Boolean Functions for Cryptography and Error Correcting Codes

Claude Carlet*

October 10, 2006

To appear soon as a chapter of the volume "Boolean Methods and Models", published by Cambridge University Press, Eds Yves Crama and Peter Hammer

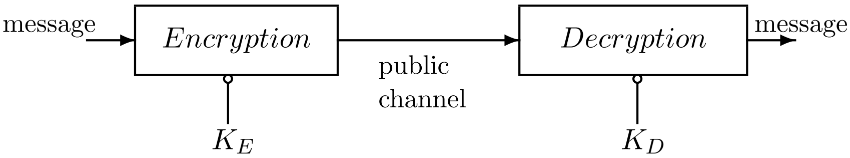
*University of Paris 8; also with INRIA, Projet CODES (address: BP 105 - 78153, Le Chesnay Cedex, FRANCE); e-mail: claude.carlet@inria.fr.

Contents

1 Introduction

A fundamental objective of *cryptography* is to enable two people 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 now 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* 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 and its receiver. If 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.

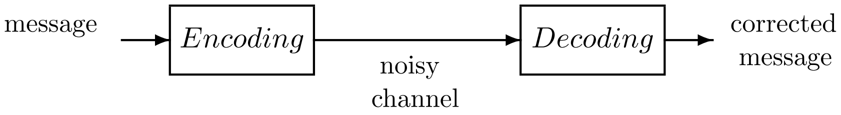


Public key cryptography is preferable to conventional cryptography, since it allows us to secretly communicate without having shared keys in a secret way: every person who wants to receive secret messages can keep secret a decryption key and publish an encryption key; if n persons want to secretly communicate pairwisely using a public key cryptosystem, they need n encryption keys and n decryption keys, when conventional cryptosystems will need at least $\binom{n}{2} = \frac{n(n-1)}{2}$ keys. But all known public key cryptosystems are much less efficient than conventional cryptosystems (they produce a

[1] According to principles already stated in 1883 by A. Kerckhoffs [159], who cited a still more ancient manuscript by R. du Carlet [41], only the secret keys must be kept secret – the confidentiality should not rely on the secrecy of the encryption method.

much lower data throughput, because they need much time to encrypt long messages) 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 and localized (and therefore corrected) by the receiver. This aim is achieved by using an encoding algorithm that transforms the information before sending it over the channel. In the case of block coding, the original message is treated as a list of binary words (vectors) of the same length – say k – that 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 and recover the correct message. The set of all possible codewords is called the *code*. Sending over the channel words of length n 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 length n, to the finite field with two elements $\mathbb{F}_2$ – denoted by $\mathcal{B}$ is some chapters of the present volume) play a role:
- every code whose length equals 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;

- in the case of conventional cryptography, the role of Boolean functions is even more important; cryptographic transformations (pseudo-random generators in stream ciphers, S-boxes in block ciphers) are designed by appropriate composition of nonlinear Boolean functions.

In both frameworks, n is rarely large, in practice, for the 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. However, this has changed with the algebraic attacks, see [88, 89, 111] and see below. The error correcting codes derived from n-variable Boolean functions have length 2^n; so, taking $n = 10$ already gives codes of length 1024.

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 some 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
$\|\mathcal{BF}_n\|$	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 on 6 variables, and about one hundred billions times the age of the universe to visit all those on 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 functions satisfying the desired conditions by picking up functions at random is also impossible for these values of n, since visiting a non-negligible part of all Boolean functions on 6 or more variables is not feasible. 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", 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 (coders say words) of length n will be viewed as a $\mathbb{F}_2$-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 a binary vector of length n. 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 $\{1, \dots, n\}$, 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 its support $\{x \in \mathbb{F}_2^n/\ f(x) \neq 0\}$. The *Hamming distance* $d_H(f, g)$ between two functions f and g is the size of the set $\{x \in \mathbb{F}_2^n/\ f(x) \neq g(x)\}$, that is, the *support of the function*. Thus it equals $w_H(f \oplus g)$.

Note. Some additions of bits will be considered in $\mathbb{Z}$ and denoted then by $+$, and some will be computed modulo 2 and denoted by $\oplus$. All the multiple sums computed in characteristic 0 will be denoted by $\sum_i$ and all the sums computed modulo 2 will be denoted by $\bigoplus_i$. For simplicity and because there will be no ambiguity, we shall denote by $+$ the addition of vectors (words) of $\mathbb{F}_2^n$ or of elements of $\mathbb{F}_{2^n}$.

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 $N = \{1, \dots, n\}$. 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 $\mathbb{F}_2[x_1, \dots, x_n]/(x_1^2 \oplus x_1, \dots, x_n^2 \oplus x_n)$. 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 $(1 \oplus x_1)(1 \oplus x_2)\, x_3 \oplus x_1(1 \oplus x_2)\, x_3 \oplus x_1x_2x_3 = x_1x_2x_3 \oplus x_2x_3 \oplus x_3$. $\diamond$

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 $P \in \mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 \oplus x_1, \ldots, x_n^2 \oplus x_n)$ 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 $\mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 \oplus x_1, \ldots, x_n^2 \oplus x_n)$, this correspondence is one to one[2]. 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 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

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

[2] 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.

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*). For every $u = (u_1, \ldots, u_n) \in \mathbb{F}_2^n$, the coefficient a_u of x^u in the ANF of f equals

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

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

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 (for instance), the table of the ANF equals the concatenation of those of the $(n-1)$-variable functions $f(x_1, \ldots, x_{n-1}, 0)$ and $f(x_1, \ldots, x_{n-1}, 0) \oplus f(x_1, \ldots, x_{n-1}, 1)$. We deduce the following recursive 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 be the restriction of f to $\mathbb{F}_2^{n-1} \times \{0\}$ and f_1 the restriction of f to $\mathbb{F}_2^{n-1} \times \{1\}$; the truth-table of f_0 (resp. f_1) corresponds to the upper (resp. lower) half of the table of f; 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 on one variable each), the global table gives the values of the ANF of f. The complexity of this algorithm is $\mathrm{O}(n2^n)$.

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), $d^\circ f$ 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 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. The algebraic degree being a linear (moreover, an affine) invariant and its value, for a function f, equalling the maximum dimension of the linear subspaces $\{x \in \mathbb{F}_2^n/\, supp(x) \subseteq I\}$ on which f takes value 1 an odd number of times, it equals the maximum dimension of all the linear (resp. affine) subspaces of $\mathbb{F}_2^n$ on which f takes value 1 an odd number of times.

Remarks.
1. Every atomic function (*i*.e. every function of weight 1) has algebraic degree n, since its ANF equals $(x_1 \oplus \epsilon_1)(x_2 \oplus \epsilon_2) \dots (x_n \oplus \epsilon_n)$, where $\epsilon_i \in \mathbb{F}_2, \forall i$. Thus, a Boolean function f has algebraic degree n if and only if, in its decomposition as a sum of atomic functions (see above), the number of these atomic functions is odd, that is, if and only if $w_H(f)$ is odd. This property will be useful at Section 3.
2. If we know that the algebraic degree of an n-variable Boolean function f is upper bounded 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 of the Boolean function f (of algebraic degree at most $d < n$) 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[3]. This happens with the set E of all words of Hamming weights smaller than or equal to d, since Relation (3) gives the values of a_I, where $I \in \mathcal{P}(N)$ and $|I| \leq d$. Notice that Relation (2) permits then to recover the value of $f(x)$ for every $x \in \mathbb{F}_2^n$, from the values taken by f at all words of Hamming weights smaller than or equal to d.

The same property happens if we replace "Boolean" by "pseudo-Boolean" (that is, real-valued) and if we consider the numerical degree (see below) instead of the the algebraic degree, *cf.* [257]. ◇

The simplest functions, from the viewpoint of the ANF, are those Boolean functions of s 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 $\mathbb{F}_2^n$, 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 permit to define 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 representations 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 [178]). 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 a $\mathbb{F}_2$-basis $(\alpha_1, \ldots, \alpha_n)$ 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 also below) that every mapping from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$ admits a

[3] Note that taking f null leads to determining the so-called annihilators of the indicator of E; this is the core analysis of Boolean functions from the viewpoint of algebraic attacks, see Subsection 4.1.

(unique) representation as a polynomial over $\mathbb{F}_{2^n}$ on 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 itself and admits therefore such a unique representation.

2. For every $u, v \in \mathbb{F}_{2^n}$ we have $(u+v)^2 = u^2 + v^2$ and $u^{2^n} = u$ (*i.*e. $u^{2^n-1} = 1$ if $u \neq 0$). Consequently, 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 $(tr^n(u))^2 = tr^n(u)$; it is therefore valued in $\mathbb{F}_2$. When there will be no ambiguity, we shall write tr instead of tr^n. This function is called the *trace function* from $\mathbb{F}_{2^n}$ to its prime field $\mathbb{F}_2$. The function $(u,v) \mapsto tr(u\,v)$ is an inner product in $\mathbb{F}_{2^n}$. Every Boolean function can be written in the form $f(x) = tr(F(x))$ where F is a mapping from $\mathbb{F}_{2^n}$ into $\mathbb{F}_{2^n}$ (an example of such mapping F is defined by $F(x) = 0$ if $f(x) = 0$ and $F(x) = \lambda$ where $tr(\lambda) = 1$ if $f(x) = 1$). Thus, every Boolean function can be also represented in the form $tr\left(\sum_{i=0}^{2^n-1} \beta_i\, x^i\right)$, where $\beta_i \in \mathbb{F}_{2^n}$. Such a representation is not unique. Now, thanks to the fact that $tr(u^2) = tr(u)$ 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 $2^n - 1$. Trace representations and the algebraic normal form are closely related. It is shown in the chapter "Vectorial Boolean Functions for Cryptography" how a representation can be obtained from the other.

3. We come back now to the representation introduced above in 1. 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. Assuming that $f(0) = 0$ (otherwise, we can apply the method to the function $f(x) \oplus f(0)$), and denoting by α a primitive element of the field $\mathbb{F}_{2^n}$ (that is, an element such that $\mathbb{F}_{2^n} = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^{2^n-2}\}$), the *Mattson-Solomon polynomial* of the vector $(f(1), f(\alpha), f(\alpha^2), \ldots, f(\alpha^{2^n-2}))$ is the polynomial

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

with:

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

Note that the Mattson Solomon transformation is a discrete Fourier transform. We have then $A(\alpha^i) = \sum_{j=1}^{2^n-1} A_j \alpha^{-ij} = \sum_{j=1}^{2^n-1} \sum_{k=0}^{2^n-2} f(\alpha^k)\alpha^{(k-i)j} = f(\alpha^i)$ for every i (since $\sum_{j=1}^{2^n-1} \alpha^{(k-i)j}$ equals 0 if $k \neq i$), and A is therefore the trace representation of f seen in 1. Note that $A_{2j} = A_j^2$. This allows representing $A(x)$ in the form $\sum_{k\in\Gamma(n)} tr^{n_k}(A_k x^k) + A_{2^n-1}x^{2^n-1}$, where $\Gamma(n)$ is 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), and where n_k is the size of the cyclotomic class containing k. Note that, for every $k \in \Gamma(n)$ and every $x \in \mathbb{F}_{2^n}$, we have $A_k \in F_{2^{n_k}}$ (since $A_k^{2^{n_k}} = A_k$) and $x^k \in F_{2^{n_k}}$ as well.

The algebraic degree of functions written in the trace representations can be determined. For instance, let i be a positive integer. Then i [mod 2^n-1] can be written in the form $\sum_{j\in A} 2^j$, where $A \subseteq \{0, 1, \ldots, n-1\}$. The size of A (say r) is often called the *2-weight* of i [mod $2^n - 1$]. Let $a \in \mathbb{F}_{2^n}$ and $f(x) = tr(ax^i)$. Then, as shown in [42], if f is not the null function, it has algebraic degree r. Indeed, it is a simple matter to show that f has algebraic degree at most r; to show that it has degree exactly r, 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_i's; $\phi(x_1, ..., x_r)$ equals the sum, for all bijective mappings σ from $\{1, \ldots, r\}$ onto A of $tr(a \prod_{i=1}^r x_i^{2^{\sigma(i)}})$; proving that f has degree r is equivalent to proving that ϕ is not null; and it is a simple matter to prove that, if ϕ is null, then f is null.

The representation over the reals has recently proved itself to be useful for characterizing several cryptographic criteria [54, 73, 74] (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 $\mathbb{R}[x_1, \ldots, x_n]/(x_1^2 - x_1, \ldots, x_n^2 - x_n)$ (or of $\mathbb{Z}[x_1, \ldots, x_n]/(x_1^2 - x_1, \ldots, x_n^2 - x_n)$ 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 $\mathbb{R}[x_1, \ldots, x_n]/(x_1^2 - x_1, \ldots, x_n^2 - x_n)$ 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*. It is shown in [211] that, if a Boolean function f has no ineffective variable, 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 [74]; see also [144]) 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:

$$f(x) = \sum_{I\subseteq supp(x)} \lambda_I. \tag{4}$$

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

Proposition 2 *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:*

$$\forall I \in \mathcal{P}(N),\ \lambda_I = (-1)^{|I|} \sum_{x\in\mathbb{F}_2^n \,|\, supp(x)\subseteq I} (-1)^{w_H(x)} f(x). \tag{5}$$

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. ◇

Notice 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

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

Expanding this last equality gives the NNF of $f(x)$ and we have [73]:

$$\lambda_I = \sum_{k=1}^{2^n} (-2)^{k-1} \sum_{\substack{\{I_1,\dots,I_k\}\,|\\ I_1\cup\ldots\cup I_k = I}} a_{I_1}\ldots a_{I_k}. \qquad (6)$$

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 $\mathbb{R}\,[x_1,\dots,x_n]/(x_1^2 - x_1,\dots,x_n^2 - x_n)$), or equivalently, denoting $supp(x)$ by I:

$$\forall I \in \mathcal{P}(N),\ \left(\sum_{J\subseteq I} \lambda_J\right)^2 = \sum_{J\subseteq I} \lambda_J. \qquad (7)$$

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 $\sum_{J\subseteq I\,|\,J\neq I} \lambda_J$ by μ. Relation (7) 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 $P(x) = 1$.

$\diamond$

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 one equation only.

Proposition 3 *Any polynomial $P \in \mathbb{R}[x_1, \ldots, x_n]/(x_1^2 - x_1, \ldots, x_n^2 - x_n)$ 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 (4) and (5). 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 $\sum_{x\in\mathbb{F}_2^n} P^2(x) = \sum_{x\in\mathbb{F}_2^n} P(x)$. $\diamond$

The translation of this characterization in terms of the coefficients of P is given in Relation (29) 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 therefore 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

$$\widehat{\varphi}(u) = \sum_{x\in\mathbb{F}_2^n} \varphi(x) \, (-1)^{x\cdot u} \tag{8}$$

(we recall that $x \cdot u$ denotes the usual inner product).

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

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

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 (for instance), the table of $\widehat{\varphi}$ equals the concatenation of those of the discrete Fourier transforms of the $(n-1)$-variable functions $\psi_0(x)=\varphi(x_1,\dots,x_{n-1},0)+\varphi(x_1,\dots,x_{n-1},1)$ and $\psi_1(x)=\varphi(x_1,\dots,x_{n-1},0)-\varphi(x_1,\dots,x_{n-1},1)$. We deduce the following recursive 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 – say – in lexicographic order;

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 $\mathbb{F}_2^{n-1}\times\{1\}$; the table of values of φ_0 (resp. φ_1) corresponds to the upper (resp. lower) half of the table of φ; 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 on one variable each), the global table gives the values of $\widehat{\varphi}$. The complexity of this algorithm is $O(n2^n)$.

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*[4]) instead of f itself.

[4] 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 will be easily done, we also 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	$f(x)$	f_χ(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 have
$$\widehat{\mathrm{f}_\chi}(u) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x) \oplus x \cdot u}.$$
We shall call *Walsh transform*[5] 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
$$\widehat{\mathrm{f}_\chi} = 2^n\, \delta_0 - 2\widehat{f} \tag{9}$$
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 5 for a proof of the relation $\widehat{1} = 2^n\, \delta_0$. Relation (9) gives conversely

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

$\widehat{f} = 2^{n-1}\delta_0 - \frac{\widehat{\mathrm{f}_\chi}}{2}$ and in particular:

$$w_H(f) = 2^{n-1} - \frac{\widehat{\mathrm{f}_\chi}(0)}{2}. \tag{10}$$

Relation (10) applied to $f \oplus \ell_a$, where $\ell_a(x) = a_1x_1 \oplus \cdots \oplus a_nx_n = a \cdot x$, gives:

$$d_H(f, \ell_a) = w_H(f \oplus \ell_a) = 2^{n-1} - \frac{\widehat{\mathrm{f}_\chi}(a)}{2}. \tag{11}$$

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

$$\mathcal{F}(f) = \widehat{\mathrm{f}_\chi}(0) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)}. \tag{12}$$

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[6] 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[7]. Thus, $\sum_{x \in E}(-1)^{\ell(x)}$ being the sum of 1's and -1's in equal numbers, it is null. ◇

Proposition 4 *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 $(-1)^{b \cdot (a+u)}\, \widehat{\varphi}(a + u)$. ◇

[6] Lemma 1 allows proving a nice property on the Walsh transform of composed vectorial functions, see the remark in the introduction of Subsection 2.1 in the chapter "Vectorial Boolean Functions for Cryptography".

[7] 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 $\ell^{-1}(1)$.

Proposition 5 *Let E be any vector subspace of $\mathbb{F}_2^n$. Denote by 1_E its* indicator *(also called characteristic function), defined by $1_E(u) = 1$ if $u \in E$ and $1_E(u) = 0$ otherwise. Then:*

$$\widehat{1_E} = |E| \, 1_{E^\perp}, \tag{13}$$

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 $\widehat{1} = 2^n \, \delta_0$.

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 5 the *Poisson summation formula*, which has been used to prove many cryptographic properties in [177], [184], [45] and later in [32, 33], 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:*

$$\sum_{u \in a+E} (-1)^{b \cdot u} \, \widehat{\varphi}(u) = |E| \, (-1)^{a \cdot b} \sum_{x \in b+E^\perp} (-1)^{a \cdot x} \, \varphi(x). \tag{14}$$

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 5:

$$\sum_{u \in E} \widehat{\varphi}(u) = |E| \sum_{x \in E^\perp} \varphi(x). \tag{15}$$

We apply this last equality to the function $(-1)^{a \cdot x} \, \varphi(x+b)$, whose Fourier transform is $(-1)^{b \cdot (a+u)} \, \widehat{\varphi}(a+u)$, according to Proposition 4. 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 (14). ◇

Relation (14) with $a = 0$ and $E = \mathbb{F}_2^n$ gives:

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

$$\widehat{\widehat{\varphi}} = 2^n \, \varphi. \tag{16}$$

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 permits 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 5, the Fourier transform of any constant function takes null value at every nonzero vector; according to Corollary 2, this is a necessary and sufficient condition. 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, into this same set, endowed with the usual (Hadamard) product of functions. 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 6 *Let φ and ψ be any pseudo-Boolean functions on $\mathbb{F}_2^n$. We have:*

$$\widehat{\varphi \otimes \psi} = \widehat{\varphi} \times \widehat{\psi}. \tag{17}$$

Consequently:

$$\widehat{\varphi} \otimes \widehat{\psi} = 2^n \, \widehat{\varphi \times \psi}. \tag{18}$$

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 (18). ◇

Relation (18) applied at 0 gives

$$\widehat{\varphi}\otimes\widehat{\psi}(0) = 2^n\,\widehat{\varphi\times\psi}(0) = 2^n \sum_{x\in\mathbb{F}_2^n} \varphi(x)\psi(x) = 2^n\,\varphi\otimes\psi(0). \tag{19}$$

Taking $\psi=\varphi$ in (19), 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:

$$\sum_{u\in\mathbb{F}_2^n} \widehat{\varphi}^{\,2}(u) = 2^{2n}. \tag{20}$$

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 (17) 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 with respect to the direction b the Boolean function $D_bf(x) = f(x)\oplus f(x+b)$.*

For instance, the derivative with respect to the vector $(0,\dots,0,1)$ of a function of the form $g(x_1,\dots,x_{n-1})\oplus x_n\,h(x_1,\dots,x_{n-1})$ equals $h(x_1,\dots,x_{n-1})$. Relation (17) applied with $\psi=\varphi=\mathrm{f}_{\chi}$ implies the so-called *Wiener-Khintchine Theorem*:

$$\widehat{\mathrm{f}_{\chi}\otimes\mathrm{f}_{\chi}} = \widehat{\mathrm{f}_{\chi}}^{\,2}. \tag{21}$$

We have $(\mathrm{f}_{\chi}\otimes\mathrm{f}_{\chi})(b) = \sum_{x\in\mathbb{F}_2^n}(-1)^{D_bf(x)} = \mathcal{F}(D_bf)$ (the notation $\mathcal{F}$ was defined at Relation (12)). Thus Relation (21) 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 [44]):

$$\forall u \in \mathbb{F}_2^n,\ \sum_{b\in\mathbb{F}_2^n} \mathcal{F}(D_b f)(-1)^{u\cdot b} = \widehat{\mathrm{f}_\chi}^{\,2}(u). \tag{22}$$

Applied at vector 0, this gives

$$\sum_{b\in\mathbb{F}_2^n} \mathcal{F}(D_b f) = \mathcal{F}^2(f). \tag{23}$$

Corollary 1 and Relation (22) imply that, for every vector subspace E of $\mathbb{F}_2^n$ and every vectors a and b (cf. [33]):

$$\sum_{u\in a+E} (-1)^{b\cdot u}\, \widehat{\mathrm{f}_\chi}^{\,2}(u) = |E|(-1)^{a\cdot b} \sum_{e\in b+E^\perp} (-1)^{a\cdot e}\mathcal{F}(D_e f)\ . \tag{24}$$

Another interesting relation has been also shown in [33] (see also [180]):

Proposition 7 *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*

$$\sum_{u\in E^\perp} \widehat{\mathrm{f}_\chi}^{\,2}(u) = |E^\perp| \sum_{a\in E'} \mathcal{F}^2(h_a)\ . \tag{25}$$

Proof. Every element of $\mathbb{F}_2^n$ can be written in a unique way in 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 (24) 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 (23) 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 8 *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 $L' : \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto M' \times \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Then*

$$\widehat{\varphi \circ L} = \widehat{\varphi} \circ L'. \tag{26}$$

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 [171] (see also [45]):

Proposition 9 *Let f be an n-variable Boolean function, and let $1 \leq k \leq n$. Assume that its Walsh transform takes values divisible by 2^k (*i.e.*, according to Relation (9), that its Fourier transform takes values divisible by 2^{k-1}, or equivalently, according to Relation (11), 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. Relation (15) 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 $\{u \in \mathbb{F}_2^n / \ \forall i \notin I, \ u_i = 0\}$. The restriction of f to $E^\perp$, viewed as a function on $\mathbb{F}_2^d$, has an ANF of degree d (because all the monomials different from $\prod_{i \in I} x_i$ in the ANF of f give monomials of degrees strictly less than d when we set the coordinates x_i, $i \notin I$). Thus, any such restriction has an odd weight (see Remark 1 of Subsection 2.1), and $\sum_{x\in E^\perp} \mathrm{f}_\chi(x)$ 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. $\diamond$

The converse of Proposition 9 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. 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 Pareseval'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 which Walsh transforms take values $\pm 2^{n/2}$ (n even ≥ 4) and $\pm 2^{(n+1)/2}$ (n odd ≥ 5). Besides, it is possible to show that the non-affine quadratic functions which Walsh transform values are divisible by 2^{n-1} are those sums of an indicator of a flat (*i.e.* an affine space) of co-dimension 2 and of an affine function. It is then an easy task to deduce that the converse of Proposition 9 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, which Walsh transform value at 0 equals 2^2, that is, which 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 $k = n - (l - 2) + 1 = n - l + 3$ (the range of k being $3 \leq k \leq n - 2$).

Determining those Boolean functions which Walsh transform is divisible by 2^k seems to be an open problem for $3 \leq k \leq n-2$ (partial results are given in [63]). This problem is interesting because of the result on resilient functions recalled in Proposition 32.

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 (14).

