mathbb{F}_{2^n}$ (n even) such that $f(x^i)$ is bent for every i co-prime with $2^n - 1$. These functions are called *hyper-bent functions*. The condition seems difficult to satisfy. However, A. Youssef and G. Gong show in [356] that hyper-bent functions exist. Their result is equivalent to the following (see [82]):

Proposition 25 *All the functions of class $\mathcal{PS}_{ap}^{\#}$ are hyper-bent.*

Let us give here a direct proof of this fact.
Proof. We can restrict ourselves without loss of generality to the functions of class $\mathcal{PS}_{ap}$. Let ω be any element in $\mathbb{F}_{2^n} \setminus \mathbb{F}_{2^{n/2}}$. The pair $(1, \omega)$ is a basis of the $\mathbb{F}_{2^{n/2}}$-vectorspace $\mathbb{F}_{2^n}$. Hence, we have $\mathbb{F}_{2^n} = \mathbb{F}_{2^{n/2}} + \omega\mathbb{F}_{2^{n/2}}$. Moreover, every element y of $\mathbb{F}_{2^{n/2}}$ satisfies $y^{2^{n/2}} = y$ and therefore $tr_n(y) = y + y^2 + \cdots + y^{2^{n/2-1}} + y + y^2 + \cdots + y^{2^{n/2-1}} = 0$. Consider the inner product in $\mathbb{F}_{2^n}$ defined by: $y \cdot y' = tr_n(y\,y')$; the subspace $\mathbb{F}_{2^{n/2}}$ is then its own orthogonal; hence, according to Relation (16), any sum of the form $\sum_{y\in\mathbb{F}_{2^{n/2}}}(-1)^{tr_n(\lambda y)}$ is null if $\lambda \notin \mathbb{F}_{2^{n/2}}$ and equals $2^{n/2}$ if $\lambda \in \mathbb{F}_{2^{n/2}}$.
Let us consider any element of the class $\mathcal{PS}_{ap}$, choosing a balanced Boolean function g on $\mathbb{F}_2^{n/2}$, vanishing at 0, and defining $f(y' + \omega\, y) = g\left(\frac{y'}{y}\right)$, with

$\frac{y'}{y} = 0$ if $y = 0$. For every $a \in \mathbb{F}_{2^n}$, we have

$$\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x) \oplus tr_n(a\, x^i)} = \sum_{y,y' \in \mathbb{F}_{2^{n/2}}} (-1)^{g\left(\frac{y'}{y}\right) \oplus tr_n(a\,(y' + \omega y)^i)}.$$

Denoting $\frac{y'}{y}$ by z, we see that:

$$\sum_{y \in \mathbb{F}^*_{2^{n/2}}, y' \in \mathbb{F}_{2^{n/2}}} (-1)^{g\left(\frac{y'}{y}\right) \oplus tr_n(a\,(y' + \omega y)^i)} = \sum_{z \in \mathbb{F}_{2^{n/2}}, y \in \mathbb{F}^*_{2^{n/2}}} (-1)^{g(z) \oplus tr_n(a\, y^i (z+\omega)^i)}.$$

The remaining sum $\sum_{y' \in \mathbb{F}_{2^{n/2}}} (-1)^{g(0) \oplus tr_n(a\, y'^i)} = \sum_{y' \in \mathbb{F}_{2^{n/2}}} (-1)^{tr_n(a\, y')}$ (this equality being due to the fact that the mapping $x \to x^i$ is one-to-one) equals $2^{n/2}$ if $a \in \mathbb{F}_{2^{n/2}}$ and is null otherwise.
Thus, $\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x) \oplus tr_n(a\, x^i)}$ equals:

$$\sum_{z \in \mathbb{F}_{2^{n/2}}} (-1)^{g(z)} \sum_{y \in \mathbb{F}_{2^{n/2}}} (-1)^{tr_n(a(z+\omega)^i\, y)} - \sum_{z \in \mathbb{F}_{2^{n/2}}} (-1)^{g(z)} + 2^{n/2} 1_{\mathbb{F}_{2^{n/2}}}(a).$$

The sum $\sum_{z \in \mathbb{F}_{2^{n/2}}} (-1)^{g(z)}$ is null since g is balanced.
The sum $\sum_{z \in \mathbb{F}_{2^{n/2}}} (-1)^{g(z)} \sum_{y \in \mathbb{F}_{2^{n/2}}} (-1)^{tr_n(a(z+\omega)^i\, y)}$ equals $\pm 2^{n/2}$ if $a \notin \mathbb{F}_{2^{n/2}}$, since we prove in the next Lemma that there exists then exactly one $z \in \mathbb{F}_{2^{n/2}}$ such that $a(z+\omega)^i \in \mathbb{F}_{2^{n/2}}$; and this sum is null if $a \in \mathbb{F}_{2^{n/2}}$ (this can be checked, if $a = 0$ thanks to the balancedness of g, and if $a \neq 0$ because y ranges over $\mathbb{F}_{2^{n/2}}$ and $a(z+\omega)^i \notin \mathbb{F}_{2^{n/2}}$). This completes the proof. □

Lemma 4 *Let n be any positive integer. Let a and ω be two elements of the set $\mathbb{F}_{2^n} \setminus \mathbb{F}_{2^{n/2}}$ and let i be co-prime with $2^n - 1$. There exists a unique element $z \in \mathbb{F}_2^{n/2}$ such that $a(z+\omega)^i \in \mathbb{F}_2^{n/2}$.*

Proof. Let j be the inverse of i modulo $2^n - 1$. We have $a(z+\omega)^i \in \mathbb{F}_2^{n/2}$ if and only if $z \in \omega + a^{-j} \times \mathbb{F}_2^{n/2}$. The sets $\omega + a^{-j} \times \mathbb{F}_2^{n/2}$ and $\mathbb{F}_2^{n/2}$ are two flats whose directions $a^{-j} \times \mathbb{F}_2^{n/2}$ and $\mathbb{F}_2^{n/2}$ are subspaces whose sum is direct and equals $\mathbb{F}_{2^n}$. Hence, they have a unique vector in their intersection. □

Relationships between the notion of hyper-bent function and cyclic codes are studied in [82]. It is proved that every hyper-bent function $f : \mathbb{F}_{2^n} \to \mathbb{F}_2$, can be represented as: $f(x) = \sum_{i=1}^{r} tr_n(a_i x^{t_i}) \oplus \epsilon$, where $a_i \in \mathbb{F}_{2^n}, \epsilon \in \mathbb{F}_2$ and $w_2(t_i) = n/2$. Consequently, all hyper-bent functions have algebraic degree $n/2$.

In [102] is proved that, for every even n, every $\lambda \in \mathbb{F}^*_{2^{n/2}}$ and every $r \in]0; \frac{n}{2}[$ such that the cyclotomic cosets of 2 modulo $2^{n/2}+1$ containing respectively $2^r - 1$ and $2^r + 1$ have size n and such that the function $tr_{\frac{n}{2}}\left(\lambda x^{2^r+1}\right)$ is balanced on $\mathbb{F}_{2^{n/2}}$, the function $tr_n\left(\lambda\left(x^{(2^r-1)(2^{n/2}-1)} + x^{(2^r+1)(2^{n/2}-1)}\right)\right)$ is bent (*i.e.* hyper-bent) if and only if the function $tr_{\frac{n}{2}}\left(x^{-1} + \lambda x^{2^r+1}\right)$ is also balanced on $\mathbb{F}_{2^{n/2}}$.

Remark. In [56] have been determined those Boolean functions on $\mathbb{F}_2^n$ such that, for a given even integer k ($2 \leq k \leq n-2$), any of the Boolean functions on $\mathbb{F}_2^{n-k}$, obtained by keeping constant k coordinates among $x_1, \cdots, x_n$, is bent (*i.e.* those functions which satisfy the propagation criterion of degree $n-k$ and order k, see Section 8). These are the four symmetric bent functions (see Section 10). They were called hyper-bent in [56] but we keep this term for the notion introduced by Youssef and Gong.

6.8 Superclasses: partially-bent functions, partial bent functions and plateaued functions

We have seen that bent functions can never be balanced, which makes them improper for a direct cryptographic use. This has led to a research of superclasses of the class of bent functions, whose elements can have high nonlinearities, but can also be balanced (and possibly, be m-resilient with large m or satisfy $PC(l)$ with large l). A first super-class having these properties has been obtained as the set of those functions which achieve a bound expressing some trade-off between the number of non-balanced derivatives (*i*.e. of nonzero auto-correlation coefficients) of a Boolean function and the number of nonzero values of its Walsh transform. This bound, given in the next proposition, had been conjectured in [301] and has been proved later in [53].

Proposition 26 *Let n be any positive integer. Let f be any Boolean function on $\mathbb{F}_2^n$. Let us denote the cardinalities of the sets $\{b \in \mathbb{F}_2^n \mid \mathcal{F}(D_b f) \neq 0\}$ and $\left\{b \in \mathbb{F}_2^n \mid \widehat{\mathrm{f}_\chi}(b) \neq 0\right\}$ by N_{Δ_f} and $N_{\widehat{\mathrm{f}_\chi}}$, respectively. Then:*

$$N_{\Delta_f} \times N_{\widehat{\mathrm{f}_\chi}} \geq 2^n. \tag{53}$$

Moreover, $N_{\Delta_f} \times N_{\widehat{\mathrm{f}_\chi}} = 2^n$ *if and only if, for every* $b \in \mathbb{F}_2^n$*, the derivative* $D_b f$ *is either balanced or constant. This property is also equivalent to the fact that there exist two linear subspaces* E *(of even dimension) and* E' *of* $\mathbb{F}_2^n$*, whose direct sum equals* $\mathbb{F}_2^n$*, and Boolean functions* g*, bent on* E*, and* h*, affine on* E'*, such that:*

$$\forall x \in E, \forall y \in E', f(x+y) = g(x) \oplus h(y). \tag{54}$$

Inequality (53) comes directly from Relation (25): since the value of the auto-correlation coefficient $\mathcal{F}(D_b f)$ lies between -2^n and 2^n for every b, we have $N_{\Delta_f} \geq 2^{-n} \sum_{b \in \mathbb{F}_2^n} (-1)^{u \cdot b} \mathcal{F}(D_b f) = 2^{-n}\, \widehat{\mathrm{f}_\chi}^2(u)$, for every $u \in \mathbb{F}_2^n$, and thus $N_{\Delta_f} \geq 2^{-n} \max_{u \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(u)$. And we have $N_{\widehat{\mathrm{f}_\chi}} \geq \frac{\sum_{u \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(u)}{\max_{u \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(u)} = \frac{2^{2n}}{\max_{u \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(u)}$. This proves Inequality (53). This inequality is an equality if and only if both inequalities above are equalities, that is, if and only if, for every b, the auto-correlation coefficient $\mathcal{F}(D_b f)$ equals 0 or $2^n(-1)^{u_0 \cdot b}$, where $\max_{u \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(u) = \widehat{\mathrm{f}_\chi}^2(u_0)$, and if f is plateaued. The condition that $D_b f$ is either balanced or constant, for every b, is sufficient to imply that f has the form (54): E' is the linear kernel of f and the restriction of f to E has balanced derivatives. Conversely, any function of the form (54) is such that Relation (53) is an equality. □

These functions such that $N_{\Delta_f} \times N_{\widehat{\mathrm{f}_\chi}} = 2^n$ are called *partially-bent functions*. Every quadratic function is partially-bent. Partially-bent functions share with quadratic functions almost all of their nice properties (Walsh spectrum easier to calculate, potential good nonlinearity and good resiliency order), see [53]. In particular, the values of the Walsh transform equal 0 or $\pm 2^{dim(E')+dim(E)/2}$.

A generalization of Relation (53) has been obtained in [307]:

Proposition 27 *Let* φ *be any nonzero* n*-variable pseudo-Boolean function. Let* $N_\varphi = |\{x \in \mathbb{F}_2^n / \varphi(x) \neq 0\}|$ *and* $N_{\widehat{\varphi}} = |\{u \in \mathbb{F}_2^n / \widehat{\varphi}(u) \neq 0\}|$*, then* $N_\varphi \times N_{\widehat{\varphi}} \geq 2^n$*.*
Equality occurs if and only if there exists a number λ *and a flat* F *of* $\mathbb{F}_2^n$ *such that* $\varphi(x) = \lambda(-1)^{u \cdot x}$ *if* $x \in F$ *and* $\varphi(x) = 0$ *otherwise.*

Proof. Denoting by 1_φ the indicator of the support $\{x \in \mathbb{F}_2^n / \varphi(x) \neq 0\}$ of φ, and replacing $\varphi(x)$ by $1_\varphi(x)\, \varphi(x)$ in the definition of $\widehat{\varphi}$, gives, for every $u \in$

$\mathbb{F}_2^n$: $\widehat{\varphi}(u) = \sum_{x\in\mathbb{F}_2^n} 1_\varphi(x)\varphi(x)(-1)^{u\cdot x}$. Applying then the Cauchy-Schwartz inequality gives $\widehat{\varphi}^2(u) \leq N_\varphi \sum_{x\in\mathbb{F}_2^n} \varphi^2(x) = 2^{-n} N_\varphi \sum_{v\in\mathbb{F}_2^n} \widehat{\varphi}^2(v)$ (according to Parseval's relation (3)). Hence, $\widehat{\varphi}^2(u) \leq 2^{-n} N_\varphi \times N_{\widehat{\varphi}} \max_{v\in\mathbb{F}_2^n} \widehat{\varphi}^2(v)$. Choosing u such that $\widehat{\varphi}^2(u)$ is maximum gives the desired inequality, since, according to Parseval's inequality, and φ being nonzero, this maximum cannot be null.
Equality occurs if and only if all of the inequalities above are equalities, that is, $\widehat{\varphi}^2(v)$ takes only one nonzero value (say μ) and there exists a number λ such that, for every u such that $\widehat{\varphi}^2(u) = \mu$, we have $\varphi(x) \neq 0 \Rightarrow \varphi(x) = \lambda(-1)^{u\cdot x}$. This is equivalent to the condition stated at the end of Proposition 27. □

Partially-bent functions must not be mistaken for *partial bent functions*, studied by P. Guillot in [172]. By definition, the Fourier transforms of partial bent functions take exactly two values[42] λ and $\lambda + 2^{n/2}$ on $\mathbb{F}_2^{n*}$ (n even). Rothaus' bound on the degree generalizes to partial bent functions. The dual $\widetilde{f}$ of f, defined by $\widetilde{f}(u) = 0$ if $\widehat{f}(u) = \lambda$ and $\widetilde{f}(u) = 1$ if $\widehat{f}(u) = \lambda + 2^{n/2}$, is also partial bent; and its dual is f. Two kinds of partial bent functions f exist: those such that $\widehat{f}(0) - f(0) = -\lambda(2^{n/2} - 1)$ and those such that $\widehat{f}(0) - f(0) = (2^{n/2} - \lambda)(2^{n/2} + 1)$. This can be proved by applying Parseval's Relation (23). The sum of two partial bent functions of the same kind, whose supports have at most the zero vector in common, is partial bent. A potential interest of partial bent functions is in the possibility of using them as building blocks for constructing bent functions.

In spite of their good properties, partially-bent functions, when they are not bent, have by definition nonzero linear structures and so do not give full satisfaction. The class of *plateaued* functions, already encountered above in Subsection 4.1 (and sometimes called *three-valued functions*) is a natural extension of that of partially-bent functions. They have been first studied by Zheng and Zhang in [363]. A function is called plateaued if its squared Walsh transform takes at most one nonzero value, that is, if its Walsh transform takes at most three values 0 and $\pm\lambda$ (where λ is some positive integer, that we call the *amplitude* of the plateaued function). Bent functions are plateaued and, according to Parseval's Relation (23), a plateaued function is bent if and only if its Walsh transform never takes the value 0.
Note that, according to Parseval's relation again, denoting as above by $N_{\widehat{f_\chi}}$

[42] Partial bent functions are the indicators of partial difference sets.

the cardinality of the support $\{a \in \mathbb{F}_2^n / \widehat{f_\chi}(a) \neq 0\}$ of the Walsh transform of a given n-variable Boolean function f, we have $N_{\widehat{f_\chi}} \times \max_{a \in \mathbb{F}_2^n} \widehat{f_\chi}^2(a) \geq 2^{2n}$ and therefore, according to Relation (35) relating the nonlinearity to the Walsh transform: $nl(f) \leq 2^{n-1}\left(1 - \frac{1}{\sqrt{N_{\widehat{f_\chi}}}}\right)$. Equality is achieved if and only if f is plateaued.

Still because of Parseval's relation, the amplitude λ of any plateaued function must be of the form 2^r where $r \geq n/2$ (since $N_{\widehat{f_\chi}} \leq 2^n$). Hence, the values of the Walsh transform of a plateaued function are divisible by $2^{n/2}$ if n is even and by $2^{(n+1)/2}$ if n is odd. The class of plateaued functions contains those functions which achieve the best possible trade-offs between resiliency, nonlinearity and algebraic degree: the order of resiliency and the nonlinearity of any Boolean function are bounded by Sarkar et al.'s bound (see Section 7 below) and the best compromise between those two criteria is achieved by plateaued functions only; the third criterion – the algebraic degree – is then also optimum. Other properties of plateaued functions can be found in [42].

Plateaued functions can be characterized by second-order covering sequences (see [93]):

Proposition 28 *A Boolean function f on $\mathbb{F}_2^n$ is plateaued if and only if there exists λ such that, for every $x \in \mathbb{F}_2^n$:*

$$\sum_{a,b \in \mathbb{F}_2^n} (-1)^{D_a D_b f(x)} = \lambda^2. \tag{55}$$

The proof is very similar to that of Proposition 52 and λ is necessarily the amplitude of the plateaued function. Indeed, a function f is plateaued with amplitude λ if and only if, for every $u \in \mathbb{F}_2^n$, we have $\widehat{f_\chi}(u)\left(\widehat{f_\chi}^2(u) - \lambda^2\right) = 0$, that is, $\widehat{f_\chi}^3(u) - \lambda^2 \widehat{f_\chi}(u) = 0$. Applying the Fourier transform to both terms of this equality and using Relation (20), we see that this is equivalent to the fact that, for every $a \in \mathbb{F}_2^n$, we have:

$$\sum_{x,y \in \mathbb{F}_2^n} (-1)^{f(x) \oplus f(y) \oplus f(x+y+a)} = \lambda^2 (-1)^{f(a)}.$$

The fact that quadratic functions are plateaued is a direct consequence of Proposition 28, since their second-order derivatives are constant; and Proposition 28 gives more insight on the relationship between the nonlinearity of

a quadratic function and the number of its nonzero second-order derivatives.

P. Langevin proved in [230] that, if f is a plateaued function, then the coset $f \oplus R(1,n)$ of the Reed-Muller code of order 1, is an *orphan of* $R(1,n)$. The notion of orphan has been introduced in [178] (with the term "urcoset" instead of orphan) and studied in [31]. A coset of $R(1,n)$ is an orphan if it is maximum with respect to the following partial order relation: $g \oplus R(1,n)$ is smaller than $f \oplus R(1,n)$ if there exists in $g \oplus R(1,n)$ an element g_1 of weight $nl(g)$ (that is, of minimum weight in $g \oplus R(1,n)$), and in $f \oplus R(1,n)$ an element f_1 of weight $nl(f)$, such that $supp(g_1) \subseteq supp(f_1)$. Clearly, if f is a function of maximum nonlinearity, then $f \oplus R(1,n)$ is an orphan of $R(1,n)$ (the converse is false, since plateaued functions with non-optimum nonlinearity exist). The notion of orphan can be used in algorithms searching for functions with high nonlinearities.

6.9 Normal and non-normal bent functions

As observed in [54] (see Theorem 9 above), if a bent function f is normal (resp. weakly-normal), that is, constant (resp. affine) on an $n/2$-dimensional flat $b+E$ (where E is a subspace of $\mathbb{F}_2^n$), then its dual $\widetilde{f}$ is such that $\widetilde{f}(u) \oplus b \cdot u$ is constant on $E^\perp$ (resp. on $a+E^\perp$, where a is a vector such that $f(x) \oplus a \cdot x$ is constant on E). Thus, $\widetilde{f}$ is weakly-normal. Moreover, we have already seen that f (resp. $f(x) \oplus a \cdot x$) is balanced on each of the other cosets of the flat. H. Dobbertin used this idea to construct balanced functions with high nonlinearities from normal bent functions (see Subsection 7.5.1).

The existence of non-normal (and even non-weakly-normal) bent functions, *i*.e. bent functions which are non-constant (resp. non-affine) on every $n/2$-dimensional flat, has been shown, contradicting a conjecture made by several authors that such bent function did not exist. It is proved in [139] that the so-called Kasami function defined over $\mathbb{F}_{2^n}$ by $f(x) = tr_n\left(ax^{2^{2k}-2^k+1}\right)$, with $gcd(k,n) = 1$, is bent if n is not divisible by 3 and if $a \in \mathbb{F}_{2^n}$ is not a cube. As shown in [45], if $a \in \mathbb{F}_4 \setminus \mathbb{F}_2$ and $k = 3$, then for $n = 10$, the function $f(x) \oplus tr_n(b)$ is non-normal for some b, and for $n = 14$, the function f is not weakly normal. Cubic bent functions on 8 variables are all normal, as shown in [101].

The direct sum (see definition in Subsection 6.4) of two normal functions is obviously a normal function, while the direct sum of two non-normal functions can be normal. What about the sum of a normal bent function and of a non-normal bent function? This question has been studied in [79].

To this aim, a notion more general than normality has been introduced as follows:

Definition 8 *Let $U \subseteq V$ be two vector spaces over $\mathbb{F}_2$. Let $\beta : U \to \mathbb{F}_2$ and $f : V \to \mathbb{F}_2$ be bent functions. Then we say that f is a* normal extension *of β, in symbols $\beta \preceq f$, if there is a direct decomposition $V = U \oplus W_1 \oplus W_2$ such that*

(i) $\beta(u) = f(u + w_1)$ *for all* $u \in U$, $w_1 \in W_1$,

(ii) $\dim W_1 = \dim W_2$.

The relation $\preceq$ is transitive and if $\beta \preceq f$ then the same relation exists between the duals: $\widetilde{\beta} \preceq \widetilde{f}$.
A bent function is normal if and only if $\epsilon \preceq f$, where $\epsilon \in \mathbb{F}_2$ is viewed as a Boolean functions over the vector space $\mathbb{F}_2^0 = \{0\}$.
Examples of normal extensions are given in [79] (including the construction of Theorem 10 and its particular cases, the indirect sum and the extension of Maiorana-McFarland type).
The clarification about the sum of a normal bent function and of a non-normal bent function comes from the two following propositions:

Proposition 29 *Let $f_i : V_i \to \mathbb{F}_2$, $i = 1, 2$, be bent functions. The direct sum $f_1 \oplus f_2$ is normal if and only if bent functions β_1 and β_2 exist such that f_i is a normal extension of β_i $(i = 1, 2)$ and either β_1 and β_2 or β_1 and $\beta_2 \oplus 1$ are linearly equivalent.*

Proposition 30 *Suppose that $\beta \preceq f$ for bent functions β and f. If f is normal, then also β is normal.*

Hence, since the direct sum of a bent function β and of a normal bent function g is a normal extension of β, the direct sum of a normal and a non-normal bent function is always non-normal.
Normal extension leads to a secondary construction of bent functions:

Proposition 31 *Let β be a bent function on U and f a bent function on $V = U \times W \times W$. Assume that $\beta \preceq f$. Let*

$$\beta' : U \to \mathbb{F}_2$$

be any bent function. Modify f by setting for all $x \in U$, $y \in W$

$$f'(x, y, 0) = \beta'(x),$$

while $f'(x, y, z) = f(x, y, z)$ for all $x \in U$, $y, z \in W$, $z \neq 0$. Then f' is bent and we have $\beta' \preceq f'$.

6.10 Kerdock codes

For every even n, the *Kerdock code* $\mathcal{K}_n$ [211] is a supercode of $R(1,n)$ (*i*.e. contains $R(1,n)$ as a subset) and is a subcode of $R(2,n)$. More precisely $\mathcal{K}_n$ is a union of cosets $f_u \oplus R(1,n)$ of $R(1,n)$, where the functions f_u are quadratic (one of them is null and all the others have algebraic degree 2). The difference $f_u \oplus f_v$ between two distinct functions f_u and f_v being bent, $\mathcal{K}_n$ has minimum distance $2^{n-1} - 2^{n/2-1}$ (n even), which is the best possible minimum distance for a code equal to a union of cosets of $R(1,n)$, according to the covering radius bound. The size of $\mathcal{K}_n$ equals 2^{2n}. This is the best possible size for such minimum distance (see [129]). We describe now how the construction of Kerdock codes can be simply stated.

6.10.1 Construction of the Kerdock code

The function

$$f(x) = \bigoplus_{1 \le i < j \le n} x_i x_j \tag{56}$$

(which can also be defined as $f(x) = \binom{w_H(x)}{2}$ [mod 2]) is bent[43] because the kernel of it associated symplectic form $\varphi(x,y) = \bigoplus_{1 \le i \ne j \le n} x_i y_j$ equals $\{0\}$. Thus, the linear code $R(1,n) \cup (f \oplus R(1,n))$ has minimum distance $2^{n-1} - 2^{n/2-1}$.

We want to construct a code of size 2^{2n} with this same minimum distance. We use the structure of field to this aim. We have recalled in Subsection 2.1 some properties of the field $\mathbb{F}_{2^m}$ (where m is any positive integer). In particular, we have seen that there exists $\alpha \in \mathbb{F}_{2^m}$ (called a primitive element) such that $\mathbb{F}_{2^m} = \{0, \alpha, \alpha^2, \cdots, \alpha^{2^m-1}\}$. Moreover, there exists α, primitive element, such that $(\alpha, \alpha^2, \alpha^{2^2}, \cdots, \alpha^{2^{m-1}})$ is a basis of the vectorspace $\mathbb{F}_{2^m}$. Such basis is called a *normal basis*. If m is odd, then there exists a self-dual normal basis, that is, a normal basis such that: $tr_m(\alpha^{2^i+2^j}) = 1$ if $i = j$ and $tr_m(\alpha^{2^i+2^j}) = 0$ otherwise, where tr_m is the trace function over $\mathbb{F}_{2^m}$.
Consequence: for all $x = x_1\alpha + \cdots + x_m\alpha^{2^{m-1}}$ in $\mathbb{F}_{2^m}$, we have

$$tr_m(x) = \bigoplus_{i=1}^{m} x_i \qquad tr_m(x^{2^j+1}) = \bigoplus_{i=1}^{m} x_i x_{i+j},$$

(where x_{i+j} is replaced by x_{i+j-m} if $i + j > m$).

