Cryptographic Functions and Design Criteria for Block Ciphers

Anne Canteaut

INRIA – projet CODES,
BP 105, 78153 Le Chesnay, France
Anne.Canteaut@inria.fr

Abstract. Most last-round attacks on iterated block ciphers provide some design criteria for the round function. Here, we focus on the links between the underlying properties. Most notably, we investigate the relations between the functions which oppose a high resistance to linear cryptanalysis and to differential cryptanalysis.

1 Introduction

The development of cryptanalysis in the last ten years has led to the definition of some design criteria for block ciphers. These criteria correspond to some mathematical properties of the round function which is used in an iterated block cipher. They essentially concern the confusion part of the round function, usually named S-box. Most notably, the use of a highly nonlinear round function ensures a high resistance to linear attacks. Similarly, the resistance to differential attacks is related to some properties of the derivatives of the round function. The functions which are optimal regarding these criteria are respectively called almost bent and almost perfect nonlinear. For instance, such functions are used in the block cipher MISTY [26]. However, these functions present some particular properties which may introduce other weaknesses in the cipher (e.g. see [17]).

This paper describes the link between the design criteria related to differential attacks, linear attacks and higher order differential attacks. We provide some tools for establishing a general relationship between the nonlinearity of a function and its resistance to differential attacks. Most notably, we give a characterization of almost bent functions using some divisibility property of their Walsh coefficients. We also show that this structure is specific of optimal functions. Most results in this paper rely on a joined work with P. Charpin and H. Dobbertin [6,4,5].

The following section reviews the design criteria associated to some classical last-round attacks. Section 3 focuses on the functions which ensure the best resistance to differential attacks, to linear attacks and to higher order differential attacks. We show in Section 4 that these optimal functions are related to other optimal objects which appear in different areas of telecommunications. For example, almost bent functions correspond to particular error-correcting codes and to pairs of m-sequences with preferred crosscorrelation. Section 5 presents the links between the previous design criteria, especially for the case of optimal functions.

C. Pandu Rangan, C. Ding (Eds.): INDOCRYPT 2001, LNCS 2247, pp. 1–16, 2001.

2 Last-Round Attacks on Iterated Block Ciphers

In an iterated block cipher, the ciphertext is obtained by iteratively applying a keyed round function F to the plaintext. In an r-round iterated cipher, we have

$$x_i = F(x_{i-1}, K_i) \text{ for } 1 \leq i \leq r \ ,$$

where x_0 is the plaintext, x_r is the ciphertext and the r-round keys $(K_1, \ldots, K_r)$ are usually derived from a unique secret key by a key schedule algorithm. For any fixed round key K, the round function $F_K : x \mapsto F(x, K)$ is a permutation of the set of n-bit vectors, $\mathbf{F}_2^n$, where n is the block size.

Most attacks on iterated block ciphers consist in recovering the last round key K_r from the knowledge of some pairs of plaintexts and ciphertexts. For this purpose, we consider the *reduced cipher*, i.e., the cipher obtained by removing the final round of the original cipher. The reduced cipher corresponds to the function $G = F_{K_{r-1}} \circ \ldots \circ F_{K_1}$. The key point in a last-round attack is to be able to distinguish the reduced cipher from a random permutation for all round keys $K_1, \ldots, K_{r-1}$. If such a *discriminator* can be found, some information on K_r can be recovered by checking whether, for a given value k_r, the function

$$x_0 \mapsto F_{k_r}^{-1}(x_r)$$

satisfies this property or not, where x_0 (resp. x_r) denotes the plaintext (resp. the ciphertext). The values of k_r for which the expected statistical bias is observed are candidates for the correct last-round key.

Different discriminators can be exploited. Most notably, a last-round attack can be performed when the reduced cipher satisfies one of the following properties:

- The reduced cipher G has a derivative, $D_a G : x \mapsto G(x+a) + G(x)$, which is not uniformly distributed. This discriminator leads to a differential attack [1];
- There exists a linear combination of the n output bits of the reduced cipher which is close to an affine function. This leads to a linear attack [24,25];
- The reduced cipher has a constant k-th derivative for a small k. This leads to a higher order differential attack [20];
- The reduced cipher, seen as a univariate polynomial in $\mathbf{F}_{2^n}[X]$, is close to a low-degree polynomial. This leads to an interpolation attack [17] or to an improved version using Sudan's algorithm [16].

In most cases, such a property on the reduced cipher can be detected only if the round function presents a similar weakness. Therefore, a necessary condition for an iterated cipher to resist these attacks is to use a round function which does not present any of the previous characteristics. Then, the round function should satisfy the following properties for any round key K:

(i) For any $a \in \mathbf{F}_2^n$, $a \neq 0$, the output distribution of $D_a F_K : x \mapsto F_K(x+a) + F_K(x)$ should be close to the uniform distribution;

(ii) For any $a \in \mathbf{F}_2^n$, $a \neq 0$, the Boolean function $x \mapsto a \cdot F_K(x)$ should be far away from all affine functions;
(iii) The Boolean functions $x \mapsto a \cdot F_K(x)$ should have a high degree;
(iv) The function F_K, seen as a univariate polynomial in $\mathbf{F}_{2^n}[X]$, should be far away from all low-degree polynomials.

Some of these conditions may be sufficient in particular cases to guarantee that the iterated cipher resists the corresponding attack (e.g. see [31]).

Note that the first three properties are invariant under both right and left composition by a linear permutation of $\mathbf{F}_2^n$. Then, they only concern the *confusion part* of the round function. In the following, we only investigate the first three properties, since the mathematical nature of the last criterion is quite different.

3 Almost Perfect Round Functions

A *Boolean function f of n variables* is a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2$. It can be expressed as a polynomial in $x_1, \ldots, x_n$, called its *algebraic normal form*. The *degree* of f, denoted by $deg(f)$, is the degree of its algebraic normal form.

