CRYPTOLOGICAL VIEWPOINT OF BOOLEAN FUNCTIONS

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

OF

THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

SERHAT SAĞDıÇOĞLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

THE DEPARTMENT OF MATHEMATICS

SEPTEMBER 2003

Approval of the Graduate School of Natural and Applied Sciences.

Prof. Dr. Canan Özgen
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Ersan Akyıldız
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Ali Doğanaksoy
Supervisor

Examining Committee Members

Prof. Dr. Ersan Akyıldız

Assoc. Prof. Dr. Ali Doğanaksoy

Assoc. Prof. Dr. Ferruh Özbudak

Assoc. Prof. Dr. Melek D. Yücel

Dr. Emrah Çakçak

ABSTRACT

CRYPTOLOGICAL VIEWPOINT OF BOOLEAN FUNCTIONS

Sağdıçoğlu, Serhat

M. Sc., Department of Mathematics

Supervisor: Assoc. Prof. Dr. Ali Doğanaksoy

September 2003, 99 pages

Boolean functions are the main building blocks of most cipher systems. Various aspects of their cryptological characteristics are examined and investigated by many researchers from different fields. This thesis has no claim to obtain original results but consists in an attempt at giving a unified survey of the main results of the subject. In this thesis, the theory of boolean functions is presented in details, emphasizing some important cryptological properties such as balance, nonlinearity, strict avalanche criterion and propagation criterion. After presenting many results about these criteria with detailed proofs, two upper bounds and two lower bounds on the nonlinearity of a boolean function due to Zhang and Zheng are proved. Because of their importance in the theory of boolean functions, construction of Sylvester-Hadamard matrices are shown and most of their properties used in cryptography are proved. The Walsh transform is investigated in detail by proving many properties. By using a property of Sylvester-Hadamard matrices, the fast Walsh transform is presented and its application in finding the nonlinearity of a boolean function is demonstrated. One of the most important classes of boolean functions, so called bent functions, are presented with many properties and by giving several examples, from the paper of Rothaus. By using bent functions, relations between balance, nonlinearity and propagation criterion are presented and it

is shown that not all these criteria can be simultaneously satisfied completely. For this reason, several constructions of functions optimizing these criteria which are due to Seberry, Zhang and Zheng are presented.

Keywords: Cryptography, Boolean functions, Hadamard matrices, Sylvester-Hadamard matrices, Nonlinearity, Strict avalanche criterion, Propagation criterion, Walsh transform, Fast Walsh transform, Bent function.

ÖZ

KRİPTOLOJİK BAKIŞ AÇISIYLA BOOLE FONKSİYONLARI

Sağdıçoğlu, Serhat
Yüksek Lisans, Matematik Bölümü
Tez Yöneticisi: Assoc. Prof. Dr. Ali Doğanaksoy

Eylül 2003, 99 sayfa

Boole fonksiyonları bir çok şifre sisteminin ana yapı taşıdır. Bunların muhtelif kriptolojik karakteristikleri farklı alanlardan bir çok araştırmacı tarafından ele alınmış ve incelenmiştir. Bu tez hiç bir özgün sonuç elde etme iddiasında bulunmamakta, sadece konunun ana sonuçlarının bütünleşik bir mütalaasını vermeye teşebbüs etmektedir. Bu tezde Boole fonksiyonlarının teorisi dengelilik, doğrusal olmama, tam çığ ölçütü ve yayılma ölçütü gibi bazı önemli kriptolojik özellikler vurgulanarak detayları ile sunulmaktadır. Bu ölçütler hakkındaki birçok sonucu detaylı ispatlar ile sunduktan sonra bir Boole fonksiyonunun doğrusal olmaması üzerinde Zhang ve Zheng'e ait iki üst sınır ve iki alt sınır ispatlanmıştır. Boole fonksiyonlar teorisindeki önemlerinden dolayı, Sylvester-Hadamard matrislerinin inşaası gösterilmiş ve kriptografide kullanılan birçok özellikleri ispatlanmıştır. Walsh dönüşümü birçok özellikleri ispatlanarak detayları ile incelenmiştir. Sylvester-Hadamard matrislerinin bir özelliğini kullanarak hızlı Walsh dönüşümü sunulmuş ve bir Boole fonksiyonunun doğrusal olmama değerinin bulunmasındaki uygulaması gösterilmiştir. Bükük (bent) fonksiyonlar olarak anılan Boole fonksiyonlarının en önemli sınıflarından birisi Rothaus'un makalesinden birçok özellikler ve çeşitli örnekler vererek sunulmuştur. Bükük fonksiyonları kullanarak dengelilik, doğrusal olmama ve yayılma ölçütü arasındaki ilişkiler sunulmuş ve

bu kriterlerin hepsinin aynı zamanda tamamen sağlanamayacağı gösterilmiştir. Bu nedenden dolayı, Seberry, Zhang ve Zheng'e ait olan ve bu kriterleri optimize eden birçok fonksiyon inşaası sunulmuştur.