[43] We shall see in Section 10 that it is, up to the addition of affine functions, the sole symmetric bent function.

The function f of Relation (56), viewed as a function $f(x, x_n)$ on $\mathbb{F}_{2^m} \times \mathbb{F}_2$, where $m = n - 1$ is odd – say $m = 2t + 1$ – can now be written as[44]:

$$f(x, x_n) = tr_m \left(\sum_{j=1}^{t} x^{2^j+1} \right) \oplus x_n tr_m (x) .$$

Notice that the associated symplectic form associated to f equals $tr_m(x)tr_m(y) \oplus tr_m(xy) \oplus x_n tr_m(y) \oplus y_n tr_m(x)$.

Let us denote $f(ux, x_n)$ by $f_u(x, x_n)$ ($u \in \mathbb{F}_{2^m}$), then $\mathcal{K}_n$ is defined as the union, when u ranges over $\mathbb{F}_{2^m}$, of the cosets $f_u \oplus R(1, n)$.
$\mathcal{K}_n$ contains 2^{n+1} affine functions and $2^{2n} - 2^{n+1}$ quadratic bent functions. Its minimum distance equals $2^{n-1} - 2^{n/2-1}$ because the sum of two distinct functions f_u and f_v is bent. Indeed, the kernel of the associated symplectic form equals the set of all ordered pairs (x, x_n) verifying $tr_m(ux)tr_m(uy) \oplus tr_m(u^2xy) \oplus x_n tr_m(uy) \oplus y_n tr_m(ux) = tr_m(vx)tr_m(vy) \oplus tr_m(v^2xy) \oplus x_n tr_m(vy) \oplus y_n tr_m(vx)$ for every $(y, y_n) \in \mathbb{F}_{2^m} \times \mathbb{F}_2$, that is, $utr_m(ux) + u^2x + x_n u = vtr_m(vx) + v^2x + x_n v$ and $tr_m(ux) = tr_m(vx)$; it is a simple matter to show that it equals $\{(0, 0)\}$.

Open problem: Other examples of codes having the same parameters exist [205]. All are equal to subcodes of the Reed-Muller code of order 2, up to affine equivalence. We do not know how to obtain the same parameters with non-quadratic functions. This would be useful for cryptographic purposes as well as for the design of sequences for code division multiple access (CDMA) in telecommunications.

Remark.
The Kerdock codes are not linear. However, they share some nice properties with linear codes: the distance distribution between any codeword and all the other codewords does not depend on the choice of the codeword (we say that the Kerdock codes are distance-invariant; this results in the fact that their distance enumerators are equal to their weight enumerators); and, as proved by Semakov and Zinoviev [329], the weight enumerators of the Kerdock codes satisfy a relation similar to Relation (33), in which C is replaced by $\mathcal{K}_n$ and $C^\perp$ is replaced by the so-called Preparata code of the same length (we say that the Kerdock codes and the Preparata codes are formally dual). An explanation of this astonishing property has been recently obtained [175]:

[44] Obviously, this expression can be taken as the definition of f.

the Kerdock code is stable under an addition inherited of the addition in $\mathbb{Z}_4 = \mathbb{Z}/4\mathbb{Z}$ (we say it is $\mathbb{Z}_4$-linear) and the Mac Williams identity still holds in this different framework. Such an explanation had been an open problem for two decades.

7 Resilient functions

We have seen in Subsection 4.1 that the combining functions in stream ciphers must be m-resilient with large m. As any cryptographic functions, they must also have high algebraic degrees and high nonlinearities.
Notation: by an $(n, m, d, \mathcal{N})$- function, we mean an n-variable, m-resilient function having algebraic degree at least d and nonlinearity at least $\mathcal{N}$.
There are necessary trade-offs between n, m, d and $\mathcal{N}$.

7.1 Bound on algebraic degree

Siegenthaler's bound states that any m-resilient function ($0 \leq m < n-1$) has algebraic degree smaller than or equal to $n-m-1$ and that any $(n-1)$-resilient function is affine[45]. This can be proved directly by using Relation (3) and the original definition of resiliency given by Siegenthaler (Definition 3), since the bit $\bigoplus_{x \in \mathbb{F}_2^n / \, supp(x) \subseteq I} f(x)$ equals the parity of the weight of the restriction of f obtained by setting to 0 the coordinates of x which lie outside I. Note that instead of using the original Siegenthaler's definition in the proof of Siegenthaler's bound, we can also use the characterization by Xiao and Massey, recalled in Theorem 3, together with the Poisson summation formula (18) applied to $\varphi = f$ and with $E^{\perp} = \{x \in \mathbb{F}_2^n \,|\, supp(x) \subseteq I\}$, where I has size strictly greater than $n-m-1$. But this gives a less simple proof. Siegenthaler's bound is also a direct consequence of a characterization of resilient functions[46] through their NNFs and of the fact that the algebraic degrees of Boolean functions are smaller than or equal to their numerical degrees:

Proposition 32 *[88] Let n be any positive integer and $m < n$ a non-negative integer. A Boolean function f on $\mathbb{F}_2^n$ is m-resilient if and only if the NNF of the function $f(x) \oplus x_1 \oplus \cdots \oplus x_n$ has degree at most $n-m-1$.*

[45] Siegenthaler also proved that any n-variable m-th order correlation-immune function has degree at most $n-m$. This can be shown by using similar methods as for resilient functions. Moreover, if such function has weight divisible by 2^{m+1} then it satisfies the same bound as m-resilient functions.

[46] A similar characterization of correlation-immune functions can be found in [63].

Proof. Let us denote by $g(x)$ the function $f(x) \oplus x_1 \oplus \cdots \oplus x_n$. For each vector $a \in \mathbb{F}_2^n$, we denote by $\overline{a}$ the componentwise complement of a equal to $a + (1, \cdots, 1)$. We have $\widehat{f_\chi}(a) = \widehat{g_\chi}(\overline{a})$. Thus, f is m-resilient if and only if, for each vector u of weight greater than or equal to $n - m$, the number $\widehat{g_\chi}(u)$ is null. Consider the NNF of g:

$$g(x) = \sum_{I \in \mathcal{P}(N)} \lambda_I \, x^I.$$

According to Relations (30), (31) and (12) applied to g, we have for nonzero u:

$$\widehat{g_\chi}(u) = (-1)^{w_H(u)+1} \sum_{I \in \mathcal{P}(N) \,|\, supp(u) \subseteq I} 2^{n-|I|+1} \lambda_I,$$

and for nonempty I:

$$\lambda_I = 2^{-n}(-2)^{|I|-1} \sum_{u \in \mathbb{F}_2^n \,|\, I \subseteq supp(u)} \widehat{g_\chi}(u).$$

We deduce that $\widehat{g_\chi}(u)$ is null for every vector u of weight greater than or equal to $n - m$ if and only if the NNF of g has degree at most $n - m - 1$.□

Proposition 32 can also be proved by using the Xiao-Massey characterization (again) and Relation (8) relating the values of the coefficients of the NNF to the values of the function, applied to $g(x) = f(x) \oplus x_1 \oplus \cdots \oplus x_n$.

Proposition 32 has been used by X.-D. Hou in [193] for constructing resilient functions. Siegenthaler's bound gives an example of the trade-offs which must be accepted in the design of combiner generators[47]. Sarkar and Maitra showed in [321] that the values of the Walsh Transform of an n-variable, m-resilient (resp. m-th order correlation-immune) function are divisible by 2^{m+2} (resp. 2^{m+1}) if $m \leq n - 2$ (a proof of a slightly more precise result is given in the next subsection, at Theorem 13)[48]. This *Sarkar-Maitra's divisibility* bound (which implies in particular that the weight of any m-th order correlation-immune function is divisible by 2^m) allows also to deduce Siegenthaler's bound, thanks to Proposition 11 applied with $k = m+2$ (resp. $k = m + 1$).

[47] One approach to avoid such trade-off is to allow memory in the nonlinear combination generator, that is, to replace the combining function by a finite state machine, see [277].

[48] More is proved in [63, 94]; in particular: if the weight of an m-th order correlation-immune is divisible by 2^{m+1}, then the values of its Walsh Transform are divisible by 2^{m+2}.

7.2 Bounds on the nonlinearity

Sarkar-Maitra's divisibility bound, recalled at the end of the previous subsection, has provided a nontrivial upper bound on the nonlinearity of resilient functions, independently obtained by Tarannikov [345] and by Zheng and Zhang [366]: the nonlinearity of any m-resilient function ($m \leq n-2$) is bounded above by $2^{n-1} - 2^{m+1}$. This bound is tight, at least when $m \geq 0.6\ n$, see [345, 346][49]. We shall call it *Sarkar et al.'s bound*. Notice that, if an m-resilient function f achieves nonlinearity $2^{n-1} - 2^{m+1}$, then f is plateaued. Indeed, the distances between f and affine functions lie then between $2^{n-1} - 2^{m+1}$ and $2^{n-1} + 2^{m+1}$ and must be therefore equal to $2^{n-1} - 2^{m+1}$, 2^{n-1} and $2^{n-1} + 2^{m+1}$ because of the divisibility result of Sarkar and Maitra. Thus, the Walsh transform of f takes three values 0 and $\pm 2^{m+2}$. Moreover, it is proved in [345] that such function f also achieves Siegenthaler's bound (and as proved in [261], achieves minimum sum-of-squares indicator). These last properties can also be deduced from a more precise divisibility bound shown later in [63]:

Theorem 13 *Let f be any n-variable m-resilient function ($m \leq n-2$) and let d be its algebraic degree. The values of the Walsh transform of f are divisible by $2^{m+2+\lfloor \frac{n-m-2}{d} \rfloor}$. Hence the nonlinearity of f is divisible by $2^{m+1+\lfloor \frac{n-m-2}{d} \rfloor}$.*

The approach for proving this result was first to use the numerical normal form (see [63]). Later, a second proof using only the properties of the Fourier transform was given in [94]:

Proof. The Poisson summation formula (18) applied to $\varphi = \mathrm{f}_\chi$ and to the vectorspace $E = \{u \in \mathbb{F}_2^n / \ \forall i \in N,\ u_i \leq v_i\}$ where v is some vector of $\mathbb{F}_2^n$, whose orthogonal equals $E^\perp = \{u \in \mathbb{F}_2^n / \ \forall i \in N,\ u_i \leq v_i \oplus 1\}$, gives $\sum_{u \in E} \widehat{\mathrm{f}_\chi}(u) = 2^{w_H(v)} \sum_{x \in E^\perp} \mathrm{f}_\chi(x)$. It is then a simple matter to prove the result by induction on the weight of v, starting with the vectors of weight $m+1$ (since it is obvious for the vectors of weights at most m), and using McEliece's divisibility property (see Subsection 3.1). □

A similar proof shows that the values of the Walsh transform of any m-th order correlation-immune function are divisible by $2^{m+1+\lfloor \frac{n-m-1}{d} \rfloor}$ (and by $2^{m+2+\lfloor \frac{n-m-2}{d} \rfloor}$ if its weight is divisible $2^{m+1+\lfloor \frac{n-m-2}{d} \rfloor}$, see [94]).

[49] Also Zheng and Zhang [366], showed that the upper bound on the nonlinearity of correlation-immune functions of high orders is the same as the upper bound on the nonlinearity of resilient functions of the same orders. The distances between resilient functions and Reed-Muller codes of orders greater than 1 have also been studied by Kurosawa et al. and by Borissov at al. [221, 23].

Theorem 13 gives directly a more precise upper bound on the nonlinearity of any m-resilient function of degree d: this nonlinearity is bounded above by $2^{n-1} - 2^{m+1+\left\lfloor \frac{n-m-2}{d} \right\rfloor}$. This gives a simpler proof that it can be equal to $2^{n-1} - 2^{m+1}$ only if $d = n - m - 1$, *i*.e. if Siegenthaler's bound is achieved. Moreover, the proof above also shows that the nonlinearity of any m-resilient n-variable Boolean function is bounded above by $2^{n-1} - 2^{m+1+\left\lfloor \frac{n-m-2}{d} \right\rfloor}$ where d is the minimum algebraic degree of the restrictions of f to the subspaces $\{u \in \mathbb{F}_2^n / \ \forall i \in N,\ u_i \leq v_i \oplus 1\}$ such that v has weight $m+1$ and $\widehat{\mathrm{f}_\chi}(v) \neq 0$.

If $2^{n-1} - 2^{m+1}$ is greater than the best possible nonlinearity of all balanced functions (and in particular if it is greater than the covering radius bound) then, obviously, a better bound exists. In the case of n even, the best possible nonlinearity of all balanced functions being strictly smaller than $2^{n-1} - 2^{n/2-1}$, Sarkar and Maitra deduce that $nl(f) \leq 2^{n-1} - 2^{n/2-1} - 2^{m+1}$ for every m-resilient function f with $m \leq n/2 - 2$. In the case of n odd, they state that $nl(f)$ is smaller than or equal to the highest multiple of 2^{m+1}, which is less than or equal to the best possible nonlinearity of all Boolean functions. But a potentially better upper bound can be given, whatever is the parity of n. Indeed, Sarkar-Maitra's divisibility bound shows that $\widehat{\mathrm{f}_\chi}(a) = \varphi(a) \times 2^{m+2}$ where $\varphi(a)$ is integer-valued. But Parseval's Relation (23) and the fact that $\widehat{\mathrm{f}_\chi}(a)$ is null for every vector a of weight $\leq m$ imply

$$\sum_{a \in \mathbb{F}_2^n / \ w_H(a) > m} \varphi^2(a) = 2^{2n-2m-4}$$

and, thus,

$$\max_{a \in \mathbb{F}_2^n} |\varphi(a)| \geq \sqrt{\frac{2^{2n-2m-4}}{2^n - \sum_{i=0}^{m} \binom{n}{i}}} = \frac{2^{n-m-2}}{\sqrt{2^n - \sum_{i=0}^{m} \binom{n}{i}}}.$$

Hence, we have $\max_{a \in \mathbb{F}_2^n} |\varphi(a)| \geq \left\lceil \frac{2^{n-m-2}}{\sqrt{2^n - \sum_{i=0}^{m} \binom{n}{i}}} \right\rceil$, and this implies:

$$nl(f) \leq 2^{n-1} - 2^{m+1} \left\lceil \frac{2^{n-m-2}}{\sqrt{2^n - \sum_{i=0}^{m} \binom{n}{i}}} \right\rceil. \tag{57}$$

When n is even and $m \leq n/2 - 2$, this number is always less than or equal to the number $2^{n-1} - 2^{n/2-1} - 2^{m+1}$ (given by Sarkar and Maitra),

because $\frac{2^{n-m-2}}{\sqrt{2^n-\sum_{i=0}^{m}\binom{n}{i}}}$ is strictly greater than $2^{n/2-m-2}$ and $2^{n/2-m-2}$ is an integer, and, thus, $\left\lceil \frac{2^{n-m-2}}{\sqrt{2^n-\sum_{i=0}^{m}\binom{n}{i}}} \right\rceil$ is at least $2^{n/2-m-2}+1$. And when n increases, the right hand-side of Relation (57) is strictly smaller than $2^{n-1}-2^{n/2-1}-2^{m+1}$ for an increasing number of values of $m \leq n/2-2$ (but this improvement does not appear when we compare the values we obtain with this bound to the values indicated in the table given by Sarkar and Maitra in [321], because the values of n they consider in this table are small).

When n is odd, it is difficult to say if Inequality (57) is better than the bound given by Sarkar and Maitra, because their bound involves a value which is unknown for $n \geq 9$ (the best possible nonlinearity of all balanced Boolean functions). In any case, this makes (57) better usable.

We know (see [258], page 310) that $\sum_{i=0}^{m}\binom{n}{i} \geq \frac{2^{nH_2(m/n)}}{\sqrt{8m(1-m/n)}}$, where $H_2(x) = -x\log_2(x)-(1-x)\log_2(1-x)$ is the so-called *binary entropy function* and satisfies $H_2(\frac{1}{2}-x) = 1-2x^2\log_2 e + o(x^2)$. Thus, we have

$$nl(f) \leq 2^{n-1} - 2^{m+1}\left\lceil \frac{2^{n-m-2}}{\sqrt{2^n - \frac{2^{nH_2(m/n)}}{\sqrt{8m(1-m/n)}}}} \right\rceil . \tag{58}$$

7.3 Bound on the maximum correlation with subsets of N

An upper bound on the maximum correlation of m-resilient functions with respect to subsets I of N can be directly deduced from Relation (40) and from Sarkar et al.'s bound. Note that we get an improvement by using that the support of $\widehat{f_\chi}$, restricted to the set of vectors $u \in \mathbb{F}_2^n$ such that $u_i = 0$, $\forall i \notin I$, contains at most $\sum_{i=m+1}^{|I|}\binom{|I|}{i}$ vectors. In particular, if $|I| = m+1$, the maximum correlation of f with respect to I equals $2^{-n}|\widehat{f_\chi}(u)|$, where u is the vector of support I, see [38, 47, 358]. The optimal number of LFSRs which should be considered together in a correlation attack on a cryptosystem using an m-resilient combining function is $m+1$, see [38].

7.4 Relationship with other criteria

The relationships between resiliency and other criteria have been studied in [105, 261, 348, 365]. For instance, m-resilient $PC(l)$ functions can exist only if $m+l \leq n-1$. This is a direct consequence of Relation (27), relating

the values of the Walsh transform of a function on a flat $a+E$ to the autocorrelation coefficients of the function on a flat $b+E^\perp$, applied with $a=b=0$, $E=\{x\in\mathbb{F}_2^n;\ x_i=0,\ \forall i\in I\}$ and $E^\perp=\{x\in\mathbb{F}_2^n;\ x_i=0,\ \forall i\notin I\}$, where I has size $n-m$: if $l\geq n-m$ then the right-hand term of (27) is non-zero while the left-hand term is null. Equality $m+l=n-1$ is possible only if $l=n-1$, n is odd and $m=0$ [365, 105]. The known upper bounds on the nonlinearity (see Section 7) can then be improved with the same argument.

The definition of resiliency has been weakened in [27, 89, 222] in order to relax some of the trade-offs recalled above without weakening the cryptosystem against the correlation attack.

Resiliency is related to the notion of corrector (useful for the generation of random sequences having good statistical properties) introduced by P. Lacharme in [224].

7.5 Constructions

High order resilient functions with high algebraic degrees, high nonlinearities and good immunity to algebraic attacks are needed for applications in stream ciphers using the combiner model. But designing constructions of Boolean functions meeting all these cryptographic criteria is still a challenge nowadays (while we would need numerous such functions in order to be able choosing among them functions satisfying additional design criteria). The primary constructions (which allow designing resilient functions without using known ones) lead potentially to wider classes of functions than secondary constructions (recall that the number of Boolean functions on $n-1$ variables is only equal to the square root of the number of n-variable Boolean functions). But the known primary constructions of such Boolean functions do not lead to very large classes of functions. In fact, only one reasonably large class of Boolean functions is known, whose elements can be analyzed with respect to the cryptographic criteria recalled in Subsection 4.1. So we observe some imbalance in the knowledge on cryptographic functions for stream ciphers: much is known on the properties of resilient functions, but little is known on how constructing them. Examples of m-resilient functions achieving the best possible nonlinearity $2^{n-1}-2^{m+1}$ (and thus the best algebraic degree) have been obtained for $n\leq 10$ in [292, 320, 321] and for every $m\geq 0.6\ n$ [345, 346] (n being then not limited). But $n\leq 10$ is too small for applications and $m\geq 0.6\ n$ is too large (because of Siegenthaler's bound) and almost nothing is known on the immunity of these functions to algebraic attacks. Moreover, these examples give very limited numbers of

functions (they are often defined recursively or obtained after a computer search) and many of these functions have cryptographic weaknesses such as linear structures (see [105, 261]). Balanced Boolean functions with high nonlinearities have been obtained by C. Fontaine in [154] and by E. Filiol and C. Fontaine in [152], who made a computer investigation - but for $n = 7$, 9 which is too small - on the corpus of *idempotent functions*. These functions, whose ANFs are invariant under the cyclic shifts of the coordinates x_i, have been called later *rotation symmetric* (see Subsection 10.6).

7.5.1 Primary constructions

Maiorana-McFarland's construction: An extension of the class of bent functions that we called above the Maiorana-McFarland original class has been given in [35], based on the same principle of concatenating affine functions[50] (we have already met this generalization in Section 6, that we shall call the *Maiorana-McFarland general construction*): let r be a positive integer smaller than n; we denote $n - r$ by s; let g be any Boolean function on $\mathbb{F}_2^s$ and let ϕ be a mapping from $\mathbb{F}_2^s$ to $\mathbb{F}_2^r$. Then, we define the function:

$$f_{\phi,g}(x, y) = x \cdot \phi(y) \oplus g(y) = \bigoplus_{i=1}^{r} x_i \phi_i(y) \oplus g(y),\ x \in \mathbb{F}_2^r,\ y \in \mathbb{F}_2^s \qquad (59)$$

where $\phi_i(y)$ is the i-th coordinate function of $\phi(y)$.

For every $a \in \mathbb{F}_2^r$ and every $b \in \mathbb{F}_2^s$, we have seen in Subsection 6.4 that

$$\widehat{\mathrm{f}_{\phi,\mathrm{g}_\chi}}(a, b) = 2^r \sum_{y \in \phi^{-1}(a)} (-1)^{g(y) \oplus b \cdot y}. \qquad (60)$$

This can be used to design resilient functions: if every element in $\phi(\mathbb{F}_2^s)$ has Hamming weight strictly greater than k, then $f_{\phi,g}$ is m-resilient with $m \geq k$ (in particular, if $\phi(\mathbb{F}_2^s)$ does not contain the null vector, then $f_{\phi,g}$ is balanced). Indeed, if $w_H(a) \leq k$ then $\phi^{-1}(a)$ is empty in Relation (60); hence, if $w_H(a) + w_H(b) \leq k$ then $\widehat{\mathrm{f}_{\phi,\mathrm{g}_\chi}}(a, b)$ is null. The k-resiliency of $f_{\phi,g}$ under this hypothesis can also be deduced from the facts that any affine function $x \in \mathbb{F}_2^r \mapsto a \cdot x \oplus \varepsilon$ ($a \in \mathbb{F}_2^r$ nonzero, $\varepsilon \in \mathbb{F}_2$) is $(w_H(a) - 1)$-resilient, and that any Boolean function equal to the concatenation of k-resilient functions is a k-resilient function (see secondary construction 3 below).

[50] These functions have also been studied under the name of linear-based functions in [1, 355].

Degree: The algebraic degree of $f_{\phi,g}$ is at most $s+1 = n-r+1$. It equals $s+1$ if and only if ϕ has algebraic degree s (*i*.e. if at least one of its coordinate functions has algebraic degree s). If we assume that every element in $\phi(\mathbb{F}_2^s)$ has Hamming weight strictly greater than k, then ϕ can have algebraic degree s only if $k \leq r-2$, since if $k = r-1$ then ϕ is constant. Thus, if $m = k$ then the algebraic degree of $f_{\phi,g}$ reaches Siegenthaler's bound $n-k-1$ if and only if either $k = r-2$ and ϕ has algebraic degree $s = n-k-2$ or $k = r-1$ and g has algebraic degree $s = n-k-1$. There are cases where $m > k$ (see [118, 64, 65]). An obvious one is when each set $\phi^{-1}(a)$ has even size and the restriction of g to this set is balanced: then $m \geq k+1$.

Nonlinearity: Relations (35), relating the nonlinearity to the Walsh transform, and (60) lead straightforwardly to a general lower bound on the nonlinearity of Maiorana-McFarland's functions (first observed in [327]):

$$nl(f_{\phi,g}) \geq 2^{n-1} - 2^{r-1} \max_{a \in \mathbb{F}_2^r} |\phi^{-1}(a)| \tag{61}$$

(where $|\phi^{-1}(a)|$ denotes the size of $\phi^{-1}(a)$). A more recent upper bound

$$nl(f_{\phi,g}) \leq 2^{n-1} - 2^{r-1} \left\lceil \sqrt{\max_{a \in \mathbb{F}_2^r} |\phi^{-1}(a)|} \right\rceil \tag{62}$$

obtained in [64] strengthens a bound previously obtained in [107, 108] which stated $nl(f_{\phi,g}) \leq 2^{n-1} - 2^{r-1}$.

Proof of (62): The sum

$$\sum_{b \in \mathbb{F}_2^s} \left(\sum_{y \in \phi^{-1}(a)} (-1)^{g(y)+b \cdot y} \right)^2 = \sum_{b \in \mathbb{F}_2^s} \left(\sum_{y,z \in \phi^{-1}(a)} (-1)^{g(y)+g(z)+b \cdot (y+z)} \right)$$

equals $2^s |\phi^{-1}(a)|$ (since the sum $\sum_{b \in \mathbb{F}_2^s} (-1)^{b \cdot (y+z)}$ is null if $y \neq z$). The maximum of a set of values being always greater than or equal to its mean, we deduce

$$\max_{b \in \mathbb{F}_2^s} \left| \sum_{y \in \phi^{-1}(a)} (-1)^{g(y)+b \cdot y} \right| \geq \sqrt{|\phi^{-1}(a)|}$$

and thus, according to Relation (60):

$$\max_{a \in \mathbb{F}_2^r; b \in \mathbb{F}_2^s} |\widehat{\mathrm{f}_{\chi}}_{\phi,g}(a,b)| \geq 2^r \left\lceil \sqrt{\max_{a \in \mathbb{F}_2^r} |\phi^{-1}(a)|} \right\rceil .$$

Relation (35) completes the proof. □

This new bound allowed characterizing the Maiorana-McFarland's functions $f_{\phi,g}$ such that $w_H(\phi(y)) > k$ for every y and achieving nonlinearity $2^{n-1} - 2^{k+1}$: the inequality $nl(f_{\phi,g}) \leq 2^{n-1} - \frac{2^{r+\frac{s}{2}-1}}{\sqrt{\sum_{i=k+1}^{r}\binom{r}{i}}}$ implies either that $r = k+1$ or $r = k+2$.
If $r = k+1$, then ϕ is the constant $(1, \cdots, 1)$ and $n \leq k+3$. Either $s = 1$ and $g(y)$ is then any function in one variable, or $s = 2$ and g is then any function of the form $y_1y_2 \oplus \ell(y)$ where ℓ is affine (thus, f is quadratic).
If $r = k+2$, then ϕ is injective, $n \leq k + 2 + \log_2(k+3)$, g is any function on $n-k-2$ variables and $d^\circ f_{\phi,g} \leq 1 + \log_2(k+3)$.

