BOOLEAN FUNCTIONS

A *Boolean Function* is a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ for which we define the *Hamming wieght* and the *Hamming distance* from another function. Every BF can be written in a polynomial form, called *Algebraic Normal Form* (ANF). The *degree* of $f \in BF_n$ is the degree of its ANF. The non linearity is defined as the minimum Hamming distance of f from any affine function $\alpha \in An$. Obviously :

$$d_H(f,0) = w_H(f) \tag{1}$$

$$d_H(f,1) = 2^n - w_H(f) \tag{2}$$

Therefore:

$$nl(f) \leq \min\{w_H(f), 2^n - w_H(f)\} \tag{3}$$

It is easy to prove using linear algebra that every affine function α has $w_H(\alpha) = 2^{n-1}$ and is therefore *balanced*.
The *Fourier Transform* of f is the function:

$$F_f(a) = \sum_x (-1)^{a \cdot x} f(x). \tag{4}$$

The *Walsh Transform* is the Fourier transform on the *sign function* of f, i.e. $\hat{f}(x) = (-1)^{f(x)}$, so:

$$W_f(a) = \sum_x (-1)^{a \cdot x + f(x)}. \tag{5}$$

which is easily seen to take value $1 - 2f$. Thus we can also seen that:

$$W_f(a) = 2^n \delta_0(a) - 2F_f(a) \tag{6}$$

with δ_0 the *Dirac* symbol.
And therefore since $w_H(f) = F_f(0) = 2^{n-1} - \frac{W_f(0)}{2}$ we get that

$$\begin{gathered} d_H(f,\alpha) = w_H(f+\alpha) = 2^{n-1} - \frac{W_f(a)}{2}, \\ d_H(f,\alpha+1) = w_H(f+\alpha+1) = 2^{n-1} + \frac{W_f(a)}{2} \end{gathered} \tag{7}$$

So we end up with the conclusion that:

$$\begin{gathered} nl(f) \leq \min\{2^{n-1} - \frac{W_f(a)}{2}, 2^{n-1} + \frac{W_f(a)}{2}\}, \\ \Longrightarrow nl(f) = 2^{n-1} - \max_{a \in \mathbb{F}_2^n} \left\{ \frac{|W_f(a)|}{2} \right\} \end{gathered} \tag{8}$$

Now a great result is given by *Parseval's Relation* which says:

$$\sum_{a \in \mathbb{F}_2^n} W_f(a)^2 = 2^{2n} \tag{9}$$

From this we derive the *covering radius bound*:

$$\max_{a \in \mathbb{F}_2^n} W_f(a) \geq 2^{\frac{n}{2}} \tag{10}$$

LFSR weakness

LFSR are cryptographically weak because of the *Berlekamp-Massey* algorithm. Given a minimum length LFSR, say L (this length is called the *linear complexity* of the sequence), then if we know at LEAST 2L consecutive bits, Berlekamp Massey algorithm recovers the length L, the coefficients of the feedback polynomial and furthermore the initialization vector (the secret key basically) of the LFSR in O(L^2).
A possible way to overcome this attack is by using Boolean functions. Cryptosystems have to rely on two principles: *confusion* and *diffusion*. Confusion is closely related to the cryptographic complexity of Boolean functions. Diffusion consits in spreding out the influence of any minor modification of the plaintext or the key among the bits of all outputs. The resistance of the cryptosystem can be quantified through some characteristics of the Boolean functions. Some of those characteristics are affine invariant while some are not.

Algebraic Degree

First of all Bf must have high degree as otherwise we will have a low linear complexity of the LFSR and therefore we can easily factorize such value to recover sequences length. Obviously affine invariant.

Nonlinearity

In order to provide confusion, cryptographic functions must lie at large Hamming distance to all affine functions. We say that there is a correlation between a Boolean function f and a linear (therefore affine) function l if $d_H(f,l)$ is different from 2^{n-1} (if it was 2^n this means that $f = l + 1$). As we've seen in Parseval's Relation any Boolean function has some correlation with a linear function. Correlation and therefore affine approximations allows mounting attacks such as the *fast correlation attack*.