3.1 Resistance against Differential Attacks

The resistance of an iterated cipher with round function F_K against differential cryptanalysis can be quantified by some properties of the derivatives (or differentials) of F_K.

Definition 1. *[22] Let F be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$. For any $a \in \mathbf{F}_2^n$, the* derivative of F with respect to a *is the function*

$$D_a F(x) = F(x + a) + F(x) \ .$$

For any k-dimensional subspace V of $\mathbf{F}_2^n$, the k-th derivative of F with respect to V *is the function*

$$D_V F = D_{a_1} D_{a_2} \ldots D_{a_k} F \ ,$$

where $(a_1, \ldots, a_k)$ is any basis of V.

It is clear that an iterated cipher is vulnerable to a differential attack if there exists two nonzero elements a and b in $\mathbf{F}_2^n$ such that, for any round key K, the number of $x \in \mathbf{F}_2^n$ satisfying

$$F_K(x + a) + F_K(x) = b \tag{1}$$

is high. Therefore, a necessary security condition is that, for any K,

$$\delta_{F_K} = \max_{a, b \neq 0} \#\{x \in \mathbf{F}_2^n, \ F_K(x + a) + F_K(x) = b\}$$

should be small. It clearly appears that the number of solutions of Equation (1) is even (because x_0 is a solution if and only if $x_0 + a$ is a solution). Then, we deduce

Proposition 1. *[31] For any function F from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$, we have*

$$\delta_F \geq 2 \ .$$

In case of equality, F is said to be almost perfect nonlinear (APN).

Note that the terminology APN comes from the general bound

$$\delta_F \geq 2^{n-m}$$

for a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$, where the functions achieving this bound are called *perfect nonlinear functions* [28]. Such functions only exist when n is even and $n \geq 2m$ [29].

The definition of APN functions can be expressed in terms of second derivatives:

Proposition 2. *A function F from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ is APN if and only if, for any nonzero elements a and b in $\mathbf{F}_2^n$, with $a \neq b$, we have*

$$D_a D_b F(x) \neq 0 \text{ for all } x \in \mathbf{F}_2^n \ .$$

All known APN functions are functions of an odd number of variables. Actually, it is conjectured that, for any function F from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ with n even, we have

$$\delta_F \geq 4 \ .$$

This statement is proved for some particular cases, most notably for power functions [2,10].

3.2 Resistance against Linear Attacks

The resistance against linear attacks involves the Walsh spectrum of the round function.

In the following, the usual dot product between two vectors x and y is denoted by $x \cdot y$. For any $\alpha \in \mathbf{F}_2^n$, φ_α is the linear function of n variables: $x \mapsto \alpha \cdot x$. For any Boolean function f of n variables, we denote by $\mathcal{F}(f)$ the following value related to the Walsh (or Fourier) transform of f:

$$\mathcal{F}(f) = \sum_{x \in \mathbf{F}_2^n} (-1)^{f(x)} = 2^n - 2wt(f) \ ,$$

where $wt(f)$ is the Hamming weight of f, i.e., the number of $x \in \mathbf{F}_2^n$ such that $f(x) = 1$.

Definition 2. *The* Walsh spectrum *of a Boolean function f of n variables f is the multiset*

$$\{\mathcal{F}(f + \varphi_\alpha), \alpha \in \mathbf{F}_2^n\} \ .$$

The Walsh spectrum *of a vectorial function F from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ consists of the Walsh spectra of all Boolean functions $\varphi_\alpha \circ F : x \mapsto \alpha \cdot F(x)$, $\alpha \neq 0$. Therefore, it corresponds to the multiset*

$$\{\mathcal{F}(\varphi_\alpha \circ F + \varphi_\beta), \ \ \alpha \in \mathbf{F}_2^n \setminus \{0\}, \beta \in \mathbf{F}_2^n\} \ .$$

The security criterion corresponding to linear cryptanalysis is that all functions $\varphi_\alpha \circ F_K$, $\alpha \neq 0$ should be far away from all affine functions. This requirement is related to the nonlinearity of the functions F_K.

Definition 3. *The* nonlinearity *of a function F from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ is the Hamming distance between all $\varphi_\alpha \circ F, \alpha \in \mathbf{F}_2^n$, $\alpha \neq 0$, and the set of affine functions. It is given by*

$$2^{n-1} - \frac{1}{2}\mathcal{L}(F) \quad \textit{where} \quad \mathcal{L}(F) = \max_{\alpha \in \mathbf{F}_2^n} \max_{\beta \in \mathbf{F}_2^n} |\mathcal{F}(\varphi_\alpha \circ F + \varphi_\beta)| \ .$$

Proposition 3. *[33,9] For any function $F : \mathbf{F}_2^n \to \mathbf{F}_2^n$,*

$$\mathcal{L}(F) \geq 2^{\frac{n+1}{2}} \ .$$

In case of equality F is called almost bent (AB).

For a function F from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$, we have

$$\mathcal{L}(F) \geq 2^{\frac{n}{2}}$$

where the functions achieving this bound are called *bent functions*. It was proved that a function is bent if and only if it is perfect nonlinear [28,29].

The minimum value of $\mathcal{L}(F)$ where F is a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ can only be achieved when n is odd. For even n, some functions with $\mathcal{L}(F) = 2^{\frac{n}{2}+1}$ are known and it is conjectured that this value is the minimum [32,12].

3.3 Resistance against Higher Order Differential Attacks

In a higher order differential attack, the attacker exploits the existence of a k-dimensional subspace $V \subset \mathbf{F}_2^n$ such that the reduced cipher G satisfies

$$D_V G(x) = c \quad \text{for all } x \in \mathbf{F}_2^n$$

where c is a constant which does not depend on the round keys $K_1, \ldots K_{r-1}$. A natural candidate for V arises when the degree of the reduced cipher is known.