A simple example of k-resilient Maiorana-McFarland's functions such that $nl(f_{\phi,g}) = 2^{n-1} - 2^{k+1}$ (and thus achieving Sarkar et al.'s bound) can be given for any $r \geq 2^s - 1$ and for $k = r - 2$ (see [64]). And, for every even $n \leq 10$, Sarkar et al.'s bound with $m = n/2 - 2$ can be achieved by Maiorana-McFarland's functions. Also, functions with high nonlinearities but achieving not Sarkar et al.'s bound exist in Maiorana-McFarland's class (for instance, for every $n \equiv 1 \,[\bmod 4]$, there exist such $\frac{n-1}{4}$-resilient functions on $\mathbb{F}_2^n$ with nonlinearity $2^{n-1} - 2^{\frac{n-1}{2}}$).

Generalizations of Maiorana-McFarland's construction have been introduced in [64] and [93]; the latter generalization has been further generalized into a class introduced in [68]. A motivation for introducing such generalizations is that Maiorana-McFarland's functions have the weakness that $x \mapsto f_{\phi,g}(x,y)$ is affine for every $y \in \mathbb{F}_2^s$ and have high divisibilities of their Fourier spectra (indeed, if we want to ensure that f is m-resilient with large value of m, then we need to choose r large; then the Walsh spectrum of f is divisible by 2^r according to Relation (60); there is also a risk that this property can be used in attacks, as it is used in [48] to attack block ciphers). The functions constructed in [64, 93] are concatenations of quadratic functions instead of affine functions. This makes them harder to study than Maiorana-McFarland's functions. But they are more numerous and more general. Two classes of such functions have been studied:
- the functions of the first class are defined as:

$$f_{\psi,\phi,g}(x,y) = \bigoplus_{i=1}^{t} x_{2i-1}x_{2i}\,\psi_i(y) \oplus x \cdot \phi(y) \oplus g(y),$$

with $x \in \mathbb{F}_2^r$, $y \in \mathbb{F}_2^s$, where $n = r + s$, $t = \left\lfloor \frac{r}{2} \right\rfloor$, and where $\psi : \mathbb{F}_2^s \to \mathbb{F}_2^t$, $\phi : \mathbb{F}_2^s \to \mathbb{F}_2^r$ and $g : \mathbb{F}_2^s \to \mathbb{F}_2$ can be chosen arbitrarily;
- the functions of the second class are defined as:

$$f_{\phi_1,\phi_2,\phi_3,g}(x,y) = (x \cdot \phi_1(y))\,(x \cdot \phi_2(y)) \oplus x \cdot \phi_3(y) \oplus g(y),$$

with $x \in \mathbb{F}_2^r$, $y \in \mathbb{F}_2^s$, where ϕ_1, ϕ_2 and ϕ_3 are three functions from $\mathbb{F}_2^s$ into $\mathbb{F}_2^r$ and g is any Boolean function on $\mathbb{F}_2^s$. The size of this class roughly equals $\left[(2^r)^{2^s}\right]^3 \times 2^{2^s} = 2^{(3r+1)2^s}$ (the exact number, which is unknown, is smaller since a same function can be represented in this form in several ways) and is larger than the size of the first class, roughly equal to $(2^t)^{2^s} \times (2^r)^{2^s} \times 2^{2^s} = 2^{(t+r+1)2^s}$.
The second construction has been generalized in [68]. The functions of this generalized class are the concatenations of functions equal to the sums of r-variable affine functions and of flat-indicators:

$$\forall (x,y) \in \mathbb{F}_2^r \times \mathbb{F}_2^s, f(x,y) = \prod_{i=1}^{\varphi(y)} (x \cdot \phi_i(y) \oplus g_i(y) \oplus 1) \oplus x \cdot \phi(y) \oplus g(y),$$

where φ is a function from $\mathbb{F}_2^s$ into $\{0, 1, \cdots, r\}$, $\phi_1, \cdots, \phi_r$ and ϕ are functions from $\mathbb{F}_2^s$ into $\mathbb{F}_2^r$ such that, for every $y \in \mathbb{F}_2^s$, the vectors $\phi_1(y), \cdots, \phi_{\varphi(y)}(y)$ are linearly independent, and $g_1, \cdots, g_r$ and g are Boolean functions on $\mathbb{F}_2^s$. There exist formulae for the Walsh transforms of the functions of these classes, which result in sufficient conditions for their resiliency and in bounds on their nonlinearities (see [64, 68]).

Other constructions: We first make a *preliminary observation.* Let $k < n$. For any k-variable function g, any surjective linear mapping $L : \mathbb{F}_2^n \to \mathbb{F}_2^k$ and any element s of $\mathbb{F}_2^n$; the function $f(x) = g \circ L(x) \oplus s \cdot x$ is $(d-1)$-resilient, where d is the Hamming distance between s and the linear code C whose generator matrix equals the matrix of L. Indeed, for any vector $a \in \mathbb{F}_2^n$ of Hamming weight at most $d - 1$, the vector $s + a$ does not belong to C. This implies that the Boolean function $f(x) \oplus a \cdot x$ is linearly equivalent to the function $g(x_1, \cdots, x_k) \oplus x_{k+1}$, since we may assume without loss of generality that L is systematic (*i.e.* has the form $[Id_k|N]$); it is therefore balanced. But such function f having nonzero linear structures, it does not give full satisfaction.

A construction derived from $\mathcal{PS}_{ap}$ construction is introduced in [58] to obtain resilient functions: let k and r be positive integers and $n \geq r$; we denote

$n-r$ by s; the vectorspace $\mathbb{F}_2^r$ is identified to the Galois field $\mathbb{F}_{2^r}$. Let g be any Boolean function on $\mathbb{F}_{2^r}$ and ϕ an $\mathbb{F}_2$-linear mapping from $\mathbb{F}_2^s$ to $\mathbb{F}_{2^r}$; set $a \in \mathbb{F}_{2^r}$ and $b \in \mathbb{F}_2^s$ such that, for every y in $\mathbb{F}_2^s$ and every z in $\mathbb{F}_{2^r}$, $a+\phi(y)$ is nonzero and $\phi^*(z)+b$ has weight greater than k, where ϕ^* is the adjoint of ϕ (satisfying $u \cdot \phi(x) = \phi^*(u) \cdot x$ for every x and u, that is, having for matrix the transpose of that of ϕ). Then, the function

$$f(x,y) = g\left(\frac{x}{a+\phi(y)}\right) \oplus b \cdot y, \text{ where } x \in \mathbb{F}_{2^r}, y \in \mathbb{F}_2^s, \tag{63}$$

is m-resilient with $m \geq k$. There exist bounds on the nonlinearities of these functions (see [65]), similar to those existing for Maiorana-McFarland's functions. But this class has much fewer elements than Maiorana-McFarland's class, because ϕ is linear.

Dobbertin's construction: in [141] is given a nice generalization of a method, introduced by Seberry et al. in [328], for modifying bent functions into balanced functions with high nonlinearities. He observes that most known bent functions on $\mathbb{F}_2^n$ (n even) are normal (that is, constant on at least one $n/2$-dimensional flat). Up to affine equivalence, we can then assume that $f(x,y)$, $x \in \mathbb{F}_2^{n/2}$, $y \in \mathbb{F}_2^{n/2}$ is such that $f(x,0) = \varepsilon$ ($\varepsilon \in \mathbb{F}_2$) for every $x \in \mathbb{F}_2^{n/2}$ and that $\varepsilon = 0$ (otherwise, consider $f \oplus 1$).

Proposition 33 *Let $f(x,y)$, $x \in \mathbb{F}_2^{n/2}$, $y \in \mathbb{F}_2^{n/2}$ be any bent function such that $f(x,0) = 0$ for every $x \in \mathbb{F}_2^{n/2}$ and let g be any balanced function on $\mathbb{F}_2^{n/2}$. Then the Walsh transform of the function $h(x,y) = f(x,y) \oplus \delta_0(y)g(x)$, where δ_0 is the Dirac symbol, satisfies:*

$$\widehat{\mathrm{h}_\chi}(u,v) = 0 \text{ if } u = 0 \text{ and } \widehat{\mathrm{h}_\chi}(u,v) = \widehat{\mathrm{f}_\chi}(u,v) + \widehat{\mathrm{g}_\chi}(u) \text{ otherwise.} \tag{64}$$

Proof. We have $\widehat{\mathrm{h}_\chi}(u,v) = \widehat{\mathrm{f}_\chi}(u,v) - \sum_{x \in \mathbb{F}_2^{n/2}} (-1)^{u \cdot x} + \sum_{x \in \mathbb{F}_2^{n/2}} (-1)^{g(x) \oplus u \cdot x} = \widehat{\mathrm{f}_\chi}(u,v) - 2^{n/2}\delta_0(u) + \widehat{\mathrm{g}_\chi}(u)$. The function g being balanced, we have $\widehat{\mathrm{g}_\chi}(0) = 0$. And $\widehat{\mathrm{f}_\chi}(0,v)$ equals $2^{n/2}$ for every v, since f is null on $\mathbb{F}_2^{n/2} \times \{0\}$ and according to Relation (46) applied to $E = \{0\} \times \mathbb{F}_2^{n/2}$ and $a = b = 0$ (or see the remark after Theorem 9). □

We deduce that:

$$\max_{u,v \in \mathbb{F}_2^{n/2}} |\widehat{\mathrm{h}_\chi}(u,v)| \leq \max_{u,v \in \mathbb{F}_2^{n/2}} |\widehat{\mathrm{f}_\chi}(u,v)| + \max_{u \in \mathbb{F}_2^{n/2}} |\widehat{\mathrm{g}_\chi}(u)|,$$

i.e. that $2^n - 2nl(h) \leq 2^n - 2nl(f) + 2^{n/2} - 2nl(g)$, that is:

$$nl(h) \geq nl(f) + nl(g) - 2^{n/2-1} = 2^{n-1} - 2^{n/2} + nl(g).$$

Applying recursively this principle (if $n/2$ is even, g can be constructed in the same way), we see that if $n = 2^k\, n'$ (n' odd), Dobbertin's method allows reaching the nonlinearity $2^{n-1} - 2^{n/2-1} - 2^{\frac{n}{4}-1} - \cdots - 2^{n'-1} - 2^{\frac{n'-1}{2}}$ since we know that, for every odd n', the nonlinearity of functions on $\mathbb{F}_2^{n'}$ can be as high as $2^{n'-1} - 2^{\frac{n'-1}{2}}$, and that balanced (quadratic) functions can achieve this value. If $n' \leq 7$ then this value is the best possible and $2^{n-1} - 2^{n/2-1} - 2^{\frac{n}{4}-1} - \cdots - 2^{n'-1} - 2^{\frac{n'-1}{2}}$ is therefore the best known nonlinearity of balanced functions in general. For $n' > 7$, the best nonlinearity of balanced n'-variable functions is larger than $2^{n'-1} - 2^{\frac{n'-1}{2}}$ (see the paragraph devoted to nonlinearity in Section 4.1) and $2^{n-1} - 2^{n/2-1} - 2^{\frac{n}{4}-1} - \cdots - 2^{2n'-1} - 2^{n'} + nl(g)$, where g is an n'-variable balanced function, can therefore reach higher values.

Unfortunately, according to Relation (64), Dobbertin's construction cannot produce m-resilient functions with $m > 0$ since, g being a function defined on $\mathbb{F}_2^{n/2}$, there cannot exist more than one vector a such that $\widehat{\mathrm{g}_\chi}(a)$ equals $\pm 2^{n/2}$.

7.5.2 Secondary constructions

There exist several simple secondary constructions, which can be combined to obtain resilient functions achieving the bounds of Sarkar et al. and Siegenthaler. We list them below in chronological order. As we shall see in the end, they all are particular cases of a single general one.

I Direct sum of functions

A. *Adding a variable*

Let f be an r-variable t-resilient function. The Boolean function on $\mathbb{F}_2^{r+1}$:

$$h(x_1, \cdots, x_r, x_{r+1}) = f(x_1, \cdots, x_r) \oplus x_{r+1}$$

is $(t+1)$-resilient [336]. If f is an $(r, t, r-t-1, 2^{r-1} - 2^{t+1})$ function[51], then h is an $(r+1, t+1, r-t-1, 2^r - 2^{t+2})$ function, and thus achieves Siegenthaler's and Sarkar et al.'s bounds. But h has the linear structure $(0, \cdots, 0, 1)$.

[51] Recall that, by an $(n, m, d, \mathcal{N})$- function, we mean an n-variable, m-resilient function having algebraic degree at least d and nonlinearity at least $\mathcal{N}$.

B. *Generalization*
If f is an r-variable t-resilient function ($t \geq 0$) and if g is an s-variable m-resilient function ($m \geq 0$), then the function:

$$h(x_1, \cdots, x_r, x_{r+1}, \cdots, x_{r+s}) = f(x_1, \cdots, x_r) \oplus g(x_{r+1}, \cdots, x_{r+s})$$

is $(t+m+1)$-resilient. This comes from the easily provable relation $\widehat{\mathrm{h}_\chi}(a,b) = \widehat{\mathrm{f}_\chi}(a) \times \widehat{\mathrm{g}_\chi}(b)$, $a \in \mathbb{F}_2^r$, $b \in \mathbb{F}_2^s$. We have also $d^\circ h = \max(d^\circ f, d^\circ g)$ and, thanks to Relation (35) relating the nonlinearity to the Walsh transform, $nl(h) = 2^{r+s-1} - \frac{1}{2}(2^r - 2nl(f))(2^s - 2nl(g)) = 2^r nl(g) + 2^s nl(f) - 2nl(f)nl(g)$. Such decomposable function does not give full satisfaction. Moreover, h has low algebraic degree, in general. And if $nl(f) = 2^{r-1} - 2^{t+1}$ ($t \leq r-2$) and $nl(g) = 2^{s-1} - 2^{m+1}$ ($m \leq s-2$), *i.e.* if $nl(f)$ and $nl(g)$ have maximum possible values, then $nl(h) = 2^{r+s-1} - 2^{t+m+3}$ and h does not achieve Sarkar's and Maitra's bound (note that this is not in contradiction with the properties of the construction recalled in **I.A**, since the function $g(x_{r+1}) = x_{r+1}$ is 1-variable, 0-resilient and has null nonlinearity).
Function h has no nonzero linear structure if and only if f and g both have no nonzero linear structure.

II. Siegenthaler's construction
Let f and g be two Boolean functions on $\mathbb{F}_2^r$. Let us consider the function

$$h(x_1, \cdots, x_r, x_{r+1}) = (x_{r+1} \oplus 1) f(x_1, \cdots, x_r) \oplus x_{r+1} g(x_1, \cdots, x_r)$$

on $\mathbb{F}_2^{r+1}$. Note that the truth-table of h can be obtained by concatenating the truth-tables of f and g. Then:

$$\widehat{\mathrm{h}_\chi}(a_1, \cdots, a_r, a_{r+1}) = \widehat{\mathrm{f}_\chi}(a_1, \cdots, a_r) + (-1)^{a_{r+1}} \widehat{\mathrm{g}_\chi}(a_1, \cdots, a_r). \qquad (65)$$

Thus:
1. If f and g are m-resilient, then h is m-resilient [336]; moreover, if for every $a \in \mathbb{F}_2^r$ of Hamming weight $m+1$, we have $\widehat{\mathrm{f}_\chi}(a) + \widehat{\mathrm{g}_\chi}(a) = 0$, then h is $(m+1)$-resilient. Note that the construction recalled in **I.A** corresponds to $g = f \oplus 1$ and satisfies this condition. Another possible choice of a function g satisfying this condition (first pointed out in [35]) is $g(x) = f(x_1 \oplus 1, \cdots, x_r \oplus 1) \oplus \epsilon$, where $\epsilon = m \,[\bmod 2]$, since $\widehat{\mathrm{g}_\chi}(a) = \sum_{x \in \mathbb{F}_2^r} (-1)^{f(x) \oplus \epsilon \oplus (x \oplus (1, \cdots, 1)) \cdot a} = (-1)^{\epsilon + w_H(a)} \widehat{\mathrm{f}_\chi}(a)$. It leads to a function h having also a nonzero linear structure (namely, the vector $(1, \cdots, 1)$);
2. The value $\max_{a_1, \cdots, a_{r+1} \in \mathbb{F}_2} |\widehat{\mathrm{h}_\chi}(a_1, \cdots, a_r, a_{r+1})|$ is bounded above by the number $\max_{a_1, \cdots, a_r \in \mathbb{F}_2} |\widehat{\mathrm{f}_\chi}(a_1, \cdots, a_r)| + \max_{a_1, \cdots, a_r \in \mathbb{F}_2} |\widehat{\mathrm{g}_\chi}(a_1, \cdots, a_r)|$; this implies $2^{r+1} -$

$2nl(h) \leq 2^{r+1} - 2nl(f) - 2nl(g)$, that is $nl(h) \geq nl(f) + nl(g)$;
a. if f and g achieve maximum possible nonlinearity $2^{r-1} - 2^{m+1}$ and if h is $(m+1)$-resilient, then the nonlinearity $2^r - 2^{m+2}$ of h is the best possible;
b. if f and g are such that, for every vector a, at least one of the numbers $\widehat{\mathrm{f}_\chi}(a)$, $\widehat{\mathrm{g}_\chi}(a)$ is null (in other words, if the supports of the Walsh transforms of f and g are disjoint), then we have $\max_{a_1,\cdots,a_{r+1}\in\mathbb{F}_2} |\widehat{\mathrm{h}_\chi}(a_1,\cdots,a_r,a_{r+1})| = \max\left(\max_{a_1,\cdots,a_r\in\mathbb{F}_2} |\widehat{\mathrm{f}_\chi}(a_1,\cdots,a_r)|; \max_{a_1,\cdots,a_r\in\mathbb{F}_2} |\widehat{\mathrm{g}_\chi}(a_1,\cdots,a_r)|\right)$. Hence we have $2^{r+1} - 2nl(h) = 2^r - 2\min(nl(f), nl(g))$ and $nl(h)$ equals therefore $2^{r-1} + \min(nl(f), nl(g))$; thus, if f and g achieve best possible nonlinearity $2^{r-1} - 2^{m+1}$, then h achieves best possible nonlinearity $2^r - 2^{m+1}$;
3. If the monomials of highest degree in the algebraic normal forms of f and g are not all the same, then $d^\circ h = 1 + \max(d^\circ f, d^\circ g)$. Note that this condition is not satisfied in the two cases indicated above in **1**, for which h is $(m+1)$-resilient.
4. For every $a = (a_1, \cdots, a_r) \in \mathbb{F}_2^r$ and every $a_{r+1} \in \mathbb{F}_2$, we have, denoting $(x_1, \cdots, x_r)$ by x: $D_{(a,a_{r+1})}h(x, x_{r+1}) = D_a f(x) \oplus a_{r+1}(f \oplus g)(x) \oplus x_{r+1}D_a(f \oplus g)(x) \oplus a_{r+1}D_a(f \oplus g)(x)$. If $d^\circ(f \oplus g) \geq d^\circ f$, then $D_{(a,1)}h$ is non-constant, for every a. And if, additionally, there does not exist $a \neq 0$ such that $D_a f$ and $D_a g$ are constant and equal to each other, then h admits no nonzero linear structure.
This construction allows obtaining:
- from any two m-resilient functions f and g having disjoint Walsh spectra, achieving nonlinearity $2^{r-1} - 2^{m+1}$ and such that $d^\circ(f \oplus g) = r - m - 1$, an m-resilient function h having algebraic degree $r - m$ and having nonlinearity $2^r - 2^{m+1}$, that is, achieving Siegenthaler's and Sarkar et al.'s bounds; note that this construction increases (by 1) the algebraic degrees of f and g;
- from any m-resilient function f achieving algebraic degree $r - m - 1$ and nonlinearity $2^{r-1} - 2^{m+1}$, a function h having resiliency order $m + 1$ and nonlinearity $2^r - 2^{m+2}$, that is, achieving Siegenthaler's and Sarkar et al.'s bounds and having same algebraic degree as f (but having nonzero linear structures).
So it allows, when combining these two methods, to keep best tradeoffs between resiliency order, algebraic degree and nonlinearity, and to increase by 1 the degree and the resiliency order.
Generalization: let $(f_y)_{y\in\mathbb{F}_2^s}$ be a family of r-variable m-resilient functions; then the function on $\mathbb{F}_2^{r+s}$ defined by $f(x, y) = f_y(x)$ ($x \in \mathbb{F}_2^r$, $y \in \mathbb{F}_2^s$) is m-resilient. Indeed, we have $\widehat{\mathrm{f}_\chi}(a, b) = \sum_{y\in\mathbb{F}_2^s}(-1)^{b\cdot y}\,\widehat{\mathrm{f}_{y_\chi}}(a)$. The function f corresponds to the concatenation of the functions f_y; hence, this secondary construction can be viewed as a generalization of Maiorana-McFarland's

construction (in which the functions f_y are m-resilient affine functions).

More on the resilient functions, achieving high nonlinearities, and constructed by using, among others, the secondary constructions above (as well as algorithmic methods) can be found in [216, 291].

III. Tarannikov's elementary construction

Let g be any Boolean function on $\mathbb{F}_2^r$. We define the Boolean function h on $\mathbb{F}_2^{r+1}$ by $h(x_1, \cdots, x_r, x_{r+1}) = x_{r+1} \oplus g(x_1, \cdots, x_{r-1}, x_r \oplus x_{r+1})$. By the change of variable $x_r \leftarrow x_r \oplus x_{r+1}$, we see that the Walsh transform $\widehat{h_\chi}(a_1, \cdots, a_{r+1})$ is equal to $\sum_{x_1,\cdots,x_{r+1}\in\mathbb{F}_2} (-1)^{a\cdot x \oplus g(x_1,\cdots,x_r)\oplus a_r x_r \oplus (a_r \oplus a_{r+1} \oplus 1)x_{r+1}}$, where $a = (a_1, \cdots, a_{r-1})$ and $x = (x_1, \cdots, x_{r-1})$; if $a_{r+1} = a_r$ then this value is null and if $a_r = a_{r+1} \oplus 1$ then it equals $2\,\widehat{g_\chi}(a_1, \cdots, a_{r-1}, a_r)$. Thus:

1. $nl(h) = 2\,nl(g)$;

2. If g is m-resilient, then h is m-resilient. If, additionally, $\widehat{g_\chi}(a_1, \cdots, a_{r-1}, 1)$ is null for every vector $(a_1, \cdots, a_{r-1})$ of weight at most m, then for every such vector $\widehat{g_\chi}(a_1, \cdots, a_{r-1}, a_r)$ is null for every a_r and h is then $(m+1)$-resilient, since if $a_r = a_{r+1} \oplus 1$ then (a_r, a_{r+1}) has weight 1; note that, in such case, if g has nonlinearity $2^{r-1} - 2^{m+1}$ then the nonlinearity of h, which equals $2^r - 2^{m+2}$ achieves then Sarkar et al.'s bound too. The condition that $\widehat{g_\chi}(a_1, \cdots, a_{r-1}, 1)$ is null for every vector $(a_1, \cdots, a_{r-1})$ of weight at most m is achieved if g does not actually depend on its last input bit; but the construction is then a particular case of the construction recalled in **I.A**. The condition is also achieved if g is obtained from two m-resilient functions, by using Siegenthaler's construction (recalled in **II**), according to Relation (65).

3. $d^\circ f = d^\circ g$ if $d^\circ g \geq 1$.

4. h has the nonzero linear structure $(0, \cdots, 0, 1, 1)$.

Tarannikov combined in [345] this construction with the constructions recalled in **I** and **II**, to build a more complex secondary construction, which allows to increase in the same time the resiliency order and the algebraic degree of the functions and which leads to an infinite sequence of functions achieving Siegenthaler's and Sarkar et al.'s bounds. Increasing then, by using the construction recalled in **I.A**, the set of ordered pairs (r, m) for which such functions can be constructed, he deduced the existence of r-variable m-resilient functions achieving Siegenthaler's and Sarkar et al.'s bounds for any number of variables r and any resiliency order m such that $m \geq \frac{2r-7}{3}$ and $m > \frac{r}{2} - 2$ (but these functions have nonzero linear structures). in [292],

Pasalic et al. slightly modified this more complex Tarannikov's construction into a construction that we shall call *Tarannikov et al.'s construction*, which allowed, when iterating it together with the construction recalled in **I.A**, to relax slightly the condition on m into $m \geq \frac{2r-10}{3}$ and $m > \frac{r}{2} - 2$.

IV. Indirect sum of functions

Tarannikov et al.'s construction has been in its turn generalized into the following construction. All the secondary constructions listed above are particular cases of it.