Characterizing the Fourier transforms of integer-valued pseudo-Boolean functions and of Boolean functions According to Relation (16), 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 $\widehat{\varphi \otimes \varphi} = 2^{2n}$, that is $\varphi \otimes \varphi = 2^{2n}\ \delta_0$, according to Relation (17) applied with $\psi = \varphi$. But these characterizations are not 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 $b = (1,\ldots,1)$, we have $\varphi(x+b) = \displaystyle\sum_{I\in\mathcal{P}(N)/\ supp(x)\cap I=\emptyset} \lambda_I$ (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 4 (with $a = 0$) and 5, we deduce:

$$\widehat{\varphi}(u) = (-1)^{w_H(u)} \sum_{I\in\mathcal{P}(N)\,|\,supp(u)\subseteq I} 2^{n-|I|}\lambda_I. \tag{27}$$

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

$$\lambda_I = 2^{-n}(-2)^{|I|} \sum_{u\in\mathbb{F}_2^n\,|\,I\subseteq supp(u)} \widehat{\varphi}(u). \tag{28}$$

Note that if φ has numerical degree D, then, according to Relation (27), we have $\widehat{\varphi}(u) = 0$ for every vector u of weight strictly greater than D.

Applying Relation (27) 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 3 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

$$\sum_{I\in\mathcal{P}(N)} 2^{n-|I|} \sum_{J,J'\in\mathcal{P}(N)\,|\,I=J\cup J'} \lambda_J \, \lambda_{J'} = \sum_{I\in\mathcal{P}(N)} 2^{n-|I|}\lambda_I. \qquad (29)$$

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 [226]);
- a mapping defined by a formula similar to Relation (5), but in which $supp(x) \subseteq I$ is replaced by $I \subseteq supp(x)$; see [122];
- the inverse of this mapping. $\diamond$

2.2.2 Fourier transform and graph theory

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 [14]), 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 Relation (14): let I be any

subset of $N = \{1, \ldots, n\}$; 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 we apply this property to a Boolean function f and if we choose for g a restriction of odd weight (whose Fourier transform takes therefore nonzero values, only), we deduce (see [14]) that $N_{\widehat{f}} \geq 2^d$, where d is the algebraic degree of f (choose a monomial x^I of degree d in the ANF of f). Notice that $N_{\widehat{f}}$ equals 2^d if and only if at most one element (that is, exactly one) satisfying $\widehat{f}(u) \neq 0$ 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 $N_{\widehat{f}}$ is also upper bounded by $\sum_{i=0}^{D} \binom{n}{i}$, where D is the numerical degree of f. This is a direct consequence of Relation (27) and of the observation which follows it.

The graphic viewpoint also gives insight on the Boolean functions whose Fourier spectra have at most three values (see [14]).

A hypergraph can also be related to the ANF of a Boolean function f. A related (weak) upper bound on the nonlinearity of f has been pointed out in [270].

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) permits 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, in principle[8]. Thus, the problem of generating a good code consists in finding a set C of binary words of the same length which *minimum distance* $\min_{a,b\in C} d_H(a,b)$ (where $d_H(a,b) = |\{i/\, a_i \neq b_i\}|$) is high[9].

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. 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 lines. The code equals the set of all the vectors of the form $u \times G$, where u ranges over $\mathbb{F}_2^k$.

As explained in the introduction, every code which length equals 2^n, for some positive integer n, can be interpreted as a set of Boolean functions. This viewpoint has led to the Reed-Muller codes.

3.1 Reed-Muller codes

The existence of *Reed-Muller codes* comes from the following observation:

Proposition 10 *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 function of algebraic degree $d \leq r$ has weight at least 2^{n-r} (see above what is observed about linear codes). This can be proved by a double induction over r and n (see [187]), 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

[8] 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.

[9] 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 [187, 222])

on $\mathbb{F}_2^d$, has an ANF of degree d, because all the monomials, different from $\prod_{i\in I} x_i$ in the ANF of f, give monomials of degrees strictly less than d when we keep constant the coordinates x_i, $i \notin I$. Thus any such restriction has an odd (and hence a nonzero) weight (see Remark 1 of 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. $\diamond$

The functions of Hamming weight 2^{n-r} and degree r have been characterized, see a proof in [187]. We give below an original proof which brings a little more insight on the reasons of this characterization.

Proposition 11 *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 degree r and Hamming weight 2^{n-r}. 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 $J = \{1,\dots,n\} \setminus I$. 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 Proposition 10, 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 $\alpha, \beta, \gamma \in F_2^J$, we have $a_{\alpha+\beta+\gamma} = a_\alpha + a_\beta + a_\gamma$ (this will complete the proof of the proposition). Proving this is equivalent to proving that $f_{\alpha+\beta+\gamma} \oplus f_\alpha \oplus f_\beta \oplus f_\gamma$ has degree at most $r-2$. But more generally, for every k-dimensional flat A of 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). $\diamond$

Remark.

1. The proof of Proposition 10 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 [270]). Indeed, the Möbius transform of f° is f.
2. The d-dimensional subspace E of equations $x_i = 0$, $i \notin I$, in the proof of Proposition 10, 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 (since the restriction of f to any coset of E has odd weight). Similarly as above, it can be proved, see [270], 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 the truth-tables of these functions). Denoted by $R(r,n)$, it is a $\mathbb{F}_2$-vectorspace of dimension $1+n+\binom{n}{2}+\cdots+\binom{n}{r}$ (since this is the number of monomials of degrees at most r) 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. $\diamond$

$R(r,n)$ is a linear code, *i*.e. a $\mathbb{F}_2$-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 line corresponds to the constant function 1 and the other lines correspond to the coordinate functions $x_1, \ldots, x_4$)[10].
For a given linear code C of length m and dimension k having a generator matrix G, a possible encoding algorithm is the mapping $u \in \mathbb{F}_2^k \mapsto u \times G \in$

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

$\mathbb{F}_2^m$. Thus, the generator matrix permits to generate the codewords, but it is not well suited for checking if a received word of length m is a codeword or not. A characterization of the codewords is obtained thanks to the generator matrix H of the dual $C^\perp = \{x \in \mathbb{F}_2^m / \forall y \in C,\ x \cdot y = \bigoplus_{i=1}^m x_i\, y_i = 0\}$ (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.

The case of Reed-Muller codes is simple:

Proposition 12 *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 at Subsection 2.1 that the functions in $\mathcal{BF}_n$ whose weights are even 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 r, the product function fg has algebraic degree at most $n-1$. This is clearly equivalent to the fact that f has algebraic degree at most $n-r-1$.$\diamond$

The Reed-Muller codes are invariant under the action of 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 [133, 135, 20, 188, 249, 250].

The open problem of the weight distribution of Reed-Muller codes, 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 $R(r,n)$? 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 5.1. 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 m; consider the polynomial $W_C(X,Y) = \sum_{i=0}^m A_i X^{m-i} Y^i$ where A_i is the number of codewords of weight i. This polynomial is called the *weight enumerator* of C and describes[11] the *weight distribution* $(A_i)_{0 \leq i \leq m}$ of C. Then (see [187, 222])

$$W_C(X+Y, X-Y) = |C|\, W_C(X,Y). \tag{30}$$

We give only a sketch of proof of this *MacWilliams' identity*: we observe first that $W_C(X,Y) = \sum_{x \in C} \prod_{i=1}^m X^{1-x_i} Y^{x_i}$; we deduce $W_C(X+Y, X-Y) = \sum_{x \in C} \prod_{i=1}^m (X + (-1)^{x_i} Y)$; applying a classical method of expansion, we derive $W_C(X+Y, X-Y) = \sum_{x \in C} \sum_{b \in \mathbb{F}_2^m} \prod_{i=1}^m \left(X^{1-b_i} ((-1)^{x_i} Y)^{b_i}\right)$

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

(for $b_i = 0$, we choose X in the factor $X + (-1)^{x_i}Y$; and for $b_i = 1$, we choose $(-1)^{x_i}Y$; all the different possible choices are taken into account by considering all binary words b of length m). We obtain then $W_C(X+Y, X-Y) = \sum_{b\in\mathbb{F}_2^m}\left(X^{m-w_H(b)}Y^{w_H(b)}\sum_{x\in C}(-1)^{b\cdot x}\right)$ and we conclude by using Relation (13) with $E = C$.

The MacWilliams identity permits, 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 $(X+Y)^{m-i}(X-Y)^i$ and to simplify the obtained expression for $W_C(X+Y, X-Y)$; this is not possible for all values 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 [13], and for $n = 2r$, because the code is then self-dual, see [187, 222]). *McEliece's theorem* [200] (or *Ax's theorem* [8]) shows that the weights (and thus the distances) in $R(r, n)$ are all divisible by $2^{\lceil \frac{n}{r}\rceil -1} = 2^{\lfloor \frac{n-1}{r}\rfloor}$ (this can also be shown by using the properties of the NNF, see [73]). Moreover, if f has degree d and g has degree $d' \leq d$, then $d_H(f,g) \equiv w_H(f)\left[\bmod 2^{\left\lceil \frac{n-d'}{d}\right\rceil}\right]$ [157]. In [26], A. Canteaut gives further properties of the weights in $f \oplus R(1, n)$. Kasami and Tokura [155] 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 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: $D_C(X,Y) = \frac{1}{|C|}\sum_{i=0}^{m} B_i X^{m-i}Y^i$, 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 $D_C(X,Y) = \frac{1}{|C|}\sum_{(x,y)\in C^2}\prod_{i=1}^{m} X^{1-(x_i\oplus y_i)}Y^{x_i\oplus y_i}$; we deduce that the polynomial $D_C(X+Y, X-Y)$ equals $\frac{1}{|C|}\sum_{(x,y)\in C^2}\prod_{i=1}^{m}(X + (-1)^{x_i\oplus y_i}Y)$. Expanding these products, we obtain $\frac{1}{|C|}\sum_{(x,y)\in C^2}\sum_{b\in\mathbb{F}_2^m}\prod_{i=1}^{m}\left(X^{1-b_i}((-1)^{x_i\oplus y_i}Y)^{b_i}\right)$, that is $D_C(X+Y, X-Y) = \frac{1}{|C|}\sum_{b\in\mathbb{F}_2^m} X^{m-w_H(b)}Y^{w_H(b)}\left(\sum_{x\in C}(-1)^{b\cdot x}\right)^2$.

The minimum nonzero exponent of Y 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 [99, 100]). This property will be useful at Subsection 4.1.

It is shown in [43] (see also the remark of Subsection 5.1) 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 takes value $\widehat{\mathrm{g}_\chi}(0) = 2^m\,\widehat{\mathrm{f}_\chi}(0)$ at 0 (the null vector). This means that the weight of f is related to the weight of g in a simple way. This shows that the distances in $R(3,n)$ can be very diverse, contrary to those in $R(2,n)$. $\diamond$

4 Boolean functions and cryptography

Stream ciphers are based on the so-called *Vernam cipher* (see Figure 1) in which the plaintext (say 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 learned this at their own expenses!).

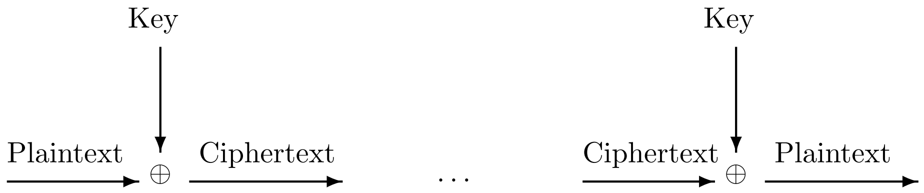


Figure 1: Vernam cipher

The Vernam cipher, which is the only known cipher offering unconditional security (see [244]) 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 (shared by the sender and the recipient) is chosen for producing long *pseudo-random sequences* from short random secret keys (only the latter are actually shared, together with the method). The pseudo-random sequence is used in the place of the key in a Vernam cipher. Stream ciphers, because they operate on data units as small as a bit or a few bits, are suitable for fast 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 (see Figure 2), at every clock-cycle, the bits $s_{n-1}, \ldots, s_{n-L}$ contained in the flip-flops of the LFSR move to the right. The left-most flip-flop is feeded with the linear combination $\bigoplus_{i=1}^{L} c_i s_{n-i}$. Thus, such an LFSR outputs a recurrent sequence satisfying the relation

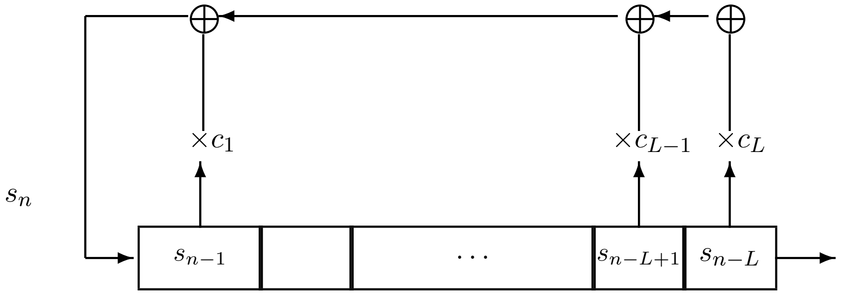


Figure 2: LFSR

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

Such sequence is always ultimately periodic[12] (if $c_L = 1$, then it is periodic; we shall assume that $c_L = 1$ in the sequel) with period at most $2^L - 1$. The short secret key gives then the initialization $s_0, \ldots, s_{L-1}$ of the LFSR and the values of the feedback coefficients c_i (these must be kept secret; otherwise, the observation of L consecutive bits of the key would permit to recover all the subsequent sequence).

But these LFSRs are cryptographically weak because of *Berlekamp-Massey algorithm* [197]: let L be the length of the minimum LFSR pro-

[12] Conversely, every ultimately periodic sequence can be generated by at least one LFSR.

ducing the same sequence (this length, called the *linear complexity* of the sequence, is assumed to be unknown from the attacker), then if we know at least $2L$ consecutive bits, Berlekamp-Massey algorithm recovers the values of L and of the c_i's and the initialization of the sequence. So, the attacker only needs in practice to know about 20 consecutive bits. The modern way of avoiding this attack is by using Boolean functions, the most usual way being with *Combining Boolean functions* (see Figure 3).

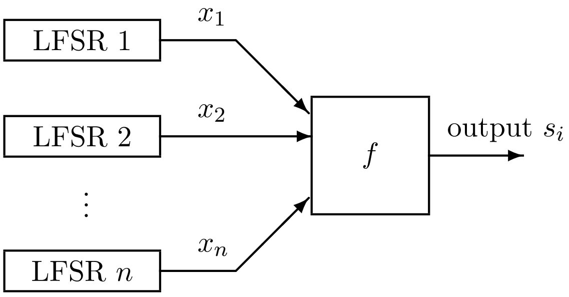


Figure 3: Combining function

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). So, this sequence is also recurrent and can therefore be produced by a single LFSR. However, 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 f such that the linear complexity and the period are large enough is not sufficient. The combining function should also not leak information about the individual LFSRs.

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 120 bit long secret key, this permits to use n LFSRs of lengths $L_1, \ldots, L_n$ such that $L_1 + \cdots + L_n = 120$.

Other ways of using Boolean functions exist. A filtered LFSR does not output the bit contained in the right-most flip-flop, but outputs $f(x_1, \ldots, x_n)$ where f is some n-variable Boolean function, called a *filtering function* and where $x_1, \ldots, x_n$ are the bits contained in some flip-flops of the LFSR, see Figure 4.

Such system is equivalent to the combining system[13]. A *Feedback Shift*

[13] However, the attacks, even when they apply to both systems – the original one and

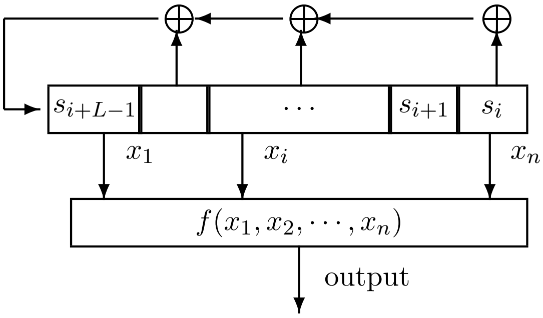


Figure 4: Filtering function

Register has the same structure as an LFSR, but the left-most flip-flop is feeded with $f(x_1, \ldots, x_n)$ where $n \leq L$, f is some n-variable Boolean function and where $x_1, \ldots, x_n$ are bits contained in the flip-flops of the FSR. The linear complexity of the produced sequence can then be potentially near 2^L. But there does not exist much published work on this subject (see [147] for general FSRs and [80] for FSRs with quadratic function f) and the linear complexity is difficult to study in general.

Boolean functions also play an important role in block ciphers. A first observation is that every block cipher admits as input a binary vector $(x_1, \ldots, x_n)$ (a block of plaintext) and outputs a binary vector $(y_1, \ldots, y_m)$; the coordinates $y_1, \ldots, y_m$ are the outputs to Boolean functions (depending on the key) whose common input is $(x_1, \ldots, x_n)$. see Figure 5.
But the number n of variables of these Boolean functions being large (most often, more than a hundred), these functions could 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 (at most 16).

We give in Figures 6 and 7 a description of these rounds for the DES and for 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 is achieved by the S-box,

the equivalent one – may not work quite similarly. Consequently, the criteria that the involved Boolean functions must satisfy because of these attacks may be different for the original system and for the equivalent one.

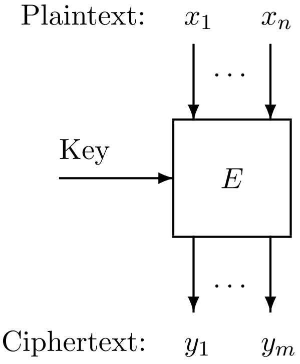


Figure 5: Block cipher

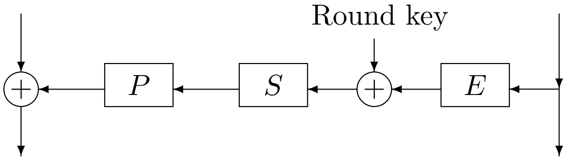


Figure 6: A DES round

which is a nonlinear transformation of a binary string of 48 bits[14] 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").

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

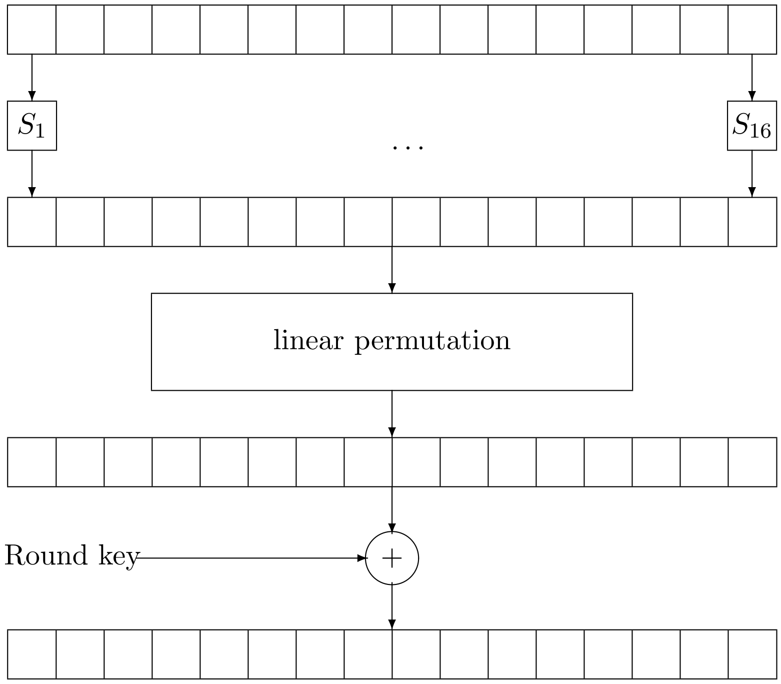


Figure 7: An AES round

4.1 Cryptographic criteria for Boolean functions

The design of conventional cryptographic systems relies on two fundamental principles introduced by Shannon [244]: *confusion* and *diffusion*. Confusion aims at concealing any algebraic structure in the system. It is closely related to the complexity[15] 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 [204, 223, 245] that the implemented cryptographic functions must satisfy. More precisely, the resistance of the cryptosystems to the known attacks can 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 con-

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

sider 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).

The algebraic degree: cryptographic functions must have high algebraic degrees. Indeed, all cryptosystems using Boolean functions for confusion (combining 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 $L_1, \dots, L_n$ 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 $N = \{1, \dots, n\}$; then (see [234]) the sequence produced by f can be obtained by a single LFSR of length

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

The algebraic degree of f (*i*.e. the largest size of I such that $a_I = 1$) has to be high so that L can have high value (the number of those nonzero coefficients a_I, in the ANF of f, such that I has large size, also plays a role, but a less important one). In the case of block ciphers, using Boolean functions of low degrees makes the higher differential attack [162, 169] 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 permit to achieve good characteristics (nonlinearity, resiliency, ...).
We have seen at Subsection 2.1 that the algebraic degree is an affine invariant.

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 (20) applied to the sign function f_χ and of Relation (11), 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 permits in various situations (block ciphers, stream ciphers) to build attacks on this system (see [124, 199]). The *nonlinearity* of f is the minimum Hamming distance between f and affine functions. It must be high (in a sense that will be clarified below). The nonlinearity criterion can be quantified through the Walsh transform: let $\ell_a(x) = a_1x_1 \oplus \cdots \oplus a_nx_n = a \cdot x$ be any linear function; according to Relation (11), 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:

$$\mathcal{NL}(f) = 2^{n-1} - \frac{1}{2} \max_{a \in \mathbb{F}_2^n} |\widehat{\mathrm{f}_\chi}(a)|. \tag{31}$$

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. Parseval's Relation (20) 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 (and we shall use below the property that 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

$$\mathcal{NL}(f) \leq 2^{n-1} - 2^{n/2-1}. \tag{32}$$

This bound, valid for every Boolean function, 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). It 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 [151]. D. Olejár and M. Stanek [214] 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 the number of functions whose nonlinearities are upper bounded by a given number, see [57]).

Equality occurs in (32) if and only if $|\widehat{\mathrm{f}_\chi}(a)|$ 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). They have the property that, for every even w, the sum $\sum_{a\in\mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^{\,w}(a)$ is minimum. Note that such sums (for even or odd w) play a role with respect to fast correlation attacks (see below for more on correlation attacks and see [38, 31] for the fact that when these sums have small magnitude, for low values of w, this contributes to a good resistance to fast correlation attacks).

For n odd, Inequality (32) 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 by quadratic functions, see Subsection 5.1) and $2^{n-1} - 2^{n/2-1}$. It has been shown in [127, 209] that it equals $2^{n-1} - 2^{\frac{n-1}{2}}$ when $n = 1,\ 3,\ 5,\ 7$, and in [220], by Patterson and Wiedemann[16] , that it is greater than $2^{n-1} - 2^{\frac{n-1}{2}}$ if $n \geq 15$ (a review on what is known on the best nonlinearities of functions on odd numbers of variables is given in [113]).

The maximum Hamming distance between a general Boolean function and $R(1,n)$, *i*.e. the maximum nonlinearity of all Boolean functions, is the *covering radius* of $R(1,n)$ (*i*.e. the minimum integer t such that every binary word of length 2^n lies at Hamming distance at most t from at least one codeword). The covering radius of a code is an important parameter [87], which can be used for analyzing and improving the decoding algorithms devoted to this code. 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 $f \oplus R(1,n)$, $f \in \mathcal{F}$, equals the minimum nonlinearity of the functions $f \oplus g$, where f and g are distinct and range over $\mathcal{F}$. This observation permits to construct 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[17] (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 func-

[16] It has been later proved (see [242, 107] and [195, 163]) 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 $n \geq 15$.