Anahtar Kelimeler: Kriptografi, Boole fonksiyonları, Hadamard matrisleri, Sylvester-Hadamard matrisleri, Doğrusal olmama, Tam çığ ölçütü, Yayılma ölçütü, Walsh dönüşümü, Hızlı Walsh dönüşümü, Bükük fonksiyon.

To my parents

ACKNOWLEDGMENTS

I express sincere appreciation to my supervisor Assoc. Prof. Dr. Ali Doğanaksoy for his guidance, insight and cooperation throughout the research without whom this work would never be finished.

I am also thankful to Prof. Dr. İsmail Güloğlu for all the courses he taught me during my undergraduate education in METU and for his excellent suggestions and comments on this thesis, especially on the presentation of Chapter 2.

I want to thank to my family for their support, encouragement and all the beautiful things that they have done for me.

I am also grateful to my friends Berrin Anıl Nalbantoğlu and Zernişan Emirleroğlu Aslan.

I would like to thank to all my colleagues at UEKAE for all they have done to me and I am thankful to my managers Önder Yetiş, Alparslan Babaoğlu and Dr. Murat Apohan for their patience and support.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The concepts completeness and the avalanche effect were first introduced by Kam and Davida [20] and Feistel [16]. Completeness means that each ciphertext bit depends on all bits of the plaintext. This means that if each ciphertext bit were written as a boolean function of each of the plaintext bits, then this function would contain all the plaintext bits, if the system is complete. Avalanche effect means that an average of one half of the output bits should change whenever a single bit of the plaintext is complemented. The concepts of completeness and avalanche effect are combined to define a new property called strict avalanche criterion [55]. Strict avalanche criterion means that each ciphertext bit should change with a probability of one half whenever a single input bit is complemented. As seen from the definitions of completeness, avalanche effect and strict avalanche criterion, they are all milestone concepts to define the theory of any cryptological function, in particular the block ciphers. These three concepts were defined to investigate block ciphers, not the S-boxes or small functions appearing in a block cipher. However, as seen from the definitions, the applicability of these statements is infeasible even for very small functions. These definitions and many successors were later defined for boolean functions. Instead of seeking for necessary properties for designing block ciphers and trying to overcome intractable amount of computations whether those properties are satisfied or not, the properties and theory are developed on core components of block ciphers, namely boolean functions and S-boxes. Then, "cryptologically strong" core functions are either looked for by exhaustive search or constructed theoretically by using this theory. Finally,

they are combined suitably to design strong block ciphers. One other advantage of this approach besides getting rid of infeasible calculations is that an important theory of boolean functions related to cryptology is obtained which is developed by mathematicians, engineers and statisticians.

After having presented some basic definitions and the situation of boolean functions in the theory, the contents of this thesis is as follows :

In Chapter 2, the theory of boolean functions with various definitions are constructed. Many well-known facts are also proved in detail for completeness. The definition of nonlinearity and cryptologically important properties of Hadamard matrices and Sylvester-Hadamard matrices are presented. The difference function of a boolean function f corresponding to a vector α is presented with its properties. The autocorrelation of f with a shift α is presented and finally a special form of the Wiener-Khintchine theorem is presented.

In Chapter 3, two upper and two lower bounds on the nonlinearity of a boolean function is presented. Moreover, for any boolean function, the nonlinearity, balance, linearity dimension and the number of vectors for which the propagation criterion is satisfied are shown to be invariant under nonsingular affine transformations on the input coordinates.

In Chapter 4, the Walsh transform of a boolean function and the Walsh transform of the sign function of a boolean function are presented with their relations to each other. The properties of the Walsh transform of the sign function of a boolean function are listed. The fast Walsh transform is given with a demonstrating example.