Theorem 14 *[67] Let r and s be positive integers and let t and m be non-negative integers such that $t < r$ and $m < s$. Let f_1 and f_2 be two r-variable t-resilient functions. Let g_1 and g_2 be two s-variable m-resilient functions. Then the function*

$$h(x, y) = f_1(x) \oplus g_1(y) \oplus (f_1 \oplus f_2)(x)\,(g_1 \oplus g_2)(y); \quad x \in \mathbb{F}_2^r, y \in \mathbb{F}_2^s$$

is an $(r+s)$-variable $(t+m+1)$-resilient function. If f_1 and f_2 are distinct and if g_1 and g_2 are distinct, then the algebraic degree of h equals $\max(d^\circ f_1, d^\circ g_1, d^\circ(f_1 \oplus f_2) + d^\circ(g_1 \oplus g_2))$; otherwise, it equals $\max(d^\circ f_1, d^\circ g_1)$. The Walsh transform of h takes value

$$\widehat{\mathrm{h}_\chi}(a, b) = \frac{1}{2}\widehat{\mathrm{f}_{1\chi}}(a)\left[\widehat{\mathrm{g}_{1\chi}}(b) + \widehat{\mathrm{g}_{2\chi}}(b)\right] + \frac{1}{2}\widehat{\mathrm{f}_{2\chi}}(a)\left[\widehat{\mathrm{g}_{1\chi}}(b) - \widehat{\mathrm{g}_{2\chi}}(b)\right]. \quad (66)$$

If the Walsh transforms of f_1 and f_2 have disjoint supports and if the Walsh transforms of g_1 and g_2 have disjoint supports, then

$$nl(h) = \min_{i,j \in \{1,2\}} \left(2^{r+s-2} + 2^{r-1} nl(g_j) + 2^{s-1} nl(f_i) - nl(f_i) nl(g_j)\right). \quad (67)$$

In particular, if f_1 and f_2 are two $(r, t, -, 2^{r-1} - 2^{t+1})$ functions with disjoint Walsh supports, if g_1 and g_2 are two $(s, m, -, 2^{s-1} - 2^{m+1})$ functions with disjoint Walsh supports, and if $f_1 \oplus f_2$ has degree $r - t - 1$ and $g_1 \oplus g_2$ has algebraic degree $s - m - 1$, then h is a $(r+s, t+m+1, r+s-t-m-2, 2^{r+s-1} - 2^{t+m+2})$ function, and thus achieves Siegenthaler's and Sarkar et al.'s bounds.

Proof. We have:

$$\widehat{\mathrm{h}_\chi}(a, b) = \sum_{y \in \mathbb{F}_2^s / \, g_1 \oplus g_2(y) = 0} \left(\sum_{x \in \mathbb{F}_2^r} (-1)^{f_1(x) \oplus a \cdot x} \right) (-1)^{g_1(y) \oplus b \cdot y}$$

$$
\begin{aligned}
&+ \sum_{y\in\mathbb{F}_2^s/\, g_1\oplus g_2(y)=1} \left(\sum_{x\in\mathbb{F}_2^r} (-1)^{f_2(x)\oplus a\cdot x}\right) (-1)^{g_1(y)\oplus b\cdot y}\\
&= \widehat{\mathrm{f}_{1_\chi}}(a) \sum_{\substack{y\in\mathbb{F}_2^s/\\ g_1\oplus g_2(y)=0}} (-1)^{g_1(y)\oplus b\cdot y} + \widehat{\mathrm{f}_{2_\chi}}(a) \sum_{\substack{y\in\mathbb{F}_2^s/\\ g_1\oplus g_2(y)=1}} (-1)^{g_1(y)\oplus b\cdot y}\\
&= \widehat{\mathrm{f}_{1_\chi}}(a) \sum_{y\in\mathbb{F}_2^s} (-1)^{g_1(y)\oplus b\cdot y} \left(\frac{1+(-1)^{(g_1\oplus g_2)(y)}}{2}\right)\\
&+ \widehat{\mathrm{f}_{2_\chi}}(a) \sum_{y\in\mathbb{F}_2^s} (-1)^{g_1(y)\oplus b\cdot y} \left(\frac{1-(-1)^{(g_1\oplus g_2)(y)}}{2}\right).
\end{aligned}
$$

We deduce Relation (66). If (a,b) has weight at most $t+m+1$ then a has weight at most t or b has weight at most m; hence we have $\widehat{\mathrm{h}_\chi}(a,b)=0$. Thus, h is $t+m+1$-resilient.
If $f_1\oplus f_2$ and $g_1\oplus g_2$ are non-constant, then the algebraic degree of h equals $\max(d^\circ f_1, d^\circ g_1, d^\circ(f_1\oplus f_2)+d^\circ(g_1\oplus g_2))$ because the terms of highest degrees in $(g_1\oplus g_2)(y)\,(f_1\oplus f_2)(x)$, in $f_1(x)$ and in $g_1(y)$ cannot cancel each others. We deduce from Relation (66) that if the supports of the Walsh transforms of f_1 and f_2 are disjoint, as well as those of g_1 and g_2, then:

$$
\max_{(a,b)\in\mathbb{F}_2^r\times\mathbb{F}_2^s} |\widehat{\mathrm{h}_\chi}(a,b)| = \frac{1}{2} \max_{i,j\in\{1,2\}} \left(\max_{a\in\mathbb{F}_2^r} |\widehat{\mathrm{f_i}}(a)| \max_{b\in\mathbb{F}_2^s} |\widehat{\mathrm{g_j}}(b)|\right)
$$

and according to Relation (35) relating the nonlinearity to the Walsh transform, this implies:

$$
2^{r+s} - 2nl(h) = \frac{1}{2} \max_{i,j\in\{1,2\}} \left((2^r - 2nl(f_i))(2^s - 2nl(g_j))\right),
$$

which is equivalent to Relation (67). □
This construction is sometimes called the *indirect sum of resilient functions*
Note that function h, defined this way, is the concatenation of the four functions f_1, $f_1\oplus 1$, f_2 and $f_2\oplus 1$, in an order controled by $g_1(y)$ and $g_2(y)$.
Examples of pairs (f_1,f_2) (or (g_1,g_2)) satisfying the hypotheses of Theorem 14 can be found in [67].

V. Constructions without extension of the number of variables
Proposition 22 leads to the following construction:

Proposition 34 *[70] Let n be any positive integer and k any non-negative integer such that $k\le n$. Let f_1, f_2 and f_3 be three k-th order correlation*

immune (resp. k-resilient) functions. Then the function $s_1 = f_1 \oplus f_2 \oplus f_3$ *is k-th order correlation immune (resp. k-resilient) if and only if the function* $s_2 = f_1 f_2 \oplus f_1 f_3 \oplus f_2 f_3$ *is k-th order correlation immune (resp. k-resilient). Moreover:*

$$nl(s_2) \geq \frac{1}{2}\left(nl(s_1) + \sum_{i=1}^{3} nl(f_i)\right) - 2^{n-1} \tag{68}$$

and if the Walsh supports of f_1, f_2 *and* f_3 *are pairwise disjoint (that is, if at most one value* $\widehat{\chi_{f_i}}(s)$, $i = 1, 2, 3$ *is nonzero, for every vector* s*), then*

$$nl(s_2) \geq \frac{1}{2}\left(nl(s_1) + \min_{1 \leq i \leq 3} nl(f_i)\right). \tag{69}$$

Proof. Relation (50) and the fact that, for every (nonzero) vector a of weight at most k, we have $\widehat{\mathrm{f}_{i_\chi}}(a) = 0$ for $i = 1, 2, 3$ imply that $\widehat{\mathrm{s}1_\chi}(a) = 0$ if and only if $\widehat{\mathrm{s}2_\chi}(a) = 0$. Relations (68) and (69) are also direct consequences of Relation (50) and of Relation (35) relating the nonlinearity to the Walsh transform. □

Note that this secondary construction is proper to allow achieving high algebraic immunity with s_2, given functions with lower algebraic immunities f_1, f_2, f_3 and s_1, since the support of s_2 can be made more complex than those of these functions. This is done without changing the number of variables and keeping similar resiliency order and nonlinearity.

Remark. Let g and h be two Boolean functions on $\mathbb{F}_2^n$ with disjoint supports and let f be equal to $g \oplus h = g + h$. Then, f is balanced if and only if $w_H(g) + w_H(h) = 2^{n-1}$. By linearity of the Fourier transform, we have: $\widehat{f} = \widehat{g} + \widehat{h}$. Thus, if g and h are m-th order correlation-immune, then f is m-resilient. For every nonzero $a \in \mathbb{F}_2^n$, we have $|\widehat{\mathrm{f}_\chi}(a)| = 2\,|\widehat{f}(a)| \leq 2\,|\widehat{g}(a)| + 2\,|\widehat{h}(a)| = |\widehat{\mathrm{g}_\chi}(a)| + |\widehat{\mathrm{h}_\chi}(a)|$. Thus, assuming that f is balanced, we have $nl(f) \geq nl(g) + nl(h) - 2^{n-1}$. The algebraic degree of f is bounded above by (and can be equal to) the maximum of the algebraic degrees of g and h.

The most part of the secondary constructions of bent functions described in Section 6.4 can be altered into constructions of correlation-immune and resilient functions, see [58].

7.6 On the number of resilient functions

It is important to ensure that the selected criteria for the Boolean functions, supposed to be used in some cryptosystems, do not restrict the choice of the functions too severely. Hence, the set of functions should be enumerated. But this enumeration is unknown for most criteria, and the case of resilient functions is not an exception in this matter. We recall below what is known.
As for bent functions, the class of balanced or resilient functions produced by Maiorana-McFarland's construction is far the widest class, compared to the classes obtained from the other usual constructions, and the number of provably balanced or resilient Maiorana-McFarland's functions seems negligible with respect to the total number of functions with the same properties. For balanced functions, this can be checked: for every positive r, the number of balanced Maiorana-McFarland's functions (59) obtained by choosing ϕ such that $\phi(y) \neq 0$, for every y, equals $(2^{r+1}-2)^{2^s}$, and is smaller than or equal to $2^{2^{n-1}}$ (since $r \geq 1$). It is quite negligible with respect to the number $\binom{2^n}{2^{n-1}} \approx \frac{2^{2^n+\frac{1}{2}}}{\sqrt{\pi 2^n}}$ of all balanced functions on $\mathbb{F}_2^n$. The number of m-resilient Maiorana-McFarland's functions obtained by choosing ϕ such that $w_H(\phi(y)) > m$ for every y equals $\left[2\sum_{i=m+1}^{r}\binom{r}{i}\right]^{2^{n-r}}$, and is probably also very small compared to the number of all m-resilient functions. But this number is unknown.
The exact numbers of m-resilient functions is known for $m \geq n-3$ (see [35], where $(n-3)$-resilient functions are characterized) and $(n-4)$-resilient functions have been characterized [75, 26].
As for bent function, an upper bound comes directly from the Siegenthaler bound on the algebraic degree: the number of m-resilient functions is bounded above by $2^{\sum_{i=0}^{n-m-1}\binom{n}{i}}$. This bound is the so-called naive bound. In 1990, Yang and Guo published an upper bound on the number of first-order correlation-immune (and thus on resilient) functions. At the same time, Denisov obtained a much stronger result (see below) but his result being published in russian, it was not known internationally. His paper was translated into english two years later but was not widely known either. This explains why several papers appeared with weaker results. Park, Lee, Sung and Kim [294] improved upon Yang-Guo's bound. Schneider [325] proved that the number of m-resilient n-variable Boolean functions is less than:

$$\prod_{i=1}^{n-m}\binom{2^i}{2^{i-1}}^{\binom{n-i-1}{m-1}}.$$

but this result was known, see [158]. A general upper bound on the number of Boolean functions whose distances to affine functions are all divisible by 2^m has been obtained in [90]. It implies an upper bound on the number of m-resilient functions which improves upon previous bounds for about half the values of (n, m) (it is better for m large). This bound divides the naive bound by approximately $2^{\sum_{i=0}^{n-m-1}\binom{m-1}{i}-1}$ if $m \geq n/2$ and by approximately $2^{2^{2m+1}-1}$ if $m < n/2$.

An upper bound on m-resilient functions ($m \geq n/2-1$) partially improving upon this latter bound was obtained for $n/2-1 \leq m < n-2$ in [84]: the number of n-variable m-resilient functions is lower than:

$$2^{\sum_{i=0}^{n-m-2}\binom{n}{i}} + \frac{\binom{n}{n-m-1}}{2^{\binom{m+1}{n-m-1}+1}} \prod_{i=1}^{n-m} \binom{2^i}{2^{i-1}}^{\binom{n-i-1}{m-1}} .$$

The expressions of these bounds seem difficult to compare mathematically. Tables have been computed in [84].

The problem of counting resilient functions is related to counting integer solutions of a system of linear equations, see [281].

An asymptotic formula for the number of m-resilient (and also for m-th order correlation-immune functions), where m is very small compared to n - namely $m = o(\sqrt{n})$ - was given by O. Denisov in [131]. This formula was not correct for $m \geq 2$ and a correction was given by the same author in [132] (as well as a simpler proof): the number of m-resilient functions is equivalent to

$$\exp_2\left(2^n - \frac{n-m}{2}\binom{n}{m} - \sum_{i=0}^{m}\binom{n}{i}\log_2\sqrt{\pi/2}\right).$$

For large resiliency orders, Y. Tarannikov and D. Kirienko showed in [347] that, for every positive integer m, there exists a number $p(m)$ such that for $n > p(m)$, any $(n-m)$-resilient function $f(x_1, \cdots, x_n)$ is equivalent, up to permutation of its input coordinates, to a function of the form $g(x_1, \cdots, x_{p(m)}) \oplus x_{p(m)+1} \oplus \cdots \oplus x_n$. It is then a simple matter to deduce that the number of $(n-m)$-resilient functions equals $\sum_{i=0}^{p(m)} A(m,i)\binom{n}{i}$, where $A(m,i)$ is the number of i-variable $(i-m)$-resilient functions that depend on all inputs x_1, x_2, ..., x_i nonlinearly. Hence, it is equivalent to $\frac{A(m,p(m))}{p(m)!}n^{p(m)}$ for m constant when n tends to infinity, and it is at most $A_m\ n^{p(m)}$, where A_m depends on m only. It is proved in [348] that $3 \cdot 2^{m-2} \leq p(m) \leq (m-1)2^{m-2}$ and in [347] that $p(4) = 10$; hence the number of $(n-4)$-resilient functions equals $(1/2)n^{10} + O(n^9)$.

8 Functions satisfying the strict avalanche and propagation criteria

In this section, we are interested in the functions (and more particularly, in the balanced functions) which achieve $PC(l)$ for some $l < n$ (the functions achieving $PC(n)$ are the bent functions and they cannot be balanced).

8.1 $PC(l)$ criterion

It is shown in [180, 60, 61] that, if n is even, then $PC(n-2)$ implies $PC(n)$; so we can find balanced n-variable $PC(l)$ functions for n even only if $l \leq n-3$. For odd $n \geq 3$, it is also known that the functions which satisfy $PC(n-1)$ are those functions of the form $g(x_1 \oplus x_n, \cdots, x_{n-1} \oplus x_n) \oplus \ell(x)$, where g is bent and ℓ is affine, and that the $PC(n-2)$ functions are those functions of a similar form, but where, for at most one index i, the term $x_i \oplus x_n$ may be replaced by x_i or by x_n (other equivalent characterizations exist [61]).
The only known upper bound on the algebraic degrees of $PC(l)$ functions is $n-1$. A lower bound on the nonlinearity of functions satisfying the propagation criterion exists [360] and can be very easily proved: if there exists an l-dimensional subspace F such that, for every nonzero $a \in F$, the derivative $D_a f$ is balanced, then $nl(f) \geq 2^{n-1} - 2^{n-\frac{1}{2}l-1}$; Relation (27), relating the values of the Walsh transform of a function on a flat $a+E$ to the autocorrelation coefficients of the function on a flat $b+E^{\perp}$, applied to any $a \in \mathbb{F}_2^n$, with $b=0$ and $E=F^{\perp}$, shows indeed that every value $\widehat{f_\chi}^2(u)$ is bounded above by 2^{2n-l}; it implies that $PC(l)$ functions have nonlinearities bounded below by $2^{n-1} - 2^{n-\frac{1}{2}l-1}$. Equality can occur only if $l = n-1$ (n odd) and $l = n$ (n even).
The maximum correlation of Boolean functions satisfying $PC(l)$ (and in particular, of bent functions) can be directly deduced from Relations (40) and (27), see [38].

8.1.1 Characterizations

There exist characterizations of the propagation criterion. A first obvious one is that, according to Relation (24), *i.e.* to the Wiener-Khintchine Theorem, f satisfies $PC(l)$ if and only if $\sum_{u \in \mathbb{F}_2^n} (-1)^{a \cdot u} \widehat{f_\chi}^2(u) = 0$ for every nonzero vector a of weight at most l. A second one is:

Proposition 35 *[61] Any n-variable Boolean function f satisfies $PC(l)$ if*

and only if, for every vector u of weight at least $n-l$, and every vector v:

$$\sum_{w \preceq u} \widehat{\mathrm{f}_\chi}^2(w+v) = 2^{n+w_H(u)}.$$

This is a direct consequence of Relation (27). A third characterization is given in Subsection 8.2 below (apply it to $k=0$).

8.1.2 Constructions

Maiorana-McFarland's construction can be used to produce functions satisfying the propagation criterion: the derivative $D_{(a,b)}(x,y)$ of a function of the form (59) being equal to $x \cdot D_b\,\phi(y) \oplus a \cdot \phi(y+b) \oplus D_b g(y)$, the function satisfies $PC(l)$ under the sufficient condition that:
1. for every nonzero $b \in \mathbb{F}_2^s$ of weight smaller than or equal to l, and every vector $y \in \mathbb{F}_2^s$, the vector $D_b\,\phi(y)$ is nonzero (or equivalently every set $\phi^{-1}(u)$, $u \in \mathbb{F}_2^r$, either is empty or is a singleton or has minimum distance strictly greater than l);
2. every linear combination of at least one and at most l coordinate functions of ϕ is balanced (this condition corresponds to the case $b=0$).
Constructions of such functions have been given in [60, 61, 223].

According to Proposition 35, Dobbertin's construction cannot produce functions satisfying $PC(l)$ with $l \geq n/2$. Indeed, if u is for instance the vector with $n/2$ first coordinates equal to 0, and with $n/2$ last coordinates equal to 1, we have, according to Relation (64): $\widehat{\mathrm{h}_\chi}^2(w) = 0$ for every $w \preceq u$.

8.2 $PC(l)$ of order k and $EPC(l)$ of order k criteria

According to the characterization of resilient functions and to the definitions of PC and EPC criteria, we have:

Proposition 36 *[302] A function f satisfies $EPC(l)$ (resp. $PC(l)$) of order k if and only if, for any vector a of Hamming weight smaller than or equal to l and any vector b of Hamming weight smaller than or equal to k, if $(a,b) \neq (0,0)$ (resp. if $(a,b) \neq (0,0)$ and if a and b have disjoint supports) then:*

$$\sum_{x \in \mathbb{F}_2^n} (-1)^{f(x) \oplus f(x+a) \oplus b \cdot x} = 0.$$

A recent paper [308] gives the following characterization:

Proposition 37 *Any n-variable Boolean function f satisfies $EPC(l)$ (resp. $PC(l)$) of order k if and only if, for every vector u of weight at least $n-l$, and every vector v of weight at least $n-k$ (resp. of weight at least $n-k$ and such that $\overline{v}$ and $\overline{u}$ have disjoint supports):*

$$\sum_{w \preceq u} \widehat{\mathrm{f}_\chi}(w)\widehat{\mathrm{g}_\chi}(w) = 2^{w_H(u)+w_H(v)},$$

where g is the restriction of f to the vectorspace $\{x \in \mathbb{F}_2^n / \, x \preceq v\}$.

This can be proved by applying Poisson summation formula (17) to the function $(a,b) \mapsto \widehat{D_a \mathrm{f}_\chi}(b)$.
Preneel showed in [300] that $SAC(k)$ functions have algebraic degrees at most $n-k-1$ (indeed, all of their restrictions obtained by fixing k input coordinates have algebraic degrees at most $n-k-1$). In [253], the criterion $SAC(n-3)$ was characterized through the ANF of the function, and its properties were further studied. A construction of $PC(l)$ of order k functions based on Maiorana-McFarland's method is given in [223] (the mapping ϕ being linear and constructed from linear codes) and generalized in [60, 61] (the mapping ϕ being not linear and constructed from nonlinear codes). A construction of n-variable balanced functions satisfying $SAC(k)$ and having algebraic degree $n-k-1$ is given, for $n-k-1$ odd, in [223] and, for $n-k-1$ even, in [320] (where balancedness and nonlinearity are also considered).
It is shown in [61] that, for every positive even $l \leq n-4$ (with $n \geq 6$) and every odd l such that $5 \leq l \leq n-5$ (with $n \geq 10$), the functions which satisfy $PC(l)$ of order $n-l-2$ are the functions of the form:

$$\bigoplus_{1 \leq i < j \leq n} x_i \, x_j \oplus h(x_1, \cdots, x_n)$$

where h is affine.

9 Algebraic immune functions

We have recalled in Section 4.1 the different algebraic attacks on stream ciphers and the related criteria of resistance for the Boolean functions used in their pseudo-random generators. We study now these criteria more in details and we describe the known functions satisfying them.

9.1 General properties of the algebraic immunity and its relationship with some other criteria

We have seen that the algebraic immunity of any n-variable Boolean function is bounded above by $\lceil n/2 \rceil$ and that the functions used in stream ciphers must have an algebraic immunity close to this maximum. Note also that for any functions f and g depending on distinct variables, we have $AI(f \oplus g) \leq AI(f) + AI(g)$. Indeed, for some $\epsilon, \eta \in \mathbb{F}_2$, let h be an annihilator of degree $AI(f)$ of $f \oplus \epsilon$ and k an annihilator of degree $AI(g)$ of $g \oplus \eta$, then the product of h and k is a nonzero annihilator of degree at most $AI(f) + AI(g)$ of $f \oplus g \oplus \epsilon \oplus \eta$.

9.1.1 Algebraic immunity of random functions

Random functions behave well with respect to algebraic immunity[52]: it has been proved in [133] that, for all $a < 1$, when n tends to infinity, $AI(f)$ is almost surely greater than $\frac{n}{2} - \sqrt{\frac{n}{2} \ln\left(\frac{n}{2a \ln 2}\right)}$.

9.1.2 Algebraic immunity of monomial functions

It has been shown in [285] that, if the number of runs $r(d)$ of 1's in the binary expansion of the exponent d of a power function $tr_n(ax^d)$ (that is, the number of full subsequences of consecutive 1's) is smaller than $\sqrt{n}/2$, then the algebraic immunity is bounded above by $r(d)\lfloor\sqrt{n}\rfloor + \left\lceil \frac{n}{\lfloor\sqrt{n}\rfloor} \right\rceil - 1$. Note that this bound is better than the general bound $\lceil n/2 \rceil$ for only a negligible part of power mappings, but it concerns however all of those whose exponents have a constant 2-weight or a constant number of runs - the power functions studied as potential S-boxes in block ciphers enter in this framework (see the chapter "Vectorial Boolean Functions for Cryptography"). Moreover, the bound is further improved when n is odd and the function is almost bent (see this same chapter for a definition): the algebraic immunity of such functions is bounded above by $2\lfloor\sqrt{n}\rfloor$.

9.1.3 Functions in odd numbers of variables with optimal algebraic immunity

In [39], A. Canteaut has observed the following property:

[52] No result is known on the behavior of random functions against fast algebraic attacks.

Proposition 38 *If an n-variable balanced function f, with n odd, admits no non-zero annihilator of algebraic degree at most $\frac{n-1}{2}$, then it has optimum algebraic immunity $\frac{n+1}{2}$.*

This means that we do not need to check also that $f \oplus 1$ has no non-zero annihilator of algebraic degree at most $\frac{n-1}{2}$ for showing that f has optimum algebraic immunity. Indeed, consider the Reed-Muller code of length 2^n and of order $\frac{n-1}{2}$. This code is self-dual (i.e. is its own dual), according to Theorem 2. Let G be a generator matrix of this code. Each column of G is labeled by the vector of $\mathbb{F}_2^n$ obtained by keeping its coordinates of indices $2, \cdots, n+1$. Saying that f has no non-zero annihilator of algebraic degree at most $\frac{n-1}{2}$ is equivalent to saying that the matrix obtained by selecting those columns of G corresponding to the elements of the support of f has full rank $\sum_{i=0}^{\frac{n-1}{2}} \binom{n}{i} = 2^{n-1}$. By hypothesis, f has weight 2^{n-1}. Since the order of the columns in G can be freely chosen, we shall assume for simplicity that the columns corresponding to the support of f are the 2^{n-1} first ones. Then we have $G = (A \,|\, B)$ where A is an invertible $2^{n-1} \times 2^{n-1}$ matrix (and the matrix $G' = A^{-1} \times G = (Id \,|\, A^{-1} \times B)$ is also a generator matrix). In terms of coding theory, the support of the function is an *information set*. Then the complement of the support of f is also an information set (*i.e.* B is also invertible): otherwise, there would exist a vector $(z \,|\, 0)$, $z \neq 0$, in the code and this is clearly impossible since G and G' are also parity-check matrices of the code.

9.1.4 Relationship between normality and algebraic immunity

If an n-variable function f is k-normal then its algebraic immunity is at most $n - k$, since the fact that $f(x) = \epsilon \in \mathbb{F}_2$ for every $x \in A$, where A is a k-dimensional flat, implies that the indicator of A is an annihilator of $f + \epsilon$. This bound is tight since, being symmetric the majority function is $\lfloor n/2 \rfloor$-normal for every n (see below) and has algebraic immunity $\lceil n/2 \rceil$. Obviously, $AI(f) \leq \ell$ does not imply conversely that f is $(n - \ell)$-normal, since when n tends to infinity, for every $a > 1$, n-variable Boolean functions are almost surely non-$(a\, log_2\, n)$-normal [66] (note that $k \sim a\, log_2\, n$ implies that $n - k \sim n$) and the algebraic immunity is always bounded above by $n/2$.