[17] 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 [9, 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 at Section 6 that highly nonlinear functions can be built from bent functions.

tions exist. The maximum nonlinearity of balanced functions is unknown for any $n \geq 8$.
Two relations have been observed in [266, 269] between the nonlinearity and the derivatives of Boolean functions: applying Relation (24) to linear hyperplanes E and with $b = 0$, we have: $\mathcal{NL}(f) \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + \max_{e \neq 0} |\mathcal{F}(D_e f)|}$. 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 $\mathcal{NL}(f) \geq 2^{n-2} - \frac{1}{4} \min_{e \neq 0} |\mathcal{F}(D_e f)|$.
Another lower bound on the nonlinearity is a consequence of Remark 2 after Proposition 10: 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 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 (however, explicit attacks using approximations by low degree functions exist for self-synchronizing stream ciphers and block ciphers more than for synchronous stream ciphers, see e.g. [164]). A relevant criterion is the *nonlinearity profile*, that is, the sequence of the Hamming distances to the Reed-Muller code of order r, for small values of r. This distance is called the r-th order nonlinearity of f and denoted $\mathcal{NL}_r(f)$. The best known asymptotic upper bound on $\mathcal{NL}_r(f)$ is $2^{n-1} - \frac{\sqrt{15}}{2} \cdot (1+\sqrt{2})^{r-2} \cdot 2^{n/2} + O(n^{r-2})$ (see [76], where a non-asymptotic - and more complex - bound is also given). Counting the number of functions whose r-th order nonlinearities are upper bounded by a given number (see [87]) allows proving that, when n tends to infinity, there exist functions with r-th order nonlinearity greater than $2^{n-1} - \sqrt{\sum_{i=0}^{r} \binom{n}{i}} \quad 2^{\frac{n-1}{2}}$.

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 input and the output (which can be used in attacks). 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 the 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 $(s_{i+\gamma_1}, \dots, s_{i+\gamma_n})$ is not uniform, where $\gamma_1, \dots, \gamma_n$ are the positions where the input bits to the filtering function are chosen. J. Golić [118] has observed that if the characteristic (or the feedback) polynomial of the LFSR is primitive and if the filtering function has the form $x_1 + g(x_2, \dots, x_n)$ or $g(x_1, \dots, x_{n-1}) + x_n$, then this property is satisfied. A. Canteaut [31] has proved that this condition on the function is also necessary. For choosing a filtering function, we shall have to choose a function g satisfying the cryptographic conditions listed above and below, 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 combination functions in stream ciphers: any such function $f(x)$ must stay balanced if we keep constant some coordinates x_i of x (at most m of them where m is as large as possible). We say that f is then an *m-resilient function*[18]. This definition of resiliency was introduced by Siegenthaler[19] in [245]; it is related to an attack on pseudo-random generators using combining functions, called *correlation attack*: if f is not m-resilient, then there exists a correlation between the output of the function and (at most) m coordinates of its input; if m is small, a divide-and-conquer attack due to Siegenthaler [246] and later improved (and also generalized to pseudo-random generators using fitering functions) by several authors [38, 148, 149, 150, 203] 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); the initializations of the other LFSRs can then be found with an independent exhaustive search. In the improved attacks (called *fast correlation attacks*), the correct initial-

[18] 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 [25]). 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").

[19] The term of resiliency was, in fact, introduced in [86], in relationship with another cryptographic problem.

ization is found in a more effective way, related to error-correcting decoding. To make stream ciphers with nonlinear filtering generators resistant against fast correlation attacks (see [115, 148, 203]), the Boolean filtering function must be highly nonlinear. In the case of stream ciphers with Boolean combining functions as well, Canteaut and Trabbia in [38] and Canteaut in [29] show that, to make fast correlation attacks as inefficient as possible, the coefficient $\widehat{\mathrm{f}_\chi}(u)$ of an m-resilient function has to be small for every vector u of Hamming weight higher than, but close to, m and this condition is satisfied by highly nonlinear Boolean m-resilient functions.
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:

Proposition 13 *[125] Any n-variable Boolean function f is m-resilient if and only if $\widehat{\mathrm{f}_\chi}(u) = 0$ 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 (25) 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; we get $\sum_{u \in E^\perp} \widehat{\mathrm{f}_\chi}^2(u) = |E^\perp| \sum_{a \in E'} \mathcal{F}^2(h_a)$; the orthogonal $E^\perp$ of E equals $\{x \in \mathbb{F}_2^n / \ x_i = 0, \ \forall i \notin I\}$ (it contains words of weight at most m, only), and $\mathcal{F}(h_a)$ is null if and only if h_a is balanced.
An alternate proof of this same result is obtained by applying Relation (14) 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 $\mathbb{F}_2^n$.
Proposition 13 shows that f is m-resilient if and only if its support has size 2^{n-1} and dual distance at least $m+1$ (see [99, 100]; see also in [198] a generalization of this result to arrays over finite fields and other related nice results); indeed, if C denotes the support of f, the dual distance of C equals the number $\min\{w_H(b); \, b \neq 0, \sum_{x \in C} (-1)^{b \cdot x} \neq 0\}$, and we have, for every vector b: $\sum_{x \in C} (-1)^{b \cdot x} = \widehat{f}(b)$. 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 [98] (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 Propositions 4 and 13. The symmetry group of the set of m-resilient functions and the orbits under its action have been studied in [142]).

Strict avalanche criterion and propagation criterion: the *Strict Avalanche Criterion* (*SAC*) was introduced by Webster and Tavares [259] and this concept was generalized into the *Propagation Criterion* (*PC*) by Bart Preneel [223] (see also [224]). 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 must be satisfied at high levels, in particular by 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 $D_af(x) = f(x) \oplus f(a+x)$ (see Definition 2) is balanced. It satisfies $PC(l)$ if it satisfies PC with respect to the set of all those nonzero vectors of weights at most l. In other words, f satisfies $PC(l)$ if the auto-correlation coefficient $\mathcal{F}(D_af)$ 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 [223], is a generalization of the strict avalanche criterion of order k, $SAC(k)$ (which is equivalent to $PC(1)$ of order k), introduced in [114]. 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 [223, 225] (see also [52]): a Boolean function satisfies the *extended propagation criterion* $EPC(l)$ of order k if every derivative D_af, 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 [166].

Non-existence of nonzero linear structure: we shall call the *linear kernel* of f the set of those vectors e such that D_ef is a constant function. The linear kernel of any Boolean function is a 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 [110]). The existence of nonzero 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* $f(x_1,\ldots,x_n)=g(x_1,\ldots,x_{n-1})\oplus\varepsilon\,x_n$ *where* $\varepsilon\in\mathbb{F}_2$.

Indeed, if we compose f on the right with a linear automorphism L such that $L(0,\ldots,0,1)=e$ is a nonzero linear structure, we have then $D_{(0,\ldots,0,1)}(f\circ L)(x)=f\circ L(x)\oplus f\circ L(x+(0,\ldots,0,1))=f\circ L(x)\oplus f(L(x)+e)=D_ef(L(x))$.
Note that, according to Proposition 14, if f admits a nonzero linear structure, then the nonlinearity of f is upper bounded 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 upper bounded by $2^{n-1}-2^{\frac{n-1}{2}-1}$.
Applying recursively this property, we deduce that $\mathcal{NL}(f)\leq 2^{n-1}-2^{\frac{n+k-2}{2}}$, where k is the dimension of the linear kernel of f [32].
Another characterization of linear structures [170, 109] (see also [34]) is a direct consequence of Relation (24), with $b=0$ and $E=\{0,e\}^\perp$, that is $\sum_{u\in a+E}\widehat{\mathrm{f}_\chi}^2(u)=2^{n-1}(2^n+(-1)^{a\cdot e}\mathcal{F}(D_ef))$.

Proposition 15 *Let f be any n-variable Boolean function. The derivative D_ef equals the constant function 1 (resp. the null function) if and only if the set $\{u\in\mathbb{F}_2^n/\widehat{\mathrm{f}_\chi}(u)=0\}$ contains the linear hyperplane $\{0,e\}^\perp$ (resp. its complement).*

Thus, D_ef 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). Notice that, if D_ef 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_ef 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 [83] (see more in [255]) 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 [204] 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, but also other functions, such as all non bent quadratic functions). This distance is always upper bounded by 2^{n-2}. More precisely, it equals $2^{n-2} - \frac{1}{4}\max_{e\in\mathbb{F}_2^{n}{}^*} |\mathcal{F}(D_ef)|$, since a function g, which admits some 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 $g(x) = g(x+e) \oplus \epsilon$ for every $x \in (e+H)$, where ϵ is chosen in $\mathbb{F}_2$; the Hamming distance between f and this function g equals $|\{x \in e+H/\, D_ef(x) = \epsilon \oplus 1\}| = \frac{1}{2}|\{x \in \mathbb{F}_2^n/\, D_ef(x) = \epsilon \oplus 1\}| = \frac{1}{2}\left(2^{n-1} - \frac{(-1)^\epsilon}{2}\mathcal{F}(D_ef)\right)$; recall that $\Delta_f(e) = \mathcal{F}(D_ef)$ is the auto-correlation function of f. We see that the distance of f to linear structures equals 2^{n-2} if and only if f is bent.

The algebraic immunity:
A new kind of attacks, called *algebraic attacks*, has been introduced recently (see [88, 89, 111]). 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 [244]. In practice, this system is too complex to be solved (its equations being highly nonlinear). However, in many situations, we can get a very overdefined system (i.e. a system with a number of independent equations much greater than the number of unknowns). Consider for instance an LFSR of length N, filtered by an n-variable Boolean function f; 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 $u_1, \dots, u_N$ the initialisation of the LFSR and by $(s_i)_{i\geq 0}$ the pseudo-random sequence output by the generator, we have, for every $i \geq 0$:

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

(this is more generally valid for every linear automata combined by a Boolean function, and in particular in the case of several LFSR combined by a Boolean function). 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 [111]), and even allows linearizing the system (i.e. obtaining a non-degenerate 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 $f * g = h$ (where $f * g$ denotes the function whose support is the intersection of the supports of f and g, we shall omit writing $*$ in the sequel). We have then, for every $i \geq 0$:

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

This equation in $u_1, \ldots, u_N$ 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 $fg = h$. More precisely, it is a simple matter to see 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, $f(g \oplus h) = 0$, 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. Clearly, the set of all annihilators is equal to the ideal of all the multiples of $f \oplus 1$. 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 (basic) *algebraic immunity* of f and denoted by $AI(f)$. This important characteristic is an affine invariant. As shown in [89], the algebraic immunity of any n-variable function is upper bounded by $\lceil n/2 \rceil$ (and consequently 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 to a function in k variables plus an affine function). 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. $g \neq 0$. It has been proved in [102] that, for all $a < 1$, when n tends to infinity, $AI(f)$ is almost surely greater than

$\frac{n}{2} - \sqrt{\frac{n}{2} \ln\left(\frac{n}{a \ln 2}\right)}$.

In [30], A. Canteaut has observed that, if a balanced function f in an odd number n of variables admits no non-zero annihilator of 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 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) [187]. Let G be a generator matrix of this code. Each column of G is labeled by a vector of F_2^n. Saying that f has no non-zero annihilator of 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 $\sum_{i=0}^{\frac{n-1}{2}} \binom{n}{i} = 2^{n-1}$. Since f has weight 2^{n-1}, this is also equivalent to saying that the support of the function is an information set, that is (assuming for simplicity that the columns corresponding to the support of f are the 2^{n-1} first ones), that we can take $G = (Id \,|\, M)$. Then the complement of the support of f is also an information set (otherwise there would exist a vector $(z \,|\, 0)$, $z \neq 0$, in the code and this is clearly impossible since G is also a parity-check matrix of the code).

Now let an n-variable function f, with algebraic immunity $\lceil n/2 \rceil$ be used as a filtering function on a linear automaton (e.g. an LFSR) with $m \geq 2k$ states, where k is the length of the key (otherwise, it is known that the system is not robust). Then the complexity of an algebraic attack using one annihilator of degree $\lceil n/2 \rceil$ is roughly $7\left(\binom{m}{0} + \ldots + \binom{m}{\lceil n/2 \rceil}\right)^{\log_2(7)} \approx 7\left(\binom{m}{0} + \ldots + \binom{m}{\lceil n/2 \rceil}\right)^{2.8}$ (see [89]). Let us choose $k = 128$ (which is usual) and $m = 256$, then the complexity of the algebraic attack is at least 2^{80} for $n \geq 13$; and it is greater than the complexity of an exhaustive search, that is 2^{128}, for $n \geq 15$. If the attacker knows several linearly independent annihilators of degree $\lceil n/2 \rceil$, then the numbers of variables must be enhanced!

It has been shown in [93] and [64] that low nonlinearity implies low algebraic immunity (but high algebraic immunity does not imply high nonlinearity). More precisely, 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 inequality must for instance be true since, otherwise, the number $w_H(f)$ of equations in system expressing that a function of degree at most $AI(f) - 1$ is an annihilator of f would have a number of equations smaller than its number of unknowns and it would therefore have non-trivial solutions, a contradiction). This implies that a function f such that $AI(f) = \frac{n+1}{2}$ (n odd) must be balanced. Since it can also be easily proved that, for every function h of degree r, we have $AI(f) - r \leq AI(f+h) \leq AI(f) + r$, we deduce

$$\mathcal{NL}(f) \geq \sum_{i=0}^{AI(f)-2} \binom{n}{i}$$

and more generally:

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

These bounds have been improved, in all cases for the first order nonlinearity into $\mathcal{NL}(f) \geq 2\sum_{i=0}^{AI(f)-2} \binom{n-1}{i}$ [182], and in most cases for the r-th order nonlinearity into $\mathcal{NL}_r(f) \geq 2\sum_{i=0}^{AI(f)-r-1} \binom{n-r}{i}$ (in fact, the improvement was slightly stronger than this, but more complex), see [62].

Note that if f is k-normal then its algebraic immunity is at most $n-k$, since the fact that $f(x) = \epsilon \in F_2$ 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 the majority function (cf. below) is $\lfloor n/2 \rfloor$-normal for every n (see [57]) and has algebraic immunity $\lceil n/2 \rceil$. But $AI(f) \leq \ell$ 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 [57] (note that $k < a\,log_2\,n$ implies that $n - k \sim n$) and the algebraic immunity is always upper bounded by $n/2$.

Balanced highly nonlinear functions in up to 20 variables (derived from the power mappings studied in the chapter "Vectorial Boolean Functions for Cryptography") with high algebraic immunities have been exhibited in [69] and [5]. However, it has been proved in [210] that, if the number of runs $r(d)$ of 1's in the binary expansion of the exponent d of a power function $tr(ax^d)$ (that is, the number of subsequences of 1's, separated by 0's) is (much) smaller than $\sqrt{n}/2$, then the algebraic immunity is low. More precisely, the algebraic immunity is upper bounded 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 majority function (first proposed by J.D. Key, T.P. McDonough and

V.C. Mavron in the context of the erasure channel [160] - rediscovered by Dalai et al. in the context of algebraic immunity [95]), $f(x) = 1$ if $w_H(x) \geq n/2$, has optimum algebraic immunity (note that 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). It is a symmetric function and its properties and structure are known. Some variants have also optimum algebraic immunity. A nice construction of an infinite class of functions with optimum algebraic immunity has been given in [94] and further studied in [64]; 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]. Indeed, a high value of $AI(f)$ is not a sufficient property for a resistance to algebraic attacks, because of fast algebraic attacks, in which h can have a greater degree than g (see [89]). 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 $d^\circ g + d^\circ h \geq n$ – there exist g 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 $fg = h$ 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 $f + 1$ and 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.

Other criteria:
- the second moment of the auto-correlation coefficients:

$$\mathcal{V}(f) = \sum_{e \in \mathbb{F}_2^n} \mathcal{F}^2(D_e f) \tag{33}$$

has been introduced by Zhang and Zheng [265] for measuring the *global avalanche criterion* (GAC). It is called the *sum-of-squares indicator* by some authors. 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 lower bounded by $\sqrt{\frac{V - 2^{2n}}{2^n - 1}}$. The functions that achieve $\mathcal{V}(f) = 2^{2n}$ are those functions whose derivatives $D_e f(x)$, $e \neq 0$, are

all balanced. We shall see at 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 is partially bent, see below).
Note that, according to Relation (23) 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 (18) 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 [33] (indeed, the Fourier transform of φ equals $\widehat{\mathrm{f}_\chi}^2$, according to Relation (22)), and thus, for $a = 0$:

$$\sum_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^4(e) = 2^n\, \mathcal{V}(f). \tag{34}$$

We have $\sum_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^4(e) \leq \left(\sum_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(e)\right)\left(\max_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(e)\right)$. According to Parseval's relation $\sum_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(e) = 2^{2n}$, we deduce, using Relation (34): $\max_{e \in \mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(e) \geq \frac{\mathcal{V}(f)}{2^n}$ (with equality if and only if f is plateaued [33], see below); thus, according to Relation (31) and to the inequality $\mathcal{V}(f) \geq 2^{2n}$, we have (as first shown in [266, 269]):

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

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 (34) also implies the following relation, first observed in [269]: $\mathcal{V}(f) \times N_{\widehat{\mathrm{f}_\chi}} \geq 2^{3n}$. Indeed, using for instance 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 (20). The functions satisfying $\mathcal{V}(f) \times N_{\widehat{\mathrm{f}_\chi}} = 2^{3n}$ 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). Constructions of balanced Boolean functions with low absolute indicators and high nonlinearities have been studied in [189].

- The *maximum correlation* of an n-variable Boolean function f with respect to a subset I of $N = \{1, \ldots, n\}$ equals by definition (see [264]) $C_f(I) = \max_{g \in \mathcal{BF}_{I,n}} \dfrac{\mathcal{F}(f \oplus g)}{2^n}$, where $\mathcal{BF}_{I,n}$ is the set of n-variable Boolean functions depending on $\{x_i, i \in I\}$ only. According to Relation (31), the distance from f to $\mathcal{BF}_{I,n}$ equals $2^{n-1}(1 - C_f(I))$. 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 $\sum_{j=1}^{2^{n-|I|}} \frac{|\mathcal{F}(h_j)|}{2^n}$, where $h_1, \ldots, h_{2^{n-|I|}}$ are the restrictions of f obtained by keeping constant the x_i's for $i \in I$, 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 Cauchy-Schwartz inequality, we have $\left(\sum_{j=1}^{2^{n-|I|}} |\mathcal{F}(h_j)|\right)^2 \leq 2^{n-|I|} \sum_{j=1}^{2^{n-|I|}} \mathcal{F}^2(h_j)$, and Relation (25) directly implies the following inequality observed in [28, 29]:

$$C_f(I) \leq 2^{-n} \left(\sum_{u \in \mathbb{F}_2^n /\, u_i = 0,\, \forall i \notin I} \widehat{\mathrm{f}_\chi}^{\,2}(u) \right)^{\frac{1}{2}} \leq 2^{-n+\frac{|I|}{2}} \left(2^n - 2\mathcal{NL}(f)\right)^{\frac{1}{2}}. \quad (35)$$

This inequality shows that the nonlinearity of any combining function should be high. An affine invariant 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 [29].

- 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 [57] and first evoked in [204]), 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*) [57] (see Subsection 5.3.2), the number of nonzero coefficients of the Walsh transform [225, 232]. It has been shown in [57, 214, 232] that (asymptotically) almost all Boolean functions have high complexities with respect to all these criteria (see also [230] for some complementary results).
For every even integer k such that $4 \leq k \leq 2^n$, the kth-order *nonhomomorphicity* [268] of a Boolean function equals the number of k-tuples $(u_1, \dots, u_k)$ 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 [268]) 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.

5 Quadratic functions and other functions whose weights, Walsh spectra or nonlinearities can be analyzed

5.1 Quadratic 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. More generally, the behavior of the functions of $R(2, n)$, called *quadratic functions*, is also peculiar. Recall that Relation (23) 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:

$$\mathcal{F}^2(f) = 2^n \sum_{b \in \mathcal{E}_f} (-1)^{D_b f(0)}, \tag{36}$$

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; 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 (10), this proves the following:

Proposition 16 *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 Proposition 16 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.

According to Relation (36) 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), 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 $\{2^{n-1}\} \cup \{2^{n-1} \pm 2^i / \, n/2-1 \leq i \leq n-1\}$. Its nonlinearity can be any element of the set $\{2^{n-1} - 2^i / \, n/2-1 \leq i \leq n-1\}$, 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 $\{2^{n-1} - 2^i, 2^{n-1}, 2^{n-1} + 2^i\}$.

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 nonsingular linear transformation, this monomial is x_1x_2. The function has then the form $x_1x_2 \oplus x_1 f_1(x_3,\dots,x_n) \oplus x_2 f_2(x_3,\dots,x_n) \oplus f_3(x_3,\dots,x_n)$ where f_1, f_2 are affine functions and f_3 is quadratic. Then, $f(x)$ equals $(x_1 \oplus f_2(x_3,\dots,x_n))(x_2 \oplus f_1(x_3,\dots,x_n)) \oplus f_1(x_3,\dots,x_n) f_2(x_3,\dots,x_n) \oplus f_3(x_3,\dots,x_n)$ and is therefore affinely equivalent to the function $x_1x_2 \oplus f_1(x_3,\dots,x_n) f_2(x_3,\dots,x_n) \oplus f_3(x_3,\dots,x_n)$. Applying this method recursively shows (see [187]):

Proposition 17 *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 permits to describe precisely the weight distribution of $R(2,n)$.

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 disappeared. This permits to show (see [43]) 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 $\widehat{\mathrm{f}_\chi}(0) = 2^m\, \widehat{\mathrm{g}_\chi}(0)$ at 0.

5.2 Indicators of flats

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 $a_1, \ldots, a_r \in \mathbb{F}_2^n$ are linearly independent and $\varepsilon_1, \ldots, \varepsilon_r \in \mathbb{F}_2$. Then f has weight 2^{n-r}. Moreover, set $a \in \mathbb{F}_2^n$. If a is linearly independent of $a_1, \ldots, a_r$, 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 $g(x_1, \ldots, x_r) \oplus x_{r+1}$. If a is linearly dependent of $a_1, \ldots, a_r$, say $a = \sum_{i=1}^r \eta_i\, a_i$, then $a \cdot x$ is clearly constant on the flat and this constant value equals $\bigoplus_{i=1}^r \eta_i\, (a_i \cdot x) = \bigoplus_{i=1}^r \eta_i\, (\epsilon_i \oplus 1)$; 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 $\widehat{\mathrm{f}_\chi}(0) = 2^n - 2^{n-r+1}$.
Note that the nonlinearity of f equals 2^{n-r} and is bad. But indicators of flats can be used to design Boolean functions with good nonlinearities (see Subsection 7.3).

Note. As recalled at 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 [155]; 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.

5.3 Other functions whose nonlinearities can be better approximated than for general functions

5.3.1 Maiorana-McFarland's functions and their generalizations

Maiorana-McFarland's functions will be defined at Sections 6 (for bent functions) and 7 (for resilient functions). The computation of their weights and Walsh spectra are easier than for general Boolean functions, and in some cases can be completely determined. Generalizations exist, sharing this same property (see Section 7). Their algebraic immunity has been studied in [69].

5.3.2 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 (25) in Proposition 7 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 (31), we deduce:

$$\forall a \in E',\ \mathcal{NL}(f) \leq 2^{n-1} - 2^{k-1} + \mathcal{NL}(h_a). \tag{37}$$

This bound was first proved (in a different way) by Zheng et al. in [270]. The present proof is from [33]. Relation (37) can also be deduced from Relation (14) 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 $\left(2^k - 2\mathcal{NL}(h_a)\right)$. Then

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

Then the mean of $(-1)^{a \cdot u} \widehat{\mathrm{f}_\chi}(u)$, when u ranges over $b \oplus E^\perp$, is equal to $\pm\left(2^k - 2\mathcal{NL}(h_a)\right)$. Thus, the maximum magnitude of $\widehat{\mathrm{f}_\chi}(u)$ is greater than or equal to $2^k - 2\mathcal{NL}(h_a)$. This implies Relation (37). These two methods, for proving (37), lead to two different necessary conditions for the case of equality (see [57]).

Relation (37) 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 $\mathcal{NL}(f) \leq 2^{n-1} - 2^{k-1}$, and that, if equality occurs, then $f \oplus \ell$ is balanced on every other coset of this flat.

Definition 3 *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 7.3.1). It is proved in [57] 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 that are nonconstant on flats of dimension $n/2$. This is not explicitly mentioned in the paper. What they actually show is that the functions (they say, the sets) are not constant on flats defined by equations $x_{i_1} = a_1, ..., x_{i_{n/2}} = a_{n/2}$. To prove that, they use however the fact that the sets have small bias with respect to linear tests. As this property is invariant w.r.t. affine transformations, it implies the result.
There are also explicit constructions that work for dimensions $(1/2 - \epsilon)\, n$, for some small $\epsilon > 0$ very recently found by Jean Bourgain [18].
Functions that are nonconstant on flats of dimensions n^δ for every $\delta > 0$ are also given in [10]. 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.