In Chapter 5, the definitions of strict avalanche criterion and propagation criterion of degree k are presented. Lower bounds on the number of functions satisfying strict avalanche criterion and asymptotics for $S(n,1)$ and $S(n,2)$ are given where $S(n,k)$ denotes the number of functions for which the output changes with probability exactly one half if any of the input variables $x_1, x_2, \ldots, x_k$ among $x = (x_1, x_2, \ldots, x_n)$ is complemented. An example of an unbalanced SAC fulfilling function is given and a method to construct SAC fulfilling functions is given.

In Chapter 6, bent functions are presented. The fact that they have the largest nonlinearity among all boolean functions is proved. An upper bound on the degree of a bent function is proved. The fact that they satisfy the propagation criterion for any nonzero vector is emphasized. Finally, some known classes of bent functions including Maiorana-McFarland construction and the relation of bent functions with difference sets are mentioned.

In Chapter 7, construction of highly nonlinear balanced functions, construction of highly nonlinear balanced functions satisfying SAC and finally construction of highly nonlinear balanced functions having good propagation characteristics are presented with several examples. Most of these constructions use bent functions by concatenating, splitting or modifying the sequence of bent functions.

CHAPTER 2

PRELIMINARIES

2.1 Boolean Functions and (0,1)-Sequences

Let V_n be the set of all n-tuples of elements of the field $GF(2)$, endowed with the natural vector space structure over $GF(2)$.

V_n possesses a natural ordering known as the lexicographic ordering defined as follows :

For $\alpha = (a_1, a_2, \ldots, a_n)$, $\beta = (b_1, b_2, \ldots, b_n)$ in V_n, set $\alpha < \beta$ if there exists k, $1 \leq k \leq n$ such that $a_1 = b_1$, $a_2 = b_2$, $\ldots$, $a_{k-1} = b_{k-1}$; $a_k = 0$ and $b_k = 1$. It follows that we can list all elements of V_n as $\alpha_0, \alpha_1, \ldots, \alpha_{2^n-1}$ so that $\alpha_0 < \alpha_1 < \ldots < \alpha_{2^n-1}$.

An element $\alpha_k = (a_1, a_2, \ldots, a_n)$ in V_n can be represented by the integer $\sum_{i=1}^{n} a_i.2^{n-i}$, where $0 \leq k \leq 2^n - 1$. With this representation of V_n, it can be shown that α_k corresponds to the integer k and the ordering defined above coincides with the natural ordering of integers. Thus, there is a correspondence between V_n and Z_{2^n} via

$$\begin{aligned} \psi : \quad V_n \quad &\longrightarrow Z_{2^n} \\ x = (x_1, x_2, \ldots, x_n) \quad &\longmapsto x_1.2^{n-1} + x_2.2^{n-2} + \ldots + x_n \end{aligned} \tag{2.1}$$

where Z_{2^n} is the ring of integers modulo 2^n.

The Hamming distance between two vectors in V_n is defined to be the number of unequal (corresponding) components. It follows that the Hamming distance $d(\alpha, \beta)$, for α, β in V_n, is the number of nonzero components of $\alpha + \beta$. The Hamming weight $w(\alpha)$ of a vector α in V_n is the Hamming distance of α

to the zero vector. In other words, $w(\alpha)$ is the number of nonzero components of α. From now on, "the distance" and "the weight" will be used instead of "the Hamming distance" and "the Hamming weight", respectively.

A $GF(2)$-valued function on V_n is referred to as a boolean function. Unless otherwise stated explicitly, by a function, we shall mean a boolean function. The set of all boolean functions will be denoted by $\mathcal{F}_n$.

$\mathcal{F}_n$ is a vector space over the field $GF(2)$ where the addition of two vectors in $\mathcal{F}_n$ and the multiplication of a vector in $\mathcal{F}_n$ with a scalar in $GF(2)$ is defined as follows :

for all f, g in $\mathcal{F}_n$, x in V_n and c in $GF(2)$

$$(f+g) \quad (x) \quad = f(x) + g(x), \tag{2.2}$$

$$(c \cdot f) \quad (x) \quad = c \cdot f(x) \ . \tag{2.3}$$

For a function f, the ordered 2^n-tuple

$$T_f = (f(\alpha_0), f(\alpha_1), \ldots, f(\alpha_{2^n-1}))$$

is called the truth table of f. It is clear that any function f can be uniquely described by its truth table T_f.