Definition 4. *The* degree *of a function F from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ is the maximum degree of its Boolean components:*

$$deg(F) = \max_{1 \leq i \leq n} deg(\varphi_{e_i} \circ F)$$

where $(e_1, \ldots, e_n)$ denotes the canonical basis of $\mathbf{F}_2^n$.

Actually, we have

Proposition 4. *[22] Let F be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ of degree d. Then, for any $(d+1)$-dimensional subspace $V \subset \mathbf{F}_2^n$, we have*

$$D_V F(x) = 0 \quad \textit{for all } x \in \mathbf{F}_2^n \ .$$

Note that the dimension of the smallest subspace V satisfying $D_V F = 0$ may be smaller than $\deg(F) + 1$.

4 Related Objects

The results concerning almost perfect functions widely apply in several areas of telecommunications: almost perfect nonlinear and almost bent functions are related to metric properties of some linear codes, especially of binary cyclic codes with two zeros. Almost bent power functions also correspond to pairs of maximum-length sequences with preferred crosscorrelation.

4.1 Links with Error-Correcting Codes

Carlet, Charpin and Zinoviev have pointed out that both APN and AB properties can be expressed in terms of error-correcting codes [8].

Since both APN and AB properties are invariant under translation, we here only consider the functions F such that $F(0,\ldots,0)=0$. We use standard notation of the algebraic coding theory (see [23]). Any k-dimensional subspace of $\mathbf{F}_2^n$ is called a binary linear code of length n and dimension k and is denoted by $[n,k]$. Any $[n,k]$-linear code $\mathcal{C}$ is associated with its dual $[n,n-k]$-code, denoted by $\mathcal{C}^\perp$:

$$\mathcal{C}^\perp = \{x \in \mathbf{F}_2^n,\ \ x\cdot c = 0\ \forall c \in \mathcal{C}\}\ .$$

Any $k\times n$ binary matrix G defines an $[n,k]$-binary linear code $\mathcal{C}$:

$$\mathcal{C} = \{xG, x\in \mathbf{F}_2^k\}$$

We then say that G is a generator matrix of $\mathcal{C}$.

Let $(\alpha_i, 1\le i\le 2^n)$ denote the set of all nonzero elements of $\mathbf{F}_2^n$. We consider the linear binary code $\mathcal{C}_F$ of length (2^n-1) and dimension $2n$ defined by the generator matrix

$$G_F = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \ldots & \alpha_{2^n} \\ F(\alpha_1) & F(\alpha_2) & F(\alpha_3) & \ldots & F(\alpha_{2^n}) \end{pmatrix}\ , \tag{2}$$

where each entry in $\mathbf{F}_2^n$ is viewed as a binary column vector of length n. It clearly appears that any codeword in $\mathcal{C}_F$ corresponds to a vector $(a\cdot\alpha_i + b\cdot F(\alpha_i), 1\le i\le 2^n)$. Therefore, its Hamming weight is given by

$$\#\{i, 1\le i\le 2^n,\ a\cdot\alpha_i + b\cdot F(\alpha_i)=1\} = 2^{n-1} - \frac{1}{2}\mathcal{F}(\varphi_b\circ F + \varphi_a)\ .$$

Moreover, a vector $(c_1,\ldots,c_{2^n})$ belongs to the dual code $\mathcal{C}_F^\perp$ if and only if

$$\sum_{i=1}^{2^n} c_i\alpha_i = 0 \text{ and } \sum_{i=1}^{2^n} c_i F(\alpha_i) = 0\ .$$

Then, we obviously have that the minimum distance of $\mathcal{C}_F^\perp$ is at least 3. Moreover, there exist three different indexes i_1, i_2, i_3 such that

$$F(\alpha_{i_1}) + F(\alpha_{i_2}) + F(\alpha_{i_3}) + F(\alpha_{i_1}+\alpha_{i_2}+\alpha_{i_3}) = 0$$

if and only if $\mathcal{C}_F^\perp$ contains a codeword of Hamming weight 4 (or 3 if $\alpha_{i_1}+\alpha_{i_2}+\alpha_{i_3}=0$).

Therefore, we obtain the following correspondence:

Theorem 1. *[8] Let F be a permutation from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ with $F(0) = 0$. Let $\mathcal{C}_F$ be the linear binary code of length $2^n - 1$ and dimension $2n$ with generator matrix G_F described by (2). Then,*

(i)

$$\mathcal{L}(F) = \max_{c \in \mathcal{C}_F, c \neq 0} |2^n - 2wt(c)| \ .$$

In particular, for odd n, F is AB if and only if for any non-zero codeword $c \in \mathcal{C}_F$,

$$2^{n-1} - 2^{\frac{n-1}{2}} \leq wt(c) \leq 2^{n-1} + 2^{\frac{n-1}{2}} \ .$$

(ii) *F is APN if and only if the code $\mathcal{C}_F^{\perp}$ has minimum distance 5.*

When the vector space $\mathbf{F}_2^n$ is identified with the finite field $\mathbf{F}_{2^n}$, the function F can be expressed as a unique polynomial of $\mathbf{F}_{2^n}[X]$. Now, we focus on power functions F, i.e., $F(x) = x^s$ over $\mathbf{F}_{2^n}$. In that case, the linear code $\mathcal{C}_F^{\perp}$ associated to $x \mapsto x^s$ is a binary cyclic code of length $(2^n - 1)$ with two zeros.