5.3.3 Partial covering sequences

The notion of covering sequence for a Boolean function has been introduced in [79].

Definition 4 *Let f be an n-variable Boolean function. An integer-valued*[20] *sequence $(\lambda_a)_{a\in\mathbb{F}_2^n}$ is called a* covering sequence *for 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*

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

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 [79] that any function admitting a non-trivial covering sequence is balanced (see Proposition 18 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 [79]: denote again by $S_{\widehat{\mathrm{f}_\chi}}$ the support $\{u\in\mathbb{F}_2^n \mid \widehat{\mathrm{f}_\chi}(u)\neq 0\}$ of $\widehat{\mathrm{f}_\chi}$; 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 $S_{\widehat{\mathrm{f}_\chi}}$. Indeed, f admits the covering sequence λ with level ρ if and only if, for every $x\in\mathbb{F}_2^n$, we have $\sum_{a\in\mathbb{F}_2^n}\lambda_a(-1)^{f(x+a)} = \left(\sum_{a\in\mathbb{F}_2^n}\lambda_a-2\rho\right)(-1)^{f(x)}$; the characterization is then a consequence of the property that the equality between two integer-valued functions is equivalent to the equality between their Fourier transforms, and of the relation $\sum_{a,x\in\mathbb{F}_2^n}\lambda_a(-1)^{f(x+a)+x\cdot b}=\left(\sum_{a\in\mathbb{F}_2^n}\lambda_a(-1)^{a\cdot b}\right)\widehat{\mathrm{f}_\chi}(b)$.
Knowing a covering sequence (trivial or not) of a function f permits to know 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 μ is this value; 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 $\widehat{\lambda}(b)=r$, where $r=\widehat{\lambda}(0)-2\rho$. 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 covering sequence playing a particular role 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 [79].
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 [60] is weakened the definition of covering sequence, so that it can help computing the (nonzero) values of the Walsh transform.

Definition 5 *Let f be a Boolean function on $\mathbb{F}_2^n$. A non-trivial* partial covering sequence *for f is an integer-valued sequence $(\lambda_a)_{a\in\mathbb{F}_2^n}$ such that the integer-valued function $\sum_{a\in\mathbb{F}_2^n}\lambda_a D_a f(x)$ takes on two values 0 and $\rho\neq 0$. The constant ρ is called the level of the partial covering sequence.*

A simple example of non-trivial partial covering sequence is as follows: let $\mathcal{E}$ be any set of derivatives of f which is not reduced to the null function. Assume that $\mathcal{E}$ is stable under addition (*i*.e. is a $\mathbb{F}_2$-vectorspace). Then $\sum_{g\in\mathcal{E}} g$ takes on values 0 and $\frac{|\mathcal{E}|}{2}$. Thus, if $\mathcal{E}=\{D_a f/\ a\in E\}$ (where 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 permit to simplify the computation of the weight of f (or the value of $\mathcal{F}(f)$, which is equivalent).

Proposition 18 *Let f be a Boolean function on $\mathbb{F}_2^n$ and let $(\lambda_a)_{a\in\mathbb{F}_2^n}$ be a non-trivial partial covering sequence for f. Denote by A the set $\{x\in\mathbb{F}_2^n/\ \sum_{a\in\mathbb{F}_2^n}\lambda_a D_a f(x)=0\}$. Then*

$$\mathcal{F}(f)=\sum_{x\in A}(-1)^{f(x)}$$

and if $A=\emptyset$, then f is balanced.

Proof. For every $a\in\mathbb{F}_2^n$, the set $(D_a f)^{-1}(1)$ is invariant under the mapping $x\mapsto x+a$. For every x in this set, we have $f(x+a)=f(x)\oplus 1$ and, thus, $(-1)^{f(x+a)}+(-1)^{f(x)}=0$. Hence, we have

$$\sum_{x\in\mathbb{F}_2^n} D_a f(x)(-1)^{f(x)}=\sum_{x\in(D_a f)^{-1}(1)}(-1)^{f(x)}=0. \tag{38}$$

We deduce that the sum $\sum_{a\in\mathbb{F}_2^n}\lambda_a\left(\sum_{x\in\mathbb{F}_2^n} D_a f(x)(-1)^{f(x)}\right)$ is null. This sum equals $\sum_{x\in\mathbb{F}_2^n}(-1)^{f(x)}\left(\sum_{a\in\mathbb{F}_2^n}\lambda_a D_a f(x)\right)=\rho\sum_{x\notin A}(-1)^{f(x)}$. Hence $\sum_{x\notin A}(-1)^{f(x)}=0$ and $\sum_{x\in\mathbb{F}_2^n}(-1)^{f(x)}=\sum_{x\in A}(-1)^{f(x)}$. ⋄

Examples are given in [60] of computations of the weights or the Walsh spectra of some Boolean functions (quadratic functions, Maiorana-McFarland's functions and their extensions, and other examples of functions), using Proposition 18.

5.3.4 Functions with low univariate degree

The following Weil's Theorem is very well-known in finite field theory (*cf.* [178, Theorem 5.38]):

Theorem 1 *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 $a \in \mathbb{F}_{2^n}$: $\left|\sum_{x \in \mathbb{F}_{2^n}} (-1)^{tr(af(x))}\right| \leq (d-1)\, 2^{n/2}$. And since adding a linear function $tr(bx)$ to the function $tr(af(x))$ corresponds to adding $(b/a)\, x$ to $f(x)$ and does not change its anivariate degree, we deduce that, if $d > 1$ is odd and $a \neq 0$, then:

$$\mathcal{NL}(tr(af)) \geq 2^{n-1} - (d-1)\, 2^{n/2-1}.$$

6 Bent functions

Bent functions have been defined, at Subsection 4.1, as those Boolean functions f on $\mathbb{F}_2^n$ (n even) whose 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 also at Subsection 4.1, f is bent if and only if $\widehat{\mathrm{f}_\chi}$ takes on values $\pm 2^{n/2}$ only. 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 $\widehat{\mathrm{f}_\chi}$ 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 $a \in \mathbb{F}_2^n$, $\widehat{\mathrm{f}_\chi}(a) \equiv 2^{n/2} \left[\ mod\ 2^{n/2+1}\right]$, or equivalently $\widehat{f}(a) \equiv 2^{n/2-1} \left[\ mod\ 2^{n/2}\right]$.*

Proof. This necessary condition is also sufficient, since, if it is satisfied, and if $\widehat{\mathrm{f}_\chi}(a) \neq \pm 2^{n/2}$ for some a, then f_χ cannot satisfy Parseval's Relation (20); a contradiction. ◇

A slightly different viewpoint is that of bent sequences[21] 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 action of the general affine group and the addition of affine functions. The notion of bent function is also *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 isomorphism, *i.*e. an isomorphism whose associated matrix is symmetric).

Thanks to Relation (22) 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 word a, the Boolean function $D_af(x) = f(x) \oplus f(x+a)$ is balanced. In other words:

Proposition 19 *Any n-variable Boolean function (n even) is bent if and only if it satisfies $PC(n)$.*

For this reason, bent functions are also called *perfect nonlinear functions*[22]. 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*[23] of the elementary Abelian 2-group $\mathbb{F}_2^n$ [104, 152]. Other types of difference sets exist (see *e.g.* [106]). This implies that the Cayley graph G_f (see Subsection 2.2.2) is strongly regular (see [14] for more precision).

Functions satisfying $PC(n)$ do not exist for odd n.

The functions whose derivatives D_af, $a \in H$, $a \neq 0$ are all balanced, where

[21] For each vector X in $Q_{2^n} = \{-1,1\}^{2^n}$, 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 Q_{2^n} are called bent sequences. They are the images by the character $\chi = (-1)^{\cdot}$ of the bent functions on $\mathbb{F}_2^n$.

[22] The characterization of Proposition 19 leads to a generalization of the notion of bent function to non-binary functions. In fact, several generalizations exist [3, 165, 183] (see [65] 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 [67]).

[23] Thus, bent functions are also related to designs, since any difference set can be used to construct a symmetric design, see [7], 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.

H is a linear hyperplane of $\mathbb{F}_2^n$, are characterized in [32, 33] for every n; they are all bent if n is even. The functions whose derivatives D_af, $a \in E$, $a \neq 0$ 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.

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 $S_f = \{u_1, \ldots, u_{w_H(f)}\}$. 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 lines of this matrix. Thus, C is the set of all the vectors $U_v = (v \cdot u_1, \ldots, v \cdot u_{w_H(f)})$, where v ranges over $\mathbb{F}_2^n$. Then:

Proposition 20 *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 (9), it equals $2^{n-2} + \frac{\widehat{f_\chi}(v) - \widehat{f_\chi}(0)}{4}$, 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 $\widehat{f_\chi}(v) = \widehat{f_\chi}(0)$ or $\widehat{f_\chi}(v) = \widehat{f_\chi}(0) + 4w_H(f) - 2^{n+1} = \widehat{f_\chi}(0) - 2\widehat{f_\chi}(0) = -\widehat{f_\chi}(0)$. If f is bent, then this condition is clearly satisfied. Conversely, according to Parseval's Relation (20), if this condition is satisfied, then $\widehat{f_\chi}(v)$ equals $\pm 2^{n/2}$ for every v, *i.e.* f is bent.

There exist two other characterizations [260] 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:

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

is also bent and its own dual is f itself. Indeed, Relation (16) applied to $\varphi = f_\chi$ (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 (19) applied with $\varphi = \mathrm{f}_\chi$ and $\psi = \mathrm{g}_\chi$ shows that

$$\mathcal{F}(\widetilde{f} \oplus \widetilde{g}) = \mathcal{F}(f \oplus g). \tag{39}$$

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 4, 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 (39), 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 $g(x) = f(x+b) \oplus \ell_a(x+b)$, it gives the following property, first observed in [52] (see also [34]):

$$\mathcal{F}(D_a\widetilde{f} \oplus \ell_b) = \mathcal{F}(D_b f \oplus \ell_a) \tag{40}$$

(from these two relations, we deduce that, if $a \cdot b = 1$, then $\mathcal{F}(D_b f \oplus \ell_a) = 0$ and thus that $D_b f$ is balanced on $a^\perp$ and on its complement; notice also that, for every a and b, $D_b f = \ell_a \oplus \epsilon$ 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 $\widetilde{f}$ (which will be useful later) can be deduced from the NNF of f. Indeed, using Relation (9) and equality $\widetilde{f} = \frac{1-(-1)^{\widetilde{f}}}{2}$, we have $\widetilde{f} = \frac{1}{2} - 2^{-n/2-1}\widehat{\mathrm{f}_\chi} = \frac{1}{2} - 2^{n/2-1}\delta_0 + 2^{-n/2}\widehat{f}$. Applying now Relation (27) to $\varphi = f$, we deduce:

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

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 [73]):

Proposition 21 *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 $2^{|I|-n/2}$.*

Using Relation (6), this property can be related to the main result of [144] (but this result by Hou was stated in a complex way).

Relation (14) applied to $\varphi = \mathrm{f}_\chi$ gives (see [45])

$$\sum_{x\in a+E}(-1)^{\widetilde{f}(x)\oplus b\cdot x} = 2^{-n/2}|E|\,(-1)^{a\cdot b}\sum_{x\in b+E^\perp}(-1)^{f(x)\oplus a\cdot x}. \qquad (41)$$

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 at Subsection 5.1, 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 $M = (m_{i,j})_{i,j\in\{1,\ldots,n\}}$ 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: $x_1x_2 \oplus x_3x_4 \oplus \cdots \oplus x_{n-1}x_n \oplus \varepsilon$ ($\varepsilon \in \mathbb{F}_2$).

Open problem: characterize the bent functions of algebraic degrees ≥ 3 (that is, classify them under the action of the general affine group). This has been done for $n \leq 6$ in [231] (see also [225] where the number of bent functions is computed for these values of n). For $n = 8$, it has been done in [140], for functions of algebraic degrees at most 3 only; all of these functions have at least one affine derivative $D_a f$, $a \neq 0$ (it has been proved in [34] that this happens for $n \leq 8$ only).

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 21 applied to $\widetilde{f}$ gives:

Proposition 22 *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 $n/2$.*

This property (which is obviously also true for $\widetilde{f}$) was first proved in [231] 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 9. This same method also permits to obtain a bound, shown in [143], relating the gaps between $n/2$ and the algebraic degrees of f and $\widetilde{f}$:

Proposition 23 *The algebraic degrees of any n-variable bent function and of its dual satisfy:*

$$n/2 - d^\circ f \geq \frac{n/2 - d^\circ \widetilde{f}}{d^\circ \widetilde{f} - 1}. \tag{42}$$

A proof of Proposition 23 and a second proof of Proposition 22. Denote by d (resp. by $\widetilde{d}$) the algebraic degree of f (resp. of $\widetilde{f}$). Consider a term x^I of degree d in the ANF of f. Relation (15) applied to $\varphi = \mathrm{f}_\chi$ (or Relation (41) 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 $2^{n/2-d+2}$. If $d > n/2$, then $\sum_{u\in E} \widehat{\mathrm{f}_\chi}(u)$ is not divisible by $2^{n/2+1}$; a contradiction with the fact that E has an even size. This proves Proposition 22. Moreover, 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 $n/2 - d + 1 \geq \frac{n-d}{\widetilde{d}}$, which is equivalent to (42). $\diamond$

Using Relation (4) instead of Relation (3) gives a more precise result than Proposition 22, first shown in [73], which will be given at Subsection 6.6.

Proposition 23 can also be deduced from Proposition 21 and from some divisibility properties, shown in [73], of the coefficients of the NNFs of Boolean functions of degree d.

6.4 Constructions

There does not exist a classification of bent functions under the action of the general affine group. In order to know as many bent functions as possible, we can try to design constructions of bent functions. Some of the known constructions lead to classes of bent functions without using known ones. We will call *primary constructions* these direct constructions. The others, leading to recursive constructions, will be called *secondary constructions*.

6.4.1 Primary constructions

1. The *Maiorana-McFarland original class* $\mathcal{M}$ (see [104, 201]) 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:

$$f(x,y) = x \cdot \pi(y) \oplus g(y) \tag{43}$$

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 the 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 for f being bent, according to Relation (44) applied with $r = n/2$. 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 at Subsection 6.2; take $\pi = id$ and g constant in (43)).

The fundamental idea of Maiorana-McFarland's construction consists in *concatenating affine functions*. Indeed, 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 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 [56]:

Proposition 24 *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):

$$\widehat{\mathrm{f}_{\phi,\mathrm{g}_\chi}}(a,b) = 2^r \sum_{y \in \phi^{-1}(a)} (-1)^{g(y) \oplus b \cdot y}, \tag{44}$$

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 (44), 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 24 is a sufficient condition for that (but it is not a necessary one). ◇

This construction is a secondary one for $r < n/2$ and a primary one for $r = n/2$. Notice that it 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 reprensented in the form $f_{\phi,g}$ for some values of $r \geq 1$ and s. It has been shown in [183] that, if a bent function has the form $f_{\phi,g}$, 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 $D_{(a,0)} f_{\phi,g}$.

2. The *Partial Spreads class $\mathcal{PS}$*, introduced in [104] 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 5 below. This is why we omit the proof of this fact here. The dual of such a function has the same form, all the $n/2$-dimensional spaces E being replaced by their orthogonals (see also Theorem 5). 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, but not all those of $\mathcal{PS}^+$ (which contains for instance all the quadratic functions, if $n/2$ is even). 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 [104] 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 $x \cdot y = tr(xy)$, where tr 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[24] 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 $g(\frac{y}{x})$. The elements of $\mathcal{PS}_{ap}$ are, equivalently, those Boolean functions f of weight $2^{n-1} - 2^{n/2-1}$ on $\mathbb{F}_{2^n}$ such that $f(0) = f(1) = 0$, and that, denoting by α a primitive element of this field, $f(\alpha^{2^{n/2}+1}x) = f(x)$ for every $x \in \mathbb{F}_{2^n}$ (see [104, 68]).
Denoting by tr the trace function from $\mathbb{F}_{2^n}$ to its prime field $\mathbb{F}_2$ (*i.e.* $tr(x) = x+x^2+x^4+\cdots+x^{2^{n-1}}$), the Boolean functions $f(x) = tr(ax^i)$, where $a \in \mathbb{F}_{2^n}$ and where i is a multiple of $2^{n/2}-1$, satisfy this last condition. Some of them belong to $\mathcal{PS}_{ap}$. Other examples of bent functions of the same form exist. For instance (see [106, 36]), if n is not divisible by 3 and if k is co-prime with n, then, for every $a \in \mathbb{F}_{2^n} \setminus \{x^3 \mid x \in \mathbb{F}_{2^n}\}$, the function $tr(ax^{2^{2k}-2^k+1})$ is bent. This gives an infinite class of bent functions (other examples of similar bent functions exist).

3. Dobbertin gives in [107] the construction of a class of bent functions that 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

[24] 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 $\mathbb{F}_{2^n} = \mathbb{F}_{2^{n/2}} + \omega\mathbb{F}_{2^{n/2}}$.

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 proof of the bentness of such functions cannot be given here because of length constraints. 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).

The bent sequences given in [262] are particular cases of the constructions given above (using also some of the secondary constructions given below). in [81] 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).

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 $\mathbb{F}_2^{n/2}$. The definition of almost bent functions, the description of the related Boolean function and the proof of its bentness are given in the chapter "Vectorial Boolean Functions for Cryptography".

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*, that is, functions of the form $tr(ax^i)$, $a \neq 0$, where tr is the trace function on $\mathbb{F}_{2^n}$ and where a and x belong to this same field. And some are defined as the sums of a few power functions; see [35, 104, 106, 175, 176]. Power functions are also called *monomial functions*. They 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). They have the peculiarity that, denoting the image $\{x^i; x \in \mathbb{F}_{2^n}\}$ of the power mapping $x \to x^i$ by U, two functions $tr(ax^i)$ and $tr(bx^i)$ such that $a/b \in U$ are linearly equivalent. In particular, if the power mapping is a permutation, *i.e.* if $gcd(i, 2^n - 1) = 1$, then all the power functions with the same exponent are linearly equivalent. It is not clear whether this is more an advantage for the designer or for the attacker.
Obviously, a power function $tr(ax^i)$ can be bent only if the mapping $x \to x^i$ is not one to one (otherwise, the function would be balanced, a contradic-

tion), that is, if i is not co-prime with $2^n - 1$. It has been proved in [31] that i must be co-prime either with $2^{n/2} - 1$ or with $2^{n/2} + 1$. The exponents of the known classes of power bent functions are multiples of $2^{n/2} - 1$ (this corresponds to the $\mathcal{PS}_{ap}$ class), or equal $2^{2k} - 2^k + 1$ with $gcd(k, n) = 1$ (this is the so-called Kasami functions), or equal $2^i + 1$ with $\frac{n}{gcd(n,i)}$ even (Gold functions), $(2^{n/4} + 1)^2$ with n divisible by 4 (Leander functions) or $2^{n/3} + 2^{n/6} + 1$ with n divisible by 6 (Canteaut-Charpin-Kyureghyan functions). The three last cases enter in fact in the Maiorana-McFarland completed class.

6.4.2 Secondary constructions

1. The first secondary construction given by J. Dillon and O. Rothaus in [104, 231] 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 $\widehat{\mathrm{h}_\chi}(a, b) = \widehat{\mathrm{f}_\chi}(a) \times \widehat{\mathrm{g}_\chi}(b)$. This construction 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 property is easy to detect and can be used for designing divide-and-conquer attacks).

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 3). No general class of bent functions has been deduced from this construction.

3. Two classes of bent functions have been derived in [45] 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 [45] that this class is not included in the completed versions of classes $\mathcal{M}$ and $\mathcal{PS}$ (*i.e.*the smallest possible classes including them). 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 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 [36], 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 [45], which has its own interest:

Theorem 2 *Let $b + E$ be any flat in $\mathbb{F}_2^n$ (E is 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 $D_a f$ 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 $\mathcal{F}(D_a f^\star)$ equals $\mathcal{F}(D_a f)$ if $a \in E$, and equals $\mathcal{F}(D_a f) - 4\sum_{x \in b+E} (-1)^{D_a f(x)}$ otherwise.

We have $\widehat{\mathrm{f}_\chi}(a) - \widehat{\mathrm{f}_\chi^\star}(a) = 2\sum_{x \in b+E} (-1)^{f(x) \oplus a \cdot x}$. Using Relation (41), applied with $E^{\perp}$ in the place of E, we deduce that for every $a \in \mathbb{F}_2^n$:

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

and $\widehat{\mathrm{f}_\chi}(a) - \widehat{\mathrm{f}_\chi^\star}(a)$ can take value 0 or $\pm 2^{n/2+1}$ if and only if Condition *2.* is satisfied. So Condition *2.* is necessary. It is also sufficient, according to

Lemma 2.
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 9.
If f is bent, if E has dimension $n/2$, and if the restriction of f to $b+E$ is affine, then the relation $\widehat{\mathrm{f}_\chi}(a) - \widehat{\mathrm{f}_\chi^\star}(a) = 2\sum_{x\in b+E}(-1)^{f(x)\oplus a\cdot x}$ shows that $f^\star$ is bent too, according to Lemma 2. $\diamond$

Remarks.
- Relation (41) 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 (41) 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 2, the function $f \oplus 1_E$ is bent if and only if $D_a D_{a'} \widetilde{f}$ 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 [34] and used in [36] to design (potentially) new bent functions.

4. Other classes of bent functions have been deduced from a construction given in [48], which generalizes the secondary constructions given in 1 and 2 above:

Theorem 3 *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, in fact, 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 by Adams and Tavares [1] under the name of bent-based functions, and later studied by J. Seberry and X.-M. Zhang in [240] in particular cases.
A particular case of this construction is nicely simple: 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 $h(x,y) = f_1(x)\oplus g_1(y)\oplus(f_1\oplus f_2)(x)\,(g_1\oplus g_2)(y)$, $x \in \mathbb{F}_2^n, y \in \mathbb{F}_2^m$ (this construction $(f_1, f_2, g_1, g_2) \mapsto h$ will appear again below to construct resilient functions; see Theorem 8). For every y, h_y equals f_1 plus a constant or f_2 plus a constant (depending on the values of y) and thus is bent; and φ_s equals g_1 plus a constant or g_2 plus a constant (depending on the values of u), and thus is bent too. According to Theorem 3, h is then bent. Its dual $\widetilde{h}$ can be obtained from $\widetilde{f}_1, \widetilde{f}_2, \widetilde{g}_1$ and $\widetilde{g}_2$ exactly in the same manner as h is obtained from f_1, f_2, g_1 and g_2. What is interesting in this particular case 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.

Several classes have been deduced from Theorem 3 in [48], and later in [143].
- 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 [48], 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

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

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.

More recently, a particular case of the construction given in Theorem 3 was pointed out in [58]: let f_1 and f_2 be two r-variable bent functions (r even) and let g_1 and g_2 be two s-variable bent functions (s even). Let us denote their duals by $\widetilde{f_1}, \widetilde{f_2}, \widetilde{g_1}$ and $\widetilde{g_2}$. Define[25]

$$f(x,y) = f_1(x) \oplus g_1(y) \oplus (f_1 \oplus f_2)(x)\,(g_1 \oplus g_2)(y),$$

for every $x \in \mathbb{F}_2^r$ and $y \in \mathbb{F}_2^s$. Then, with no extra condition on f_1, f_2, g_1 and g_2, the function f satisfies the hypothesis of Theorem 3 and its dual is

$$\widetilde{f}(a,b) = \widetilde{f_1}(a) \oplus \widetilde{g_1}(b) \oplus (\widetilde{f_1} \oplus \widetilde{f_2})(a)(\widetilde{g_1} \oplus \widetilde{g_2})(b).$$

We see that $\widetilde{f}$ can be obtained from $\widetilde{f_1}, \widetilde{f_2}, \widetilde{g_1}$ and $\widetilde{g_2}$ exactly in the same manner as f is obtained from f_1, f_2, g_1 and g_2.

5. X.-D. Hou and P. Langevin have made in [145] a very simple observation which leads to potentially new bent functions:

[25] f is then 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 $g_2(y)$.