9.1.5 Relationship between algebraic immunity, weight and nonlinearity

It can be easily shown that $\sum_{i=0}^{AI(f)-1}\binom{n}{i} \leq w_H(f) \leq \sum_{i=0}^{n-AI(f)}\binom{n}{i}$: the left-hand side inequality must be satisfied since, otherwise, the number $w_H(f)$ of equations in the linear system expressing that a function of algebraic degree at most $AI(f)-1$ is an annihilator of f would have a number of equations smaller than its number of unknowns (*i.e.* the number of coefficients in its algebraic normal form) and it would therefore have non-trivial solutions, a contradiction. The right-hand side inequality is obtained from the other one by replacing f by $f \oplus 1$. This implies that a function f such that $AI(f) = \frac{n+1}{2}$ (n odd) must be balanced.
It has been shown in [121] and [76] that low nonlinearity implies low algebraic immunity (but high algebraic immunity does not imply high nonlinearity): it can be easily proved that, for every function h of algebraic degree r, we have $AI(f) - r \leq AI(f \oplus h) \leq AI(f) + r$, and this implies:

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

and more generally:

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

These bounds have been improved in all cases for the first order nonlinearity into

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

by Lobanov [256] and in most cases for the r-th order nonlinearity into

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

(in fact, the improvement was slightly stronger than this, but more complex) in [71]. Another improvement:

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

(which always improves upon the bound of [76] and improves upon the bound of [71] for low values of r) has been subsequently obtained by Mesnager in [280].

9.2 The problem of finding functions achieving high algebraic immunity and high nonlinearity

We know that functions achieving optimal or suboptimal algebraic immunity and in the same time high algebraic degree and high nonlinearity must exist thanks to the results of [133, 311]. But knowing that almost all functions have high algebraic immunity does not mean that constructing such functions is easy.

The bounds of [71] and [280] seen above are weak[53] and Lobanov's bound, which is tight, does not assure that the nonlinearity is high enough:

• For n even and $AI(f) = \frac{n}{2}$, it gives $nl(f) \geq 2^{n-1} - 2\binom{n-1}{n/2-1} = 2^{n-1} - \binom{n}{n/2}$ which is much smaller than the best possible nonlinearity $2^{n-1} - 2^{n/2-1}$ and, more problematically, much smaller than the asymptotic almost sure nonlinearity of Boolean functions, which is, when n tends to ∞, located in the neighbourhood of $2^{n-1} - 2^{n/2-1}\sqrt{2n \ln 2}$ as we saw. Until recently, the best nonlinearity reached by the known functions with optimum AI was that of the majority function and of the iterative construction (see more details below on these functions): $2^{n-1} - \binom{n-1}{n/2} = 2^{n-1} - \frac{1}{2}\binom{n}{n/2}$ [124]. This was a little better than what gives Lobanov's bound but insufficient.

• For n odd and $AI(f) = \frac{n+1}{2}$, Lobanov's bound gives $nl(f) \geq 2^{n-1} - \binom{n-1}{(n-1)/2} \simeq 2^{n-1} - \frac{1}{2}\binom{n}{(n-1)/2}$ which is a little better than in the n even case, but still far from the average nonlinearity of Boolean functions. Until recently, the best known nonlinearity was that of the majority function and matched this bound.

Efficient algorithms have been given in [5, 134] for computing the algebraic immunity and tables are given in [5].

9.3 The functions with high algebraic immunity found so far and their parameters

Sporadic functions Balanced highly nonlinear functions in up to 20 variables (derived from power functions) with high algebraic immunities have been exhibited in [83] and [5].

[53] Their interest is to be valid for every function with given algebraic immunity.

Infinite classes of functions The majority function (first proposed by J.D. Key, T.P. McDonough and V.C. Mavron in the context of the erasure channel [213] - rediscovered by Dalai et al. in the context of algebraic immunity [124]), defined as $f(x) = 1$ if $w_H(x) \geq n/2$ and $f(x) = 0$ otherwise, has optimum algebraic immunity[54]. It is a symmetric function (which can represent a weakness) and its nonlinearity is insufficient. Some variants have also optimum algebraic immunity.
A nice iterative construction of an infinite class of functions with optimum algebraic immunity has been given in [122] and further studied in [76]; however, the functions it produces are neither balanced nor highly nonlinear.
All of these functions are weak against fast algebraic attacks, as shown in [5].
More numerous functions with optimum algebraic immunity were given in [72]. Among them are functions with better nonlinearities. However, the method of [72] did not allow to reach high nonlinearities (see [96]) and some functions constructed in [246, 247] seem still worse from this viewpoint. Hence, the question of designing infinite classes of functions achieving all the necessary criteria remained open after these papers.
A function with optimum algebraic immunity, apparently (according to computer investigations) good immunity to fast algebraic attacks, provably much better nonlinearity than the functions mentioned above and in fact, according to computer investigations, quite sufficient nonlinearity has been exhibited very recently in [151, 81]:

Theorem 15 *Let n be any positive integer and α a primitive element of the field $\mathbb{F}_{2^n}$. Let f be the balanced Boolean function on $\mathbb{F}_{2^n}$ whose support equals $\{0, 1, \alpha, \cdots, \alpha^{2^{n-1}-2}\}$. Then f has optimum algebraic immunity $\lceil n/2 \rceil$.*

Proof.
Let g be any Boolean function of algebraic degree at most $\lceil n/2 \rceil - 1$. Let $g(x) = \sum_{i=0}^{2^n-1} g_i x^i$ be its univariate representation in the field $\mathbb{F}_{2^n}$, where $g_i \in \mathbb{F}_{2^n}$ is null if the 2-weight $w_2(i)$ of i is at least $\lceil n/2 \rceil$ (which implies in particular that $g_{2^n-1} = 0$).
If g is an annihilator of f, then we have $g(\alpha^i) = 0$ for every $i = 0, \cdots, 2^{n-1} - 2$, that is, the vector $(g_0, \cdots, g_{2^n-2})$ belongs to the Reed-Solomon code over $\mathbb{F}_{2^n}$ of zeroes $1, \alpha, \cdots, \alpha^{2^{n-1}-2}$ (see [258]). According to the BCH bound, if g is non-zero, then this vector has Hamming weight at least 2^{n-1}. We

[54] Changing $w_H(x) \geq n/2$ into $w_H(x) > n/2$ or $w_H(x) \leq n/2$ or $w_H(x) < n/2$ changes the function into an affinely equivalent one, up to addition of the constant 1, and therefore does not change the AI.

briefly recall how this lower bound can be simply proved in our framework. By definition, we have:

$$\begin{pmatrix} g(1) \\ g(\alpha) \\ g(\alpha^2) \\ \vdots \\ g(\alpha^{2^n-2}) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \alpha & \alpha^2 & \cdots & \alpha^{2^n-2} \\ 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{2(2^n-2)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \alpha^{2^n-2} & \alpha^{2(2^n-2)} & \cdots & \alpha^{(2^n-2)(2^n-2)} \end{pmatrix} \times \begin{pmatrix} g_0 \\ g_1 \\ g_2 \\ \vdots \\ g_{2^n-2} \end{pmatrix}$$

which implies (since $\sum_{k=0}^{2^n-2} \alpha^{(i-j)k}$ equals 1 if $i = j$ and 0 otherwise):

$$\begin{pmatrix} g_0 \\ g_1 \\ g_2 \\ \vdots \\ g_{2^n-2} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \alpha^{-1} & \alpha^{-2} & \cdots & \alpha^{-(2^n-2)} \\ 1 & \alpha^{-2} & \alpha^{-4} & \cdots & \alpha^{-2(2^n-2)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \alpha^{-(2^n-2)} & \alpha^{-2(2^n-2)} & \cdots & \alpha^{-(2^n-2)(2^n-2)} \end{pmatrix} \times \begin{pmatrix} g(1) \\ g(\alpha) \\ g(\alpha^2) \\ \vdots \\ g(\alpha^{2^n-2}) \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha^{-(2^{n-1}-1)} & \alpha^{-2^{n-1}} & \cdots & \alpha^{-(2^n-2)} \\ \vdots & \vdots & \cdots & \vdots \\ \alpha^{-(2^{n-1}-1)(2^n-2)} & \alpha^{-2^{n-1}(2^n-2)} & \cdots & \alpha^{-(2^n-2)(2^n-2)} \end{pmatrix} \times \begin{pmatrix} g(\alpha^{2^{n-1}-1}) \\ g(\alpha^{2^{n-1}}) \\ \vdots \\ g(\alpha^{2^n-2}) \end{pmatrix}$$

Suppose that at least 2^{n-1} of the g_i's are null. Then, $g(\alpha^{2^{n-1}-1}), \cdots, g(\alpha^{2^n-2})$ satisfy a homogeneous system whose matrix is obtained from the latter matrix above by erasing $2^{n-1} - 1$ rows. This is a $2^{n-1} \times 2^{n-1}$ Vandermonde matrix and its determinant is therefore non-null. This implies that $g(\alpha^{2^{n-1}-1}), \cdots, g(\alpha^{2^n-2})$ and therefore g must then be null. Hence the vector $(g_0, \cdots, g_{2^n-2})$ has weight at least 2^{n-1}.
Moreover, suppose that the vector $(g_0, \cdots, g_{2^n-2})$ has Hamming weight 2^{n-1} exactly. Then n is odd and $g(x) = \sum_{\substack{0 \le i \le 2^n-2 \\ w_2(i) \le (n-1)/2}} x^i$; but this contradicts the fact that $g(0) = 0$. We deduce that the vector $(g_0, \cdots, g_{2^n-2})$ has Hamming weight strictly greater than 2^{n-1}, leading to a contradiction with the fact that g has algebraic degree at most $\lceil n/2 \rceil - 1$, since the number of integers of 2-weight at most $\lceil n/2 \rceil - 1$ is not strictly greater than 2^{n-1}.
Let g be now a non-zero annihilator of $f \oplus 1$. The vector $(g_0, \cdots, g_{2^n-2})$ belongs then to the Reed-Solomon code over $\mathbb{F}_{2^n}$ of zeroes $\alpha^{2^{n-1}-1}, \cdots, \alpha^{2^n-2}$. According to the BCH bound (which can be proven similarly as above), this vector has then Hamming weight strictly greater than 2^{n-1}. We arrive to

the same contradiction. Hence, there does not exist a non-zero annihilator of f or $f \oplus 1$ of algebraic degree at most $\lceil n/2 \rceil - 1$ and f has then (optimum) algebraic immunity $\lceil n/2 \rceil$. □

It is shown in [81] that the univariate representation of f equals

$$1 + \sum_{i=1}^{2^n-2} \frac{\alpha^i}{(1+\alpha^i)^{1/2}} \, x^i \tag{70}$$

where $u^{1/2} = u^{2^{n-1}}$, which shows that f has algebraic degree $n-1$ (which is optimum for a balanced function), and that we have:

$$nl(f) \geq 2^{n-1} - n \cdot \ln 2 \cdot 2^{\frac{n}{2}} - 1.$$

It could be checked, for small values of n, that the exact value of $nl(f)$ is much better than what gives this lower bound and seems quite sufficient for resisting fast correlation attacks (for these small values of n, it behaves as $2^{n-1} - 2^{n/2}$). Finally, the function seems to show good immunity against fast algebraic attacks: the computer investigations made using Algorithm 2 of [5] suggest the following properties:

- No nonzero function g of algebraic degree at most e and no function h of algebraic degree at most d exist such that $fg = h$, when $(e, d) = (1, n-2)$ for n odd and $(e, d) = (1, n-3)$ for n even. This has been checked for $n \leq 12$ and we conjecture it for every n.

- For $e > 1$, pairs (g, h) of algebraic degrees (e, d) such that $e + d < n - 1$ were never observed. Precisely, the non-existence of such pairs could be checked exhaustively for $n \leq 9$ and $e < n/2$, for $n = 10$ and $e \leq 3$ and for $n = 11$ and $e \leq 2$. This suggests that this class of functions, even if not always optimal against fast algebraic attacks, has a very good behavior.

Hence, the functions of this class gather all the properties needed for allowing the stream ciphers using them as filtering functions to resist all the main attacks (the Berlekamp-Massey and Rønjom-Helleseth attacks, fast correlation attacks, standard and fast algebraic attacks). They are the only functions found so far for which such properties could be shown. There remains at least one attack against which the resistance of the functions should be evaluated: the algebraic attack on the augmented function (this obliges to consider all possible update functions of the linear part of the pseudo-random generator).

The construction of Proposition 22 allows increasing the complexity of Boolean functions while keeping their high nonlinearities and may allow increasing their algebraic immunity as well.

10 Symmetric functions

A Boolean function is called a *symmetric function* if it is invariant under the action of the symmetric group (*i*.e. if its output is invariant under permutation of its input bits). Its output depends then only on the Hamming weight of the input. So, in other words, f is symmetric if and only if there exists a function $f^{\#}$ from $\{0, 1, \cdots, n\}$ to $\mathbb{F}_2$ such that $f(x) = f^{\#}(w_H(x))$. Such functions are of some interest to cryptography, as they allow to implement in an efficient way nonlinear functions on large numbers of variables. Let us consider for example an LFSR filtered by a 63 variable symmetric function f, whose input is the content of an interval of 63 consecutive flip-flops of the LFSR. This device may be implemented with a cost similar to that of a 6 variable Boolean function, thanks to a 6 bit counter calculating the weight of the input to f (this counter is incremented if a 1 is shifted in the interval and decremented if a 1 is shifted out). However, the pseudo-random sequence obtained this way has correlation with transitions (sums of consecutive bits), and a symmetric function should not take all its inputs in a full interval. In fact, it is not yet completely clarified whether the advantage of allowing much more variables and the cryptographic weaknesses these symmetric functions may introduce result in an advantage for the designer or for the attacker, in more sophisticated devices.

10.1 Representation

Let $r = 0, \cdots, n$ and let φ_r be the Boolean function whose support is the set of all vectors of weight r in $\mathbb{F}_2^n$. Then, according to Relation (8) relating the values of the coefficients of the NNF to the values of the function, the coefficient of x^I, $I \in \mathcal{P}(N)$, in the NNF of φ_r is: $\lambda_I = (-1)^{|I|-r} \binom{|I|}{r}$.

Any symmetric function f being equal to $\bigoplus_{r=0}^{n} f^{\#}(r)\, \varphi_r$, it is therefore equal to $\sum_{r=0}^{n} f^{\#}(r)\, \varphi_r$, since the functions φ_r have disjoint supports. The coefficient

of x^I in its NNF equals then $\sum_{r=0}^{n} f^{\#}(r)(-1)^{|I|-r} \left(\begin{array}{c} |I| \\ r \end{array} \right)$ and depends only on the size of I. The NNF of f is then

$$f(x) = \sum_{i=0}^{n} c_i \, S_i(x), \text{ where } c_i = \sum_{r=0}^{n} f^{\#}(r)(-1)^{i-r} \left(\begin{array}{c} i \\ r \end{array} \right) \tag{71}$$

and where $S_i(x)$ is the i-th elementary symmetric pseudo-Boolean function whose NNF is $\sum_{I \in \mathcal{P}(N)/\ |I|=i} x^I$. The degree of the NNF of f equals $\max\{i/\ c_i \neq 0\}$.

We have clearly $S_i(x) = \binom{w_H(x)}{i} = \frac{w_H(x)\,(w_H(x)-1)\cdots(w_H(x)-i+1)}{i!}$. According to Relation (71), we see that the univariate function $f^{\#}(z)$ admits the polynomial representation $\sum_{i=0}^{n} c_i \binom{z}{i} = \sum_{i=0}^{n} c_i \frac{z\,(z-1)\cdots(z-i+1)}{i!}$ in one variable z, whose degree equals the degree of the NNF of f. Since this degree is at most n, and the values taken by this polynomial at $n+1$ points are set, this polynomial representation is unique.

Denoting by $\sigma_i(x)$ the reduction of $S_i(x)$ modulo 2, $\sigma_i(x)$ equals 1 if and only if $\left(\begin{array}{c} w_H(x) \\ i \end{array} \right)$ is odd, that is, according to Lucas' theorem [258], if and only if the binary expansion of i is covered by that of $w_H(x)$. Reducing Relation (71) modulo 2 and writing that $j \preceq i$ when the binary expansion of i covers that of j (*i.e.* $j = \sum_{l=1}^{\log_2 n} j_l \, 2^{l-1}$, $i = \sum_{l=1}^{\log_2 n} i_l \, 2^{l-1}$, $j_l \leq i_l$, $\forall l = 1, \cdots, \log_2 n$), we deduce from Luca's theorem again that the ANF of f is

$$f(x) = \bigoplus_{i=0}^{n} \lambda_i \, \sigma_i(x), \text{ where } \lambda_i = \bigoplus_{j \preceq i} f^{\#}(j). \tag{72}$$

Conversely (since the Möbius transform is involutive as we saw) $f^{\#}(i) = \bigoplus_{j \preceq i} \lambda_j$.

Note that a symmetric Boolean function f has algebraic degree 1 if and only if it equals $\bigoplus_{i=1}^{n} x_i$ or $\bigoplus_{i=1}^{n} x_i \oplus 1$, that is, if the binary function $f^{\#}(r)$ equals r [mod 2] or $r+1$ [mod 2], and that it is quadratic if and only if it equals $\bigoplus_{1 \leq i < j \leq n} x_i x_j$ (introduced to generate the Kerdock code) plus a symmetric function of algebraic degree at most 1, that is, if the function $f^{\#}(r)$ equals $\binom{r}{2}$ [mod 2] or $\binom{r}{2} + r$ [mod 2] or $\binom{r}{2} + 1$ [mod 2] or $\binom{r}{2} + r + 1$ [mod 2]. Hence, f has algebraic degree 1 if and only if $f^{\#}$ satisfies $f^{\#}(r+1) = f^{\#}(r) \oplus 1$ and it has degree 2 if and only if $f^{\#}$ satisfies $f^{\#}(r+2) = f^{\#}(r) \oplus 1$.

As observed in [49], the algebraic degree of a symmetric function f is at most $2^t - 1$, for some positive integer t, if and only if the sequence $(f^{\#}(r))_{r \geq 0}$ is

periodic with period 2^t. This is a direct consequence of (72). Here again, it is not clear whether this is a greater advantage for the designer of a cryptosystem using such symmetric function f (since, to compute the image of a vector x by f, it is enough to compute the number of nonzero coordinates $x_1, \cdots, x_t$ only) or for the attacker.

10.2 Fourier and Walsh transforms

By linearity, the Fourier transform of any symmetric function $\sum_{r=0}^{n} f^{\#}(r)\,\varphi_r$ equals $\sum_{r=0}^{n} f^{\#}(r)\,\widehat{\varphi_r}$.

For every vector $a \in \mathbb{F}_2^n$, denoting by ℓ the Hamming weight of a, we have $\widehat{\varphi_r}(a) = \sum_{x \in \mathbb{F}_2^n \,|\, w_H(x)=r} (-1)^{a\cdot x} = \sum_{j=0}^{n} (-1)^j \binom{\ell}{j}\binom{n-\ell}{r-j}$, denoting by j the size of $supp(a) \cap supp(x)$. The polynomials $K_{n,r}(X) = \sum_{j=0}^{n} (-1)^j \binom{X}{j}\binom{n-X}{r-j}$ are called *Krawtchouk polynomials*. They are caracterized by their generating series:

$$\sum_{r=0}^{n} K_{n,r}(\ell) z^r = (1-z)^\ell (1+z)^{n-\ell}$$

and have nice resulting properties (see *e.g.* [258, 96]).

From the Fourier transform, we can deduce the Walsh transform thanks to Relation (12).

10.3 Nonlinearity

If n is even, then the restriction of every symmetric function f on $\mathbb{F}_2^n$ to the $n/2$-dimensional flat:

$$A = \{(x_1, \ldots, x_n) \in \mathbb{F}_2^n \ ;\ x_{i+n/2} = x_i \oplus 1, \forall i \leq n/2\}$$

is constant, since all the elements of A have the same weight $n/2$. Thus, f is $n/2$-normal[55] (see Definition 4). But Relation (42) gives nothing more than the covering radius bound (36). The symmetric functions which achieve this bound, *i.e.* which are bent, have been first characterized by P. Savicky in [324]: the bent symmetric functions are the four symmetric functions of algebraic degree 2 already described above: $f_1(x) = \bigoplus_{1 \leq i < j \leq n} x_i x_j$, $f_2(x) =$

[55] Obviously, this is more generally valid for every function which is constant on the set $\{x \in \mathbb{F}_2^n;\ w_H(x) = n/2\}$.

$f_1(x) \oplus 1$, $f_3(x) = f_1(x) \oplus x_1 \oplus \cdots \oplus x_n$ and $f_4(x) = f_3(x) \oplus 1$. A stronger result can be proved in a very simple way [169]:

Theorem 16 *For every positive even n, the $PC(2)$ n-variable symmetric functions are the functions f_1, f_2, f_3 and f_4 above.*

Proof. Let f be any $PC(2)$ n-variable symmetric function and let $i < j$ be two indices in the range $[1; n]$. Let us denote by x' the following vector: $x' = (x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_{j-1}, x_{j+1}, \cdots, x_n)$. Since $f(x)$ is symmetric, it has the form $x_i\, x_j g(x') \oplus (x_i \oplus x_j)\, h(x') \oplus k(x')$. Let us denote by $e_{i,j}$ the vector of weight 2 whose nonzero coordinates stand at positions i and j. The derivative $D_{e_{i,j}} f$ of f with respect to $e_{i,j}$ equals $(x_i \oplus x_j \oplus 1)g(x')$. Since this derivative is balanced, by hypothesis, then g must be equal to the constant function 1 (indeed if $g(x') = 1$ then $(x_i \oplus x_j \oplus 1)g(x')$ equals 1 for half of the inputs and otherwise it equals 1 for none). Hence, the degree-2-part of the ANF of f equals $\bigoplus_{1 \leq i < j \leq n} x_i x_j$. □

Some more results on the propagation criterion for symmetric functions can be found in [49].

If n is odd, then the restriction of any symmetric function f to the $\frac{n+1}{2}$-dimensional flat

$$A = \{(x_1, \ldots, x_n) \in \mathbb{F}_2^n \ ;\ x_{i+\frac{n-1}{2}} = x_i \oplus 1, \forall i \leq n/2\}$$

is affine, since the weight function w_H is constant on the hyperplane of A of equation $x_n = 0$ and on its complement. Thus, f is $\frac{n+1}{2}$-weakly-normal. According to Relation (42), this implies that its nonlinearity is upper bounded by $2^{n-1} - 2^{\frac{n-1}{2}}$. It also allows showing that the only symmetric functions achieving this bound are the same as the 4 functions f_1, f_2, f_3 and f_4 above, but with n odd (this has been first proved by Maitra and Sarkar [265], in a more complex way). Indeed, Relation (42) implies the following result:

Theorem 17 *[66] Let n be any positive integer and let f be any symmetric function on $\mathbb{F}_2^n$. Let l be any integer satisfying $0 < l \leq n/2$. Denote by h_l the symmetric Boolean function on $n-2l$ variables defined by $h_l(y_1, \cdots, y_{n-2l}) = f(x_1, \cdots, x_l, x_1 \oplus 1, \cdots, x_l \oplus 1, y_1, \cdots, y_{n-2l})$, where the values of $x_1, \cdots, x_l$ are arbitrary (equivalently, h_l can be defined by $h_l^{\#}(r) = f^{\#}(r + l)$, for every $0 \leq r \leq n - 2l$). Then $nl(f) \leq 2^{n-1} - 2^{n-l-1} + 2^l nl(h_l)$.*

Proof: Let $A = \{(x_1, \ldots, x_n) \in \mathbb{F}_2^n \mid x_{i+l} = x_i \oplus 1, \forall i \leq l\}$. For every element x of A, we have $f(x) = h_l(x_{2l+1}, \cdots, x_n)$. Let us consider the restriction g

of f to A as a Boolean function on $\mathbb{F}_2^{n-l}$, say $g(x_1, \cdots, x_l, x_{2l+1}, \cdots, x_n)$. Then, since $g(x_1, \cdots, x_l, x_{2l+1}, \cdots, x_n) = h_l(x_{2l+1}, \cdots, x_n)$, g has nonlinearity $2^l\, nl(h_l)$. According to Relation (42) applied with $h_a = g$, we have $nl(f) \leq 2^{n-1} - 2^{n-l-1} + 2^l nl(h_l)$. □

Then, the characterizations recalled above of the symmetric functions achieving best possible nonlinearity can be straightforwardly deduced. Moreover:
- if, for some integer l such that $0 \leq l < \left\lfloor \frac{n-1}{2} \right\rfloor$, the nonlinearity of an n-variable symmetric function f is strictly greater than $2^{n-1} - 2^{n-l-1} + 2^l \left(2^{n-2l-1} - 2^{\left\lfloor \frac{n-2l-1}{2} \right\rfloor} - 1\right) = 2^{n-1} - 2^{\left\lfloor \frac{n-1}{2} \right\rfloor} - 2^l$, then, thanks to these characterizations and to Theorem 17, the function h_l must be quadratic, and $f^\#$ satisfies $f^\#(r+2) = f^\#(r) \oplus 1$, for all $l \leq r \leq n-2-l$ (this property has been observed in [49, Theorem 6], but proved slightly differently);
- if the nonlinearity of f is strictly greater than $2^{n-1} - 2^{\left\lfloor \frac{n-1}{2} \right\rfloor} - 2^{l+1}$, then h_l either is quadratic or has odd weight, that is, either $f^\#$ satisfies $f^\#(r+2) = f^\#(r) \oplus 1$ for all $l \leq r \leq n-2-l$, or h_l has odd weight.

Further properties of the nonlinearities of symmetric functions can be found in [49, 66].