Definition 5. *A linear binary code $\mathcal{C}$ of length N is cyclic if for any codeword $(c_0, \ldots, c_{N-1})$ in $\mathcal{C}$, the vector $(c_{N-1}, c_0, \ldots, c_{N-2})$ is also in $\mathcal{C}$.*

If each vector $(c_0, \ldots, c_{N-1}) \in \mathbf{F}_2^N$ is associated with the polynomial $c(X) = \sum_{i=0}^{N-1} c_i X^i$ in $\mathcal{R}_N = \mathbf{F}_2^N[X]/(X^N - 1)$, any binary cyclic code of length N is an ideal of $\mathcal{R}_N$. Since $\mathcal{R}_N$ is a principal domain, any cyclic code $\mathcal{C}$ of length N is generated by a unique monic polynomial g having minimal degree. This polynomial is called the generator polynomial of the code and its roots are the zeros of $\mathcal{C}$. For $N = 2^n - 1$, the defining set of $\mathcal{C}$ is then the set

$$I(\mathcal{C}) = \{i \in \{0, \cdots, 2^n - 2\} |\ \alpha^i \text{ is a zero of } \mathcal{C}\} \ .$$

where α is a primitive element of $\mathbf{F}_{2^n}$. Since $\mathcal{C}$ is a binary code, its defining set is a union of 2-cyclotomic cosets modulo $(2^n - 1)$, $Cl(a)$, where $Cl(a) = \{2^j a \bmod (2^n - 1)\}$. Therefore, the defining set of a binary cyclic code of length $(2^n - 1)$ is usually identified with the representatives of the corresponding 2-cyclotomic cosets modulo $(2^n - 1)$. In this context, the linear code $\mathcal{C}_F$ associated to the power function $F : x \mapsto x^s$ on $\mathbf{F}_{2^n}$ is defined by the following generator matrix:

$$G_F = \begin{pmatrix} 1 & \alpha & \alpha^2 & \ldots & \alpha^{2^n-2} \\ 1 & \alpha^s & \alpha^{2s} & \ldots & \alpha^{(2^n-2)s} \end{pmatrix} \ .$$

Then, the dual code $\mathcal{C}_F^{\perp}$ consists of all binary vectors c of length $(2^n - 1)$ such that $c\, G_F^T = 0$. The code $\mathcal{C}_F^{\perp}$ is therefore the binary cyclic code of length $(2^n - 1)$ with defining set $\{1, s\}$.

4.2 Crosscorrelation of a Pair of Binary m-sequences

A binary sequence $(u_i)_{i \geq 0}$ generated by a linear feedback shift register (LFSR) of length n has maximal period when the feedback polynomial of the LFSR is

primitive. Such a sequence is called an *m-sequence* of length (2^n-1). A binary m-sequence of length (2^n-1) is identified with the binary vector of length (2^n-1) consisting of its first (2^n-1) bits. A further property of m-sequences is that they are almost uncorrelated with their cyclic shifts. This property is important in many communication systems (as radar communications or transmissions using spread-spectrum techniques) since it is often required that a signal can be easily distinguished from any time-shifted version of itself. It is well-known that for any m-sequence u of length (2^n-1) there exists a unique $c \in \mathbf{F}_{2^n} \setminus \{0\}$ such that

$$\forall i,\ 0 \leq i \leq 2^n-2,\ \ u_i = \mathrm{Tr}(c\alpha^i)$$

where α is a root of the feedback polynomial of the LFSR generating u (i.e., α is a primitive element of $\mathbf{F}_{2^n}$) and Tr denotes the trace function from $\mathbf{F}_{2^n}$ to $\mathbf{F}_2$.

When a communication system uses a set of several signals (usually corresponding to different users), it is also required that each of these signals can be easily distinguished from any other signal in the set and its time-shifted versions. This property is of great importance especially in code-division multiple access systems. The distance between a sequence u and all cyclic shifts of another sequence v can be computed with the crosscorrelation function:

Definition 6. *Let u and v be two different binary sequences of length N. The* crosscorrelation function between u and v, *denoted by $\theta_{u,v}$, is defined as*

$$\theta_{u,v}(\tau) = \sum_{i=0}^{N-1} (-1)^{u_i+v_{i+\tau}}\ .$$

The corresponding crosscorrelation spectrum *is the multiset*

$$\{\theta_{u,v}(\tau),\ \ 0 \leq \tau \leq N-1\}\ .$$

Since $\theta_{u,v}(\tau) = N - 2wt(u+\sigma^\tau v)$ where σ denotes the cyclic shift operator, the above mentioned applications use pairs of sequences (u,v) such that $|\theta_{u,v}(\tau)|$ is small for all $\tau \in \{0,\ldots,N-1\}$.

If u and v are two different binary m-sequences of length (2^n-1), there exists an integer s in $\{0,\ldots,2^n-2\}$ and a pair (c_1,c_2) of non-zero elements of $\mathbf{F}_{2^n}$ such that

$$\forall i,\ 0 \leq i \leq 2^n-2,\ \ u_i = \mathrm{Tr}(c_1\alpha^i) \text{ and } v_i = \mathrm{Tr}(c_2\alpha^{si})\ .$$

If $c_1 = c_2$, the sequence v is said to be a *decimation by s of u*. Writing $c_1 = \alpha^{j_1}$ and $c_2 = \alpha^{j_2}$, the crosscorrelation function for the pair (u,v) is given by:

$$\theta_{u,v}(\tau) = \sum_{i=0}^{2^n-2} (-1)^{\mathrm{Tr}(\alpha^{i+j_1}+\alpha^{si+j_2+\tau})} = \sum_{x\in\mathbf{F}_{2^m}^*} (-1)^{\mathrm{Tr}(\alpha^{\tau'}[\alpha^{j_1-\tau'}x+x^s])}\ ,$$

where $\tau' = j_2+\tau$. It follows that the corresponding crosscorrelation spectrum does not depend on the choice of j_2. It is then sufficient to study the pairs (u,v) where v is a decimation by s of u.

Now, we show that the crosscorrelation spectrum of pairs of binary m-sequences is related to the Walsh spectrum of a power function.