Proposition 25 *Let f be a Boolean function on $\mathbb{F}_2^n$, n even. Let σ be a permutation on $\mathbb{F}_2^n$. Denote its coordinate functions by $\sigma_1, \ldots, \sigma_n$. Assume that*

$$d_H(f, \bigoplus_{i=1}^{n} a_i\,\sigma_i) = 2^{n-1} \pm 2^{n/2-1} / \; \forall a \in \mathbb{F}_2^n.$$

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 $i = 1, \ldots, n$, 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 [143] 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 $\sigma_1, \ldots, \sigma_n$ 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. A binary secondary construction without extension of the number of variables was introduced in [61]. It is based on the following result:

Proposition 26 *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:*

$$\widehat{\mathrm{f_1}_\chi} + \widehat{\mathrm{f_2}_\chi} + \widehat{\mathrm{f_3}_\chi} = \widehat{\mathrm{s_1}_\chi} + 2\,\widehat{\mathrm{s_2}_\chi}. \tag{45}$$

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 (45). ⋄

Proposition 26 leads to the following double construction of bent functions:

Corollary 4 *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 at 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 (45) implies:

$$\begin{aligned}\widehat{\mathrm{s_2}_\chi}(a) &= \left[(-1)^{\tilde{f_1}(a)}+(-1)^{\tilde{f_2}(a)}+(-1)^{\tilde{f_3}(a)}-(-1)^{\tilde{f_1}(a)\oplus\tilde{f_2}(a)\oplus\tilde{f_3}(a)}\right]2^{\frac{n-2}{2}}\\ &= (-1)^{\tilde{f_1}(a)\tilde{f_2}(a)\oplus\tilde{f_1}(a)\tilde{f_3}(a)\oplus\tilde{f_2}(a)\tilde{f_3}(a)}\,2^{n/2}.\end{aligned}$$

- 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 (45), 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. ⋄

6.4.3 Decompositions of bent functions

The following theorem, proved in [33], is a direct consequence of Relation (25), applied to $f\oplus\ell$ where ℓ is linear, and to a 1-dimensional subspace 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 4 *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 [33]) 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 [33] 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 [33, 34, 82].

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 usual constructions.

The number of bent functions of the form (43) equals $(2^{n/2})! \times 2^{2^{n/2}}$, and 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, that equals $2^n(2^n-1)(2^n-2)\dots(2^n-2^{n-1})$. 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 at Subsection 6.3 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*. We know that for $n = 6$ (the highest number of variables for which the number of bent functions is known), the number of bent functions is approximately equal to 2^{32} (see [225]), which is much less than $2^{2^5+\frac{1}{2}\binom{6}{3}} = 2^{42}$. Also, it has been checked experimentally that there is no hope of obtaining a bent function on 8 variables by just picking at random a Boolean function of algebraic degree upper bounded by 4 (but more clever methods exist, see [97, 68]). An upper bound improving upon the naive bound has been found recently [75]. 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.

6.6 Characterizations

Proposition 27 *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 $N = \{1,\ldots,n\}$) 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 (27) applied with $\varphi = f$, Conditions *1.* and *2.* imply that f is bent.
Conversely, Condition *1.* is necessary, according to Proposition 21. Condition *2.* is also necessary since $\widehat{f}(1,\ldots,1) = (-1)^n\lambda_N$ (from Relation (27)). ◇

Proposition 27 and Relation (6) imply some restrictions on the coefficients of the ANFs of bent functions, observed and used in [75] (and also partially observed by Hou and Langevin in [145]).
Rothaus' bound is a direct consequence of Proposition 27 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. Proposition 27 also permits to prove the following characterization:

Theorem 5 *[71] 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 $E_1, \dots, E_k$ of $\mathbb{F}_2^n$ (there is no constraint on the number k) and integers $m_1, \dots, m_k$ (positive or negative) such that, for any element x of $\mathbb{F}_2^n$:*

$$f(x) \equiv \sum_{i=1}^{k} m_i 1_{E_i}(x) - 2^{n/2-1}\delta_0(x) \quad \left[\; mod \; 2^{n/2} \right]. \tag{46}$$

Proof (sketch of). Relation (46) is a sufficient condition for f being bent, according to Lemma 2 and to Relation (13).
Conversely, if f is bent, then Proposition 27 permits to deduce Relation (46), by expressing all the monomials x^I by means of the indicators of subspaces (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. ◇

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

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

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

The class of those functions f which satisfy the relation obtained from (46) by withdrawing "[mod $2^{n/2}$]" is denoted by $\mathcal{GPS}$. The dual $\widetilde{f}$ of such function f of $\mathcal{GPS}$ equals $\widetilde{f}(x) = \sum_{i=1}^{k} m_i 1_{E_i^\perp}(x) - 2^{n/2-1}\delta_0(x)$.

There is no uniqueness of the representation of a given bent function in the form (46). But there exists another characterization:

Theorem 6 *[72] Let f be a Boolean function on $\mathbb{F}_2^n$. Then f is bent if and only if there exist vector subspaces $E_1, \dots, E_k$ of $\mathbb{F}_2^n$ of dimensions $n/2$ or $n/2 + 1$ and integers $m_1, \dots, m_k$ (positive or negative) such that for any element x of $\mathbb{F}_2^n$:*

$$f(x) = \sum_{i=1}^{k} m_i 1_{E_i}(x) \pm 2^{n/2-1}\delta_0(x). \tag{47}$$

There is not a unique way, either, to choose these spaces E_i among all $n/2$-dimensional and $(n/2+1)$-dimensional vector subspaces of $\mathbb{F}_2^n$. 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 later [122] that, up to composition by a mapping $x \mapsto x + a$, every bent function belongs to $\mathcal{GPS}$.

A characterization of bent functions through Cayley graphs also exists, see [14].

6.7 Subclasses: hyper-bent functions

in [47] (see also [51, 52]) 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 $x_1, \ldots, x_n$, is bent (*i.e.* those functions which satisfy the propagation criterion of degree $n-k$ and order k, see Section 8). These functions (which were called hyper-bent in [47], but we will keep this term for a notion introduced by Youssef and Gong ; see below) are the four symmetric bent functions (see Section 9).

In [263], A. Youssef and G. Gong study the Boolean functions f on the field $\mathbb{F}_{2^n}$ (n even) whose Hamming distances to all functions $tr(a\,x^i) \oplus \varepsilon$ ($a \in \mathbb{F}_{2^n}$, $\varepsilon \in \mathbb{F}_2$), where tr denotes the trace function from $\mathbb{F}_{2^n}$ to $\mathbb{F}_2$ and where i is co-prime with 2^n-1, equal $2^{n-1} \pm 2^{n/2-1}$. These functions are bent, since every affine function has the form $tr(a\,x) \oplus \varepsilon$. They are called *hyper-bent functions*. The (equivalent) condition that $\sum_{x \in \mathbb{F}_{2^n}} (-1)^{f(x) \oplus tr(a\,x^i)}$ equals $\pm 2^{n/2}$ for every $a \in \mathbb{F}_{2^n}$ and every i co-prime with $2^n - 1$, seems difficult to satisfy, since it is equivalent to the fact that the function $f(x^i)$ is bent for every such i. However, A. Youssef and G. Gong show in [263] that hyper-bent functions exist. Their result is equivalent to the following (see [68]):

Proposition 28 *All the functions of class $\mathcal{PS}_{ap}$ are hyper-bent.*

Let us give here a direct proof of this fact.
Proof. 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 $tr(y) = y + y^2 + \cdots + y^{2^{n/2-1}} + y + y^2 + \cdots + y^{2^{n/2-1}} = 0$. Consider the inner product in $\mathbb{F}_{2^n}$ defined by: $y \cdot y' = tr(y\,y')$; the subspace $\mathbb{F}_{2^{n/2}}$ is then its own orthogonal;

hence, according to Relation (13), any sum of the form $\sum_{y\in\mathbb{F}_{2^{n/2}}}(-1)^{tr(\lambda y)}$ is null if $\lambda\notin\mathbb{F}_{2^{n/2}}$ and equals $2^{n/2}$ if $\lambda\in\mathbb{F}_{2^{n/2}}$.

Consider any element of the class $\mathcal{PS}_{ap}$, *i.e.* choose a balanced Boolean function g on $\mathbb{F}_2^{n/2}$, vanishing at 0, and define $f(y'+\omega\, y)=g\left(\frac{y'}{y}\right)$, with $\frac{y'}{y}=0$ if $y=0$. For every $a\in\mathbb{F}_{2^n}$, we have

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

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

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

The sum $\displaystyle\sum_{y'\in\mathbb{F}_{2^{n/2}}}(-1)^{g(0)\oplus tr(a\,y'^i)}$ equals $(-1)^{g(0)}\,2^{n/2}$ if $a\in\mathbb{F}_{2^{n/2}}$ and is null otherwise.

Thus, $\displaystyle\sum_{x\in\mathbb{F}_{2^n}}(-1)^{f(x)\oplus tr(a\,x^i)}$ equals:

$$\sum_{z\in\mathbb{F}_{2^{n/2}}}(-1)^{g(z)}\sum_{y\in\mathbb{F}_{2^{n/2}}}(-1)^{tr(a(z+\omega)^i\,y^i)}-\sum_{z\in\mathbb{F}_{2^{n/2}}}(-1)^{g(z)}+(-1)^{g(0)}\,2^{n/2}1_{\mathbb{F}_{2^{n/2}}}(a).$$

The sum $\sum_{z\in\mathbb{F}_{2^{n/2}}}(-1)^{g(z)}$ is null since g is balanced.

The sum $\sum_{z\in\mathbb{F}_{2^{n/2}}}(-1)^{g(z)}\sum_{y\in\mathbb{F}_{2^{n/2}}}(-1)^{tr(a(z+\omega)^i\,y^i)}$ equals $\pm 2^{n/2}$ if $a\notin\mathbb{F}_{2^{n/2}}$, since we prove in the next Lemma that there exists 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^i 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. ◇

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 that 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 [224] by B. Preneel and it has been proved later in [44].

Proposition 29 *Let n be any positive integer. Let f be any Boolean function on $\mathbb{F}_2^n$. 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:*

$$N_{\Delta_f} \times N_{\widehat{\mathrm{f}_\chi}} \geq 2^n. \tag{48}$$

Moreover, $N_{\Delta_f} \times N_{\widehat{\mathrm{f}_\chi}} = 2^n$ if and only if $D_b f$ is either balanced or constant, for every b. This is 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:

$$\forall x \in E, \forall y \in E', f(x+y) = g(x) \oplus h(y). \tag{49}$$

Inequality (48) comes directly from Relation (22): 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 (48). 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 in fact sufficient to imply that f has the form (49): E' is the linear kernel of f and the restriction of f

to E has balanced derivatives. Conversely, any function of the form (49) is such that Relation (48) is an equality. $\diamond$
Note that E' is then the linear kernel of the function.
The 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 [44]. In particular, the values of the Walsh transform equal 0 or $\pm 2^{dim(E')+dim(E)/2}$.

A generalization of Relation (48) has been obtained in [227]:

Proposition 30 *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 Cauchy's 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 30. $\diamond$

Partially-bent functions must not be mistaken for *partial bent functions*, studied by P. Guillot in [123]. The Fourier transforms of partial bent functions take exactly two values[26] 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 own dual is f. Two kinds of partial bent functions

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

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 (20). 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, at Subsection 4.1 (and sometimes called *three-valued functions*) is a natural extension of that of partially-bent functions, first studied by Zheng and Zhang in [269]. 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 (20), a plateaued function is bent if and only if its Walsh transform never takes the value 0. Also because of Parseval's relation, λ must be of the form 2^r where $r \geq n/2$. 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 that 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. Also, according to Parseval's relation, if we denote 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 a given n-variable Boolean function f, we have $N_{\widehat{\mathrm{f}_\chi}} \times \max_{a\in\mathbb{F}_2^n} \widehat{\mathrm{f}_\chi}^2(a) \geq 2^{2n}$ and therefore, according to Relation (31): $\mathcal{NL}(f) \leq 2^{n-1}\left(1 - \frac{1}{\sqrt{N_{\widehat{\mathrm{f}_\chi}}}}\right)$. Equality is achieved if and only if f is plateaued. Other properties of plateaued functions can be found in [33].

Plateaued functions can be characterized by second-order covering sequences (see [77]):

Theorem 7 *A Boolean function f on $\mathbb{F}_2^n$ is plateaued if and only if there exists θ such that, for every $x \in \mathbb{F}_2^n$:*

$$\sum_{a,b\in\mathbb{F}_2^n} (-1)^{D_aD_bf(x)} = \theta. \tag{50}$$

If this condition is satisfied, then the amplitude of the plateaued function f equals $\sqrt{\theta}$, and θ is therefore a power of 2 whose exponent is even and greater than or equal to n.

Proof. f satisfies (50) for a given vector x if and only if

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

Applying three times the inverse Fourier formula (16), we have

$$\sum_{a,b\in\mathbb{F}_2^n}(-1)^{f(x+a)\oplus f(x+b)\oplus f(x+a+b)}$$

$$=2^{-3n}\sum_{u,v,w,a,b\in\mathbb{F}_2^n}\widehat{\mathrm{f}_\chi}\,(u)\,\widehat{\mathrm{f}_\chi}\,(v)\,\widehat{\mathrm{f}_\chi}\,(w)\,(-1)^{(x+a)\cdot u\oplus(x+b)\cdot v\oplus(x+a+b)\cdot w}$$

$$=2^{-3n}\sum_{u,v,w,a,b\in\mathbb{F}_2^n}\widehat{\mathrm{f}_\chi}\,(u)\,\widehat{\mathrm{f}_\chi}\,(v)\,\widehat{\mathrm{f}_\chi}\,(w)\,(-1)^{x\cdot(u+v+w)\oplus a\cdot(u+w)\oplus b\cdot(v+w)}\,.$$

Since $\sum_{a\in\mathbb{F}_2^n}(-1)^{a\cdot(u+w)}$ is null if $u\neq w$, and $\sum_{b\in\mathbb{F}_2^n}(-1)^{b\cdot(v+w)}$ is null if $v\neq w$, we deduce that $\sum_{a,b\in\mathbb{F}_2^n}(-1)^{D_aD_bf(x)}=\theta$ if and only if

$$2^{-n}\sum_{u\in\mathbb{F}_2^n}\widehat{\mathrm{f}_\chi}^{\,3}\,(u)\,(-1)^{x\cdot u}=\theta(-1)^{f(x)}.$$

Hence, according to the inverse Fourier formula (16) again, Relation (50) is satisfied for every $x\in\mathbb{F}_2^n$ if and only if:

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

that is, if $\widehat{\mathrm{f}_\chi}(u)$ equals $\pm\sqrt{\theta}$ or 0 for every $u\in\mathbb{F}_2^n$. $\diamond$

The fact that quadratic functions are plateaued is a direct consequence of Theorem 7, since their second-order derivatives are constant. And Theorem 7 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 [172] 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 [128] with the "urcoset" terminology, and studied in [21]. 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 $\mathcal{NL}(g)$ (that is, of minimum weight in $g \oplus R(1,n)$), and in $f \oplus R(1,n)$ an element f_1 of weight $\mathcal{NL}(f)$, 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

The definition of normality has been given at Definition 3. As observed in [45] (see Theorem 2 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 7.3.1).

A proof of the existence of non-(weakly)-normal bent functions, *i.*e. bent functions which are non-constant (non-affine) on every $n/2$-dimensional flat, has been obtained recently (see [36]), contradicting a conjecture made by several authors that such bent function did not exist. Other non-normal bent functions have been found in [68]. But cubic bent functions on 8 variables are all normal, as shown in [82].

The stability of the class of non-normal bent functions with respect to the construction of functions called direct sum (see Subsection 7.3.2) has been studied in [66].

6.10 Kerdock codes

For every even n, the *Kerdock code* $\mathcal{K}_n$ [158] 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). 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 $R(1,n)$. The size of $\mathcal{K}_n$ equals 2^{2n}. This is the best possible size for such minimum distance (see [99]). We recall now briefly how the construction of Kerdock codes can be simply described.

6.10.1 Construction of the Kerdock code

The function

$$f(x) = \bigoplus_{1\leq i<j\leq n} x_i x_j \tag{51}$$

(which can also be defined as $f(x) = \binom{w_H(x)}{2}$ [mod 2]) is bent because the kernel of it associated symplectic form $\varphi(x,y) = \bigoplus_{1\leq i\neq j\leq n} x_i y_j$ is $\{0\}$. 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 at Subsection 2.1 some properties of the field $\mathbb{F}_{2^m}$ with 2^m elements. Other properties of this field are the following:

- there exists $\alpha\in\mathbb{F}_{2^m}$ such that $\mathbb{F}_{2^m} = \{0,\alpha,\alpha^2,\ldots,\alpha^{2^m-1}\}$ (α is called a *primitive element*);
- moreover, there exists α, primitive element, such that $(\alpha,\alpha^2,\alpha^{2^2},\ldots,\alpha^{2^{m-1}})$ is a basis of the vectorspace $\mathbb{F}_{2^m}$ (called a *normal basis*);
- if m is odd, then there exists a self-dual normal basis, that is a normal basis such that: $tr(\alpha^{2^i+2^j})=1$ if $i=j$; and $tr(\alpha^{2^i+2^j})=0$ otherwise, where tr is the trace function.

Consequence: $\forall x = x_1\alpha+\cdots+x_m\alpha^{2^{m-1}}\in\mathbb{F}_{2^m}$,

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

(where the indices are computed modulo m).

The function f of Relation (51), viewed as a function on $\mathbb{F}_{2^m}\times\mathbb{F}_2$, where $m=n-1$ is odd – say $m=2t+1$ – can now be written as:

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

Notice that the associated symplectic form associated to f equals $f(x,x_n)\oplus f(y,y_n)\oplus f(x+y,x_n\oplus y_n) = tr(\sum_{j=1}^t(x^{2^j}y+xy^{2^j}))\oplus x_n tr(y)\oplus y_n tr(x) = tr(\sum_{j=1}^t(x^{2^j}y+x^{2^{m-j}}y))\oplus x_n tr(y)\oplus y_n tr(x) = tr(x)tr(y)\oplus tr(xy)\oplus x_n tr(y)\oplus$

$y_n tr(x)$.

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 $tr(ux)tr(uy) \oplus tr(u^2xy) \oplus x_n tr(uy) \oplus y_n tr(ux) = tr(vx)tr(vy) \oplus tr(v^2xy) \oplus x_n tr(vy) \oplus y_n tr(vx)$ for every y, that is, $utr(ux) + u^2x + x_n u = vtr(vx) + v^2x + x_n v$ and $tr(ux) = tr(vx)$; it is a simple matter to show that it equals $\{(0, 0)\}$.

Open problem: Other examples of codes having the same parameters exist [153]. 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.

Remarks.
1. 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 [243], the weight enumerators of the Kerdock codes satisfy a relation similar to Relation (30), 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 [126]: 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). Such an explanation had been an open problem for two decades.
2. Another example of quadratic bent function whose definition uses two trace functions, the trace function tr_n on the whole field $\mathbb{F}_{2^n}$ and the trace function $tr_{n/2}$ on the subfield $\mathbb{F}_{2^{n/2}}$, is: $f(x) = tr_n(\sum_{i=1}^t x^{2^i+1}) \oplus tr_{n/2}(x^{2^{n/2}+1})$, $t = n/2 - 1$.

7 Resilient functions

We have seen at Subsection 4.1 that combining functions in stream ciphers must be m-resilient with large m. But, 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 the number of variables, the algebraic degree, the nonlinearity and the resiliency order of a function.

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[27]. This can be proved by using Relation (3) and the original definition of resiliency given by Siegenthaler, 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. Instead of this original Siegenthaler's definition, we can also use its characterization by Xiao and Massey, recalled in Proposition 13, together with Relation (15) 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 Siegenthaler's bound is also a direct consequence of a characterization of resilient functions[28] through their NNFs and of the fact that the algebraic degrees of Boolean functions are smaller than or equal to their numerical degrees:

Proposition 31 *[74] 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 $n-m-1$.*

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 $a + (1, \ldots, 1)$. We have $\widehat{\mathrm{f}_\chi}(a) = \widehat{\mathrm{g}_\chi}(\overline{a})$. Thus, f is m-resilient if and only if, for each word u of weight greater than or equal to $n-m$, the number $\widehat{\mathrm{g}_\chi}(u)$

[27] 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.

[28] A similar characterization of correlation-immune functions can be found in [54].

is null. Consider the NNF of g:

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

According to Relations (27) and (28) applied to g, we have for nonzero u:

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

and for nonempty I:

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

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

Thus, according to Relation (5), f is m-resilient if and only if the function $g(x) = f(x) \oplus x_1 \oplus \cdots \oplus x_n$ satisfies $\sum_{x\in\mathbb{F}_2^n \,|\, supp(x)\subseteq I} (-1)^{w_H(x)} g(x) = 0$, for all $I \in \mathcal{P}(N)$ of size at least $n-m$.

Proposition 31 has been used by X.-D. Hou in [141] for constructing resilient functions. Siegenthaler's bound gives an example of the trade-offs which must be accepted in the design of combiner generators[29]. Sarkar and Maitra showed in [236] 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 Proposition 32)[30]. 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) permits also to deduce Siegenthaler's bound, thanks to Proposition 9 applied with $k = m+2$ (resp. $k = m+1$).

7.2 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

[29] 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 [205].

[30] More is proved in [54, 78]; in particular that, 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}.

functions, independently obtained by Tarannikov [253] and by Zheng and Zhang [272]: their nonlinearity is upper bounded by $2^{n-1} - 2^{m+1}$. This bound is tight, at least when $m \geq 0.6\ n$, see [253, 254][31]. 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 [253] that such function f also achieves Siegenthaler's bound (and as proved in [190], achieves minimum sum-of-squares indicator). These last properties can also be deduced from a more precise divisibility bound shown later in [54]:

Proposition 32 *Let f be any n-variable m-resilient function 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 [54]). Later, a second proof using only the properties of the Fourier transform was given in [78]:
Proof. Relation (15) 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 words of weight $m+1$ (since it is obvious for the words 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 [78]).
Proposition 32 gives directly a more precise upper bound on the nonlinearity of any m-resilient function of degree d: this nonlinearity is upper bounded by $2^{n-1} - 2^{m+1+\lfloor \frac{n-m-2}{d} \rfloor}$. 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

[31] Also Zheng and Zhang [272], 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. [168].

n-variable Boolean function is upper bounded by $2^{n-1} - 2^{m+1+\lfloor \frac{n-m-2}{d} \rfloor}$ 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 best possible nonlinearity $2^{n-1} - 2^{n/2-1}$ of all Boolean functions) then, obviously, a better bound exists. In the case of n even, the best possible nonlinearity of all balanced functions being smaller than $2^{n-1} - 2^{n/2-1}$, Sarkar and Maitra deduce that $\mathcal{NL}(f) \leq 2^{n-1} - 2^{n/2-1} - 2^{m+1}$ for every m-resilient function f with $m \leq n/2 - 2$. In the case of n odd, they state that $\mathcal{NL}(f)$ 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 evenness 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 (20) and the fact that $\widehat{\mathrm{f}_\chi}(a)$ is null for every word a of weight $\leq m$ imply

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

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 $\max_{a \in \mathbb{F}_2^n} |\varphi(a)| \geq \left\lceil \frac{2^{n-m-2}}{\sqrt{2^n - \sum_{i=0}^m \binom{n}{i}}} \right\rceil$ (where $\lceil u \rceil$ denotes the smallest integer greater than or equal to u), and this implies:

$$\mathcal{NL}(f) \leq 2^{n-1} - 2^{m+1} \left\lceil \frac{2^{n-m-2}}{\sqrt{2^n - \sum_{i=0}^m \binom{n}{i}}} \right\rceil. \qquad (52)$$

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 (52) is 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 [236], because the values of n they consider in this table are small).

When n is odd, it is difficult to say if Inequality (52) 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 (52) better usable than their bound.