Suppose g is a linear approximation of the Boolean function f so $d_H(f,g) < 2^{n-1}$. The probability

$$p = Pr(f(x) \neq g(x)) = \frac{d_H(f,g)}{2^n} = \frac{1}{2} - \epsilon \tag{11}$$

with $\epsilon > 0$. Now the pseudo random sequence s generated with g is the transmission with errors of the sequence σ generated with f. Basically attacking the cipher can be done by correcting the errors in transmission over a noisy channel. Therefore the larger is the nonlinearity the larger is the probability p defined above, and so the less efficient is the attack. It must be high, and this is one of the most important cryptographic criteria. It is affine invariant since $d_H(f \circ L, l \circ L) = d_H(f,l)$ since L is an affine automorphism $L(x)$ gives every element of the vector space (field).
Maximum value in (9) can be achieved $\iff$ $W_f(l) = 2^{\frac{n}{2}}\ \forall\ l$. Function reaching this equality are called *bent functions* whic obviously exist only for even values of n since $2^{n-1} - 2^{\frac{n}{2}-1}$ must be an integer. For n odd we can achieve nonlinearity $2^{n-1} - 2^{\frac{n-1}{2}}$, these quadratic functions are called *semi-bent*, and their Walsh spectra only contain 0 and $2^{\frac{n+1}{2}}$. The maximum algebraic degree for a bent function f is $\frac{n}{2}$. Moreover we have that nonlinearity is larger for low degree functions. Indeed we have a great bound that says:

$$deg(f) \leq n - k + 1 \tag{12}$$

where k is the biggest such that $2^k | W_f(a) \forall a$

Balancedness

Cryptographic functions must be *balanced* in order to avoid statistical dependence between plaintext and ciphertext. It is easily seen that f is balanced if $w_H(f) = 2^{n-1} = \sum\limits_{x \in F} f(x) = \sum\limits_{x \in F} (-1)^0 f(x) = F_f(0)$ or else if $W_f(0) = F(f) = 0$. In the case of a combinig cipher the Boolean combining function must remain balanced if we fix some coordinates x_i. We say that an $n-variable$ function is $m-resilient$ if fixing m variables it is still balanced. This is related to the *correlation attack* .

VECTORIAL BOOLEAN FUNCTIONS

Also called $S - boxes$. They are part of iterative *blockciphers* and determine their robustness. Every round of such ciphers consist of vectorial Boolean functions ($v.B.f.$) combined in different ways. There are some attacks on v.B.f. that will define cryptographic criteria.

Differential cryptanalysis

The *differential attack* assumes the existence of ordered pairs (α, β) such that a block m of plaintext being randomly chosen and c and c' being the ciphertexts related to m and $m + \alpha$, the bitwise difference $c + c'$ has larger probability of being equal to β than if c and c' were randomly chosen. (α, β) is called a differential. The larger the probability of the differential the more efficient the attack is. The related criterion on a v.B.f. F used as S-box is that the output of the derivative $D_a(F) = F(x) + F(x + a)$ must be as uniformly distributed as possible.

Linear cryptanalysis

The *linear attack* is based on the distinguisher triple (α, β, γ) of binary strings such that, a block m of plaintext and a key k being randomly chosen, the bit $\alpha \cdot m + \beta \cdot c + \gamma \cdot k$ has probability different from $\frac{1}{2}$ of being null. The more distant from $\frac{1}{2}$ the probability is the more efficient is the attack. What we come up with is that the *component functions* (i.e. $v \cdot F = \sum_{i=1}^{m} v_i f_i$ where f_i are *coordinate functions* of F and $v \in \mathbb{F}_2^m$) must have the highest nonlinearity possible.

Balancedness

As for standard Boolean function *balancedness* is important fro v.B.f. in cryptography. We say that F is balanced if it takes every value of $\mathbb{F}_2^m$ the same number of times, i.e. 2^{n-m}. This means that $|F^{-1}(b)|$ must be equal $\forall b \in \mathbb{F}_2^m$. Obviously balanced $(n - n)$-functions are permutations over $\mathbb{F}_2^n$. Balancedness can be characterized through the component functions by saying that F is balanced if every component function is so.

Nonlinearity

The *nonlinearity* $nl(F)$ of a v.B.f. is the minimum nonlineatity of all the component functions of F. This quantifies the resistance of the S-box against linear attacks. It is an affine invariant. The *covering radius bound* is therefore still valid for v.B.f., i.e. $nl(F) \leq 2^{n-1} - 2^{\frac{n}{2}-1}$. F is said to be *bent* if it achieves the covering radius bound with equality. We notice that F is bent $\iff$ every component function is bent. This is because since $nl(F) = 2^{n-1} - 2^{\frac{n}{2}-1}$ for a bent F and $\max_{l \in \mathbb{F}_2^n}\{|W_f(l)|\} \geq 2^{\frac{n}{2}}$ this is true only for equality in this disequality and therefore it is true for ever l.

Hence the *algebraic degree* of a bent v.B.f. is at most $\frac{n}{2}$.

We also have that a standard Boolean function is bent $\iff$ all of its derivatives $D_a(f) = f(x) + f(x+a)$ is balanced, therefore F is bent if all $v \cdot (F(x) + F(x + a))$ are balanced. Therefore we deduce that F is bent $\iff$ every of its derivative are balanced.