Proposition 5. *Let n and s be two positive integers such that $\gcd(s, 2^n-1) = 1$ and s is not a power of 2. Let $\{\theta_s(\tau),\ 0 \le \tau \le 2^n - 2\}$ be the crosscorrelation spectrum between an m-sequence of length $(2^n - 1)$ and its decimation by s. Let F be the power function $x \mapsto x^s$ over $\mathbf{F}_{2^n}$. Then, for any $\alpha \in \mathbf{F}_2^n$, $\alpha \neq 0$, we have*

$$\{\theta_s(\tau),\ 0 \le \tau \le 2^n - 2\} = \{\mathcal{F}(\varphi_\alpha \circ F + \varphi_\beta) - 1, \beta \in \mathbf{F}_2^n \setminus \{0\}\} \ .$$

Most notably,

$$\max_{0 \le \tau \le 2^n - 2} |\theta_s(\tau) + 1| = \mathcal{L}(F) \ .$$

In particular when n is odd, the lowest possible value for $\max_\tau |\theta_s(\tau) + 1|$ is $2^{\frac{n+1}{2}}$.

Definition 7. *The crosscorrelation $\theta_{u,v}$ between two m-sequences u and v of length $(2^n - 1)$ is said to be* preferred *if it satisfies*

$$\max_\tau |\theta_{u,v}(\tau) + 1| = 2^{\frac{n+1}{2}} \ .$$

Therefore, the decimations s which lead to a preferred crosscorrelation exactly correspond to the exponents s such that $x \mapsto x^s$ is an almost bent permutation over $\mathbf{F}_{2^n}$.

5 Relations between the Security Criteria

Now, we establish the links between both APN and AB properties. Chabaud and Vaudenay [9] proved that any AB function is APN. Here, we refine this result, since we give a necessary and sufficient condition for an APN function to be AB. We use the following relation involving the Walsh coefficients of a function.

Proposition 6. *Let F be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$. Then, we have*

$$\sum_{\alpha \in \mathbf{F}_2^n \setminus \{0\}} \sum_{\beta \in \mathbf{F}_2^n} \mathcal{F}^4(\varphi_\alpha \circ F + \varphi_\beta) = 2^{3n+1}(2^n - 1) + 2^{2n}\Delta \ ,$$

where $\Delta = \#\{(x, a, b) \in (\mathbf{F}_2^n)^3,\ a \neq 0, b \neq 0, a \neq b,\ \text{such that } D_a D_b F(x) = 0\}$.

Most notably, we have

$$\sum_{\alpha \in \mathbf{F}_2^n \setminus \{0\}} \sum_{\beta \in \mathbf{F}_2^n} \mathcal{F}^4(\varphi_\alpha \circ F + \varphi_\beta) \ge 2^{3n+1}(2^n - 1),$$

with equality if and only if F is APN.

Proof. For any Boolean function f of n variables, we have [3, Prop. II.1]

$$\sum_{\beta\in\mathbf{F}_2^n} \mathcal{F}^4(f+\varphi_\beta) = 2^n \sum_{a,b\in\mathbf{F}_2^n} \mathcal{F}(D_aD_bf) \ .$$

By applying this relation to all $\varphi_\alpha \circ F$, we deduce

$$\begin{aligned} S &= \sum_{\alpha\in\mathbf{F}_2^n\setminus\{0\}} \sum_{\beta\in\mathbf{F}_2^n} \mathcal{F}^4(\varphi_\alpha\circ F+\varphi_\beta) \\ &= 2^n \sum_{\alpha\in\mathbf{F}_2^n\setminus\{0\}} \sum_{a,b\in\mathbf{F}_2^n} \mathcal{F}(D_aD_b(\varphi_\alpha\circ F)) \\ &= 2^n \sum_{\alpha\in\mathbf{F}_2^n\setminus\{0\}} \sum_{a,b\in\mathbf{F}_2^n} \mathcal{F}(\varphi_\alpha\circ D_aD_bF) \\ &= 2^n \sum_{a,b\in\mathbf{F}_2^n} \sum_{\alpha\in\mathbf{F}_2^n} \mathcal{F}(\varphi_\alpha\circ D_aD_bF) - 2^{4n} \end{aligned}$$

where the last equality is obtained by adding the terms corresponding to $\alpha = 0$ in the sum. Now, for any $a, b \in \mathbf{F}_2^n$, we have

$$\sum_{\alpha\in\mathbf{F}_2^n} \mathcal{F}(\varphi_\alpha\circ D_aD_bF) = \sum_{\alpha\in\mathbf{F}_2^n} \sum_{x\in\mathbf{F}_2^n} (-1)^{\alpha\cdot D_aD_bF(x)} \ .$$

Using that

$$\sum_{\alpha\in\mathbf{F}_2^n} (-1)^{\alpha\cdot y} = 2^n \text{ if } y=0 \text{ and } 0 \text{ otherwise,}$$

we obtain

$$\sum_{\alpha\in\mathbf{F}_2^n} \mathcal{F}(\varphi_\alpha\circ D_aD_bF) = 2^n \#\{x\in\mathbf{F}_2^n,\ D_aD_bF(x)=0\} \ .$$

Therefore,

$$S = 2^{2n}\#\{x,a,b\in\mathbf{F}_2^n,\ D_aD_bF(x)=0\} - 2^{4n} \ .$$

Since $D_aD_bF = 0$ when either $a = 0$ or $b = 0$ or $a = b$, we get

$$\begin{aligned} S &= 2^{2n}\left[2^n(3(2^n-1)+1)+\Delta\right] - 2^{4n} \\ &= 2^{3n+1}(2^n-1) + 2^{2n}\Delta \ . \end{aligned}$$

Since $\Delta \geq 0$ with equality if and only if F is APN (see Proposition 2), we obtain the expected result.

We then derive the following theorem.

Theorem 2. *Let F be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$. Let*

$$\Delta = \#\{(x,a,b)\in(\mathbf{F}_2^n)^3,\ a\neq 0, b\neq 0, a\neq b,\ \text{such that } D_aD_bF(x)=0\} \ .$$

Then, we have

(i)

$$\Delta \leq (2^n - 1)(\mathcal{L}(F)^2 - 2^{n+1}) \ ,$$

where equality holds if and only if the values occurring in the Walsh spectrum of F belong to $\{0, \pm\mathcal{L}(F)\}$.