So far, we have seen some definitions and properties of V_n. Now, we will see analogies of some definitions made for elements of V_n with elements of $\mathcal{F}_n$. Recall how the distance $d(\alpha, \beta)$ of two vectors α, β in V_n and the weight $w(\alpha)$ of a vector α in V_n were made.

For any function f, the weight of f is the weight of its truth table T_f in V_n. Namely, it is the number of vectors α in V_n such that $f(\alpha) = 1$.

The distance between two functions f, g in $\mathcal{F}_n$ is the distance between their truth tables T_f and T_g in V_n. It is denoted by $d(f, g)$. It follows that $d(f, g) = |\{x \in V_n | f(x) \neq g(x)\}|$. Note that, as $w(f+g)$ is the number of unequal (corresponding) components of T_f and T_g, we have $d(f, g) = w(f+g)$.

A $(0,1)$-sequence $\alpha = (a_1, a_2, \ldots, a_n)$ is called 0,1 balanced or simply balanced if it has an equal number of 0's and 1's. Similarly, a function f in $\mathcal{F}_n$ is balanced if $w(f) = 2^{n-1}$. Note that a function f is balanced if and only if

$w(f) = w(f+1)$ where 1 denotes the all-one constant function in $\mathcal{F}_n$. Namely, it is the function having the property $w(f) = 2^n$.

For a function f the support of f, denoted by $Supp(f)$, is defined as $Supp(f) = \{x \in V_n | f(x) = 1\}$. Note that $|Supp(f)| = w(f)$. Now, consider two functions f, g. Then,

$$\begin{aligned} Supp(f+g) &= \{x \in V_n | (f+g)(x) = 1\} \\ &= \{x \in V_n | f(x) = 1, g(x) = 0\} \cup \{x \in V_n | f(x) = 0, g(x) = 1\} \\ &= [Supp(f) \cap (V_n \setminus Supp(g))] \cup [Supp(g) \cap (V_n \setminus Supp(f))] \\ &= (Supp(f) \setminus Supp(g)) \cup (Supp(g) \setminus Supp(f)) \\ &= Supp(f) \triangle Supp(g) \end{aligned} \tag{2.4}$$

where $A \setminus B$ denotes the difference and $A \triangle B$ denotes the symmetric difference of the sets A and B.

Let F be a finite field. f is a polynomial in the indeterminates $x_1, x_2, \ldots, x_n$ over the field F, denoted by $f \in F[x_1, x_2, \ldots, x_n]$, means that f is a formal sum of the form

$$\sum_{\vec{i}=(i_1,i_2,\ldots,i_n)\in\mathbb{N}^n} c_{\vec{i}} \, . \, x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n} = \sum_{\vec{i}\in\mathbb{N}^n} c_{\vec{i}} \, X^{\vec{i}} \tag{2.5}$$

where $c_{\vec{i}} \in F$ and all but finite $c_{\vec{i}}$'s are equal to 0_F. By $\mathbb{N}$ and $\mathbb{N}^n$ we mean the set of nonnegative integers and n copies of the set $\mathbb{N}$, respectively.

Given a polynomial $f \in F[x_1, x_2, \ldots, x_n]$ of the form (2.5), any $X^{\vec{i}}$ is called a term of f if $c_{\vec{i}} \neq 0_F$. $c_{\vec{i}}$ is called the coefficient of the corresponding term. For any term $X^{\vec{i}} = x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$ of f, the degree of this term is the number of nonzero k's where $i_k \neq 0$ for $1 \leq k \leq n$. In other words, it is the number of indeterminates (variables) appearing in that term. The degree of f denoted by $deg(f)$ is the degree of the highest degree term appearing in f. The degree of a variable x_i in f, denoted by $deg(f, x_i)$, is the degree of the highest degree term among all terms in which x_i appears. Note that the degree of f and the degree of any variable x_i in f can take values from 0 to n.

For any term $X^{\vec{i}} = x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}$ of f, the power of a variable x_k in this term, denoted by $pow(X^{\vec{i}}, x_k)$, is the nonnegative integer i_k where $1 \leq k \leq n$.

Similarly, the power of any variable x_k in f, denoted by $pow(f, x_k)$, is the largest integer among all $pow(X^{\vec{i}}, x_k)$'s where $\vec{i}$ runs through all vectors in $\mathbb{N}^n$ for which $c_{\vec{i}} \neq 0_F$.