We know (see [187], page 310) that $\sum_{i=0}^{m}\binom{n}{i} \geq \frac{2^{nH_2(m/n)}}{\sqrt{8m(1-n/m)}}$, where $H_2(x) = -x\log_2(x)-(1-x)\log_2(1-x)$ is the so-called entropy function and satisfies $H_2(\frac{1}{2}-x) = 1-2x^2\log_2 e + o(x^2)$. Thus, we have

$$\mathcal{NL}(f) \leq 2^{n-1}-2^{m+1}\left\lceil \frac{2^{n-m-2}}{\sqrt{2^n-\frac{2^{nH_2(m/n)}}{\sqrt{8m(1-m/n)}}}} \right\rceil. \tag{53}$$

Maximum correlation An upper bound on the maximum correlation of m-resilient functions with respect to subsets I of N can be directly deduced from Relation (35) and from Sarkar et al.'s bound. Note that we get an improvement by using that the support of $\widehat{\mathrm{f}_\chi}$, 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 $2^{-n}\,|\widehat{\mathrm{f}_\chi}(u)|$, where u is the vector of support I, see [28, 29, 38, 264]. The optimal number of LFSRs that should be considered together in a correlation attack on a cryptosystem using an m-resilient combining function is $m+1$, see [28, 29].

Other criteria The relationships between resiliency and other criteria have been studied in [83, 190, 256, 271]. For instance, m-resilient $PC(l)$ functions can exist only if $m+l \leq n-1$. This is a direct consequence of Relation (24), 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 $n-m$. And equality is possible only if $l = n-1$, n is odd and $m = 0$ [271, 83]. 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 (or maybe should we write "specified") in [19]. This has the advantage of relaxing some of the trade-offs recalled above.

7.3 Constructions

High order resilient functions with high degrees and high nonlinearities are needed for applications in stream ciphers. But designing constructions of Boolean functions meeting these cryptographic criteria is still a crucial challenge nowadays. The primary constructions (which permit to design resilient functions without using known ones) lead potentially to wider classes of functions than secondary (*i*.e. recursive) 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). Unfortunately, the known primary constructions of such Boolean functions [49] 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 at Subsection 4.1. So we observe some imbalance in the knowledge on cryptographic functions for stream ciphers: after the results recently published [235, 236, 54, 78], 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 [217, 235, 236] and for every $m \geq 0.6\ n$ [253, 254] (n being then not limited). But 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 [83, 190]). Numerous examples of (balanced) Boolean functions with high nonlinearities have been obtained by C. Fontaine in [113] and by E. Filiol and C. Fontaine in [112], who made a computer investigation, for $n = 7$, 9, on the corpus of *idempotent functions*. These functions are those whose ANFs are invariant under the cyclic shifts of the coordinates x_i. They found new weight distributions of cosets of $R(1,7)$, with (optimum) minimum weight 56. They also obtained numerous weight distributions of cosets of $R(1,9)$, with (best known) minimum weight 240. Other works are also interesting, see *e.g.* [195, 191, 218].
But designing constructions leading to large numbers of functions achieving good trade-offs between the nonlinearity, the algebraic degree and the resiliency order (if possible, on any numbers of variables) are still necessary for permitting to choose in applications cryptographic functions satisfying

specific constraints.

7.3.1 Primary constructions

Maiorana-McFarland's class: An extension of Maiorana-McFarland's original class of bent functions has been given in [25], based on the same principle of concatenating affine functions (we have already met this generalization at Section 6): Let r be a positive integer smaller than n; 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, define the function:

$$f_{\phi,g}(x,y) = x \cdot \phi(y) \oplus g(y) = \bigoplus_{i=1}^{r} x_i \phi_i(y) \oplus g(y),\ x \in \mathbb{F}_2^r,\ y \in \mathbb{F}_2^s \qquad (54)$$

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

Remark. These functions have also been studied under the name of linear-based functions in [1, 262].
For every $a \in \mathbb{F}_2^r$ and every $b \in \mathbb{F}_2^s$, we have seen at Subsection 6.4 that

$$\widehat{\mathrm{f}_{\chi\,\phi,g}}(a,b) = 2^r \sum_{y \in \phi^{-1}(a)} (-1)^{g(y) \oplus b \cdot y}. \qquad (55)$$

The extension of *Maiorana-McFarland construction* 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. This is a direct consequence of Relation (55). It 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).

Degree: The degree of $f_{\phi,g}$ is at most $s+1 = n-r+1$. It equals $s+1$ if and only if ϕ has degree s (*i*.e. if at least one of its coordinate functions has 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 degree s only if $k \leq r-2$, since if $k = r-1$ then ϕ is constant. Thus, if $m = k$ then the degree of $f_{\phi,g}$ reachs Siegenthaler's bound $n-m-1$ if and only if either $m = r-2$ and ϕ has degree $s = n-m-2$ or $m = r-1$ and g has degree $s = n-m-1$. There are cases where $m > k$ (see [90, 56]).

Nonlinearity: Relations (31) and (55) lead straightforwardly to a general lower bound on the nonlinearity of Maiorana-McFarland's functions (first observed in [241]):

$$\mathcal{NL}(f_{\phi,g}) \geq 2^{n-1} - 2^{r-1} \max_{a\in\mathbb{F}_2^r} |\phi^{-1}(a)| \tag{56}$$

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

$$\mathcal{NL}(f_{\phi,g}) \leq 2^{n-1} - 2^{r-1} \left\lceil \sqrt{\max_{a\in\mathbb{F}_2^r} |\phi^{-1}(a)|} \right\rceil \tag{57}$$

obtained in [55] strengthens the bound $\mathcal{NL}(f_{\phi,g}) \leq 2^{n-1} - 2^{r-1}$ previously obtained in [84, 85].

Proof. The sum $\sum_{b\in\mathbb{F}_2^s} \left(\sum_{y\in\phi^{-1}(a)} (-1)^{g(y)+b\cdot y}\right)^2$, that is clearly equal to $\sum_{b\in\mathbb{F}_2^s} \left(\sum_{y,z\in\phi^{-1}(a)} (-1)^{g(y)+g(z)+b\cdot(y+z)}\right)$, 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

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

and thus, according to Relation (55):

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

Hence, according to Relation (31): $\mathcal{NL}(f_{\phi,g}) \leq 2^{n-1} - 2^{r-1} \left\lceil \sqrt{\max_{a\in\mathbb{F}_2^r} |\phi^{-1}(a)|} \right\rceil$. $\diamond$

This new bound permitted to characterize 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 $\mathcal{NL}(f_{\phi,g}) \leq 2^{n-1} - \frac{2^{r+\frac{s}{2}-1}}{\sqrt{\sum_{i=k+1}^{r}\binom{r}{i}}}$ implies either that $r = k+1$ or $r = k+2$.

If $r = k+1$, then ϕ is the constant $(1,\dots,1)$ and $n \leq k+3$. Either $s = 1$ and $g(y)$ is then any function on 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 $\mathcal{NL}(f_{\phi,g}) = 2^{n-1} - 2^{k+1}$ (and thus achieving Sarkar et al.'s bound) can be given for any $r \geq 2^s - 1$ and for $k = r - 2$ (see [55]). 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 $2^{n-1} - 2^{\frac{n-1}{2}}$).
in [55] are also studied functions $f_{\phi,g}$, such that $\phi(\mathbb{F}_2^s)$ is included in $\{x \in \mathbb{F}_2^n;\, w_H(x) > k\}$, whose resiliency orders are strictly greater than k.

Generalizations of Maiorana-McFarland's construction have been introduced in [55], [59] and [77]. 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 (55); there is also a risk that this property can be used in attacks, as it is used in [39] to attack block ciphers). The functions constructed in [55, 77] 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 $n = r + s$, $t = \lfloor \frac{r}{2} \rfloor$, 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 $n = r+s$, 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 equals $\left[(2^r)^{2^s}\right]^3 \times 2^{2^s} = 2^{(3r+1)2^s}$ and is larger than the size $(2^t)^{2^s} \times (2^r)^{2^s} \times 2^{2^s} = 2^{(t+r+1)2^s}$ of the first class.
There exist formulae for the Walsh transforms of the functions of these two classes, which result in sufficient conditions for their resiliency and in bounds

on their nonlinearities.
The second construction has been generalized in [59]. The functions of this generalized class are the following 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 $n=r+s$, φ is a function from $\mathbb{F}_2^s$ into $\{0,1,\ldots,r\}$, $\phi_1,\ldots,\phi_r$ 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 $\phi_1(y),\ldots,\phi_{\varphi(y)}(y)$ are linearly independent, and $g_1,\ldots,g_r$ and g are Boolean functions on $\mathbb{F}_2^s$.

Other constructions: We first make a *preliminary observation.* Let $k<n$ and let g be any k-variable function, $L:\mathbb{F}_2^n\rightarrow\mathbb{F}_2^k$ any surjective linear mapping and s any element 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 $g(x_1,\ldots,x_k)\oplus x_{k+1}$, 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 [49] to obtain resilient functions: let k and r be positive integers and $n\geq r$; 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 a $\mathbb{F}_2$-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. Then, the function

$$f(x,y)=g\left(\frac{x}{a+\phi(y)}\right)\oplus b\cdot y,\ \text{where } x\in\mathbb{F}_{2^r}, y\in\mathbb{F}_2^s, \tag{58}$$

is m-resilient with $m\geq k$. There exist bounds on the nonlinearities of these functions (see [56]), similar to those existing for Maiorana-McFarland's functions. But this class has much fewer elements than Maiorana-McFarland's class, because ϕ must be linear.

Dobbertin's construction: in [107]is given a nice generalization of a method, introduced by Seberry et al. in [242], 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 Wlash transform of the function $h(x,y) = f(x,y) \oplus \delta_0(y)g(x)$, where δ_0 is the Dirac symbol, satisfies:*

$$\widehat{\mathrm{h}_\chi}(u,v) = 0 \textit{ if } u = 0 \textit{ and } \widehat{\mathrm{h}_\chi}(u,v) = \widehat{\mathrm{f}_\chi}(u,v) + \widehat{\mathrm{g}_\chi}(u) \textit{ otherwise.} \qquad (59)$$

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 (41) applied to $E = \{0\} \times \mathbb{F}_2^{n/2}$ and $a = b = 0$ (or see the remark after Theorem 2). ◇

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 $2^n - 2\mathcal{NL}(h) \leq 2^n - 2\mathcal{NL}(f) + 2^{n/2} - 2\mathcal{NL}(g)$, that is:

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

Applying recursively this principle (if $n/2$ is even, g can be constructed in the same way), we see that if $n = 2^k\, n'$ ($n' \leq 13$ odd), the best known (but perhaps not the best possible) nonlinearity that can be obtained by using Dobbertin's method is $2^{n-1} - 2^{n/2-1} - 2^{\frac{n}{4}-1} - \ldots - 2^{\frac{n}{2^k}-1} - 2^{\frac{n'-1}{2}}$. Indeed, for every odd n', there exists a balanced (quadratic) function on $\mathbb{F}_2^{n'}$ with nonlinearity $2^{n'-1} - 2^{\frac{n'-1}{2}}$, and no balanced function with better nonlinearity is known if $n' \leq 13$.

Unfortunately, according to Relation (59), 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 $\pm 2^{n/2}$.

7.3.2 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.

I Direct sums of functions
A. *Adding a variable*
Let f be an r-variable t-resilient function. The Boolean function on $\mathbb{F}_2^{r+1}$:

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

is $(t+1)$-resilient [245]. If f is an $(r, t, r-t-1, 2^{r-1} - 2^{t+1})$ function[32], 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 $(0, \dots, 0, 1)$.
B. *Generalization*
If f is an r-variable t-resilient function and if g is an s-variable m-resilient function, then the function:

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

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 (31), $\mathcal{N}_h = 2^{r+s-1} - \frac{1}{2}(2^r - 2\mathcal{N}_f)(2^s - 2\mathcal{N}_g) = 2^r\mathcal{N}_g + 2^s\mathcal{N}_f - 2\mathcal{N}_f\mathcal{N}_g$. Such function does not give full satisfaction (J. Dillon already pointed out in [105] that such *decomposable functions* have weaknesses; their property can be used for designing divide-and-conquer attacks). Moreover, h has low degree, in general. And if $\mathcal{N}_f = 2^{r-1} - 2^{t+1}$ and $\mathcal{N}_g = 2^{s-1} - 2^{m+1}$ (*i.e.* if $\mathcal{N}_f$ and $\mathcal{N}_g$ have maximum possible values), then $\mathcal{N}_h = 2^{r+s-1} - 2^{t+m+3}$ 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 0-resilient, that is, balanced, but has nonlinearity 0, which is greater than $2^0 - 2^1$).
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 $\mathbb{F}_2^r$. Consider the function

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

[32] 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 $\mathcal{N}$.

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

1. If f and g are m-resilient, then h is m-resilient [245]; 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 [25]) is $g(x) = f(x_1 \oplus 1, \ldots, x_r \oplus 1) \oplus \epsilon$, where $\epsilon = m\,[\bmod 2]$, since $\widehat{\mathrm{g}_\chi}(a) = \sum_{x \in \mathbb{F}_2^r} (-1)^{f(x) \oplus \epsilon \oplus (x \oplus (1,\ldots,1)) \cdot a} = (-1)^{\epsilon + w_H(a)} \widehat{\mathrm{f}_\chi}(a)$. It leads to a function h having also a nonzero linear structure (namely, the vector $(1, \ldots, 1)$);

2. The maximum $\max_{a_1,\ldots,a_{r+1} \in \mathbb{F}_2} |\widehat{\mathrm{h}_\chi}(a_1, \ldots, a_r, a_{r+1})|$ is upper bounded by $\max_{a_1,\ldots,a_r \in \mathbb{F}_2} |\widehat{\mathrm{f}_\chi}(a_1, \ldots, a_r)| + \max_{a_1,\ldots,a_r \in \mathbb{F}_2} |\widehat{\mathrm{g}_\chi}(a_1, \ldots, a_r)|$; this implies $2^{r+1} - 2\mathcal{N}_h \leq 2^{r+1} - 2\mathcal{N}_f - 2\mathcal{N}_g$, that is $\mathcal{N}_h \geq \mathcal{N}_f + \mathcal{N}_g$;

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 word 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 $\max_{a_1,\ldots,a_{r+1} \in \mathbb{F}_2} |\widehat{\mathrm{h}_\chi}(a_1, \ldots, a_r, a_{r+1})| = \max\Big(\max_{a_1,\ldots,a_r \in \mathbb{F}_2} |\widehat{\mathrm{f}_\chi}(a_1, \ldots, a_r)|; \max_{a_1,\ldots,a_r \in \mathbb{F}_2} |\widehat{\mathrm{g}_\chi}(a_1, \ldots, a_r)|\Big)$. Hence we have $2^{r+1} - 2\mathcal{N}_h = 2^r - 2\min(\mathcal{N}_f, \mathcal{N}_g)$ and $\mathcal{N}_h$ equals therefore $2^{r-1} + \min(\mathcal{N}_f, \mathcal{N}_g)$; thus, if f and g achieve maximum 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 $a = (a_1, \ldots, a_r) \in \mathbb{F}_2^r$ and every $a_{r+1} \in \mathbb{F}_2$, we have, denoting $(x_1, \ldots, x_r)$ 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 permits to obtain:

- 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 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 degrees of f and g;
- from any m-resilient function f achieving 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 degree as f (but having nonzero linear structures).
So it permits, when combining these two methods, to keep best tradeoffs between resiliency order, 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).

III. Tarannikov's elementary construction
Let g be any Boolean function on $\mathbb{F}_2^r$. Define the Boolean function h on $\mathbb{F}_2^{r+1}$ by $h(x_1,\dots,x_r,x_{r+1})=x_{r+1}\oplus g(x_1,\dots,x_{r-1},x_r\oplus x_{r+1})$. For every $(a_1,\dots,a_{r+1})\in\mathbb{F}_2^{r+1}$, the Walsh transform $\widehat{\mathrm{h}_\chi}(a_1,\dots,a_{r+1})$ is equal to
$$\sum_{x_1,\dots,x_{r+1}\in\mathbb{F}_2}(-1)^{a\cdot x\oplus g(x_1,\dots,x_r)\oplus a_rx_r\oplus(a_r\oplus a_{r+1}\oplus 1)x_{r+1}},$$ where $a=(a_1,\dots,a_{r-1})$ and $x=(x_1,\dots,x_{r-1})$; it is null if $a_{r+1}=a_r$ and it equals $2\,\widehat{\mathrm{g}_\chi}(a_1,\dots,a_{r-1},a_r)$ if $a_r=a_{r+1}\oplus 1$. Thus:
1. $\mathcal{N}_h=2\,\mathcal{N}_g$;
2. If g is m-resilient, then h is m-resilient. If, additionally, $\widehat{\mathrm{g}_\chi}(a_1,\dots,a_{r-1},1)$ is null for every vector $(a_1,\dots,a_{r-1})$ of weight at most m, then h is $(m+1)$-resilient; 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. The condition that $\widehat{\mathrm{g}_\chi}(a_1,\dots,a_{r-1},1)$ is null for every vector $(a_1,\dots,a_{r-1})$ 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**).
3. $d^\circ f=d^\circ g$ if $d^\circ g\geq 1$.
4. h has the nonzero linear structure $(0,\dots,0,1,1)$.

Tarannikov combined in [253] this construction with the constructions recalled in **I** and **II**, to build a more complex secondary construction, which permits to increase in the same time the resiliency order and the 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 the use of Construction **I.A** gives then functions with nonzero linear structures). in [217], Pasalic et al. slightly modified this more complex Tarannikov's construction into a construction that we shall call *Tarannikov et al.'s construction*, which permitted, 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 $m > \frac{r}{2} - 2$.

Tarannikov et al.'s construction has been in its turn generalized (see [58]):

Theorem 8 *Let r, s, t and m be positive 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 $h(x, y) = f_1(x) \oplus g_1(y) \oplus (f_1 \oplus f_2)(x)\,(g_1 \oplus g_2)(y)$, $x \in \mathbb{F}_2^r, y \in \mathbb{F}_2^s$ 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*

$$\widehat{\mathrm{h}_\chi}(a, b) = \frac{1}{2}\widehat{\mathrm{f}_{1\chi}}(a)\left[\widehat{\mathrm{g}_{1\chi}}(b) + \widehat{\mathrm{g}_{2\chi}}(b)\right] + \frac{1}{2}\widehat{\mathrm{f}_{2\chi}}(a)\left[\widehat{\mathrm{g}_{1\chi}}(b) - \widehat{\mathrm{g}_{2\chi}}(b)\right]. \quad (60)$$

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

$$\mathcal{N}_h = \min_{i,j \in \{1,2\}} \left(2^{r+s-2} + 2^{r-1}\mathcal{N}_{g_j} + 2^{s-1}\mathcal{N}_{f_i} - \mathcal{N}_{f_i}\mathcal{N}_{g_j}\right). \quad (61)$$

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 $f_1 + f_2$ has degree $r - t - 1$ and $g_1 + g_2$ has 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.

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 $g_2(y)$. The proof of this theorem and examples of such pairs (f_1, f_2) (or (g_1, g_2)) can be found in [58].

IV. 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 $w_H(g) + w_H(h) = 2^{n-1}$. We assume now that this condition is satisfied. 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, $\mathcal{NL}(f) \geq \mathcal{NL}(g) + \mathcal{NL}(h) - 2^{n-1}$. The algebraic degree of f is upper bounded by (and can be equal to) the maximum of the algebraic degrees of g and h.

V. 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 [49].

VI. Proposition 26 leads to the following construction:

Corollary 5 *Let n be any positive integer and k any non-negative integer such that $k \leq n$. 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).*

Proof. Relation (45) 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 $\widehat{\mathrm{s1}_\chi}(a) = 0$ if and only if $\widehat{\mathrm{s2}_\chi}(a) = 0$. $\diamond$

Note that this secondary construction is proper to permit 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.

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 [163, 216].

7.4 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 (54) 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 k-resilient Maiorana-McFarland's functions obtained by choosing ϕ such that $w_H(\phi(y)) > k$ for every y equals $\left[2\sum_{i=k+1}^{2^r}\binom{2^r}{i}\right]^{2^{n-r}}$, and is probably also very small compared to the number of all k-resilient functions. But this number is unknown.

The exact numbers of m-resilient functions is known for $m \geq n-3$ (see [25], where $(n-3)$-resilient functions are characterized) and $(n-4)$-resilient functions have been partially characterized [63]. Asymptotic formulae for the numbers of k-resilient and k-th order correlation-immune functions, where $k = o(\sqrt{n})$, were given by O. Denisov in [101]. Also, Y. Tarannikov and D. Kirienko showed in [255] that, for every positive integer k, there exists a number $p(k)$ such that any $(n-k)$-resilient function $f(x_1,\ldots,x_n)$ is equivalent, up to permutation of its input coordinates, to a function of the form $g(x_1,\ldots,x_{p(k)}) \oplus x_{p(k)+1} \oplus \cdots \oplus x_n$. It is then a simple matter to deduce that the number of $(n-k)$-resilient functions is at most $A_k\, n^{p(k)}$, where A_k depends on k only. It is proved in [256] that $3 \cdot 2^{k-2} \leq p(k) \leq (k-1)2^{k-2}$ and in [255] that $p(4) = 10$.

in 1990, Yang and Guo published the first upper bound on first-order correlation-immune (and thus on resilient) functions. Park, Lee, Sung and Kim [219] proceeded further and improved upon Yang-Guo's bound. in 1995, Schneider [239] used a new idea to improve upon previous bounds and to generalize them to every m. He proved that the number of m-resilient n-variable Boolean functions is less than:

$$\prod_{i=1}^{n-m} \binom{2^i}{2^{i-1}}^{\binom{n-i-1}{m-1}}.$$

He also obtained bounds for the number of mth-order correlation-immune functions. 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 [75]. It implies an upper bound on the number of m-resilient functions which improves upon Schneider's bound for about half the values of (n, m) (it is better for m large). This bound divides the naive bound $2^{\sum_{i=0}^{n-m-1}\binom{n}{i}}$ 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 ($m \geq n/2 - 1$) improving upon Schneider's bound and partially improving upon this latter bound was obtained for $n/2 - 1 \leq m < n - 2$ in [70]: 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 [70].

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 [51, 52, 130] 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 $g(x_1 \oplus x_n, \dots, x_{n-1} \oplus x_n) \oplus \ell(x)$, 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 [52]).

The only known upper bound on the degrees of $PC(l)$ functions is $n - 1$. A lower bound on the nonlinearity of functions satisfying the propagation criterion exists [266] 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 $\mathcal{NL}(f) \geq 2^{n-1} - 2^{n-\frac{1}{2}l-1}$; Relation (24), applied to any $a \in \mathbb{F}_2^n$, with $b = 0$ and $E = F^\perp$, shows indeed that every value $\widehat{\mathrm{f}_\chi}^2(u)$ is upper bounded by 2^{2n-l}; it implies that $PC(l)$ functions have nonlinearities lower bounded 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 (35) and (24), see [28, 29].

8.1.1 Characterizations

There exist characterizations of the propagation criterion. A first obvious one is that, according to Relation (21), f satisfies $PC(l)$ if and only if $\sum_{u\in\mathbb{F}_2^n}(-1)^{a\cdot u}\,\widehat{\mathrm{f}_\chi}^2(u) = 0$ for every nonzero vector a of weight at most l. A second one is:

Proposition 34 *[52] Any n-variable Boolean function f satisfies $PC(l)$ if and only if, for every vector u of weight at least $n-\ell$, 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 (24). A third characterization is given at 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 (54) 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.
Constructions of such functions have been given in [51, 52, 167].

According to Proposition 34, 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 (59): $\widehat{\mathrm{h}_\chi}^2(w) = 0$ 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 35 *[225] A function f satisfies $EPC(l)$ (resp. $PC(l)$) of order k if and only if, for any word a of Hamming weight smaller than or equal to l and any word 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 [228] gives the following characterization:

Proposition 36 *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 (14) to the function $(a,b) \mapsto \widehat{D_a\mathrm{f}_\chi}(b)$.
Preneel showed in [223] that $SAC(k)$ functions have algebraic degrees at most $n-k-1$ (indeed, all of their restrictions have degrees at most $n-k-1$). In [184], 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 [167] (the mapping ϕ being linear and constructed from linear codes) and generalized in [51, 52] (the mapping ϕ being not linear and constructed from nonlinear codes). A construction of n-variable balanced functions satisfying $SAC(k)$ and having degree $n-k-1$ is given, for $n-k-1$ odd, in [167] and, for $n-k-1$ even, in [235] (where balancedness and nonlinearity are conjointly considered).
It is shown in [52] 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 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 $\{0, 1, \ldots, n\}$ 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, which 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 it is not clear whether a balance, between the advantage of allowing much more variables and the cryptographic weaknesses these symmetric functions may introduce, can be found in more sophisticated devices.