(ii) *For any positive integer ℓ such that all nonzero Walsh coefficients of F satisfy*

$$|\mathcal{F}(\varphi_\alpha \circ F + \varphi_\beta)| \geq \ell \ ,$$

we have

$$\Delta \geq (2^n - 1)(\ell^2 - 2^{n+1}) \ ,$$

where equality holds if and only if the values occurring in the Walsh spectrum of F belong to $\{0, \pm\ell\}$.

Proof. Let ℓ be a positive integer. Let $\mathcal{I}(\ell)$ denote the following quantity

$$\begin{aligned}\mathcal{I}(\ell) &= \sum_{\alpha \in \mathbf{F}_2^n \setminus \{0\}} \sum_{\beta \in \mathbf{F}_2^n} \left[\mathcal{F}^4(\varphi_\alpha \circ F + \varphi_\beta) - \ell^2 \mathcal{F}^2(\varphi_\alpha \circ F + \varphi_\beta)\right] \\ &= \sum_{\alpha \in \mathbf{F}_2^n \setminus \{0\}} \sum_{\beta \in \mathbf{F}_2^n} \mathcal{F}^2(\varphi_\alpha \circ F + \varphi_\beta) \left[\mathcal{F}^2(\varphi_\alpha \circ F + \varphi_\beta) - \ell^2\right] \ .\end{aligned}$$

By combining Proposition 6 and Parseval's relation, we obtain that

$$\begin{aligned}\mathcal{I}(\ell) &= 2^{3n+1}(2^n - 1) + 2^{2n}\Delta - 2^{2n}(2^n - 1)\ell^2 \\ &= 2^{2n}(2^n - 1)(2^{n+1} - \ell^2) + 2^{2n}\Delta \ .\end{aligned}$$

Now, any term in the sum defining $\mathcal{I}(\ell)$ satisfies

$$\begin{aligned}\mathcal{F}^2(\varphi_\alpha \circ F + \varphi_\beta) \left[\mathcal{F}^2(\varphi_\alpha \circ F + \varphi_\beta) - \ell^2\right] &< 0 \ \text{ if } 0 < |\mathcal{F}(\varphi_\alpha \circ F + \varphi_\beta)| < \ell \\ &= 0 \ \text{ if } |\mathcal{F}(\varphi_\alpha \circ F + \varphi_\beta)| \in \{0, \pm\ell\} \\ &> 0 \ \text{ if } |\mathcal{F}(\varphi_\alpha \circ F + \varphi_\beta)| > \ell\end{aligned}$$

This implies that all terms appearing in $\mathcal{I}(\mathcal{L}(F))$ are negative. Then, we have

$$\Delta \leq (2^n - 1)(\mathcal{L}(F)^2 - 2^{n+1}) \ ,$$

with equality if and only if all terms in the sum are zero. This situation only occurs if the values occurring in the Walsh spectrum of F belong to $\{0, \pm\mathcal{L}(F)\}$.

Similarly, if all nonzero Walsh coefficients of F satisfy

$$|\mathcal{F}(\varphi_\alpha \circ F + \varphi_\beta)| \geq \ell \ ,$$

then all terms appearing in $\mathcal{I}(\ell)$ are positive. Therefore,

$$\Delta \geq (2^n - 1)(\ell^2 - 2^{n+1}) \ ,$$

with equality if and only if all terms in the sum are zero.

Another proof of this result can be obtained by using the error-correcting code corresponding to F [6]. In that case, the proof is based on Pless identities and on some techniques due to Kasami [18]. As a direct application of the previous theorem, we derive a characterization of almost bent functions.

Corollary 1. *Let n be an odd integer and let F be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$. Then, F is AB if and only if F is APN and all its Walsh coefficients are divisible by $2^{\frac{n+1}{2}}$.*

Proof. F is AB if and only if $\mathcal{L}(F) = 2^{(n+1)/2}$. Using Theorem 2 (i), we obtain that $\Delta \leq 0$. Since Δ is a non-negative integer, it follows that $\Delta = 0$, i.e., F is APN. Moreover, the upper bound given in Theorem 2 (i) is achieved. Therefore, the values occurring in the Walsh spectrum of F belong to $\{0, \pm 2^{(n+1)/2}\}$. This implies that all Walsh coefficients are divisible by $2^{(n+1)/2}$.

Conversely, if all Walsh coefficients are divisible by $2^{(n+1)/2}$, then all nonzero Walsh coefficients satisfy

$$|\mathcal{F}(\varphi_\alpha \circ F + \varphi_\beta)| \geq 2^{(n+1)/2} \ .$$

From Theorem 2 (ii) applied to $\ell = 2^{(n+1)/2}$, we obtain $\Delta \geq 0$. If F is APN, we have $\Delta = 0$ and the lower bound given in Theorem 2 (ii) is reached. Therefore, the values occurring in the Walsh spectrum of F belong to $\{0, \pm 2^{(n+1)/2}\}$. This implies that F is AB.

Note that both properties of AB functions derived from the sufficient condition in the previous corollary have been proved in [9].

A first consequence of the divisibility of the Walsh coefficients of an AB function is the following upper bound on its degree. This bound can be derived from [7, Lemma 3].

Corollary 2. *[8] Let n be an odd integer and F be an AB function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$. Then,*

$$\deg(F) \leq \frac{n+1}{2} \ .$$

Therefore, there exists a trade-off between the security criteria involved by linear cryptanalysis and by higher order differential attacks.

When F is a power function, $F : x \mapsto x^s$, the corresponding code $\mathcal{C}_F$ is the dual of the binary cyclic code of length $(2^n - 1)$ with defining set $\{1, s\}$ (see Section 4.1). The weight divisibility of a cyclic code can be obtained by applying McEliece's theorem:

Theorem 3. *[27] The weights of all codewords in a binary cyclic code $\mathcal{C}$ are exactly divisible by 2^ℓ if and only if ℓ is the smallest number such that $(\ell + 1)$ nonzeros of $\mathcal{C}$ (with repetitions allowed) have product 1.*

This leads to the following characterization of AB power functions.