For any $f = \sum_{\vec{i} \in \mathbb{N}^n} c_{\vec{i}} \,.\, x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n} \in F[x_1, x_2, \ldots, x_n]$ and any $\alpha = (a_1, a_2, \ldots, a_n) \in F^n$, we write $f(\alpha)$ for the element $\sum_{\vec{i} \in \mathbb{N}^n} c_{\vec{i}} \,.\, a_1^{i_1} a_2^{i_2} \ldots a_n^{i_n}$ of F. This correspondence $\alpha \longmapsto f(\alpha)$ defines a function on F^n which is determined by the polynomial f. It is called the polynomial function on F^n determined by the polynomial f.

Now, we claim that for any function $f : V_n \longrightarrow GF(2)$, there exists a polynomial $\tilde{f} \in GF(2)[x_1, x_2, \ldots, x_n]$ such that $\tilde{f}(\alpha) = f(\alpha)$ for all α in V_n. This claim can easily be seen to be true if one can show this result for the functions χ_α in $\mathcal{F}_n$ which attains the value 1 at exactly one point α in V_n. In other words,

$$\chi_\alpha(\beta) = \begin{cases} 1 & \text{if} \quad \beta = \alpha, \\ 0 & \text{if} \quad \beta \neq \alpha. \end{cases}$$

χ_α is called the characteristic function of α in V_n.

It is clear that we can express any function f as a linear combination of these functions :

$$f = \sum_{\alpha \in V_n} f(\alpha) . \chi_\alpha \,. \tag{2.6}$$

Let $\alpha = (a_1, a_2, \ldots, a_n)$ be in V_n. Consider $\tilde{\chi}_\alpha$, the following polynomial in $GF(2)[x_1, x_2, \ldots, x_n]$:

$$\tilde{\chi}_{(a_1, a_2, \ldots, a_n)}(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} (x_i + a_i + 1). \tag{2.7}$$

Clearly, $\tilde{\chi}_{(a_1, a_2, \ldots, a_n)}(b_1, b_2, \ldots, b_n) = 1$ if and only if $a_i = b_i$ for all $i = 1, 2, \ldots n$. In other words, $\tilde{\chi}_\alpha(\beta) = \chi_\alpha(\beta)$ for all β in V_n. Thus, $\chi_{(a_1, a_2, \ldots, a_n)}$ can be represented by the polynomial $\tilde{\chi}_{(a_1, a_2, \ldots, a_n)}$. Note that any variable x_i has power at most one in $\tilde{\chi}_\alpha$ for all $i = 1, 2, \ldots, n$ and for all $\alpha \in V_n$.

By straightforward computation, we obtain that f is represented by the

polynomial $\tilde{f}$, which is equal to

$$\begin{aligned}
\tilde{f}(x_1, x_2, \ldots, x_n) &= \sum_{\alpha=(a_1,a_2,\ldots,a_n)\in V_n} f(\alpha) \prod_{i=1}^{n}(x_i + a_i + 1) && (2.8)\\
&= \sum_{\alpha=(a_1,a_2,\ldots,a_n)\in V_n} f(\alpha) \sum_{B\subseteq I} \prod_{i\notin B}(a_i + 1) \prod_{i\in B} x_i && (2.9)\\
&= \sum_{B\subseteq I} \left(\sum_{\alpha\in V_n} f(\alpha) \prod_{i\notin B}(a_i + 1) \right) \prod_{i\in B} x_i && (2.10)\\
&= \sum_{B\subseteq I} \left(\sum_{\substack{\alpha \in V_n \\ Supp(\alpha)\subseteq B}} f(\alpha) \right) \prod_{i\in B} x_i && (2.11)\\
&= \sum_{\beta=(b_1,b_2,\ldots,b_n)\in V_n} c_\beta \; . \; x_1^{b_1} x_2^{b_2} \ldots x_n^{b_n} && (2.12)\\
&= \sum_{\beta\in V_n} c_\beta \; X^\beta \; . && (2.13)
\end{aligned}$$

$\tilde{f}$ is a polynomial in which each variable appears with power at most one. Now, a few remarks and explanations will certainly clarify the above expressions greatly.

First of all let us explain the notations and definitions used above. By I, we mean the set $I = \{1, 2, \ldots, n\}$. For any $\alpha = (a_1, a_2, \ldots, a_n)$ in V_n the support of α, denoted by $Supp(\alpha)$, is defined to be the set of i's where $a_i \neq 0$. Namely $Supp(\alpha) = \{\ i \mid a_i \neq 0\}$. It is clear that the function $Supp : V_n \longrightarrow I$ defined as above is a bijection.