9.1 Representation

Let $r = 0, \ldots, n$ and let φ_r be the Boolean function whose support is the set of all words of weight r in $\mathbb{F}_2^n$. Then, according to Relation (5), the coefficient of x^I, $I \in \mathcal{P}(N)$ in the NNF of φ_r is:

$$\lambda_I = (-1)^{|I|-r} \left(\begin{array}{c} |I| \\ r \end{array} \right). \tag{62}$$

Note that the coefficient a_I of x^I in the ANF of φ_r equals then 1 if and only if $\left(\begin{array}{c} |I| \\ r \end{array} \right)$ is odd, that is, according to Lucas' theorem [187], if and only if the binary expansion of r is covered by the binary expansion of $|I|$.

The symmetric function f being equal to $\displaystyle\sum_{r=0}^{n} f^\#(r)\,\varphi_r$, its NNF is easy to compute. It can be also written in the form $\displaystyle\sum_{i=0}^{n} c_i\,\sigma_i(x)$ where $c_i \in \mathbb{Z}$ and

$\sigma_i(x)$ is the i-th elementary symmetric pseudo-Boolean function whose NNF is $\sum_{I\in\mathcal{P}(N)/\,|I|=i} x^I$. Hence, $\sigma_i(x)$ equals 1 if and only if $\begin{pmatrix} w_H(x) \\ i \end{pmatrix}$ is odd, that is, according to Lucas' theorem again, if and only if the binary expansion of i is covered by x. Notice that the degree of the NNF of σ_i being equal to i, the degree of the NNF of $\displaystyle\sum_{i=0}^{n} c_i\,\sigma_i(x)$ equals $\max\{i/\,c_i\neq 0\}$. We have clearly $\sigma_i(x) = \binom{w_H(x)}{i} = \frac{w_H(x)\,(w_H(x)-1)\dots(w_H(x)-i+1)}{i!}$. We see that $f^{\#}$ admits the polynomial representation $\sum_{i=0}^{n} c_i\binom{j}{i} = \sum_{i=0}^{n} c_i\frac{j\,(j-1)\dots(j-i+1)}{i!}$ on one variable j over $\mathbb{Z}$, 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.

Note that a symmetric function f has degree 1 if and only if the function $f^{\#}(r)$ equals r [mod 2] or $r+1$ [mod 2], and that it is quadratic if and only 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], that is, satisfies $f^{\#}(r+2) = f^{\#}(r)\oplus 1$.

It has been proved in [40] that 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. 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 $x_1,\dots,x_t$ only) or for the attacker.

9.2 Fourier and Walsh transforms

Since the functions φ_r have disjoint supports, the Fourier transform of any symmetric function $\displaystyle\sum_{r=0}^{n} f^{\#}(r)\,\varphi_r$ equals $\displaystyle\sum_{r=0}^{n} f^{\#}(r)\,\widehat{\varphi_r}$.
For every vector $a\in\mathbb{F}_2^n$, denoting by ℓ the Hamming weight of a, we have $\widehat{\varphi_r}(a) = \displaystyle\sum_{x\in\mathbb{F}_2^n\,|\,w_H(x)=r} (-1)^{a\cdot x} = \sum_{j=0}^{n}(-1)^j\binom{\ell}{j}\binom{n-\ell}{r-j}$. 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:

$$\sum_{r=0}^{n} K_{n,r}(\ell) z^r = (1-z)^\ell (1+z)^{n-\ell}.$$

From the Fourier transform, we can deduce the Walsh transform thanks to Relation (9).

9.3 Nonlinearity

If n is even, then the restriction of every symmetric function f on $\mathbb{F}_2^n$ to the $n/2$-dimensional flat:

$$A = \{(x_1, \dots, x_n) \in \mathbb{F}_2^n \; ; \; x_{i+n/2} = x_i \oplus 1, \forall i \leq n/2\}$$

is constant, since all the elements of A have the same weight $n/2$. Thus, f is $n/2$-normal[33] (see Definition 3). But Relation (37) gives nothing more than the covering radius bound (32). The symmetric functions which achieve this bound, *i.e.* which are bent, have been first characterized by P. Savicky in [238]: the bent symmetric functions are the function $f_1(x) = \bigoplus_{1 \leq i < j \leq n} x_i x_j$ (introduced to generate the Kerdock code), and the functions $f_2(x) = f_1(x) \oplus 1$, $f_3(x) = f_1(x) \oplus x_1 \oplus \cdots \oplus x_n$ and $f_4(x) = f_3(x) \oplus 1$. A stronger result can be proved in a very simple way [121]:

Theorem 9 *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: $x' = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$. 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. 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 [40].

If n is odd, then the restriction of any symmetric function f to the $\frac{n+1}{2}$-dimensional flat

$$A = \{(x_1, \dots, x_n) \in \mathbb{F}_2^n \; ; \; x_{i+\frac{n-1}{2}} = x_i \oplus 1, \forall i \leq n/2\}$$

[33] This is more generally valid for every function which is constant on the set $\{x \in \mathbb{F}_2^n ; \; w_H(x) = n/2\}$.

is affine, since the weight function w_H is constant on the hyperplane of A of equation $x_n = 0$ and on its complement[34]. Thus, f is $\frac{n+1}{2}$-weakly-normal. According to Relation (37), this implies that its nonlinearity is upper bounded by $2^{n-1} - 2^{\frac{n-1}{2}}$. It also permits to show [57] 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 [193], in a more complex way). Indeed, Relation (37) implies the following result:

Theorem 10 *[57] 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 $h_l(y_1, \ldots, y_{n-2l}) = f(x_1, \ldots, x_l, x_1 \oplus 1, \ldots, x_l \oplus 1, y_1, \ldots, y_{n-2l})$, where the values of $x_1, \ldots, x_l$ are arbitrary (equivalently, h_l can be defined by $h_l^{\#}(r) = f^{\#}(r+l)$, for every $0 \leq r \leq n-2l$). Then $\mathcal{NL}(f) \leq 2^{n-1} - 2^{n-l-1} + 2^l \mathcal{NL}(h_l)$.*

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 $x \in A$, we have $f(x) = h_l(x_{2l+1}, \ldots, x_n)$. Let us consider the restriction g of f to A as a Boolean function on $\mathbb{F}_2^{n-l}$, say $g(x_1, \ldots, x_l, x_{2l+1}, \ldots, x_n)$. Then, since $g(x_1, \ldots, x_l, x_{2l+1}, \ldots, x_n) = h_l(x_{2l+1}, \ldots, x_n)$, g has nonlinearity $2^l \mathcal{NL}(h_l)$. According to Relation (37) applied with $h_a = g$, we have $\mathcal{NL}(f) \leq 2^{n-1} - 2^{n-l-1} + 2^l \mathcal{NL}(h_l)$. $\diamond$

Then, the characterizations recalled above of those symmetric functions achieving best possible nonlinearity can be straightforwardly deduced. Note that these characetrizations imply that, if f is symmetric and not quadratic, then $\mathcal{NL}(f) \leq 2^{n-1} - 2^{\lfloor \frac{n-1}{2} \rfloor} - 1$. Moreover, if additionally, f has degree strictly smaller than n, then $\mathcal{NL}(f) \leq 2^{n-1} - 2^{\lfloor \frac{n-1}{2} \rfloor} - 2$ (indeed, since we have necessarily $n \geq 3$, the number $2^{n-1} - 2^{\lfloor \frac{n-1}{2} \rfloor}$ is even, and we know that $\mathcal{NL}(f)$ is then also even and is strictly smaller than this number). These properties applied to the function h_l of Theorem 10 imply that:

- if, for some integer l such that $0 \leq l < \lfloor \frac{n-1}{2} \rfloor$, the nonlinearity of an n-variable symmetric function f is strictly greater than $2^{n-1} - 2^{\lfloor \frac{n-1}{2} \rfloor} - 2^l$, then $f^{\#}$ satisfies $f^{\#}(r+2) = f^{\#}(r) \oplus 1$, for all $l \leq r \leq n-2-l$ (this property has been observed in [40, Theorem 6], but proved slightly differently);
- if the nonlinearity of f is strictly greater than $2^{n-1} - 2^{\lfloor \frac{n-1}{2} \rfloor} - 2^{l+1}$, then either $f^{\#}$ satisfies $f^{\#}(r+2) = f^{\#}(r) \oplus 1$ for all $l \leq r \leq n-2-l$, or h_l has

[34] This is more generally valid for every function which is constant on the sets $\{x \in \mathbb{F}_2^n;\ w_H(x) = \frac{n-1}{2}\}$ and $\{x \in \mathbb{F}_2^n;\ w_H(x) = \frac{n+1}{2}\}$.

odd weight.

Further properties of the nonlinearities of symmetric functions can be found in [57].

9.4 Resiliency

There exists a joint conjecture on symmetric Boolean functions and on functions defined over $\{0, 1, \dots, n\}$ and valued in $\mathbb{F}_2$: if f is a non-constant symmetric Boolean function, then the degree of the polynomial representation on one variable of $f^\#$ (which equals the numerical degree 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 31, 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, according to Relation (62) and to the equality $f = \sum_{r=0}^{n} f^\#(r)\, \varphi_r$, the numerical degree of f is upper bounded by d if and only if:

$$\forall k > d,\ k \leq n,\ \sum_{r=0}^{k} (-1)^r \binom{k}{r} f^\#(r) = 0. \tag{63}$$

Hence, the conjecture is equivalent to saying that Relation (63), with $d = n-4$, has no binary solution $f^\#(0), \dots, f^\#(n)$.
J. von zur Gathen and J. R. Roche [116] 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 [120]). They proved 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, all non-affine n-variable symmetric Boolean functions are unbalanced): indeed, the binomial coefficient $\binom{n}{r}$ being congruent with $\frac{(-1)(-2)\dots(-r)}{1\cdot 2 \dots r} = (-1)^r$, 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 (63) with $k = n$ implies then that $f^\#$ must be constant. Notice that, applying Relation (63) with $k = p-1$ where p is a prime less than or equal

to $n+1$, shows that the degree of any symmetric non-constant Boolean function is greater than or equal to $p-1$, where p is the largest prime less than or equal to $n+1$ (or equivalently, no symmetric non-affine Boolean function is $(n-p+1)$-resilient): otherwise, the string $f^{\#}(0), \ldots, f^{\#}(k)$ would be constant, and $f^{\#}$ having degree less than or equal to k, the function $f^{\#}$, and thus f itself, would be constant.
Some results on symmetric functions with sub-optimal nonlinearity and on the balancedness and resiliency of symmetric functions can be found in [40].

A super-class of symmetric functions, called idempotent or rotation symmetric functions, has been investigated with respect to the criteria of bentness and correlation immunity (see e.g. [112, 247]).

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 Math. Appl.* vol 4, no. 4, pp. 341-350, 1994.

[4] F. Armknecht. Improving fast algebraic attacks. *In Fast Software Encryption* 2004, no. 3017 in LNCS, pp. 65-82, 2004.

[5] F. Armknecht, C. Carlet, P. Gaborit, S. Knzli, W. Meier and O. Ruatta. Efficient computation of algebraic immunity for algebraic and fast algebraic attacks. Advances in Cryptology, EUROCRYPT 2006, Lecture Notes in Computer Science 4004 , pp. 147-164, 2006.

[6] E.F. Assmus. On the Reed-Muller codes. *Discrete Mathematics* 106/107, pp. 25-33, 1992.

[7] E.F. Assmus and J. D. Key. *Designs and their Codes*, Cambridge Univ. Press., Cambridge, 1992.

[8] J. Ax. Zeroes of polynomials over finite fields. *Amer. J. Math.* no. 86, pp. 255-261, 1964.

[9] T. Baignères, P. Junod and S. Vaudenay. How far can we go beyond linear cryptanalysis? *Proceedings of ASIACRYPT 2002*, Advances in Cryptology, LNCS 3329, pp. 432-450, 2004.

[10] B. Barak, G. Kindler, R. Shaltiel, B. Sudakov and A. Wigderson. Simulating Independence: New Constructions of Condensers, Ramsey Graphs, Dispersers, and Extractors. Preprint available at http://www.math.ias.edu/ boaz/Papers/BKSSW.html

[11] E. Berlekamp, *Algebraic Coding Theory,* McGraw-Hill, New York, 1968.

[12] 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.

[13] E.R. Berlekamp and L.R. Welch. Weight distributions of the cosets of the (32,6) Reed-Muller code. *IEEE Trans. Inform. Theory,* 18(1), pp. 203-207, 1972.

[14] 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.

[15] A. Bernasconi and I. Shparlinski. Circuit complexity of testing square-free numbers. *Pro. of STACS 99, 16th Annual Symposium on Theoretical Aspects of Computer Science*, 1999, Lecture Notes in Computer Science 1563, Springer, pp. 47-56, 1999.

[16] J. Bierbrauer, K. Gopalakrishnan and D.R. Stinson. Bounds for resilient functions and orthogonal arrays. *Advances in Cryptology - CRYPTO'94*, Lecture Notes in Computer Science 839, pp. 247-256, 1994.

[17] 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.

[18] J. Bourgain. On the construction of affine extractors. Preprint 2005.

[19] 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.

[20] E. Brier and P. Langevin. Classification of cubic Boolean functions of 9 variables. Proceedings of 2003 IEEE Information Theory Workshop, Paris, France, 2003.

[21] 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.

[22] P. Camion and A. Canteaut. Construction of t-resilient functions over a finite alphabet, *Advances in Cryptology, EUROCRYPT'96, Lecture Notes in Computer Sciences, Springer Verlag* no. 1070, pp. 283-293, 1996.

[23] P. Camion and A. Canteaut. Generalization of Siegenthaler inequality and Schnorr-Vaudenay multipermutations. *Advances in Cryptology - CRYPTO'96*, Lecture Notes in Computer Science no. 1109, pp. 372–386 Springer-Verlag, 1996.

[24] P. Camion and A. Canteaut. Correlation-immune and resilient functions over finite alphabets and their applications in cryptography. *Designs, Codes and Cryptography* 16, 1999.

[25] P. Camion, C. Carlet, P. Charpin, N. Sendrier. On correlation-immune functions, *Advances in Cryptology: Crypto '91, Proceedings, Lecture Notes in Computer Science*, vol. 576, pp. 86-100, 1991.

[26] 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.

[27] A. Canteaut. Cryptographic functions and design criteria for block ciphers. *Progress in Cryptology - INDOCRYPT 2001*, LNCS 2247, Springer-Verlag, pp. 1-16, 2001.

[28] A. Canteaut. Approximations of nonlinear combining functions in stream ciphers. Preprint, 2002

[29] A. Canteaut. On the correlations between a combining function and functions of fewer variables. Proceedings of the Information Theory Workshop'02, Bangalore, 2002.

[30] A. Canteaut. Open problems related to algebraic attacks on stream ciphers. Proceedings of WCC 2005, pp. 1-10, 2005.

[31] A. Canteaut. Analysis and design of symmetric ciphers. Habilitation for directing Theses, University of Paris 6, 2006.

[32] A. Canteaut, C. Carlet, P. Charpin and C. Fontaine. Propagation characteristics and correlation-immunity of highly nonlinear Boolean functions. *Proceedings of EUROCRYPT'2000, Advances in Cryptology, Lecture Notes in Computer Science* n 187, pp. 507-522 (2000)

[33] 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.

[34] A. Canteaut and P. Charpin. Decomposing bent functions. *In Proceedings 2002 IEEE International Symposium on Information Theory, Lausanne*, 2002. And *IEEE Transactions on Information Theory* 49, pp. 2004-2019, 2003.

[35] A. Canteaut and P. Charpin and G. Kyureghyan. A new class of monomial bent functions. Proceedings of the 2006 IEEE International Symposium on Information Theory - ISIT 2006, Seattle, USA, 2006.

[36] 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.

[37] A. Canteaut and E. Filiol. Ciphertext only reconstruction of stream ciphers based on combination generators. *In Fast Software Encryption* 2000, no. 1978 in LNCS, pp. 165-180. Springer-Verlag, 2001.

[38] 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.

[39] A. Canteaut and M. Videau. Degree of Composition of Highly Nonlinear Functions and Applications to Higher Order Differential Crypt-

analysis, *Advances in Cryptology, EUROCRYPT2002, Lecture Notes in Computer Science* 2332, Springer Verlag, pp. 518-533, 2002.

[40] A. Canteaut and M. Videau. Symmetric Boolean functions. *IEEE Transactions on Information Theory* 51(8), pp. 2791-2811, 2005.

[41] 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.

[42] 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.

[43] C. Carlet. A transformation on Boolean functions, its consequences on some problems related to Reed-Muller codes, *EUROCODE'90, G. Cohen, P. Charpin eds*, LNCS 514, Springer Verlag, pp. 42-50, 1991.

[44] C. Carlet. Partially-bent functions, *Designs Codes and Cryptography*, 3, pp. 135-145 (1993) and proceedings of CRYPTO' 92, Advances in Cryptology, Lecture Notes in Computer Science 740, Springer Verlag, pp. 280-291, 1993.

[45] C. Carlet. Two new classes of bent functions. In *Advances in Cryptology - EUROCRYPT'93*, no. 765 in Lecture Notes in Computer Science, pp. 77-101. Springer-Verlag, 1994.

[46] C. Carlet. Generalized Partial Spreads, *IEEE Transactions on Information Theory*, vol. 41, no. 5, pp. 1482-1487, 1995.

[47] C. Carlet. Hyper-bent functions. *PRAGOCRYPT'96, Czech Technical University Publishing House*, pp. 145-155, 1996.

[48] C. Carlet. A construction of bent functions. *Finite Fields and Applications, London Mathematical Society*, Lecture Series 233, Cambridge University Press, pp. 47-58, 1996.

[49] 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, Springer Verlag pp. 422-433, 1997.

[50] C. Carlet. On Kerdock codes, American Mathematical Society (Proceedings of the conference Finite Fields and Applications Fq4) Contemporary Mathematics 225, pp. 155-163, 1999.

[51] C. Carlet. On the propagation criterion of degree ℓ and order k. *Advances in Cryptology - EUROCRYPT'98*, no. 1403 in Lecture Notes in Computer Science, pp. 462-474. Springer-Verlag, 1998.

[52] C. Carlet. On cryptographic propagation criteria for Boolean functions. *Information and Computation* , vol. 151, Academic Press pp. 32-56, 1999.

[53] 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.

[54] 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, Springer, pp. 131-144, 2001.

[55] C. Carlet. A larger Class of Cryptographic Boolean Functions via a Study of the Maiorana-McFarland Construction. *Advances in Cryptology - CRYPT0 2002, no. 2442 in Lecture Notes in Computer Science*, pp. 549-564, 2002.

[56] 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.

[57] C. Carlet. On the degree, nonlinearity, algebraic thickness and nonnormality of Boolean functions, with developments on symmetric functions. *IEEE Transactions on Information Theory*, vol. 50, pp. 2178-2185, 2004.

[58] 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, K. Feng, H. Niederreiter and C. Xing Eds., pp. 3-28, 2004.

[59] C. Carlet. Concatenating indicators of flats for designing cryptographic functions. *Design, Codes and Cryptography* volume 36, Number 2, pp.189 - 202, 2005.

[60] C. Carlet. Partial covering sequences. Preprint.

[61] C. Carlet. On bent and highly nonlinear balanced/resilient functions and their algebraic immunities. Proceedings of AAECC 16, LNCS 3857, pp. 1-28, 2006.

[62] C. Carlet. On the higher order nonlinearities of algebraic immune functions. *Advances in cryptology–CRYPTO 2006, Lecture Notes in Computer Science* 4117, pp. 584-601, 2006..

[63] C. Carlet and P. Charpin. Cubic Boolean functions with highest resiliency. *IEEE Transactions on Information Theory*, vol. 51, no. 2, pp. 562-571, 2005.

[64] 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, July 2006.

[65] 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.

[66] 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.

[67] C. Carlet and S. Dubuc. On generalized bent and q-ary perfect nonlinear functions. D. Jungnickel and H. Niederreiter Eds. Proceedings of Finite Fields and Applications Fq5, Augsburg, Germany, Springer, pp. 81-94, 2000.

[68] C. Carlet and P. Gaborit. Hyper-bent functions and cyclic codes. To appear in the Journal of Combinatorial Theory, Series A, 2005.

[69] 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.

[70] C. Carlet and A. Gouget. An upper bound on the number of m-resilient Boolean functions. *Proceedings of ASIACRYPT 2002*, Advances in Cryptology, LNCS 2501, pp. 484-496, 2002.

[71] 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.

[72] 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.

[73] 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.

[74] 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.

[75] C. Carlet and A. Klapper. Upper bounds on the numbers of resilient functions and of bent functions. *Springer-Verlag, Lecture Notes dedicated to Philippe Delsarte* (to appear). A shorter version has appeared in the *Proceedings of the 23rd Symposium on Information Theory in the Benelux*, Louvain-La-Neuve, Belgian, 2002.

[76] C. Carlet and S. Mesnager. Improving the upper bounds on the covering radii of binary Reed-Muller codes. C. Carlet et S. Mesnager. To appear in IEEE Transactions on Information Theory, 2006.

[77] 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.

[78] C. Carlet and P. Sarkar. Spectral Domain Analysis of Correlation Immune and Resilient Boolean Functions. *Finite fields and Applications* 8, pp. 120-130, 2002.

[79] C. Carlet and Y. V. Tarannikov. Covering sequences of Boolean functions and their cryptographic significance. *Designs, Codes and Cryptography*, 25, pp. 263-279, 2002.

[80] 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.

[81] C. Charnes, M. Rötteler and T. Beth. Homogeneous bent functions, invariants, and designs. *Designs, Codes and Cryptography*, 26, pp. 139-154, 2002.

[82] 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.

[83] P. Charpin and E. Pasalic. On propagations characteristics of resilient functions. *Advances in Cryptology - SAC 2002, Lecture Notes in Computer Science* 2595, pages 356–365. Springer-Verlag, 2002.

[84] 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.

[85] S. Chee, S. Lee, D. Lee and S. H. Sung. On the correlation immune functions and their nonlinearity. *Proceedings of Asiacrypt'96*, LNCS 1163, pp. 232-243.

[86] B. Chor, O. Goldreich, J. Hastad, J. Freidmann, S. Rudich and R. Smolensky. The bit extraction problem or t-resilient functions. *Proc. 26th IEEE Symposium on Foundations of Computer Science*, pp. 396-407, 1985.

[87] G. Cohen, I. Honkala, S. Litsyn and A. Lobstein. *Covering codes.* North-Holland, 1997.

[88] N. Courtois. Fast Algebraic Attacks on Stream Ciphers with Linear Feedback. *Advances in cryptology–CRYPTO 2003, Lecture Notes in Computer Science* 2729, pp. 177-194, Springer, 2003.

[89] N. Courtois and W. Meier. Algebraic Attacks on Stream Ciphers with Linear Feedback. *Advances in cryptology–EUROCRYPT 2003, Lecture Notes in Computer Science* 2656, pp. 346-359, Springer, 2002.

[90] T. W. Cusick. On constructing balanced correlation immune functions. *Proceedings of SETA'98* (Sequences and their Applications 1998), Dis-

crete Mathematics and Theoretical Computer Science, Springer, pp. 184-190, 1999.

[91] T. W. Cusick, C. Ding and A. Renvall, *Stream Ciphers and Number Theory,* North-Holland Mathematical Library 55. Amsterdam: North-Holland/Elsevier, 1998.

[92] D.M. Cvetkovic, M. Doob and H. Sachs. *Spectra of graphs.* Academic Press, 1979.

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

[94] 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.

[95] 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. Cryptology ePrint Archive, http://eprint.iacr.org/, No. 2005/229, 15 July, 2005.

[96] 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.

[97] M. Daum, H. Dobbertin and G. Leander. Short description of an algorithm to create bent functions. Private communication.

[98] E. Dawson and C.-K. Wu. Construction of correlation immune Boolean functions. *Proceedings of ICICS 1997*, pp. 170-180, 1997.