Corollary 3. *Let n be an odd integer and let $F : x \mapsto x^s$ be a power function over $\mathbf{F}_{2^n}$. Then, F is AB if and only if F is APN and*

$$\forall u,\ \ 1 \leq u \leq 2^n - 1,\ \ w_2(us \bmod (2^n - 1)) \leq \frac{n-1}{2} + w_2(u)$$

where $w_2(u)$ corresponds to the number of 1s in the 2-adic expansion of u.

Thanks to McEliece's theorem, the determination of the values of s such that $x \mapsto x^s$ is almost bent on $\mathbf{F}_{2^n}$ is reduced to a combinatorial problem. Most notably, this technique was directly used to prove that some power functions are AB [5,15]. Moreover, it leads to a very efficient method for proving that a given power function is not AB. For example, the APN power function $x \mapsto x^s$ over $\mathbf{F}_{2^{5g}}$ with $s = 2^{4g} + 2^{3g} + 2^{2g} + 2^g - 1$ does not satisfy the condition of Corollary 3 [6].

These recent results lead to the following list (up to equivalence) of known AB permutations (Table 1). All these functions are power functions. Here, we only give one exponent per cyclotomic coset modulo $(2^n - 1)$. We do not mention the exponent corresponding to the inverse permutation (which is AB too).

Table 1. Known AB power permutations x^s on $\mathbf{F}_{2^n}$

exponents s	condition on n	
$2^i + 1$ with $\gcd(i, n) = 1$ and $1 \leq i \leq (n-1)/2$		[13,30]
$2^{2i} - 2^i + 1$ with $\gcd(i, n) = 1$ and $2 \leq i \leq (n-1)/2$		[19]
$2^{\frac{n-1}{2}} + 3$		[5]
$2^{\frac{n-1}{2}} + 2^{\frac{n-1}{4}} - 1$	$n \equiv 1 \bmod 4$	[15]
$2^{\frac{n-1}{2}} + 2^{\frac{3n-1}{4}} - 1$	$n \equiv 3 \bmod 4$	[15]

When n is even, the smallest known value of $\mathcal{L}(F)$ for a function F from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ is $\mathcal{L}(F) = 2^{n/2+1}$. The only known functions (up to equivalence) achieving this bound are power functions. Since power permutations cannot be APN, it clearly appears that the security criteria corresponding to differential cryptanalysis and to linear cryptanalysis are not so strongly related. Moreover, the divisibility of the Walsh coefficients of these highly nonlinear functions varies. In particular, the degree of such a function is not upper-bounded since there is no requirement on the divisibility of the Walsh coefficients. Table 2 gives all known power functions achieving the highest known nonlinearity and the divisibility of their Walsh coefficients.

6 Conclusion

The functions which opposes the best resistance to linear cryptanalysis possess a very strong algebraic structure. The AB property appears very restrictive. In

Table 2. Known power permutations x^s on $\mathbf{F}_{2^n}$ with the highest nonlinearity and highest divisibility of their Walsh coefficients

exponents s	condition on n	divisibility	
$2^{n-1}-1$		2^2	[21]
2^i+1 with $\gcd(i,n)=2$	$n \equiv 2 \bmod 4$	$2^{\frac{n}{2}+1}$	[13,30]
$2^{2i}-2^i+1$ with $\gcd(i,n)=2$	$n \equiv 2 \bmod 4$	$2^{\frac{n}{2}+1}$	[19]
$\sum_{i=0}^{n/2} 2^{ik}$ with $\gcd(k,n)=1$	$n \equiv 0 \bmod 4$	$2^{\frac{n}{2}}$	[12]
$2^{\frac{n}{2}}+2^{\frac{n+2}{4}}+1$	$n \equiv 2 \bmod 4$	$2^{\frac{n}{2}+1}$	[11]
$2^{\frac{n}{2}}+2^{\frac{n}{2}-1}+1$	$n \equiv 2 \bmod 4$	$2^{\frac{n}{2}+1}$	[11]
$2^{\frac{n}{2}}+2^{\frac{n}{4}}+1$	$n \equiv 4 \bmod 8$	$2^{\frac{n}{2}}$	[12]

particular, AB functions also guarantee the highest possible resistance against differential cryptanalysis. But, besides the APN property, they can be characterized by the divisibility of their Walsh coefficients. This particular structure leads to an upper-bound on their degree (it then limits their resistance against higher order differential attacks) and it may introduce some other weaknesses. Therefore, it seems preferable to use as round function a function whose nonlinearity is high but not optimal. Most notably, the functions of an even number of variables which have the highest known nonlinearity do not present any similar properties. As an example, the inverse function over a finite field $\mathbf{F}_{2^n}$ with n even (used in AES) offers a very high resistance against differential, linear and higher order differential attacks. Moreover, its Walsh coefficients are divisible by 4 only (which is the lowest possible divisibility).

References

1. E. Biham and A. Shamir. Differential cryptanalysis of DES-like cryptosystems. *Journal of Cryptology*, 4(1):3–72, 1991.
2. A. Canteaut. Differential cryptanalysis of Feistel ciphers and differentially uniform mappings. In *Selected Areas on Cryptography, SAC'97*, pages 172–184, Ottawa, Canada, 1997.
3. A. Canteaut, C. Carlet, P. Charpin, and C. Fontaine. On cryptographic properties of the cosets of $R(1,m)$. *IEEE Trans. Inform. Theory*, 47(4):1494–1513, 2001.
4. A. Canteaut, P. Charpin, and H. Dobbertin. A new characterization of almost bent functions. In *Fast Software Encryption 99*, number 1636 in Lecture Notes in Computer Science, pages 186–200. Springer-Verlag, 1999.
5. A. Canteaut, P. Charpin, and H. Dobbertin. Binary m-sequences with three-valued crosscorrelation: A proof of Welch conjecture. *IEEE Trans. Inform. Theory*, 46(1):4–8, 2000.
6. A. Canteaut, P. Charpin, and H. Dobbertin. Weight divisibility of cyclic codes, highly nonlinear functions on GF(2^m) and crosscorrelation of maximum-length sequences. *SIAM Journal on Discrete Mathematics*, 13(1):105–138, 2000.