10.4 Resiliency

There exists a conjecture on symmetric Boolean functions and, equivalently, on functions defined over $\{0, 1, \cdots, n\}$ and valued in $\mathbb{F}_2$: if f is a non-constant symmetric Boolean function, then the numerical degree of f (that is, the degree of the polynomial representation in one variable of $f^\#$) is greater than or equal to $n-3$. It is a simple matter to show that this numerical degree is greater than or equal to $n/2$ (otherwise, the polynomial ${f^\#}^2 - f^\#$ would have degree at most n, and being null at $n+1$ points, it would equal the null polynomial, a contradiction with the fact that f is assumed not to be constant), but the gap between $n/2+1$ and $n-3$ is open. According to Proposition 32, the conjecture is equivalent to saying that there does not exist any symmetric 3-resilient function. And proving this conjecture is also a problem on binomial coefficients since the numerical degree of f is bounded above by d if and only if, for every k such that $d < k \leq n$:

$$\sum_{r=0}^{k} (-1)^r \binom{k}{r} f^\#(r) = 0. \tag{73}$$

Hence, the conjecture is equivalent to saying that Relation (73), with $d = n - 4$, has no binary solution $f^{\#}(0), \cdots, f^{\#}(n)$.
J. von zur Gathen and J. R. Roche [161] observed that all symmetric n-variable Boolean functions have numerical degrees greater than or equal to $n - 3$, for any $n \leq 128$ (they exhibited Boolean functions with numerical degree $n - 3$; see also [166]).
The same authors observed also that, if the number $m = n + 1$ is a prime, then all non-constant n-variable symmetric Boolean functions have numerical degree n (and therefore, considering the function $g(x) = f(x) \oplus x_1 \oplus \cdots \oplus x_n$ and applying Proposition 32, all non-affine n-variable symmetric Boolean functions are not 0-resilient, that is, are unbalanced): indeed, the binomial coefficient $\binom{n}{r}$ being congruent with $\frac{(-1)(-2)\cdots(-r)}{1\cdot 2\cdots r} = (-1)^r$, modulo m, the sum $\sum_{r=0}^{n}(-1)^r\binom{n}{r}f^{\#}(r)$ is congruent with $\sum_{r=0}^{n} f^{\#}(r)$, modulo m; and Relation (73) with $k = n$ implies then that $f^{\#}$ must be constant.
Notice that, applying Relation (73) with $k = p - 1$, where p is the largest prime less than or equal to $n + 1$, shows that the numerical degree of any symmetric non-constant Boolean function is greater than or equal to $p - 1$ (or equivalently that no symmetric non-affine Boolean function is $(n-p+1)$-resilient): otherwise, reducing (73) modulo p, we would have that the string $f^{\#}(0), \cdots, f^{\#}(k)$ is constant, and $f^{\#}$ having univariate degree less than or equal to k, the function $f^{\#}$, and thus f itself, would be constant.
More results on the balancedness and resiliency/correlation immunity of symmetric functions can be found in [21, 283, 354] and more recent ones in [49, 323].

10.5 Algebraic immunity

We have seen in Section 4.1 that, for every n-variable Boolean function f, there exist $g \neq 0$ and h of algebraic degrees at most $\lceil n/2 \rceil$ such that $f\,g = h$ (and equivalently, there exists a nonzero annihilator of degree at most $\lceil n/2 \rceil$ of f or of $f \oplus 1$). The same property can be proven when restricting ourselves to symmetric functions: the elementary symmetric functions of degrees at most $\lceil n/2 \rceil$ and their products with f give a family of $2\,(\lceil n/2 \rceil + 1) > n + 1$ symmetric functions, which must be linearly dependent since they live in a vectorspace of dimension $n + 1$. However, given an n-variable symmetric function f, there do not necessarily exist symmetric functions $g \neq 0$ and h of algebraic degrees at most $AI(f)$ such that $f\,g = h$.
We have seen that the majority function, which is symmetric, has optimum algebraic immunity. In the case n is odd, it is the only symmetric function having such property, up to the addition of a constant (see [304] which

completed a partial result of [245]). In the case n is even, other symmetric functions exist (up to the addition of a constant and to the transformation $x \to \overline{x} = (x_1 \oplus 1, \cdots, x_n \oplus 1)$) with this property; more precisions and more results on the algebraic immunity of symmetric functions can be found in [28, 252, 303, 304, 305] and the references therein.

10.6 The super-classes of rotation symmetric and Matriochka symmetric functions

A super-class of symmetric functions, called idempotent or rotation symmetric functions (see Subsection 7.5 above), has been investigated from the viewpoints of bentness and correlation immunity (see e.g. [152, 338]). Recently, it could be proved in [210], thanks to a further investigation on these functions, that the best nonlinearity of Boolean functions in odd numbers of variables is strictly greater than the quadratic bound if and only if $n > 7$. Indeed, a function of nonlinearity 241 could be found (while the quadratic bound gives 240, and the covering radius bound 244), and using direct sum with quadratic functions, it gave then 11-variable functions of nonlinearity 994 (while the quadratic bound gives 992 and the covering radius bound 1000), and 13-variable functions of nonlinearity 4036 (while the quadratic bound gives 4032 and the covering radius bound 4050). Still more recently, it was checked that 241 is the best nonlinearity of 9-variable rotation symmetric functions, but that 9-variable functions whose truth-tables (or equivalently ANFs) are invariant under cyclic shifts by 3 steps and under inversion of the order of the input bits can reach nonlinearity 242, which led to 11-variables functions of nonlinearity 996 and 13-variable functions of nonlinearity 4040. Balanced functions in 13 variables beating the quadratic bound could also be found. However, this construction gives worse nonlinearity than the Patterson-Widemann functions for 15 variables (whose nonlinearity is 16276).

In [238] is introduced the class of Matriochka symmetric functions, which are the sums of symmetric functions whose sets of variables are different and nested. Contrary to symmetric functions, they do not depend on the single weight of the input but on the sequence of the weights of the corresponding subinputs, and contrary to rotation symmetric functions, they are not invariant under cyclic shifts of the input coordinates. They can be almost as fastly computable as symmetric functions. Their cryptographic parameters will have to be further studied.

Acknowledgement

We thank Caroline Fontaine for her careful reading of a previous draft of this chapter.

References

[1] C.M. Adams and S.E. Tavares. Generating and Counting Binary Bent Sequences, *IEEE Trans. Inf. Theory*, vol 36, no. 5, pp. 1170-1173, 1990.

[2] N. Alon, O. Goldreich, J. Hastad and R. Peralta. Simple constructions of almost k-wise independent random variables. *Random Stuctures and Algorithms*, Vol 3, No 3, pp 289-304, 1992.

[3] A.S. Ambrosimov. Properties of bent functions of q-valued logic over finite fields. *Discrete Mathematics Appl.* vol 4, no. 4, pp. 341-350, 1994.

[4] F. Armknecht. Improving fast algebraic attacks. *Proceedings of Fast Software Encryption 2004, Lecture Notes in Computer Science* 3017, pp. 65-82, 2004.

[5] F. Armknecht, C. Carlet, P. Gaborit, S. Künzli, W. Meier and O. Ruatta. Efficient computation of algebraic immunity for algebraic and fast algebraic attacks. *Proceedings of EUROCRYPT 2006, Lecture Notes in Computer Science* 4004 , pp. 147-164, 2006.

[6] F. Armknecht and M. Krause. Algebraic Attacks on combiners with memory. *Proceedings of CRYPTO 2003, Lecture Notes in Computer Science* 2729, pp. 162–175, 2003.

[7] F. Armknecht and M. Krause. Constructing single- and multi-output boolean functions with maximal immunity. *Proceedings of ICALP 2006, Lecture Notes of Computer Science* 4052, pp. 180-191, 2006.

[8] F. Arnault and T. P. Berger. Design and properties of a new pseudo-random generator based on a filtered FCSR automaton. *IEEE Trans. Computers* 54 (11), pp. 1374-1383, 2005.

[9] G. Ars and J.-C. Faugère. Algebraic immunities of functions over finite fields. *Proceedings of the conference BFCA 2005*, Publications des universités de Rouen et du Havre, pp. 21-38, 2005.

[10] E.F. Assmus. On the Reed-Muller codes. *Discrete Mathematics* 106/107, pp. 25-33, 1992.

[11] E.F. Assmus and J. D. Key. *Designs and their Codes*, Cambridge Univ. Press., Cambridge, 1992.

[12] J. Ax. Zeroes of polynomials over finite fields. *American Journal on Mathematics* no. 86, pp. 255-261, 1964.

[13] T. Baignères, P. Junod and S. Vaudenay. How far can we go beyond linear cryptanalysis? *Proceedings of ASIACRYPT 2004, Lecture Notes in Computer Science* 3329, pp. 432-450, 2004.

[14] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov and A. Wigderson. Simulating Independence: New Constructions of Condensers, Ramsey Graphs, Dispersers, and Extractors. *Proceedings of the 37-th ACM STOC*, 2005. Preprint available at http://www.math.ias.edu/ boaz/Papers/BKSSW.html

[15] E. Berlekamp, *Algebraic Coding Theory,* McGraw-Hill, New York, 1968.

[16] E.R. Berlekamp and N.J.A. Sloane. Restrictions on the weight distributions of the Reed-Muller codes. *Information and Control* 14, pp. 442-446, 1969.

[17] E.R. Berlekamp and L.R. Welch. Weight distributions of the cosets of the (32,6) Reed-Muller code. *IEEE Transactions on Information Theory*, 18(1), pp. 203-207, 1972.

[18] A. Bernasconi and B. Codenotti. Spectral analysis of Boolean functions as a graph eigenvalue problem. *IEEE Transactions on computers* 48 (3), pp. 345-351, 1999.

[19] A. Bernasconi and I. Shparlinski. Circuit complexity of testing square-free numbers. *Proceedings of STACS 99, 16th Annual Symposium on Theoretical Aspects of Computer Science, Lecture Notes in Computer Science* 1563, pp. 47-56, 1999.

[20] A. Bhattacharyya, S. Kopparty, G. Shoenebeck, M. Sudan and D. Zuckerman. Optimal testing of Reed-Muller codes. Electronic Colloquium on Computational Complexity, report no. 86, 2009.

[21] J. Bierbrauer, K. Gopalakrishnan and D.R. Stinson. Bounds for resilient functions and orthogonal arrays. *Proceedings of CRYPTO'94, Lecture Notes in Computer Science* 839, pp. 247-256, 1994.

[22] Y. Borissov, N. Manev and S. Nikova. On the non-minimal codewords of weight $2d_{min}$ in the binary Reed-Muller code. *Proceedings of the Workshop on Coding and Cryptography* 2001, published by *Electronic Notes in Discrete Mathematics*, Elsevier, vo. 6, pp. 103-110, 2001. A revised version has been published in *Discrete Applied Mathematics* 128 (Special Issue "International Workshop on Coding and Cryptography (2001)"), pp. 65-74, 2003.

[23] Y. Borissov, A. Braeken, S. Nikova, B. Preneel. On the Covering Radii of Binary Reed-Muller Codes in the Set of Resilient Boolean Functions. *IEEE Transactions on Information Theory*, vol. 51, No.3, pp. 1182-1189, 2005.

[24] J. Bourgain. On the construction of affine extractors. *Geometric & Functional Analysis GAFA*, Vol. 17, No. 1, pp. 33-57, 2007.

[25] A. Braeken, Y. Borisov, S. Nikova and B. Preneel. Classification of Boolean functions of 6 variables or less with respect to cryptographic properties. *Proceedings of ICALP 2005, Lecture Notes in Computer Science* 3580, pp. 324-334, 2005.

[26] A. Braeken, Y. Borisov, S. Nikova and B. Preneel. Classification of cubic $(n-4)$-resilient boolean functions. *IEEE transactions on information theory* 52, no. 4, pp. 1670-1676, 2006.

[27] A. Braeken, V. Nikov, S. Nikova and B. Preneel. On Boolean functions with generalized cryptographic properties. *Proceedings of INDOCRYPT'2004, Lecture Notes in Computer Science* 3348, pp. 120-135, 2004.

[28] A. Braeken and B. Preneel. On the Algebraic Immunity of Symmetric Boolean Functions. *Proceedings of Indocrypt 2005, Lecture Notes in Computer Science* 3797, pp. 35–48, 2005. Some false results of this reference have been corrected in Braeken's PhD thesis entitled "Cryptographic properties of Boolean functions and S-boxes" and available at URL http://homes.esat.kuleuven.be/ abraeken/thesisAn.pdf.

[29] J. Bringer, V. Gillot and P. Langevin. Exponential sums and Boolean functions. *Proceedings of the conference BFCA 2005*, Publications des universités de Rouen et du Havre, pp. 177-185, 2005.

[30] E. Brier and P. Langevin. Classification of cubic Boolean functions of 9 variables. *Proceedings of 2003 IEEE Information Theory Workshop*, Paris, France, 2003.

[31] R. A. Brualdi, N. Cai and V. S. Pless. Orphans of the first order Reed-Muller codes. *IEEE Transactions on Information Theory* 36, pp. 399-401, 1990.

[32] P. Camion and A. Canteaut. Construction of t-resilient functions over a finite alphabet, *Proceedings of EUROCRYPT'96, Lecture Notes in Computer Sciences* 1070, pp. 283-293, 1996.

[33] P. Camion and A. Canteaut. Generalization of Siegenthaler inequality and Schnorr-Vaudenay multipermutations. *Proceedings of CRYPTO'96, Lecture Notes in Computer Science* 1109, pp. 372–386, 1996.

[34] P. Camion and A. Canteaut. Correlation-immune and resilient functions over finite alphabets and their applications in cryptography. *Designs, Codes and Cryptography* 16, 1999.

[35] P. Camion, C. Carlet, P. Charpin, N. Sendrier. On correlation-immune functions, *Proceedings of CRYPTO'91, Lecture Notes in Computer Science*, vol. 576, pp. 86-100, 1991.

[36] A. Canteaut. On the weight distributions of optimal cosets of the first-order Reed-Muller code. *IEEE Transactions on Information Theory*, 47(1), pp. 407-413, 2001.

[37] A. Canteaut. Cryptographic functions and design criteria for block ciphers. *Proceedings of INDOCRYPT 2001, Lecture Notes in Computer Science* 2247, pp. 1-16, 2001.

[38] A. Canteaut. On the correlations between a combining function and functions of fewer variables. *Proceedings of the Information Theory Workshop'02*, Bangalore, 2002.

[39] A. Canteaut. Open problems related to algebraic attacks on stream ciphers. *Proceedings of Workshop on Coding and Cryptography* WCC 2005, pp. 1-10, 2005. See also a revised version in *Lecture Notes in Computer Science* 3969, pp. 120134, 2006. Paper available on the web http://www-rocq.inria.fr/codes/Anne.Canteaut/Publications/canteaut06a.pdf

[40] A. Canteaut. Analysis and design of symmetric ciphers. Habilitation for directing Theses, University of Paris 6, 2006.

[41] A. Canteaut, C. Carlet, P. Charpin and C. Fontaine. Propagation characteristics and correlation-immunity of highly nonlinear Boolean functions. *Proceedings of EUROCRYPT'2000, Lecture Notes in Computer Science* 187, pp. 507-522 (2000)

[42] A. Canteaut, C. Carlet, P. Charpin and C. Fontaine. On cryptographic properties of the cosets of $R(1,m)$. *IEEE Transactions on Information Theory* vol. 47, no 4, pp. 1494-1513, 2001.

[43] A. Canteaut and P. Charpin. Decomposing bent functions. *IEEE Transactions on Information Theory* 49, pp. 2004-2019, 2003.

[44] A. Canteaut and P. Charpin and G. Kyureghyan. A new class of monomial bent functions. *Finite Fields and Application* 14(1), pp.221-241, 2008.

[45] A. Canteaut, M. Daum, H. Dobbertin and G. Leander. Normal and Non-Normal Bent Functions. *Proceedings of the Workshop on Coding and Cryptography 2003*, pp. 91-100, 2003.

[46] A. Canteaut and E. Filiol. Ciphertext only reconstruction of stream ciphers based on combination generators. *Proceedings of Fast Software Encryption 2000, Lecture Notes in Computer Science* 1978, pp. 165-180, 2001.

[47] A. Canteaut and M. Trabbia. Improved fast correlation attacks using parity-check equations of weight 4 and 5, *Advanced in Cryptology-EUROCRYPT 2000. Lecture notes in computer science* 1807, pp. 573-588, 2000.

[48] A. Canteaut and M. Videau. Degree of Composition of Highly Nonlinear Functions and Applications to Higher Order Differential Cryptanalysis, *Advances in Cryptology, EUROCRYPT2002, Lecture Notes in Computer Science* 2332, pp. 518-533, 2002.

[49] A. Canteaut and M. Videau. Symmetric Boolean functions. *IEEE Transactions on Information Theory* 51(8), pp. 2791-2811, 2005.

[50] Jean Robert Du Carlet. La Cryptographie, contenant une très subtile manière descrire secrètement, composée par Maistre Jean Robert Du

Carlet, 1644. A manuscript exists at the Bibliothèque Nationale (Très Grande Bibliothèque), Paris, France.

[51] C. Carlet. *Codes de Reed-Muller, codes de Kerdock et de Preparata*. PhD thesis. Publication of LITP, Institut Blaise Pascal, Université Paris 6, 90.59, 1990.

[52] C. Carlet. A transformation on Boolean functions, its consequences on some problems related to Reed-Muller codes, *Proceedings of EUROCODE'90, Lecture Notes in Computer Science* 514, pp. 42-50, 1991.

[53] C. Carlet. Partially-bent functions, *Designs Codes and Cryptography*, 3, pp. 135-145, 1993, and *Proceedings of CRYPTO' 92, Lecture Notes in Computer Science* 740, pp. 280-291, 1993.

[54] C. Carlet. Two new classes of bent functions. In *Proceedings of EUROCRYPT'93, Lecture Notes in Computer Science* 765, pp. 77-101, 1994.

[55] C. Carlet. Generalized Partial Spreads, *IEEE Transactions on Information Theory*, vol. 41, no. 5, pp. 1482-1487, 1995.

[56] C. Carlet. Hyper-bent functions. *PRAGOCRYPT'96, Czech Technical University Publishing House*, pp. 145-155, 1996.

[57] C. Carlet. A construction of bent functions. *Finite Fields and Applications, London Mathematical Society*, Lecture Series 233, Cambridge University Press, pp. 47-58, 1996.

[58] C. Carlet. More correlation-immune and resilient functions over Galois fields and Galois rings. *Advances in Cryptology, EUROCRYPT' 97, Lecture Notes in Computer Science* 1233, pp. 422-433, 1997.

[59] C. Carlet. On Kerdock codes, American Mathematical Society. *Proceedings of the conference Finite Fields and Applications Fq4, Contemporary Mathematics* 225, pp. 155-163, 1999.

[60] C. Carlet. On the propagation criterion of degree ℓ and order k. *Proceedings of EUROCRYPT'98, Lecture Notes in Computer Science* 1403, pp. 462-474, 1998.

[61] C. Carlet. On cryptographic propagation criteria for Boolean functions. *Information and Computation* , vol. 151, Academic Press pp. 32-56, 1999.

[62] C. Carlet. Recent results on binary bent functions. *Proceedings of the International Conference on Combinatorics, Information Theory and Statistics; Journal of Combinatorics, Information and System Sciences*, Vol. 25, Nos. 1-4, pp. 133-149, 2000.

[63] C. Carlet. On the coset weight divisibility and nonlinearity of resilient and correlation-immune functions, *Proceedings of SETA'01 (Sequences and their Applications 2001), Discrete Mathematics and Theoretical Computer Science*, pp. 131-144, 2001.

[64] C. Carlet. A larger Class of Cryptographic Boolean Functions via a Study of the Maiorana-McFarland Construction. *Proceedings of CRYPT0 2002, Lecture Notes in Computer Science* 2442, pp. 549-564, 2002.

[65] C. Carlet. On the confusion and diffusion properties of Maiorana-McFarland's and extended Maiorana-McFarland's functions. *Special Issue "Complexity Issues in Coding and Cryptography", dedicated to Prof. Harald Niederreiter on the occasion of his 60th birthday, Journal of Complexity* 20, pp. 182-204, 2004.

[66] C. Carlet. On the degree, nonlinearity, algebraic thickness and non-normality of Boolean functions, with developments on symmetric functions. *IEEE Transactions on Information Theory*, vol. 50, pp. 2178-2185, 2004.

[67] C. Carlet. On the secondary constructions of resilient and bent functions. *Proceedings of the Workshop on Coding, Cryptography and Combinatorics 2003, published by Birkhäuser Verlag*, pp. 3-28, 2004.

[68] C. Carlet. Concatenating indicators of flats for designing cryptographic functions. *Design, Codes and Cryptography* volume 36, Number 2, pp.189 - 202, 2005.

[69] C. Carlet. Partial covering sequences: a method for designing classes of cryptographic functions. *Proceedings of the conference "The first Symposium on Algebraic Geometry and its Applications" (SAGA'07), Tahiti, 2007, World Scientific, Series on Number Theory and its Applications*, Vol. 5, pp. 366-387, 2008.

[70] C. Carlet. On bent and highly nonlinear balanced/resilient functions and their algebraic immunities. *Proceedings of AAECC 16, Lecture*

Notes in Computer Science 3857, pp. 1-28, 2006. This paper is an extended version of the paper entitled "Improving the algebraic immunity of resilient and nonlinear functions and constructing bent function", IACR ePrint Archive http://eprint.iacr.org/ 2004/276.

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

[72] C. Carlet. A method of construction of balanced functions with optimum algebraic immunity. To appear in the Proceedings of the International Workshop on Coding and Cryptography, The Wuyi Mountain, Fujiang, China, June 11-15, 2007, World Scientific Publishing Co. in its series of Coding and Cryptology, 2008. A preliminary version is available on IACR ePrint Archive, http://eprint.iacr.org/, 2006/149.

[73] C. Carlet. Recursive lower bounds on the nonlinearity profile of Boolean functions and their applications. *IEEE Transactions on Information Theory*, vol.54, No. 3, pp. 1262-1272, 2008.

[74] C. Carlet. On the higher order nonlinearities of Boolean functions and S-boxes, and their generalizations. *Proceedings of SETA 2008, Lecture Notes in Computer Science* 5203, pp. 345-367, 2008.

[75] C. Carlet and P. Charpin. Cubic Boolean functions with highest resiliency. *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 562-571, 2005.

[76] C. Carlet, D. Dalai, K. Gupta and S. Maitra. Algebraic Immunity for Cryptographically Significant Boolean Functions: Analysis and Construction. *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 3105-3121, 2006.

[77] C. Carlet, L.E. Danielsen, M.G. Parker and P. Solé Self dual bent functions. Proceedings of the conference BFCA 2008, Copenhagen, to appear in Lecture Notes in Computer Science.

[78] C. Carlet and C. Ding. Highly Nonlinear Mappings. *Special Issue "Complexity Issues in Coding and Cryptography", dedicated to Prof. Harald Niederreiter on the occasion of his 60th birthday, Journal of Complexity* 20, pp. 205-244, 2004.

[79] C. Carlet, H. Dobbertin and G. Leander. Normal extensions of bent functions. *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2880-2885, 2004.

[80] C. Carlet and S. Dubuc. On generalized bent and q-ary perfect nonlinear functions. *Proceedings of Finite Fields and Applications Fq5*, Augsburg, Germany, Springer, pp. 81-94, 2000.

[81] C. Carlet and K. Feng. An infinite class of balanced functions with optimum algebraic immunity, good immunity to fast algebraic attacks and good nonlinearity. *Proceedings of ASIACRYPT 2008, Lecture Notes in Computer Science* 5350, pp. 425-440, 2008.

[82] C. Carlet and P. Gaborit. Hyper-bent functions and cyclic codes. *Journal of Combinatorial Theory, Series A*, 113, no. 3, 466-482, 2006.

[83] C. Carlet and P. Gaborit. On the construction of balanced Boolean functions with a good algebraic immunity. *Proceedings of International Symposium on Information Theory, ISIT*, Adelaide, Australia, 2005. A longer version of this paper has been published in the *Proceedings of BFCA 2005*, Publications des universités de Rouen et du Havre, pp. 1-20, 2005.

[84] C. Carlet and A. Gouget. An upper bound on the number of m-resilient Boolean functions. *Proceedings of ASIACRYPT 2002, Lecture Notes in Computer Science* 2501, pp. 484-496, 2002.

[85] C. Carlet and P. Guillot. A characterization of binary bent functions, *Journal of Combinatorial Theory, Series A*, vol. 76, No. 2, pp. 328-335, 1996.

[86] C. Carlet and P. Guillot. An alternate characterization of the bentness of binary functions, with uniqueness. *Designs, Codes and Cryptography*, 14, pp. 133-140, 1998.

[87] C. Carlet and P. Guillot. A new representation of Boolean functions, *Proceedings of AAECC'13, Lecture Notes in Computer Science* 1719, pp. 94-103, 1999.

[88] C. Carlet and P. Guillot. Bent, resilient functions and the Numerical Normal Form. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science,* 56, pp. 87-96, 2001.

[89] C. Carlet, P. Guillot and S. Mesnager. On immunity profile of Boolean functions. *Proceedings of SETA 2006 (International Conference on Sequences and their Applications), Lecture Notes in Computer Science* 4086, pp. 364-375, 2006.

[90] C. Carlet and A. Klapper. Upper bounds on the numbers of resilient functions and of bent functions. This paper was meant to appear in an issue of Lecture Notes in Computer Sciences dedicated to Philippe Delsarte, Editor Jean-Jacques Quisquater. But this issue finally never appeared. A shorter version has appeared in the *Proceedings of the 23rd Symposium on Information Theory in the Benelux*, Louvain-La-Neuve, Belgium, 2002.

[91] C. Carlet and S. Mesnager. On the supports of the Walsh transforms of Boolean functions. *Proceedings of the conference BFCA 2005*, Publications des universités de Rouen et du Havre, pp. 65-82, 2005.

[92] C. Carlet and S. Mesnager. Improving the upper bounds on the covering radii of binary Reed-Muller codes. *IEEE Transactions on Information Theory* 53, pp. 162-173, 2007.

[93] C. Carlet and E. Prouff. On plateaued functions and their constructions. *Proceedings of Fast Software Encryption 2003, Lecture notes in computer science* 2887, pp. 54-73, 2003.

[94] C. Carlet and P. Sarkar. Spectral Domain Analysis of Correlation Immune and Resilient Boolean Functions. *Finite fields and Applications* 8, pp. 120-130, 2002.

[95] C. Carlet and Y. V. Tarannikov. Covering sequences of Boolean functions and their cryptographic significance. *Designs, Codes and Cryptography*, 25, pp. 263-279, 2002.

[96] C. Carlet, X. Zeng, C. Lei and L. Hu. Further properties of several classes of Boolean functions with optimum algebraic immunity. *Proceedings of the First International Conference on Symbolic Computation and Cryptography* SCC 2008, LMIB, pp. 42-54, 2008

[97] L. Carlitz and S. Uchiyama. Bounds for exponential sums. *Duke Math. Journal* 1, pp. 37-41, 1957.