Turning back to the equations above, note that (2.8) follows directly from (2.6) and (2.7). (2.9) follows from the observation that for any subset B of I, the coefficient of $\prod_{i\in B} x_i$ in $\prod_{i=1}^{n}(x_i + a_i + 1)$ is $\prod_{i\notin B}(a_i + 1)$. (2.10) is just the reordering of (2.9). (2.11) follows from (2.10) by the following :

In (2.10), $\prod_{i\notin B}(a_i + 1)$ is zero if and only if $a_i = 1$ for some $i \in I \setminus B$. In other words, $\prod_{i\notin B}(a_i + 1) = 1$ if and only if $Supp(\alpha) \cap (I \setminus B) = \emptyset$ if and only if $Supp(\alpha) \subseteq B$. (2.12) follows from (2.11) easily by using the correspondence

between a subset B of I and an element β of V_n via

$$B = \{\ i \mid i \in I\} \longmapsto \beta = (b_1, b_2, \ldots, b_n)$$

where $b_i = 1$ if and only if $i \in B$. Note that this function is in fact the inverse of $Supp : V_n \longrightarrow I$. Finally, (2.13) trivially follows from (2.12) by using (2.5).

Thus, it is proved that for any $f : V_n \longrightarrow GF(2)$, there exists a polynomial $\tilde{f} \in GF(2)[x_1, x_2, \ldots, x_n]$ such that $\tilde{f}(\alpha) = f(\alpha)$ for any α in V_n and each variable appears with power at most one in $\tilde{f}$.

Lemma 2.1.1 *Let $\tilde{f} \in GF(2)[x_1, x_2, \ldots, x_n]$ be a polynomial which vanishes for all α in V_n and each variable appears with power at most one in $\tilde{f}$. Then, $\tilde{f}$ is the zero function. Namely, if $\tilde{f}(X) = \sum_{\alpha \in V_n} c_\alpha\ X^\alpha$, then all c_α's are zero for all α in V_n.*

Proof. We have,

$$\begin{aligned} \tilde{f}(X) &= \sum_{\alpha \in V_n} c_\alpha\ X^\alpha \\ \tilde{f}(x_1, x_2, \ldots, x_n) &= \sum_{\alpha = (b_1, b_2, \ldots, b_n) \in V_n} c_\alpha\ .\ x_1^{b_1} x_2^{b_2} \ldots x_n^{b_n}\ . \end{aligned}$$

Since each variable appears with power at most one in $\tilde{f}$, it can be written as

$$\tilde{f}(x_1, x_2, \ldots, x_n) = x_n.\tilde{f}_1(x_1, x_2, \ldots, x_{n-1}) + \tilde{f}_2(x_1, x_2, \ldots, x_{n-1}) \quad (2.14)$$

where $\tilde{f}_1$ and $\tilde{f}_2$ are functions in $\mathcal{F}_{n-1}$. The proof will be made by induction on n.

For $n = 1$, the lemma clearly holds. Suppose that the lemma is true for $n-1$, that is, for any polynomial $\tilde{f} \in GF(2)[x_1, x_2, \ldots, x_{n-1}]$ which vanishes for all α in V_{n-1} and in which each variable appears with power at most one, $\tilde{f}$ is the zero function, namely if $\tilde{f}(X) = \sum_{\alpha \in V_{n-1}} c_\alpha\ X^\alpha$, then all c_α's are zero for all α in V_{n-1}.

Let $\tilde{f}$ be a polynomial in $GF(2)[x_1, x_2, \ldots, x_n]$ which satisfies all hypotheses in the lemma. Using (2.14), we get that $\tilde{f}(x_1, x_2, \ldots, x_{n-1}, x_n) = 0$ for all

$(x_1, x_2, \ldots, x_{n-1})$ in V_{n-1} and for all x_n in $GF(2)$. It follows that

$$\tilde{f}(x_1, x_2, \ldots, x_{n-1}, 0) = 0 \quad \text{and} \quad \tilde{f}(x_1, x_2, \ldots, x_{n-1}, 1) = 0 \tag{2.15}$$

for all $(x_1, x_2, \ldots, x_{n-1})$ in V_{n-1}.