7. C. Carlet. Two new classes of bent functions. In *Advances in Cryptology - EUROCRYPT'93*, number 765 in Lecture Notes in Computer Science, pages 77–101. Springer-Verlag, 1994.
8. C. Carlet, P. Charpin, and V. Zinoviev. Codes, bent functions and permutations suitable for DES-like cryptosystems. *Designs, Codes and Cryptography*, 15:125–156, 1998.
9. F. Chabaud and S. Vaudenay. Links between differential and linear cryptanalysis. In *Advances in Cryptology - EUROCRYPT'94*, number 950 in Lecture Notes in Computer Science, pages 356–365. Springer-Verlag, 1995.
10. P. Charpin, A. Tietäväinen, and V. Zinoviev. On binary cyclic codes with minimum distance $d = 3$. *Problems of Information Transmission*, 33(4):287–296, 1997.
11. T. Cusick and H. Dobbertin. Some new 3-valued crosscorrelation functions of binary m-sequences. *IEEE Transactions on Information Theory*, 42:1238–1240, 1996.
12. H. Dobbertin. One-to-one highly nonlinear power functions on $GF(2^n)$. *Appl. Algebra Engrg. Comm. Comput.*, 9(2):139–152, 1998.
13. R. Gold. Maximal recursive sequences with 3-valued recursive crosscorrelation functions. *IEEE Transactions on Information Theory*, 14:154–156, 1968.
14. T. Helleseth and P. Vijay Kumar. *Handbook of Coding Theory*, volume II, chapter 21 - Sequences with low correlation, pages 1765–1853. Elsevier, 1998.
15. H. Hollman and Q. Xiang. A proof of the Welch and Niho conjectures on crosscorrelations of binary m-sequences. *Finite Fields and Their Applications*, 7(2):253–286, 2001.
16. T. Jakobsen. Cryptanalysis of block ciphers with probabilistic non-linear relations of low degree. In *Advances in Cryptology - CRYPTO'98*, number 1462 in Lecture Notes in Computer Science, pages 212–222. Springer-Verlag, 1998.
17. T. Jakobsen and L.R. Knudsen. The interpolation attack on block ciphers. In *Fast Software Encryption 97*, number 1267 in Lecture Notes in Computer Science. Springer-Verlag, 1997.
18. T. Kasami. Weight distributions of Bose-Chaudhuri-Hocquenghem codes. In *Proceedings of the conference on combinatorial mathematics and its applications*, pages 335–357. The Univ. of North Carolina Press, 1968.
19. T. Kasami. The weight enumerators for several classes of subcodes of the second order binary Reed-Muller codes. *Information and Control*, 18:369–394, 1971.
20. L. R. Knudsen. Truncated and higher order differentials. In *Fast Software Encryption - Second International Workshop*, number 1008 in Lecture Notes in Computer Science, pages 196–211. Springer-Verlag, 1995.
21. G. Lachaud and J. Wolfmann. The weights of the orthogonal of the extended quadratic binary Goppa codes. *IEEE Transactions on Information Theory*, 36(3):686–692, 1990.
22. X. Lai. Higher order derivatives and differential cryptanalysis. In *Proc. "Symposium on Communication, Coding and Cryptography", in honor of J. L. Massey on the occasion of his 60'th birthday*, 1994.
23. F.J. MacWilliams and N.J.A. Sloane. *The theory of error-correcting codes*. North-Holland, 1977.
24. M. Matsui. Linear cryptanalysis method for DES cipher. In *Advances in Cryptology - EUROCRYPT'93*, number 765 in Lecture Notes in Computer Science. Springer-Verlag, 1994.
25. M. Matsui. The first experimental cryptanalysis of the Data Encryption Standard. In *Advances in Cryptology - CRYPTO'94*, number 839 in Lecture Notes in Computer Science. Springer-Verlag, 1995.

26. M. Matsui. New Block Encryption Algorithm MISTY. In *Proceedings of the Fourth International Workshop of Fast Software Encryption*, number 1267 in Lecture Notes in Computer Science, pages 54–68. Springer-Verlag, 1997.
27. R.J. McEliece. Weight congruence for p-ary cyclic codes. *Discrete Mathematics*, 3:177–192, 1972.
28. W. Meier and O. Staffelbach. Nonlinearity criteria for cryptographic functions. In *Advances in Cryptology - EUROCRYPT'89*, number 434 in Lecture Notes in Computer Science, pages 549–562. Springer-Verlag, 1990.
29. K. Nyberg. Perfect nonlinear S-boxes. In *Advances in Cryptology - EUROCRYPT'91*, number 547 in Lecture Notes in Computer Science, pages 378–385. Springer-Verlag, 1991.
30. K. Nyberg. Differentially uniform mappings for cryptography. In *Advances in Cryptology - EUROCRYPT'93*, number 765 in Lecture Notes in Computer Science, pages 55–64. Springer-Verlag, 1993.
31. K. Nyberg and L.R. Knudsen. Provable security against differential cryptanalysis. In *Advances in Cryptology - CRYPTO'92*, number 740 in Lecture Notes in Computer Science, pages 566–574. Springer-Verlag, 1993.
32. D.V. Sarwate and M.B. Pursley. Crosscorrelation properties of pseudorandom and related sequences. *Proceedings of the IEEE*, 68(5):593–619, 1980.
33. V.M. Sidelnikov. On mutual correlation of sequences. *Soviet Math. Dokl.*, 12:197–201, 1971.