[98] A.H. Chan and R.A. Games. On the quadratic spans of De Bruijn sequences. *IEEE Transactions on Information Theory*, vol. 36, no. 4, pp. 822-829, 1990.

[99] C. Charnes, M. Rötteler and T. Beth. Homogeneous bent functions, invariants, and designs. *Designs, Codes and Cryptography*, 26, pp. 139-154, 2002.

[100] P. Charpin. *Open problems on cyclic codes*, In "Handbook of Coding Theory", Part 1, chapter 11, V. S. Pless, W. C. Huffman, Eds, R. A. Brualdi, assistant editor, Amsterdam, the Netherlands: Elsevier, 1998.

[101] P. Charpin. Normal Boolean functions. *Special Issue "Complexity Issues in Coding and Cryptography", dedicated to Prof. Harald Niederreiter on the occasion of his 60th birthday, Journal of Complexity* 20, pp. 245-265, 2004.

[102] P. Charpin and G. Gong. Hyperbent functions, Kloosterman sums and Dickson polynomials. *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 4230-4238, 2008.

[103] P. Charpin, T. Helleseth, V. Zinoviev. Propagation characteristics of $x \rightarrow 1/x$ and Kloosterman sums. *Finite Fields and Applications* 13 Issue 2, 366-381, 2007.

[104] P. Charpin and G. Kyureghyan. Cubic monomial bent functions: A subclass of $\mathcal{M}$. *SIAM, J. Discr. Math., Vol.22, no.2*, pp. 650–665, 2008.

[105] P. Charpin and E. Pasalic. On propagation characteristics of resilient functions. *Proceedings of SAC 2002, Lecture Notes in Computer Science* 2595, pp. 356–365, 2002.

[106] P. Charpin, E. Pasalic and C. Tavernier. On bent and semi-bent quadratic Boolean functions. *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4286-4298, 2005.

[107] S. Chee, S. Lee, K. Kim and D. Kim. Correlation immune functions with controlable nonlinearity. *ETRI Journal*, vol 19, no 4, pp. 389-401, 1997.

[108] S. Chee, S. Lee, D. Lee and S. H. Sung. On the correlation immune functions and their nonlinearity. *Proceedings of Asiacrypt'96, Lecture Notes in Computer Science* 1163, pp. 232-243.

[109] V. Chepyzhov and B. Smeets. On a fast correlation attack on certain stream ciphers. *Proceedings of EUROCRYPT'91, Lecture Notes in Computer Science* 547, pp. 176-185, 1992.

[110] B. Chor, O. Goldreich, J. Hastad, J. Freidmann, S. Rudich and R. Smolensky. The bit extraction problem or t-resilient functions. *Proceedings of the 26th IEEE Symposium on Foundations of Computer Science*, pp. 396-407, 1985.

[111] G. Cohen, I. Honkala, S. Litsyn and A. Lobstein. *Covering codes*. North-Holland, 1997.

[112] Higher Order Correlation Attacks, XL algorithm and Cryptanalysis of Toyocrypt. *Proceedings of ICISC 2002, Lecture notes in computer science* 2587, pp. 182-199, 2003.

[113] N. Courtois. Fast Algebraic Attacks on Stream Ciphers with Linear Feedback. *Proceedings of CRYPTO 2003, Lecture Notes in Computer Science* 2729, pp. 177-194, 2003.

[114] N. Courtois. Algebraic attacks on combiners with memory and several outputs. *Proceedings of ICISC 2004, Lecture notes in computer science* 3506, pp. 3-20, 2005.

[115] N. Courtois. Cryptanalysis of SFINKS. *Proceedings of ICISC 2005*. Also available at IACR ePrint Archive http://eprint.iacr.org/, Report 2005/243, 2005.

[116] N. Courtois. General Principles of Algebraic Attacks and New Design Criteria for Components of Symmetric Ciphers. *AES 4 Conference, Lecture Notes in Computer Science* 3373, 2004.

[117] N. Courtois and W. Meier. Algebraic Attacks on Stream Ciphers with Linear Feedback. *Proceedings of EUROCRYPT 2003, Lecture Notes in Computer Science* 2656, pp. 346-359.

[118] T. W. Cusick. On constructing balanced correlation immune functions. *Proceedings of SETA'98, (Sequences and their Applications 1998), Discrete Mathematics and Theoretical Computer Science*, pp. 184-190, 1999.

[119] T. W. Cusick, C. Ding and A. Renvall, *Stream Ciphers and Number Theory,* North-Holland Mathematical Library 55. Amsterdam: North-Holland/Elsevier, 1998.

[120] D.M. Cvetkovic, M. Doob and H. Sachs. *Spectra of graphs*. Academic Press, 1979.

[121] D. K. Dalai, K. C. Gupta and S. Maitra. Results on Algebraic Immunity for Cryptographically Significant Boolean Functions. *Proceedings of Indocrypt 2004, Lecture Notes in Computer Science* 3348, pp. 92-106, 2004

[122] D. K. Dalai, K. C. Gupta and S. Maitra. Cryptographically Significant Boolean functions: Construction and Analysis in terms of Algebraic Immunity. *Fast Software Encryption 2005, Lecture Notes in Computer Science* 3557, pp. 98-111, 2005.

[123] D. K. Dalai, K. C. Gupta and S. Maitra. Notion of algebraic immunity and its evaluation related to fast algebraic attacks. *Proceedings of the conference BFCA 2006*, Publications des universités de Rouen et du Havre, pp. 107-124, 2006.

[124] D. K. Dalai, S. Maitra and S. Sarkar. Basic Theory in Construction of Boolean Functions with Maximum Possible Annihilator Immunity. *Designs, Codes and Cryptography*, Volume 40, Number 1, Pages 41–58, July 2006. IACR ePrint Archive http://eprint.iacr.org/, No. 2005/229, 15 July, 2005.

[125] D. Dalai and S. Maitra. Balanced Boolean functions with (more than) maximum algebraic immunity. *Proceedings of the Workshop on Coding and Cryptography* (in the memory of Hans Dobbertin) WCC 2007, pp. 99-108, 2007.

[126] M. Daum, H. Dobbertin and G. Leander. An algorithm for checking normality of Boolean functions. *Proceedings of the Workshop on Coding and Cryptography 2003*, pp. 133-142, 2003.

[127] M. Daum, H. Dobbertin and G. Leander. Short description of an algorithm to create bent functions. Private communication.

[128] E. Dawson and C.-K. Wu. Construction of correlation immune Boolean functions. *Proceedings of ICICS 1997*, pp. 170-180, 1997.

[129] P. Delsarte. *An algebraic approach to the association schemes of coding theory.* PhD thesis. Université Catholique de Louvain (1973)

[130] P. Delsarte. Four fundamental parameters of a code and their combinatorial significance. *Information and Control*, vol. 23 (5), pp. 407-438, 1973.

[131] O. Denisov. An asymptotic formula for the number of correlation-immune of order k Boolean functions. *Discrete Mathematics Appl.*, Vol. 2, No. 4; pp. 407-426, 1992. Translation of a russian article in Diskretnaya Matematika 3, pp. 25-46, 1990.

[132] O. Denisov. A local limit theorem for the distribution of a part of the spectrum of a random binary function. *Discrete Mathematics and Applications*, V. 10, No 1, pp. 87-102, 2000.

[133] F. Didier. A new upper bound on the block error probability after decoding over the erasure channel. *IEEE Transactions on Information Theory*, vol. 52, pp. 4496- 4503, 2006.

[134] F. Didier. Using Wiedemann's algorithm to compute the immunity against algebraic and fast algebraic attacks. *Proceedings of Indocrypt 2006, Lecture Notes in Computer Science* 4329, pp. 236-250.

[135] J. Dillon. A survey of bent functions. *NSA Technical Journal Special Issue*, pp. 191-215, 1972.

[136] J. F. Dillon. *Elementary Hadamard Difference sets.* Ph. D. Thesis, Univ. of Maryland, 1974.

[137] J. F. Dillon. Elementary Hadamard Difference sets, *Proceedings of the Sixth S-E Conf. Comb. Graph Theory and Comp.*, Winnipeg Utilitas Math, pp. 237-249, 1975.

[138] J. Dillon. More DD difference sets. To appear in *Designs, Codes and Cryptography* (on line), 2008.

[139] J. F. Dillon and H. Dobbertin. New cyclic difference sets with Singer parameters. *Finite Fields and Their Applications* 10, pp. 342-389, 2004.

[140] J. F. Dillon and G. McGuire. Near bent functions on a hyperplane. *Finite Fields and Their Applications* Vol. 14, Issue 3, pp. 715-720, 2008.

[141] H. Dobbertin. Construction of bent functions and balanced Boolean functions with high nonlinearity. *Proceedings of Fast Software Encryption, Second International Workshop*, Lecture Notes in Computer Science 1008, pp. 61-74, 1995.

[142] H. Dobbertin, P. Felke, T. Helleseth and P. Rosenthal. Niho type cross-correlation functions via Dickson polynomials and Kloosterman sums. *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 613-627, 2006.

[143] H. Dobbertin and G. Leander. Bent functions embedded into the recursive framework of $\mathbb{Z}$-bent functions. *Designs, Codes and Cryptography*, Vol. 49, no. 1-3, pp. 3-22, 2008.

[144] H. Dobbertin, G. Leander. A survey of some recent results on bent functions. *Proceeding of SETA 2004, Lecture Notes in Computer Science* 3486, pp. 1-29, 2005.

[145] H. Dobbertin, G. Leander, A. Canteaut, C. Carlet, P. Felke and P. Gaborit. Construction of Bent Functions via Niho Power Functions. *Journal of Combinatorial Theory, Series A*, Volume 113, Issue 5, pp. 779-798, 2006.

[146] S. Dubuc. Characterization of linear structures. *Designs, Codes and Cryptography* vol. 22, pp. 33-45, 2001.

[147] I. Dumer, G. Kabatiansky and C. Tavernier. List decoding of Reed-Muller codes up to the Johnson bound with almost linear complexity. *Proceedings of ISIT 2006*. Seattle, USA.

[148] J. H. Evertse. Linear structures in block ciphers. *Proceedings of EUROCRYPT'87, Lecture Notes in Computer Science* 304, pp. 249-266, 1988.

[149] J. C. Faugère. "Fast Gröbner. Algebraic cryptanalysis of HFE and filter generators".*Proceedings of the Workshop on Coding and Cryptography 2003*, pp. 175-176, 2003.

[150] J.-C. Faugère and G. Ars. An Algebraic Cryptanalysis of Nonlinear Filter Generators using Gröbner bases. *Rapport de Recherche INRIA* 4739, 2003.

[151] K. Feng, Q. Liao and J. Yang. Maximal values of generalized algebraic immunity. *Designs, Codes and Cryptography* 50, pp. 243-252, 2009.

[152] E. Filiol and C. Fontaine. Highly nonlinear balanced Boolean functions with a good correlation-immunity. *Proceedings of EUROCRYPT'98, Lecture Notes in Computer Science* 1403, pp. 475-488, 1998.

[153] S. Fischer and W. Meier. Algebraic Immunity of S-boxes and Augmented Functions. *Proceedings of Fast Software Encryption 2007. Lecture Notes in Computer Science* 4593, pp. 366-381, 2007.

[154] C. Fontaine. On some cosets of the First-Order Reed-Muller code with high minimum weight. *IEEE Transactions on Information Theory*, vol. 45, no. 4, pp. 1237-1243, 1999.

[155] R. Forré. The strict avalanche criterion: spectral properties of Boolean functions and an extended definition. *Proceedings of CRYPTO'88, Lecture Notes in Computer Science* 403, pp. 450-468, 1989.

[156] R. Forré. A fast correlation attack on nonlinearly feedforward filtered shift register sequences. *Proceedings of EUROCRYPT '89, Lecture Notes in Computer Science* 434, pp. 586-595, 1990.

[157] R. Fourquet and C. Tavernier. List decoding of second order Reed-Muller and its covering radius implications. *Proceedings of the Workshop on Coding and Cryptography 2007* WCC, pp. 147-156, 2007.

[158] J. Friedman. On the bit extraction problem. *Proceedings of the 33rd IEEE Symposium on Foundations of Computer Science*, pp. 314-319, 1992.

[159] THE BOOLEAN PLANET. Webpage on the equivalence classes of Boolean functions in at most 6 variables, maintained by J. Fuller at URL http://www.booleanfunction.com/

[160] R. G. Gallager. *Low density parity check codes*. Cambridge, MA: MIT Press, 1963.

[161] J. von zur Gathen and J. R. Roche. Polynomials with two values. *Combinatorica* 17(3), pp. 345-362, 1997.

[162] J. Golić. Fast low order approximation of cryptographic functions. *Proceedings of EUROCRYPT'96, Lecture Notes in Computer Science* 1070, pp. 268-282, 1996.

[163] J. Golić. On the security of nonlinear filter generators. *Proceedings of Fast Software Encryption'96, Lecture Notes in Computer Science* 1039, pp. 173-188, 1996.

[164] F. Göloglu and A. Pott. Results on the crosscorrelation and autocorrelation of sequences. *Proceedings of Sequences and their Applications - SETA 2008 - Lecture Notes in Computer Science* 5203, pp. 95-105, 2008.

[165] S.W. Golomb. *Shift Register Sequences*. Aegean Park Press, 1982.

[166] K. Gopalakrishnan, D. G. Hoffman and D. R. Stinson. A Note on a Conjecture Concerning Symmetric Resilient Functions. *Information Processing Letters* 47 (3), pp. 139-143, 1993.

[167] M. Goresky and A. Klapper. Fibonacci and Galois representation of feedback with carry shift registers. *IEEE Transactions on Information Theory*, vol. 48, pp. 2826-2836, 2002.

[168] M. Goresky and A. Klapper. Periodicity and distribution properties of combined FCSR sequences. *Proceedings of Sequences and their Applications - SETA 2006 - Lecture Notes in Computer Science* 4086, pp. 334-341, 2006.

[169] A. Gouget. On the propagation criterion of Boolean functions. *Proceedings of the Workshop on Coding, Cryptography and Combinatorics 2003, Birkhäuser Verlag*, pp. 153-168, 2004.

[170] A. Gouget and H. Sibert. Revisiting correlation-immunity in filter generators. *Proceedings of SAC 2007, Lecture Notes in Computer Science* 4876, pp. 378-395, 2007.

[171] P. Guillot. Completed GPS Covers All Bent Functions. *Journal of Combinatorial Theory*, Series A 93, pp. 242-260, 2001.

[172] P. Guillot. Partial bent functions. *Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics, SCI 2000*, 2000.

[173] Xiao Guo-Zhen, C. Ding and W. Shan. *The stability theory of stream ciphers*, Lecture Notes in Computer Science 561, 1991.

[174] X. Guo-Zhen and J. L. Massey. A Spectral Characterization of Correlation-Immune Combining Functions. *IEEE Trans. Inf. Theory*, vol. 34, no. 3, pp. 569-571, 1988.

[175] A. R. Hammons Jr., P. V. Kumar, A. R. Calderbank, N. J. A. Sloane and P. Solé. The Z_4-linearity of Kerdock, Preparata, Goethals and related codes. *IEEE Transactions on Information Theory*, vol. 40, pp. 301-319, 1994.

[176] P. Hawkes and G. Rose. Rewriting Variables: The Complexity of Fast Algebraic Attacks on Stream Ciphers. *Proceedings of CRYPTO 2004, Lecture Notes in Computer Science* 3152, pp. 390–406, 2004.

[177] T. Helleseth, T. Kløve, and J. Mykkelveit. On the covering radius of binary codes. *IEEE Transactions on Information Theory*, vol. 24, no. 5, pp. 627-628, 1978.

[178] T. Helleseth and H.F. Mattson Jr. On the cosets of the simplex code. *Discrete Mathematics* 56, pp. 169-189, 1985.

[179] S. Hirose and K. Ikeda. Nonlinearity criteria of Boolean functions. *KUIS Technical Report*, KUIS-94-0002, 1994.

[180] S. Hirose and K. Ikeda. Complexity of Boolean functions satisfying the propagation criterion. *Proceedings of the 1995 Symposium on Cryptography and Information Security*, SCIS95-B3.3, 1995.

[181] I. Honkala and A. Klapper. Bounds for the multicovering radii of Reed-Muller codes with applications to stream ciphers. *Designs, Codes and Cryptography* 23, pp. 131-145, 2001.

[182] X.-D. Hou. Some results on the covering radii of Reed-Muller codes. *IEEE Transactions on Information Theory*, vol. 39, no. 2, pp. 366-378, 1993.

[183] X.-D. Hou. Classification of cosets of the Reed-Muller code $R(m-3, m)$. *Discrete Mathematics*, 128, pp. 203-224, 1994.

[184] X.-D. Hou. The covering radius of $R(1, 9)$ in $R(4, 9)$. *Designs, Codes and Cryptography* 8 (3), pp. 285-292, 1995.

[185] X.-D. Hou. $AGL(m, 2)$ acting on $R(r, m)/R(s, m)$. *Journal of Algebra* 171, pp. 921-938, 1995.

[186] X.-D. Hou. Covering radius of the Reed-Muller code $R(1, 7)$ - a simpler proof. *J. Combin. Theory*, Series A 74, pp. 337-341, 1996.

[187] X.-D. Hou. $GL(m, 2)$ acting on $R(r, m)/R(r-1, m)$. *Discrete Mathematics* 149, pp. 99-122, 1996.

[188] X.-D. Hou. On the covering radius of $R(1, m)$ in $R(3, m)$. *IEEE Transactions on Information Theory*,vol. 42, no. 3, pp. 1035-1037, 1996.

[189] X.-D. Hou. The Reed-Muller code $R(1, 7)$ is normal. *Designs, Codes and Cryptography* 12, pp. 75-82, 1997.

[190] X.-D. Hou. Cubic bent functions. *Discrete Mathematics* vol. 189, pp. 149-161, 1998.

[191] X.D. Hou. On the coefficients of binary bent functions. *Proceedings of the American American Society*, Vol. 128, No. 4, pp. 987-996, 2000.

[192] X.-D. Hou. New Constructions of Bent Functions. *Proceedings of the International Conference on Combinatorics, Information Theory and Statistics; Journal of Combinatorics, Information and System Sciences*, Vol. 25, Nos. 1-4, pp. 173-189, 2000.

[193] X.-D. Hou. On Binary Resilient Functions. *Des. Codes Cryptography* 28(1), pp. 93-112, 2003.

[194] X.-D. Hou. Group Actions on Binary Resilient Functions. *Appl. Algebra Eng. Commun. Comput.* 14(2), pp. 97-115, 2003.

[195] X.-D. Hou. A note on the proof of a theorem of Katz. *Finite Fields and their Applications* Volume 11, pp. 316-319, 2005.

[196] X.-D. Hou and P. Langevin. Results on bent functions, *Journal of Combinatorial Theory, Series A*, 80, pp. 232-246, 1997.

[197] H. Hu and D. Feng. On quadratic bent functions in polynomial forms. *IEEE Trans. Info. Theory*, vol. 53, pp. 2610-2615, 2007.

[198] T. Iwata and K. Kurosawa. Probabilistic higher order differential attack and higher order bent functions. *Proceedings of ASIACRYPT'99, Lecture Notes in Computer Science* 1716, pp. 62-74, 1999.

[199] C.J.A. Jansen and D.E. Boekee. The shortest feedback shift register that can generate a given sequence. *Proceedings of CRYPTO'89, Lecture Notes in Computer Science* 435, pp. 90-99,1990 (this paper refers to the classified PhD thesis of C.J.A. Jansen entitled "Investigations on nonlinear streamcipher systems: construction and evaluation methods", Philips).

[200] T. Johansson and F. Jönsson. Improved fast correlation attack on stream ciphers via convolutional codes. *Proceedings of EUROCRYPT'99, Lecture Notes in Computer Science* 1592, pp. 347-362, 1999.

[201] T. Johansson and F. Jönsson. Fast correlation attacks based on turbo code techniques. *Advances in Cryptology - CRYPTO'99, no. 1666 in Lecture Notes in Computer Science*, pp. 181-197, 1999.

[202] T. Johansson and F. Jönsson. Fast correlation attacks through reconstruction of linear polynomials. *Advances in Cryptology - CRYPTO 2000, no. 1880 in Lecture Notes in Computer Science*, pp. 300-315, 2000.

[203] F. Jönsson. *Some results on fast correlation attacks*. PhD thesis. Lund University. 2002.

[204] D. Jungnickel. *Difference sets.* Contemporary Design Theory: A Collection of Surveys, J. Dinitz and D. R. Stinson eds. John Wiley & Sons, 1992.

[205] W. Kantor. An Exponential Number of Generalized Kerdock Codes. *Inf. and Contr.* 53, pp. 74-80, 1982.

[206] T. Kasami. The weight enumerators for several classes of subcodes of the second order binary Reed-Muller codes. *Information and Control*, 18, pp. 369-394, 1971.

[207] T. Kasami and N. Tokura. On the weight structure of the Reed Muller codes, *IEEE Trans. Info. Theory*, vol. 16, pp. 752-759, 1970.

[208] T. Kasami, N. Tokura, and S. Azumi. On the Weight Enumration of Weights Less than 2.5d of Reed-Muller Codes. *Information and Control*, 30:380–395, 1976.

[209] N. Katz. On a theorem of Ax. *American Journal of Mathematics* 93, pp. 485-499, 1971.

[210] S. Kavut, S. Maitra and M. D. Yücel. Search for Boolean functions with excellent profiles in the rotation symmetric class. *IEEE Transactions on Information Theory*, vol. 53, no. 5, pp. 1743-1751, 2007.

[211] A. M. Kerdock. A class of low-rate non linear codes. *Information and Control* 20, pp. 182-187, 1972.

[212] A. Kerckhoffs. La Cryptographie Militaire. *Journal des Sciences Militaires*, 1883.

[213] J.D. Key, T.P. McDonough and V.C. Mavron. Information sets and partial permutation decoding for codes from finite geometries. *Finite Fields and their Applications* Volume 12, Issue 2, pp. 232-247, 2006.

[214] J. Khan, G. Kalai and N. Linial. The influence of variables on Boolean functions. *IEEE 29th Symp. on foundations of Computer Science*, pp. 68-80, 1988.

[215] L.R. Knudsen. Truncated and higher order differentials. *Proceedings of Fast Software Encryption, Second International Workshop, Lecture Notes in Computer Science* 1008, pp. 196-211, 1995.

[216] K. Khoo and G. Gong. New constructions for resilient and highly nonlinear Boolean functions. *Proceedings of 8th Australasian Conference, ACISP 2003, Wollongong, Austrialia, Lecture Notes in Computer Science* 2727, pp. 498-509, 2003.

[217] K. Khoo, G. Gong and D. Stinson. A New Characterization of Semibent and Bent Functions on Finite Fields. *Designs, Codes and Cryptography*, Volume 38, Number 2, pp. 279-295, 2006.

[218] A. Klapper and M. Goresky. Feedback Shift Registers, 2-Adic Span, and Combiners with Memory. *Journal of Cryptology*, vol. 10, pp. 111-147. 1997.

[219] L.R. Knudsen and M.P.J. Robshaw. Non-linear approximations in linear cryptanalysis. *Proceedings of EUROCRYPT'96, Lecture Notes in Computer Science* 1070, pp. 224-236, 1996.

[220] P.V. Kumar, R.A. Scholtz and L.R. Welch. Generalized bent functions and their properties, *Journal of Combinatorial Theory, Series A* 40, pp. 90-107, 1985.

[221] K. Kurosawa, T. Iwata and T. Yoshiwara. New covering radius of Reed-Muller codes for t-resilient functions. *Proceedings of Selected Areas in Cryptography, 8th Annual International Workshop, Lecture Notes in Computer Science* 2259, pp. 75 ff, 2001.

[222] K. Kurosawa and R. Matsumoto. Almost security of cryptographic Boolean functions. *IEEE Transactions on Information Theory*, vol.50, no. 11, pp. 2752-2761, 2004.

[223] K. Kurosawa and T. Satoh. Design of $SAC/PC(\ell)$ of order k Boolean functions and three other cryptographic criteria. *Advances in Cryptology, EUROCRYPT' 97, Lecture Notes in Computer Science* 1233, pp. 434-449, 1997.

[224] P. Lacharme. Post processing functions for a physical random number generator. *Proceedings of Fast Software Encryption 2008, Lecture Notes in Computer Science* 5086, p 334-342, 2008.

[225] G. Lachaud and J. Wolfmann. The Weights of the Orthogonals of the Extended Quadratic Binary Goppa Codes. *IEEE Trans. Inform. Theory*, vol. 36, pp. 686-692, 1990.

[226] J. Lahtonen, G. McGuire and H. Ward. Gold and Kasami-Welch functions, quadratic forms and bent functions. *Advances of Mathematics of Communication*, vol. 1, pp. 243-250, 2007.

[227] X. Lai. Higher order derivatives and differential cryptanalysis. *Proceedings of the "Symposium on Communication, Coding and Cryptography", in honor of J. L. Massey on the occasion of his 60'th birthday.* 1994.

[228] X. Lai. Additive and linear structures of cryptographic functions. *Proceedings of Fast Software Encryption, Second International Workshop, Lecture Notes in Computer Science* 1008, pp. 75-85, 1995.

[229] P. Langevin. Covering radius of $RM(1,9)$ in $RM(3,9)$. *Eurocode'90, Lecture Notes in Computer Science* 514, pp. 51-59, 1991.

[230] P. Langevin. On the orphans and covering radius of the Reed-Muller codes. *Proceedings of AAECC 9, Lecture Notes in Computer Science* 539, pp. 234-240, 1991.

[231] P. Langevin. On generalized bent functions. *CISM Courses and Lectures 339 (Eurocode)*, pp. 147-157, 1992.

[232] P. Langevin and G. Leander. Monomial Bent Functions and Stickelberger's Theorem. T*Finite Fields and theirs Applications* vol. 14, no. 3, pp. 727-742, 2008.