Due to the fact that bent functions are $perfect\ nonlinear$ and that derivatives are balanced they are strong against both linear and differential cryptanalysis. They exist for n even and $m \leq \frac{n}{2}$ (??).

- **Parseval's Relation**

$$\sum_{u \in \mathbb{F}_2^n, v \neq 0 \in \mathbb{F}_2^n} W_F^2(u, v) = 2^n(2^m - 1) \tag{13}$$

- **SCV bound**

For v.B.f. functions with $m > n - 1 > \frac{n}{2}$ (therefore not bent) we can find a better upper bound for $nl(F)$, i.e. :

$$nl(F) \leq 2^{n-1} - \frac{1}{2}\sqrt{3 \times 2^n - 2 - 2\frac{(2^n - 1)(2^{n-1})}{2^m - 1}} \tag{14}$$

Here $m \geq n - 1$ to avoid negative values under the square root. For $m = n - 1$ SVC is the covering radius.

A F with $m = n$ achieving SVC with equality is said *AlmostBent*

F v.B.f. and $\delta > 0$ integer, we say that F is δ-*differentially uniform* if $\forall a \in \mathbb{F}_2^n, a \neq 0$ and $\forall b \in \mathbb{F}_2^m$ the equation $D_a(F) = b$ has at most δ solutions, i.e.

$$\triangle_{a,b}(F) = |D_a(F)^{-1}(b)| \leq \delta \tag{15}$$

Obviously if F is bent (and therefore its derivatives are balanced) $\delta = 2^{n-m}$, and moreover there does not exist any 1-differentially uniform function. If $\delta \leq 2$ (0 or 2) we say that F is *APN* (*Almost Perfect Nonlinear*), therefore the smaller δ is the better is the resistance of F to differential cryptanalysis, while *AB* functions provide maximal resistance against both linear and differential. We have that every *AB* function is *APN*.

We say that F is *weakly-δ-differentially uniform* if $\forall a \neq 0$:

$$|Im(D_a(F))| > \frac{2^{n-1}}{\delta} \tag{16}$$

F is *weakly APN* if it is *weakly-2-differentially uniform*, i.e. $|Im(D_a(F))| > 2^{n-2}$.

Write

$$\hat{n}(F) = \max_{a \in \mathbb{F}_2^n, a \neq 0} |\{v \in \mathbb{F}_2^m | v \cdot D_a(F) \ is \ constant\}| \tag{17}$$

Basically, this is the number of vectors such that the component functions of the derivative in a of F is constant $\forall a$.

• n=m=4

Now we say that if F with $n = m = 4$ is *weakly APN* then $\hat{n}(F) \leq 1$ and moreover if F is *APN* then $\hat{n} \leq 1$ If F is a permutation with $n = m = 4$ then $nl(F) = nl(F^{-1})$

We have also other two kind of equivalence which are the *Extended Affine Equivalence* and the *CCZ-Equivalence*. The first one is defined as:

$$F \sim_{EA} G \iff \exists A, B, C \in AGL(\mathbb{F}_2^n) | G = AF(B(x)) + C(x). \tag{18}$$

The CCZ-equivalence istead is defined on the two functions with regards to their grapsh (i.e. tuple of the form $(x, F(x))$):

$$F \sim_{CCZ} G \iff \exists \Lambda \in AGL(\mathbb{F}_2^n) | \{(x, F(x)) | x \in \mathbb{F}_2^n\} = \{\Lambda(y, G(x)) | y \in \mathbb{F}_2^n\} \tag{19}$$

The relation is $A \subset AE \subset CCZ$.
To remember that weakly δ diff. uniformity is not a CCZ-invariant.

Consider $V = \mathbb{F}_2^n$ as the direct sum $V = V_1 \oplus V_2 \oplus ... \oplus V_b$ where V_i have same dimension m. We will call *wall* every nontrivial sum of such V_i.
We call $bricklayer\ transformation$ any $\gamma \in Sym(V)$ s.t. it acts $v\gamma = v_1\gamma_1 \oplus v_2\gamma_2 \oplus ... \oplus v_b\gamma_b$, and γ_i a brick.
$\lambda \in AGL(\mathbb{F}_2^n)$ is said to be *proper* if it has no invariant walls.

We say that a block cipher φ_k is *Translation based* if it can be written as the composition:

$$\gamma \circ \lambda \circ \delta_k \tag{20}$$

where:

- γ is a round-dependent bricklayer transformation,
- λ a round-dependent linear map,
- δ_k with $k \in V$ is the round key addition.

And for at least one round we can assume that $lambda$ is proper and the key schedule map : $K \to V$ is surjective.