[99] P. Delsarte. An algebraic approach to the association schemes of coding theory. PhD thesis. Université Catholique de Louvain (1973)

[100] P. Delsarte. Four fundamental parameters of a code and their combinatorial significance. *Information and Control*, vol. 23 (5), pp. 407-438, 1973.

[101] 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.

[102] F. Didier. A new upper bound on the block error probability after decoding over the erasure channel. Preprint available at http://www-rocq.inria.fr/codes/Frederic.Didier/
A revised version will appear in IEEE Transactions on Information Theory, 2006.

[103] J. Dillon. A survey of bent functions. *NSA Technical Journal Special Issue*, pp. 191-215, 1972.

[104] J. F. Dillon. Elementary Hadamard Difference sets. Ph. D. Thesis, Univ. of Maryland, 1974.

[105] J. F. Dillon. Elementary Hadamard Difference sets, *Proc. Sixth S-E Conf. Comb. Graph Theory and Comp.*, F. Hoffman et al. (Eds), Winnipeg Utilitas Math, pp. 237-249, 1975.

[106] J. F. Dillon and H. Dobbertin. New cyclic difference sets with Singer parameters. Finite Fields and Their Applications 10, pp. 342-389, 2004.

[107] H. Dobbertin. Construction of bent functions and balanced Boolean functions with high nonlinearity. *Fast Software Encryption, Second International Workshop*, Lecture Notes in Computer Science 1008, pp. 61-74, 1995.

[108] H. Dobbertin, G. Leander, A. Canteaut, C. Carlet, P. Felke and P. Gaborit. Construction of Bent Functions via Niho Power Functions. To appear in the *Journal of Combinatorial Theory, Series A*, 2005.

[109] S. Dubuc. Characterization of linear structures. *Designs, Codes and Cryptography* vol. 22, pp. 33-45, 2001.

[110] J. H. Evertse. Linear structures in block ciphers. In *Advances in Cryptology - EUROCRYPT' 87*, no. 304 in Lecture Notes in Computer Science, Springer Verlag, pp. 249-266, 1988.

[111] J.-C. Faugère and G. Ars. An Algebraic Cryptanalysis of Nonlinear Filter Generators using Gröbner bases. *Rapport de Recherche INRIA* 4739, 2003.

[112] E. Filiol and C. Fontaine. Highly nonlinear balanced Boolean functions with a good correlation-immunity. *Advances in Cryptology - EUROCRYPT'98*, no. 1403 in Lecture Notes in Computer Science, pp. 475-488. Springer-Verlag, 1998.

[113] C. Fontaine. On some cosets of the First-Order Reed-Muller code with high minimum weight. *IEEE Trans. Inform. Theory*, vol. 45 (4), pp. 1237-1243, 1999.

[114] R. Forré. The strict avalanche criterion: spectral properties of Boolean functions and an extended definition. *Advances in Cryptology – CRYPTO'88, LNCS 403, Springer-Verlag*, pp. 450-468, 1989.

[115] R. Forré. A fast correlation attack on nonlinearly feedforward filtered shift register sequences. *Advances in cryptology –EUROCRYPT '89, Lecture Notes in Comput. Sci.* 434, pp. 586-595, Springer, 1990.

[116] J. von zur Gathen and J. R. Roche. Polynomials with two values. *Combinatorica* 17(3), pp. 345-362, 1997.

[117] J. Golić. Fast low order approximation of cryptographic functions. *Advanced in Cryptology-EUROCRYPT'96. Lecture notes in computer science*, LNCS 1070, pp. 268-282, 1996.

[118] J. Golić. On the security of nonlinear filter generators. *Fast Software Encryption'96*, Lecture Notes in Computer Science 1039, pp. 173-188, 1996.

[119] S.W. Golomb. *Shift Register Sequences.* Aegean Park Press, 1982.

[120] 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.

[121] A. Gouget. On the propagation criterion of Boolean functions. Proceedings of the Workshop on Coding, Cryptography and Combinatorics 2003, published by Birkhäuser Verlag, K. Feng, H. Niederreiter and C. Xing Eds., pp. 153-168, 2004.

[122] P. Guillot. Completed GPS Covers All Bent Functions. *Journal of Combinatorial Theory*, Series A 93, pp. 242-260, 2001.

[123] P. Guillot. Partial bent functions. *Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics, SCI 2000*, 2000.

[124] Xiao Guo-Zhen, C. Ding and W. Shan. *The stability theory of stream ciphers*, vol. LNCS 561, Springer Verlag, 1991.

[125] Xiao Guo-Zhen and J. L. Massey. A Spectral Characterization of Correlation-Immune Combining Functions. *IEEE Trans. Inf. Theory*, vol. IT 34, no. 3, pp. 569-571, 1988.

[126] 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.

[127] T. Helleseth, T. Kløve, and J. Mykkelveit. On the covering radius of binary codes. *IEEE Trans. Inform. Theory*, IT-24(5), pp. 627-628, 1978.

[128] T. Helleseth and H.F. Mattson Jr. On the cosets of the simplex code. *Discr. Math.* 56, pp. 169-189, 1985.

[129] S. Hirose and K. Ikeda. Nonlinearity criteria of Boolean functions. *KUIS Technical Report*, KUIS-94-0002, 1994.

[130] S. Hirose and K. Ikeda. Complexity of Boolean functions satisfying the propagation criterion. *The Proc. of the 1995 Symposium on Cryptography and Information Security*, SCIS95-B3.3, (1995).

[131] 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.

[132] X.-D. Hou. Some results on the covering radii of Reed-Muller codes. *IEEE Trans. Inform. Theory*, vol. IT-39, no. 2, pp. 366-378, 1993.

[133] X.-D. Hou. Classification of cosets of the Reed-Muller code $R(m-3,m)$. *Discrete Math.*, 128, pp. 203-224, 1994.

[134] X.-D. Hou. The covering radius of $R(1,9)$ in $R(4,9)$. *Designs, Codes and Cryptography* 8 (3), pp. 285-292, 1995.

[135] X.-D. Hou. $AGL(m,2)$ acting on $R(r,m)/R(s,m)$. *Journal of Algebra* 171, pp. 921-938, 1995.

[136] 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.

[137] X.-D. Hou. $GL(m, 2)$ acting on $R(r, m)/R(r-1, m)$. *Discrete Math.* 149, pp. 99-122, 1996.

[138] X.-D. Hou. On the covering radius of $R(1, m)$ in $R(3, m)$. *IEEE Trans. Inform. Theory*, 42(3), pp. 1035-1037, 1996.

[139] X.-D. Hou. The Reed-Muller code $R(1, 7)$ is normal. *Designs, Codes and Cryptography* 12, pp. 75-82, 1997.

[140] X.-D. Hou. Cubic bent functions. *Discrete Mathematics* vol. 189, pp. 149-161, 1998.

[141] X.-D. Hou. On Binary Resilient Functions. *Des. Codes Cryptography* 28(1), pp. 93-112, 2003.

[142] X.-D. Hou. Group Actions on Binary Resilient Functions. *Appl. Algebra Eng. Commun. Comput.* 14(2), pp. 97-115, 2003.

[143] 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.

[144] X.D. Hou. On the coefficients of binary bent functions. *Proceedings of the American American Society* (electronically published) S 0002-9939(99)05146-1, 1999.

[145] X.-D. Hou and P. Langevin. Results on bent functions, *Journal of Combinatorial Theory, Series A*, 80, pp. 232-246, 1997.

[146] T. Jakobsen and L.R. Knudsen. The interpolation attack on block ciphers. *Fast Software Encryption'97*, Lecture Notes in Computer Science 1267, 1997.

[147] C.J.A. Jansen and D.E. Boekee. The shortest feedback shift register that can generate a given sequence. *Advances in Cryptology – CRYPTO'89, LNCS 435, Springer-Verlag*, 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).

[148] T. Johansson and F. Jönsson. Improved fast correlation attack on stream ciphers via convolutional codes. *Advances in Cryptology - EUROCRYPT'99, no. 1592 in Lecture Notes in Computer Science*, pp. 347-362, 1999.

[149] 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.

[150] 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.

[151] F. Jönsson. PhD thesis.Some results on fast correlation attacks. Lund University. 2002.

[152] D. Jungnickel. *Difference sets.* Contemporary Design Theory: A Collection of Surveys, J. Dinitz and D. R. Stinson eds. John Wiley & Sons, 1992.

[153] W. Kantor, *An Exponential Number of Generalized Kerdock Codes*, Inf. and Contr. 53, pp. 74-80, 1982.

[154] 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.

[155] T. Kasami and N. Tokura. On the weight structure of the Reed Muller codes, *IEEE Trans. Info. Theory* 16, pp. 752-759, 1970.

[156] 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.

[157] N. Katz. On a theorem of Ax. *American Journal of Mathematics* 93, pp. 485-499, 1971.

[158] A. M. Kerdock. A class of low-rate non linear codes. *Information and Control*, 20, 182-187 (1972).

[159] A. Kerckhoffs. La Cryptographie Militaire. *Journal des Sciences Militaires*, 1883.

[160] J.D. Key, T.P. McDonough and V.C. Mavron. Information sets and partial permutation decoding for codes from finite geometries. To appear in Finite Fields and their Applications.

[161] 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.

[162] L.R. Knudsen. Truncated and higher order differentials. *Fast Software Encryption, Second International Workshop*, Lecture Notes in Computer Science, n 1008, pp. 196-211. – Springer-Verlag, 1995.

[163] 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 Springer, 2003.

[164] L.R. Knudsen and M.P.J. Robshaw. Non-linear approximations in linear cryptanalysis. *Advances in Cryptology - EUROCRYPT'96*, Lecture Notes in Computer Science 1070, pp. 224-236. Springer-Verlag, 1996.

[165] 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.

[166] K. Kurosawa and R. Matsumoto. Almost security of cryptographic Boolean functions. *IEEE Transactions on Information Theory*, vol.50 (11), pp. 2752-2761, 2004.

[167] 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, Springer Verlag, pp. 434-449, 1997.

[168] K. Kurosawa, T. Iwata and T. Yoshiwara. New covering radius of Reed-Muller codes for t-resilient functions. Selected Areas in Cryptography, 8th Annual International Workshop, Vaudenay and Youssef Eds., LNCS 2259, pp. 75 ff, 2001.

[169] X. Lai. Higher order derivatives and differential cryptanalysis. *Proc. "Symposium on Communication, Coding and Cryptography", in honor of J. L. Massey on the occasion of his 60'th birthday.* 1994.

[170] 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.

[171] P. Langevin. Covering radius of $RM(1,9)$ in $RM(3,9)$. *Eurocode'90*, no. 514 in Lecture Notes in Computer Science, pp. 51-59. Springer-Verlag, 1991.

[172] 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.

[173] P. Langevin. On generalized bent functions. *CISM Courses and Lectures 339 (Eurocode)*, pp. 147-157, 1992.

[174] 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.

[175] G. Leander. Bent functions with 2^r Niho exponents. *Proceedings of the Workshop on Coding and Cryptography* 2005, Bergen, pp. 454-461, 2005.

[176] G. Leander. Monomial bent functions. *Proceedings of the Workshop on Coding and Cryptography* 2005, Bergen, pp. 462-470, 2005.

[177] R. J. Lechner. *Harmonic analysis of switching functions.* In Recent Developments in Switching Theory, Academic Press, New York, 1971.

[178] R. Lidl and H. Niederreiter, *Finite Fields*, Encyclopedia of Mathematics and its Applications, vol. 20, Addison-Wesley, Reading, Massachussetts (1983)

[179] S. Ling and C. Xing, *Coding Theory*, Cambridge: Cambridge University Press, 2004.

[180] 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.

[181] J. H. van Lint. *Introduction to coding theory*, Springer, New York, 1982.

[182] M. Lobanov. Tight bound between nonlinearity and algebraic immunity. Paper 2005/441 in http://eprint.iacr.org/

[183] O.A. Logachev, A.A. Salnikov and V.V. Yashchenko. Bent functions on a finite Abelian group. *Discrete Math. Appl.* vol 7, N° 6, pp. 547-564, 1997.

[184] S. Lloyd. Properties of binary functions. *Advances in Cryptology - EUROCRYPT '90, Lecture Notes in Computer Science* 473, pp. 124-139, 1991.

[185] S. Lloyd. Counting binary functions with certain cryptographic properties. *Journal of Cryptology* 5, pp. 107-131; 1992.

[186] S. Lloyd. Balance, uncorrelatedness and the strict avalanche criterion. *Discrete Applied Mathematics*, 41, pp. 223-233, 1993.

[187] F. J. MacWilliams and N. J. Sloane. *The theory of error-correcting codes*, Amsterdam, North Holland. 1977.

[188] 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.

[189] 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.

[190] S. Maitra. Autocorrelation properties of correlation immune Boolean functions. *Progress in Cryptology - INDOCRYPT 2001*, Lecture Notes in Computer Science 2247, pp. 242-253, 2001.

[191] S. Maitra and E. Pasalic. Further constructions of resilient Boolean functions with very high nonlinearity. *IEEE Transactions on Information Theory*, vol.48 (7), pp. 1825-1834, 2002.

[192] S. Maitra and P. Sarkar. Enumeration of correlation-immune Boolean functions. ACISP, pp. 12-25, 1999.

[193] 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.

[194] S. Maitra and P. Sarkar. Highly nonlinear resilient functions optimizing Siegenthaler's inequality. In *Advances in Cryptology - CRYPTO'99*, no. 1666 in Lecture Notes in Computer Science, pp. 198-215. Springer-Verlag, 1999.

[195] S. Maitra and P. Sarkar. Modifications of Patterson-Wiedemann functions for cryptographic applications. *IEEE Trans. Inform. Theory*, vol. 48, pp. 278-284, 2002.

[196] S. Maity and S. Maitra. Minimum distance between bent and 1-resilient Boolean functions. *Proceedings of Fast Software Encryption* 2004, LNCS 3017, pp. 143-160, 2004.

[197] J. L. Massey. Shift-register analysis and BCH decoding. *IEEE Trans. Inform. Theory*, vol. 15, pp. 122-127, 1969.

[198] J. L. Massey. Randomness, arrays, differences and duality. *IEEE Trans. Inform. Theory*, vol. 48, pp. 1698-1703, 2002.

[199] M. Matsui. Linear cryptanalysis method for DES cipher. *Advances in Cryptology - EUROCRYPT'93, no. 765 in Lecture Notes in Computer Science. Springer-Verlag*, pp. 386-397, 1994.

[200] R.J. McEliece. Weight congruence for p-ary cyclic codes. *Discrete Mathematics*, 3, pp. 177-192, 1972.

[201] R. L. McFarland. A family of noncyclic difference sets, *Journal of Comb. Theory, Series A*, no. 15, pp. 1-10, 1973.

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

[203] W. Meier and O. Staffelbach. Fast correlation attacks on stream ciphers. *Advances in Cryptology, EUROCRYPT'88, Lecture Notes in Computer Science 330, Springer Verlag*, pp. 301-314, 1988.

[204] W. Meier and O. Staffelbach. Nonlinearity Criteria for Cryptographic Functions. *Advances in Cryptology, EUROCRYPT' 89, Lecture Notes in Computer Science 434, Springer Verlag*, pp. 549-562, 1990.

[205] 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, Springer Verlag*, pp. 204-213, 1990.

[206] A. Menezes, P. van Oorschot and S. Vanstone. *Handbook of Applied Cryptography.* CRC Press Series on Discrete Mathematics and Its Applications, 1996.

[207] 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, Springer Verlag, 1998.

[208] C. J. Mitchell. Enumerating Boolean functions of cryptographic signifiance. Journal of Cryptology 2 (3), pp. 155-170, 1990.

[209] J. Mykkelveit. The covering radius of the [128,8] Reed-Muller code is 56. *IEEE Trans. Inform. Theory*, vol. 26 (3), pp. 359-362, 1980.

[210] Y.Nawaz, G.Gong, and K.Gupta. Upper Bounds on Algebraic Immunity of Power Functions. Proceeding of Fast Software Encryption 2006. To appear.

[211] N. Nisan and M. Szegedy. On the degree of Boolean functions as real polynomials. *Comput. Complexity* 4, pp. 301-313, 1994.

[212] K. Nyberg. Constructions of bent functions and difference sets, *EUROCRYPT'90, Advances in Cryptology, Lecture Notes in Computer Science 473, Springer Verlag*, pp. 151-160, 1991.

[213] L. O'Connor and A. Klapper. Algebraic nonlinearity and its applications to cryptography. *Journal of Cryptology* 7, pp. 213-227, 1994.

[214] 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.

[215] J. D. Olsen, R. A. Scholtz and L. R. Welch. Bent function sequences, *IEEE Trans. on Inf. Theory*, vol IT- 28, no. 6, 1982.

[216] E. Pasalic. *On Boolean functions in symmetric-key ciphers.* Ph.D. Thesis, 2003.

[217] 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.

[218] 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.

[219] 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.

[220] N.J. Patterson and D.H. Wiedemann. The covering radius of the $[2^{15}, 16]$ Reed-Muller code is at least 16276. *IEEE Trans. Inform. Theory*, IT-29, pp. 354-356, 1983.

[221] N.J. Patterson and D.H. Wiedemann. *Correction to [220].* IEEE Trans. Inform. Theory, IT-36(2), pp. 443, 1990.

[222] V. S. Pless, W. C. Huffman, Eds, R. A. Brualdi, assistant editor. *Handbook of Coding Theory*, Amsterdam, the Netherlands: Elsevier, 1998.

[223] B. Preneel, W. Van Leekwijck, L. Van Linden, R. Govaerts and J. Vandevalle. Propagation characteristics of Boolean functions, *Advances in Cryptology, EUROCRYPT'90, Lecture Notes in Computer Sciences, Springer Verlag* no. 473, pp. 161-173, 1991.

[224] B. Preneel, R. Govaerts and J. Vandevalle. Boolean functions satisfying higher order propagation criteria, *Advances in Cryptology, EUROCRYPT'91, Lecture Notes in Computer Sciences, Springer Verlag* no. 547, pp. 141-152, 1991.

[225] 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.

[226] M. Quisquater. The sum transform: a new tool to study cryptographic properties of Boolean functions. Preprint, 2002.

[227] M. Quisquater, B. Preneel and J. Vandewalle. A new inequality in discrete Fourier theory. *IEEE Trans. on Inf. Theory* 49, pp. 2038-2040, 2003.

[228] M. Quisquater, B. Preneel and J. Vandewalle. Spectral characterization of functions satisfying the (extended) propagation criterion of degree l and order k. Preprint, 2004.

[229] C. R. Rao. Factorial experiments derived from combinatorial arrangements of arrays. *J. Roy. Statist.* 9, pp. 128-139, 1947.

[230] F. Rodier. On the nonlinearity of Boolean functions. *Proceedings of the Workshop on Coding and Cryptography 2003*, pp. 397-405, 2003.

[231] O. S. Rothaus. On “bent” functions. *J. Comb. Theory*, 20A, pp. 300-305, 1976.

[232] B.V. Ryazanov. On the distribution of the spectral complexity of Boolean functions. *Discrete Math. Appl.*, vol. 4, No. 3, pp. 279-288, 1994.

[233] R. A. Rueppel *Analysis and design of stream ciphers* Com. and Contr. Eng. Series, Berlin, Heidelberg, NY, London, Paris, Tokyo 1986

[234] R. A. Rueppel and O. J. Staffelbach. Products of linear recurring sequences with maximum complexity. *IEEE Transactions on Information theory*, vol. IT-33, no. 1, 1987.

[235] P. Sarkar and S. Maitra. Construction of nonlinear Boolean functions with important cryptographic properties. *Advances in Cryptology - EUROCRYPT 2000*, no. 1807 in Lecture Notes in Computer Science, Springer Verlag, pp. 485-506, 2000.

[236] P. Sarkar and S. Maitra. Nonlinearity Bounds and Constructions of Resilient Boolean Functions. *CRYPTO 2000, LNCS*, vol. 1880, ed. Mihir Bellare, pp. 515-532, 2000.

[237] 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.

[238] P. Savicky. On the bent Boolean functions that are symmetric. *Eur. J. Combinatorics* 15, pp. 407-410, 1994.

[239] M. Schneider. A note on the construction and upper bounds of correlation-immune functions. *6th IMA Conference*, pp. 295-306, 1997.

[240] 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.

[241] J. Seberry, X-.M. Zhang and Y. Zheng. On constructions and nonlinearity of correlation immune Boolean functions. *Advances in Cryptology - EUROCRYPT’93*, LNCS 765, pp. 181-199, 1994.

[242] 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.

[243] N. V. Semakov and V. A. Zinov'ev, *Balanced codes and tactical configurations*, Problems of Info. Trans. 5(3), pp. 22-28 (1969)

[244] C.E. Shannon. Communication theory of secrecy systems. *Bell system technical journal*, 28, pp. 656-715, 1949.

[245] T. Siegenthaler. Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information theory*, vol. IT-30, No 5, pp. 776-780, 1984.

[246] T. Siegenthaler. Decrypting a Class of Stream Ciphers Using Ciphertext Only. *IEEE Transactions on Computer, vol. C-34*, No 1, pp. 81-85, 1985.

[247] P. Stanica, S. Maitra and J. Clark. Results on rotation symmetric bent and correlation immune Boolean functions. *Proceedings of Fast Software Encryption* 2004, LNCS 3017, pp. 161-177, 2004.

[248] P. Stanica and S. H. Sung. Boolean functions with five controllable cryptographic properties. *Designs, Codes and Cryptography* 31, pp. 147-157, 2004.

[249] I. Strazdins. Universal affine classification of Boolean functions. *Acta Applicandae Mathematicae* 46, pp. 147-167, 1997.

[250] 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.

[251] 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.

[252] 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.

[253] 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.

[254] 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, vol. 2355, pp. 66-77, 2001.

[255] 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 Cryptology ePrint archive http://eprint.iacr.org/).

[256] 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, Springer-Verlag, 2001

[257] S. Tsai. Lower bounds on representing Boolean functions as polynomials in $\mathbb{Z}_m^\star$.*SIAM J. Discrete Math.*, vol. 9 (1), pp. 55-62, 1996.

[258] S. F. Vinokurov and N. A. Peryazev. An expansion of Boolean function into a sum of products of subfunctions. *Discrete Math. Appl.*, vol. 3 (5), pp. 531-533, 1993.

[259] A.F. Webster and S.E. Tavares. On the design of S-boxes. In *Advances in Cryptology - CRYPTO'85*, no. 219 in Lecture Notes in Computer Science, pp. 523-534. Springer-Verlag, 1985.

[260] 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.

[261] Y. X. Yang and B. Guo. Further enumerating Boolean functions of cryptographic signifiance. Journal of Cryptology **8 (3)**, pp. 115-122, 1995.

[262] R. Yarlagadda and J.E. Hershey. Analysis and synthesis of bent sequences, *Proc. IEE*, vol. 136, Pt. E, pp. 112-123, 1989.

[263] A.M. Youssef and G. Gong. Hyper-bent functions. *Advances in Cryptology-EUROCRYPT 2001*, Lecture Notes in Computer Science, 2045, Springer-Verlag, Berlin, pp. 406-419, 2001.

[264] M. Zhang. Maximum correlation analysis of nonlinear combining functions in stream ciphers. *Journal of Cryptology* 13 (3), pp. 301-313, 2000.

[265] 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.

[266] X.-M. Zhang and Y. Zheng. Auto-correlations and new bounds on the nonlinearity of Boolean functions. *Advances in Cryptology - EUROCRYPT'96, no. 1070 in Lecture Notes in Computer Science, Springer-Verlag*, pp. 294-306, 1996.

[267] 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.

[268] X.-M. Zhang and Y. Zheng. The nonhomomorphicity of Boolean functions. *Advances in Cryptology - SAC 1998, Lecture Notes in Computer Science* 1556, pp. 280-295, 1999.

[269] Y. Zheng and X. M. Zhang. Plateaued functions. *ICICS'99, Lecture Notes in Computer Science*, Heidelberg, Ed., Springer-Verlag, vol. 1726, pp. 284-300, 1999.

[270] 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.

[271] Y. Zheng and X.-M. Zhang. On relationships among avalanche, nonlinearity and correlation immunity. *Advances in Cryptography - Asiacrypt 2000*, Lecture Notes in Computer Science, 1976, pp. 470-483, 2000.

[272] 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.

Index