[233] P. Langevin, G. Leander, P. Rabizzoni, P. Veron and J.-P. Zanotti. Web page http://langevin.univ-tln.fr/project/quartics/

[234] P. Langevin, P. Rabizzoni, P. Veron, J.-P. Zanotti. On the number of bent functions with 8 variables. *Proceedings of the conference BFCA 2006*, Publications des universités de Rouen et du Havre, pp. 125-136, 2007.

[235] P. Langevin and P. Solé. Kernels and defaults. American Mathematical Society (*Proceedings of the conference Finite Fields and Applications* Fq4) *Contemporary Mathematics* 225, pp. 77-85, 1999.

[236] P. Langevin and P. Véron. On the nonlinearity of power functions. *Designs, Codes and Cryptography* 37, pp. 31 - 43, 2005.

[237] P. Langevin and J.-P. Zanotti. Nonlinearity of some invariant Boolean functions. *Designs, Codes and Cryptography* 36, pp. 131 - 146, 2005.

[238] C. Lauradoux and M. Videau. Matriochka symmetric Boolean functions. *Proceedings of International Symposium on Information Theory, ISIT 2008.*

[239] G. Leander. Bent functions with 2^r Niho exponents. *Proceedings of the Workshop on Coding and Cryptography 2005*, pp. 454-461, 2005.

[240] G. Leander. Monomial bent functions. *Proceedings of the Workshop on Coding and Cryptography* 2005, Bergen, pp. 462-470, 2005. And *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 738-743, 2006.

[241] G. Leander. Another class of non-normal bent functions. *Proceedings of the conference BFCA 2006*, Publications des universités de Rouen et du Havre, pp. 87-98, 2006.

[242] G. Leander and A. Kholosha. Bent functions with 2^r Niho exponents. *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5529-5532, 2006.

[243] R. J. Lechner. *Harmonic analysis of switching functions.* In Recent Developments in Switching Theory, Academic Press, New York, 1971.

[244] S. Leveiller, G. Zemor, P. Guillot and J. Boutros. A new cryptanalytic attack for PN-generators filtered by a Boolean function. *Proceedings of Selected Areas of Cryptography 2002, Lecture Notes in Computer Science* 2595, pp. 232 - 249 (2003).

[245] N.Li and W.Qi. Symmetric Boolean functions depending on an odd number of variables with maximum algebraic immunity. *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2271-2273, 2006.

[246] N. Li and W.-Q. Qi. Construction and analysis of Boolean functions of $2t + 1$ variables with maximum algebraic immunity. *Proceedings of Asiacrypt 2006, Lecture Notes in Computer Science* 4284, pp. 84-98, 2006.

[247] N. Li, L. Qu, W.-F. Qi, G. Feng, C. Li and D. Xie. On the construction of Boolean functions with optimal algebraic immunity. *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 1330-1334, 2008.

[248] R. Lidl and H. Niederreiter, *Finite Fields*, Encyclopedia of Mathematics and its Applications, vol. 20, Addison-Wesley, Reading, Massachussetts (1983)

[249] S. Ling and C. Xing, *Coding Theory*, Cambridge: Cambridge University Press, 2004.

[250] N. Linial, Y. Mansour and N. Nisan. Constant depth circuits, Fourier transform, and learnability. *Journal of the Association for Computing Machinery*, vol. 40 (3), pp. 607-620, 1993.

[251] J. H. van Lint. *Introduction to coding theory*, Springer, New York, 1982.

[252] F. Liu and K. Feng. On the 2^m-variable symmetric Boolean functions with maximum algebraic immunity 2^{m-1}. *Proceedings of the Workshop on Coding and Cryptography 2007* WCC, pp. 225-232, 2007.

[253] S. Lloyd. Properties of binary functions. *Proceedings of EUROCRYPT'90, Lecture Notes in Computer Science* 473, pp. 124-139, 1991.

[254] S. Lloyd. Counting binary functions with certain cryptographic properties. *Journal of Cryptology* 5, pp. 107-131; 1992.

[255] S. Lloyd. Balance, uncorrelatedness and the strict avalanche criterion. *Discrete Applied Mathematics*, 41, pp. 223-233, 1993.

[256] M. Lobanov. Tight bound between nonlinearity and algebraic immunity. IACR ePrint Archive http://eprint.iacr.org/ 2005/441.

[257] O.A. Logachev, A.A. Salnikov and V.V. Yashchenko. Bent functions on a finite Abelian group. *Discrete Mathematics Appl.* vol 7, N° 6, pp. 547-564, 1997.

[258] F. J. MacWilliams and N. J. Sloane. *The theory of error-correcting codes*, Amsterdam, North Holland. 1977.

[259] J. A. Maiorana. A classification of the cosets of the Reed-Muller code $R(1, 6)$. *Mathematics of Computation.* vol. 57, No. 195, pp. 403-414, 1991.

[260] S. Maitra. Highly nonlinear balanced Boolean functions with very good autocorrelation property. *Proceedings of the Workshop on Coding and Cryptography 2001* published by *Electronic Notes in Discrete Mathematics*, Elsevier, vo. 6, pp. 355-364, 2001.

[261] S. Maitra. Autocorrelation properties of correlation immune Boolean functions. *Proceedings of INDOCRYPT 2001, Lecture Notes in Computer Science* 2247, pp. 242-253, 2001.

[262] S. Maitra, S. Kavut, M. Yucel. Balanced Boolean Function on 13-variables having Nonlinearity greater than the Bent Concatenation Bound. Proceedings of the conference BFCA 2008, Copenhagen, to appear.

[263] S. Maitra and E. Pasalic. Further constructions of resilient Boolean functions with very high nonlinearity. *IEEE Transactions on Information Theory*, vol.48, no. 7, pp. 1825-1834, 2002.

[264] S. Maitra and P. Sarkar. Enumeration of correlation-immune Boolean functions. *Proceedings of ACISP 1999*, pp. 12-25, 1999.

[265] S. Maitra and P. Sarkar. Maximum nonlinearity of symmetric Boolean functions on odd number of variables. *IEEE Transactions on Information Theory*, vol. 48, pp. 2626-2630, 2002.

[266] S. Maitra and P. Sarkar. Highly nonlinear resilient functions optimizing Siegenthaler's inequality. *Proceedings of CRYPTO'99, Lecture Notes in Computer Science* 1666, pp. 198-215, 1999.

[267] S. Maitra and P. Sarkar. Modifications of Patterson-Wiedemann functions for cryptographic applications. *IEEE Transactions on Information Theory*, vol. 48, pp. 278-284, 2002.

[268] S. Maity and S. Maitra. Minimum distance between bent and 1-resilient Boolean functions. *Proceedings of Fast Software Encryption 2004, Lecture Notes in Computer Science* 3017, pp. 143-160, 2004.

[269] J. L. Massey. Shift-register analysis and BCH decoding. *IEEE Transactions on Information Theory*, vol. 15, pp. 122-127, 1969.

[270] J. L. Massey. Randomness, arrays, differences and duality. *IEEE Transactions on Information Theory*, vol. 48, pp. 1698-1703, 2002.

[271] M. Matsui. Linear cryptanalysis method for DES cipher. *Proceedings of EUROCRYPT'93, Lecture Notes in Computer Science* 765, pp. 386-397, 1994.

[272] R.J. McEliece. Weight congruence for p-ary cyclic codes. *Discrete Mathematics*, 3, pp. 177-192, 1972.

[273] R. L. McFarland. A family of noncyclic difference sets, *Journal of Comb. Theory, Series A*, no. 15, pp. 1-10, 1973.

[274] W. Meier, E. Pasalic and C. Carlet. Algebraic attacks and decomposition of Boolean functions. *Advances in Cryptology, EUROCRYPT 2004, Lecture Notes in Computer Science* 3027, pp. 474-491, 2004.

[275] W. Meier and O. Staffelbach. Fast correlation attacks on stream ciphers. *Advances in Cryptology, EUROCRYPT'88, Lecture Notes in Computer Science* 330, pp. 301-314, 1988.

[276] W. Meier and O. Staffelbach. Nonlinearity Criteria for Cryptographic Functions. *Advances in Cryptology, EUROCRYPT' 89, Lecture Notes in Computer Science* 434, pp. 549-562, 1990.

[277] W. Meier and O. Staffelbach. Correlation properties of combiners with memory in stream ciphers. *Advances in Cryptology, EUROCRYPT'90, Lecture Notes in Computer Science* 473, pp. 204-213, 1990.

[278] A. Menezes, P. van Oorschot and S. Vanstone. *Handbook of Applied Cryptography*. CRC Press Series on Discrete Mathematics and Its Applications, 1996.

[279] Q. Meng, H. Zhang, M. Yang and J. Cui. On the degree of homogeneous bent functions. *Discrete Applied Mathematics* Volume 155, Issue 5, pp. 665-669, 2007.

[280] S. Mesnager. Improving the lower bound on the higher order nonlinearity of Boolean functions with prescribed algebraic immunity. *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3656 - 3662, 2008. Preliminary version available at IACR ePrint Archive http://eprint.iacr.org/, 2007/117.

[281] S. Mesnager. On the number of resilient Boolean functions. *Proceedings of the conference "The first Symposium on Algebraic Geometry and its Applications" (SAGA'07), Tahiti, 2007, published by World Scientific,*

Series on Number Theory and its Applications, Vol. 5, pp. 419-433, 2008.

[282] W. Millan, A. Clark and E. Dawson. Heuristic design of cryptographically strong balanced Boolean functions. *EUROCRYPT'98, Advances in Cryptology, Lecture Notes in Computer Science* 1403, 1998.

[283] C. J. Mitchell. Enumerating Boolean functions of cryptographic signifiance. *Journal of Cryptology* 2 (3), pp. 155-170, 1990.

[284] J. Mykkelveit. The covering radius of the [128,8] Reed-Muller code is 56. *IEEE Transactions on Information Theory*, vol. 26, no. 3, pp. 359-362, 1980.

[285] Y. Nawaz, G. Gong, and K. Gupta. Upper Bounds on Algebraic Immunity of Power Functions. *Proceeding of Fast Software Encryption 2006, Lecture Notes in Computer Science* 4047, pp. 375-389, 2006.

[286] N. Nisan and M. Szegedy. On the degree of Boolean functions as real polynomials. *Comput. Complexity* 4, pp. 301-313, 1994.

[287] K. Nyberg. Constructions of bent functions and difference sets, *EUROCRYPT'90, Advances in Cryptology, Lecture Notes in Computer Science* 473, pp. 151-160, 1991.

[288] L. O'Connor and A. Klapper. Algebraic nonlinearity and its applications to cryptography. *Journal of Cryptology* 7, pp. 213-227, 1994.

[289] D. Olejár and M. Stanek. On cryptographic properties of random Boolean functions. *Journal of Universal Computer Science*, vol. 4, No.8, pp. 705-717, 1998.

[290] J. D. Olsen, R. A. Scholtz and L. R. Welch. Bent function sequences, *IEEE Trans. on Inf. Theory*, Vol. 28, no. 6, 1982.

[291] E. Pasalic. *On Boolean functions in symmetric-key ciphers.* Ph.D. Thesis, 2003.

[292] E. Pasalic, T. Johansson, S. Maitra and P. Sarkar. New constructions of resilient and correlation immune Boolean functions achieving upper bounds on nonlinearity. *Proceedings of the Workshop on Coding and Cryptography* 2001, published by *Electronic Notes in Discrete Mathematics*, Elsevier, vo. 6, pp. 425-434, 2001.

[293] E. Pasalic and S. Maitra. A Maiorana-McFarland type construction for resilient Boolean functions on n variables (n even) with nonlinearity $> 2^{n-1} - 2^{n/2} + 2^{n/2-2}$. *Proceedings of the Workshop on Coding and Cryptography 2003*, pp. 365-374, 2003.

[294] S. M. Park, S. Lee, S. H. Sung, K. Kim. Improving bounds for the number of correlation-immune Boolean functions. *Information Processing Letters* 61, pp. 209-212, 1997.

[295] N.J. Patterson and D.H. Wiedemann. The covering radius of the $[2^{15}, 16]$ Reed-Muller code is at least 16276. *IEEE Transactions on Information Theory*, vol. 29, pp. 354-356, 1983.

[296] N.J. Patterson and D.H. Wiedemann. Correction to [295]. *IEEE Transactions on Information Theory*, vol. 36, no. 2, pp. 443, 1990.

[297] J. Pieprzyk and X.-M. Zhang. Computing Möbius Transforms of Boolean Functions and Characterizing Coincident Boolean functions. *Proceedings of the conference BFCA 2007*, Publications des universités de Rouen et du Havre, 2007.

[298] V. S. Pless, W. C. Huffman, Eds, R. A. Brualdi, assistant editor. *Handbook of Coding Theory*, Amsterdam, the Netherlands: Elsevier, 1998.

[299] A. Pott. *Finite Geometry and Character Theory*. Lecture Notes in Mathematics, vol. 1601, Berlin, Springer Verlag, 1995.

[300] B. Preneel, W. Van Leekwijck, L. Van Linden, R. Govaerts and J. Vandevalle. Propagation characteristics of Boolean functions, *Proceedings of EUROCRYPT'90, Lecture Notes in Computer Sciences* 473, pp. 161-173, 1991.

[301] B. Preneel, R. Govaerts and J. Vandevalle. Boolean functions satisfying higher order propagation criteria, *Proceedings of EUROCRYPT'91, Lecture Notes in Computer Sciences* 547, pp. 141-152, 1991.

[302] B. Preneel. *Analysis and Design of Cryptographic Hash Functions*, Ph. D. Thesis, Katholieke Universiteit Leuven, K. Mercierlaan 94, 3001 Leuven, Belgium, U.D.C. 621.391.7, 1993.

[303] L. Qu and C. Li. Weight support technique and the symmetric Boolean functions with maximum algebraic immunity on even number of vari-

ables. *Proceedings of INSCRYPT 2007, Lecture Note in Computer Science* 4990, pp. 271-282.

[304] L. Qu, C. Li and K. Feng. A note on symmetric Boolean functions with maximum algebraic immunity in odd number of variables. *IEEE Trans. on Inf. Theory*, vol. 53, pp. 2908-2910, 2007.

[305] L. Qu, K. Feng, L. Feng and L. Wang. Constructing symmetric Boolean functions with maximum algebraic immunity. *IEEE Trans. on Inf. Theory*, vol. 55, pp. 2406-2412, 2009.

[306] M. Quisquater. *Applications of character theory and the Möbius inversion principle to the study of cryptographic properties of Boolean functions*. PhD thesis, 2004.

[307] M. Quisquater, B. Preneel and J. Vandewalle. A new inequality in discrete Fourier theory. *IEEE Trans. on Inf. Theory*, vol. 49, pp. 2038-2040, 2003.

[308] M. Quisquater, B. Preneel and J. Vandewalle. Spectral characterization of cryptographic Boolean functions satisfying the (extended) propagation criterion of degree l and order k. *Inf. Process. Lett.* 93(1), pp. 25-28, 2005.

[309] C. Riera and M. G. Parker. Generalised Bent Criteria for Boolean Functions (I). *IEEE Transactions on Information theory*, vol. 52, no. 9, pp. 4142-4159, 2006.

[310] C. R. Rao. Factorial experiments derived from combinatorial arrangements of arrays. *J. Roy. Statist.* 9, pp. 128-139, 1947.

[311] F. Rodier. Asymptotic nonlinearity of Boolean functions. *Designs, Codes and Cryptography*, no 40:1 2006, pp 59-70.

[312] S. Ronjom, M. Abdelraheem and L. E. Danielsen. Online database of Boolean Functions. http://www.ii.uib.no/ mohamedaa/odbf/index.html

[313] S. Rønjom and T. Helleseth. A new attack on the filter generator. *IEEE Transactions on Information theory*, vol. 53, no. 5, pp. 1752-1758, 2007.

[314] S. Rønjom and T. Helleseth. Attacking the filter generator over $GF(2^m)$. *Proceedings of the International Workshop on the Arithmetic of Finite Fields, WAIFI 2007, Lecture Notes in Computer Science* 4547, pp. 264-275, June 2007.

[315] O. S. Rothaus. On "bent" functions. *J. Comb. Theory*, 20A, pp. 300-305, 1976.

[316] B.V. Ryazanov. On the distribution of the spectral complexity of Boolean functions. *Discrete Mathematics Appl.*, vol. 4, No. 3, pp. 279-288, 1994.

[317] R. A. Rueppel *Analysis and design of stream ciphers* Com. and Contr. Eng. Series, Berlin, Heidelberg, NY, London, Paris, Tokyo 1986

[318] R. A. Rueppel and O. J. Staffelbach. Products of linear recurring sequences with maximum complexity. *IEEE Transactions on Information theory*, vol. 33, no. 1, 1987.

[319] P. Sarkar. The Filter-Combiner Model for Memoryless Synchronous Stream Ciphers. *Proceedings of CRYPTO 2002, Lecture Notes in Computer Science* 2442, pp. 533-548, 2002.

[320] P. Sarkar and S. Maitra. Construction of nonlinear Boolean functions with important cryptographic properties. *Proceedings of EUROCRYPT 2000, Lecture Notes in Computer Science* 1807, pp. 485-506, 2000.

[321] P. Sarkar and S. Maitra. Nonlinearity Bounds and Constructions of Resilient Boolean Functions. *Proceedings of CRYPTO 2000, Lecture Notes in Computer Science* 1880, pp. 515-532, 2000.

[322] P. Sarkar and S. Maitra. Construction of nonlinear resilient Boolean functions using "small" affine functions. *IEEE Transactions on Information theory*, vol. 50, no 9, pp. 2185-2193, 2004.

[323] P. Sarkar and S. Maitra. Balancedness and correlation immunity of symmetric Boolean functions. *Discrete Mathematics* 307, pp. 2351-2358, 2007.

[324] P. Savicky. On the bent Boolean functions that are symmetric. *Eur. J. Combinatorics* 15, pp. 407-410, 1994.

[325] M. Schneider. A note on the construction and upper bounds of correlation-immune functions. *Proceedings of the 6th IMA Conference, Lecture Notes In Computer Science* 1355, pp. 295-306, 1997. An extended version appeared under the title "On the construction and upper bounds of balanced and correlation-immune functions", *Selected Areas Crypt.*, pp. 73-87, 1997.

[326] J. Seberry and X-.M. Zhang. Constructions of bent functions from two known bent functions. *Australasian Journal of Combinatorics* no. 9, pp. 21-35, 1994.

[327] J. Seberry, X-.M. Zhang and Y. Zheng. On constructions and nonlinearity of correlation immune Boolean functions. *Advances in Cryptology - EUROCRYPT'93, Lecture Notes in Computer Science* 765, pp. 181-199, 1994.

[328] J. Seberry, X-.M. Zhang and Y. Zheng. Nonlinearly balanced Boolean functions and their propagation characteristics. *Advances in Cryptology - CRYPTO'93*, pp. 49-60, 1994.

[329] N. V. Semakov and V. A. Zinoviev. Balanced codes and tactical configurations. *Problems of Info. Trans.* 5(3), pp. 22-28 (1969)

[330] A. Shanbhag, V. Kumar and T. Helleseth. An upper bound for the extended Kloosterman sums over Galois rings. *Finite Fields and their Applications* 4, pp. 218-238, 1998.

[331] C. E. Shannon. A mathematical theory of communication. *Bell system technical journal*, 27, pp. 379-423, 1948.

[332] C.E. Shannon. Communication theory of secrecy systems. *Bell system technical journal*, 28, pp. 656-715, 1949.

[333] C. E. Shannon. The synthesis of two-terminal switching circuits. *Bell system technical journal*, 28, pp. 59–98, 1949.

[334] I. Shparlinski. Bounds on the Fourier coefficients of the weighted sum function. *Inf. Process. Lett.* 103(3), pp. 83-87, 2007.

[335] I. Shparlinski and A. Winterhof. On the nonlinearity of linear recurrence sequences. *Applied Mathematics Letters* 19, pp. 340-344, 2006.

[336] T. Siegenthaler. Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information theory*, vol. 30, no 5, pp. 776-780, 1984.

[337] T. Siegenthaler. Decrypting a Class of Stream Ciphers Using Ciphertext Only. *IEEE Transactions on Computer, vol. C-34*, No 1, pp. 81-85, 1985.

[338] P. Stanica, S. Maitra and J. Clark. Results on rotation symmetric bent and correlation immune Boolean functions. *Proceedings of Fast Software Encryption 2004, Lecture Notes in Computer Science* 3017, pp. 161-177, 2004.

[339] P. Stanica and S. H. Sung. Boolean functions with five controllable cryptographic properties. *Designs, Codes and Cryptography* 31, pp. 147-157, 2004.

[340] V. Strassen. Gaussian elimination is not optimal. *Numerische Math.* 13, pp. 354-356, 1969.

[341] I. Strazdins. Universal affine classification of Boolean functions. *Acta Applicandae Mathematicae* 46, pp. 147-167, 1997.

[342] T. Sugita, T. Kasami and T. Fujiwara. Weight distributions of the third and fifth order Reed-Muller codes of length 512. Nara Inst. Sci. Tech. Report, 1996.

[343] S. H. Sung, S. Chee and C. Park. Global avalanche characteristics and propagation criterion of balanced Boolean functions. *Information Processing Letters* 69, pp. 21-24, 1999.

[344] H. Tapia-Recillas and G. Vega. An upper bound on the number of iterations for transforming a Boolean function of degree greater than or equal than 4 to as function of degree 3. *Designs, Codes and Cryptography* 24, pp. 305-312, 2001.

[345] Y. V. Tarannikov. On resilient Boolean functions with maximum possible nonlinearity. *Proceedings of INDOCRYPT 2000, Lecture Notes in Computer Science* 1977, pp. 19-30, 2000.

[346] Y. V. Tarannikov. New constructions of resilient Boolean functions with maximum nonlinearity. *Proceedings of FSE 2001, 8th International Workshop, FSE 2001, Lecture Notes in Computer Science* 2355, pp. 66-77, 2001.

[347] Y. V. Tarannikov and D. Kirienko. Spectral analysis of high order correlation immune functions. *Proceedings of 2001 IEEE International Symposium on Information Theory*, p. 69, 2001 (full preliminary version at IACR ePrint Archive http://eprint.iacr.org/).

[348] Y. V. Tarannikov, P. Korolev and A. Botev. Autocorrelation coefficients and correlation immunity of Boolean functions. *Proceedings of Asiacrypt 2001, Lecture Notes in Computer Science* 2248, pp. 460-479, 2001

[349] F.Didier and J.Tillich: Computing the Algebraic Immunity Efficiently. *Proceedings of Fast Software Encryption 2006, Lecture Notes in Computer Science* 4047, pp. 359-374, 2006.

[350] S. Tsai. Lower bounds on representing Boolean functions as polynomials in $\mathbb{Z}_m^\star$. *SIAM J. Discrete Mathematics*, vol. 9 (1), pp. 55-62, 1996.

[351] S. F. Vinokurov and N. A. Peryazev. An expansion of Boolean function into a sum of products of subfunctions. *Discrete Mathematics Appl.*, vol. 3 (5), pp. 531-533, 1993.

[352] A.F. Webster and S.E. Tavares. On the design of S-boxes. In *Proceedings of CRYPTO'85, Lecture Notes in Computer Science* 219, pp. 523-534, 1985.

[353] J. Wolfmann. Bent functions and coding theory. *Difference Sets, Sequences and their Correlation Properties,* A. Pott, P. V. Kumar, T. Helleseth and D. Jungnickel, eds., pp. 393–417. Amsterdam: Kluwer, 1999.

[354] Y. X. Yang and B. Guo. Further enumerating Boolean functions of cryptographic signifiance. *Journal of Cryptology* 8 (3), pp. 115-122, 1995.

[355] R. Yarlagadda and J.E. Hershey. Analysis and synthesis of bent sequences, *IEE proceedings. Part E. Computers and digital techniques*, vol. 136, pp. 112-123, 1989.

[356] A.M. Youssef and G. Gong. Hyper-bent functions. *Proceedings of EUROCRYPT 2001, Lecture Notes in Computer Science* 2045, Berlin, pp. 406-419, 2001.

[357] N. Y. Yu and G. Gong. Constructions of quadratic bent functions in polynomial forms. *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 3291-3299, 2006.

[358] M. Zhang. Maximum correlation analysis of nonlinear combining functions in stream ciphers. *Journal of Cryptology* 13 (3), pp. 301-313, 2000.

[359] X.-M. Zhang and Y. Zheng. GAC - the criterion for global avalanche characteristics of cryptographic functions. *Journal of Universal Computer Science*, 1(5), pp. 320-337, 1995.

[360] X.-M. Zhang and Y. Zheng. Auto-correlations and new bounds on the nonlinearity of Boolean functions. *Proceedings of EUROCRYPT'96, Lecture Notes in Computer Science* 1070, pp. 294-306, 1996.

[361] X.-M. Zhang and Y. Zheng. Characterizing the structures of cryptographic functions satisfying the propagation criterion for almost all vectors. *Designs, Codes and Cryptography*, 7(1), pp. 11-134, 1996.

[362] X.-M. Zhang and Y. Zheng. The nonhomomorphicity of Boolean functions. *Proceedings of SAC 1998, Lecture Notes in Computer Science* 1556, pp. 280-295, 1999.

[363] Y. Zheng and X. M. Zhang. Plateaued functions. *Proceedings of ICICS'99, Lecture Notes in Computer Science* 1726, pp. 284-300, 1999.

[364] Y. Zheng, X.-M. Zhang, and H. Imai. Restriction, terms and nonlinearity of Boolean functions. *Theoretical Computer Science*, 226(1-2), pp. 207-223, 1999.

[365] Y. Zheng and X.-M. Zhang. On relationships among avalanche, nonlinearity and correlation immunity. *Proceedings of Asiacrypt 2000, Lecture Notes in Computer Science* 1976, pp. 470-483, 2000.

[366] Y. Zheng and X.-M. Zhang. Improving upper bound on the nonlinearity of high order correlation immune functions. *Proceedings of Selected Areas in Cryptography 2000, Lecture Notes in Computer Science* 2012, pp. 262-274, 2001.

[367] N. Zierler and W. H. Mills. Products of linear recurring sequences. *Journal of Algebra* 27, pp. 147-157, 1973.

Index