Using (2.14) with both equations in (2.15), we get that $\tilde{f}_1(X) = 0$ and $\tilde{f}_2(X) = 0$ for all X in V_{n-1}. By induction hypothesis, all coefficients of $\tilde{f}_1$ and $\tilde{f}_2$ are zero implying that all coefficients of $\tilde{f}$ are zero, which we wanted to prove. □

Lemma 2.1.1 is in fact equivalent to the following fact :

Let f be a function in $\mathcal{F}_n$ having two representations $\tilde{f}_1$, $\tilde{f}_2$ in $GF(2)[x_1, x_2, \ldots, x_n]$ such that each variable appears with power at most one in $\tilde{f}_i$ for $i = 1, 2$. Then, $\tilde{f}_1 = \tilde{f}_2$.

This gives the following theorem :

Theorem 2.1.2 *Any function f in $\mathcal{F}_n$ can be uniquely represented by a multivariate polynomial $\tilde{f}$ in $GF(2)[x_1, x_2, \ldots, x_n]$ in which each variable appears with power at most one.*

Note that the assumption that each variable appears with power at most one in the representing polynomial function is crucial for the uniqueness of the representation. If this assumption is not satisfied, then the representation of a function f is not unique.

Given a function f, the function

$$\tilde{f}(X) = \sum_{\alpha \in V_n} c_\alpha \, X^\alpha \tag{2.16}$$

$$= \sum_{\alpha=(a_1,a_2,\ldots,a_n) \in V_n} c_\alpha \, . \, x_1^{a_1} x_2^{a_2} \ldots x_n^{a_n} \tag{2.17}$$

is called the algebraic normal form of f. By a simple rearrangement, $\tilde{f}$ can be written as

$$\tilde{f}(X) = a_0 + \left(\sum_{i=1}^{n} a_i x_i \right) + \left(\sum_{1 \le i < j \le n} a_{ij} x_i x_j \right) + \cdots + a_{12\ldots n} x_1 x_2 \ldots x_n$$

where the coefficients $a_0, a_i, a_{ij}, \ldots, a_{12\ldots n}$ are all in $GF(2)$. Recall the definition of the degree of a variable x_i in f, denoted by $deg(f, x_i)$, for any i satisfying $1 \leq i \leq n$. If $deg(f, x_i) = 1$, then f is said to depend on x_i linearly and such a term of length one is called a linear term. Analogously, if $deg(f, x_i) > 1$, then f is said to depend on x_i nonlinearly and such a term is called a nonlinear term.

Now, we give an example of obtaining the algebraic normal form of a function when its truth table is given.

Example 2.1.1

Let $f : V_4 \longrightarrow GF(2)$ be the function given by

$$T_f = (0, 1, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 1).$$

Using (2.11), we have :

$$\begin{aligned}
\tilde{f}(x_1, x_2, x_3, x_4) = (f(\alpha_0)) &+(f(\alpha_0) + f(\alpha_8)).x_1 +(f(\alpha_0) + f(\alpha_4)).x_2 \\
&+(f(\alpha_0) + f(\alpha_2)).x_3 +(f(\alpha_0) + f(\alpha_1)).x_4 \\
&+(f(\alpha_0) + f(\alpha_4) + f(\alpha_8) + f(\alpha_{12})).\ x_1x_2 \\
&+(f(\alpha_0) + f(\alpha_2) + f(\alpha_8) + f(\alpha_{10})).\ x_1x_3 \\
&+(f(\alpha_0) + f(\alpha_1) + f(\alpha_8) + f(\alpha_9)).\ x_1x_4 \\
&+(f(\alpha_0) + f(\alpha_2) + f(\alpha_4) + f(\alpha_6)).\ x_2x_3 \\
&+(f(\alpha_0) + f(\alpha_1) + f(\alpha_4) + f(\alpha_5)).\ x_2x_4 \\
&+(f(\alpha_0) + f(\alpha_1) + f(\alpha_2) + f(\alpha_3)).\ x_3x_4 \\
+\left(\sum_{i=0}^{7} f(\alpha_{2i})\right).\ x_1x_2x_3 &+\left(\sum_{i=0}^{3}(f(\alpha_{4i}) + f(\alpha_{4i+1}))\right).\ x_1x_2x_4 \\
+\left(\sum_{i=0}^{1}(f(\alpha_{8i}) + f(\alpha_{8i+1}) + f(\alpha_{8i+2}) + f(\alpha_{8i+3}))\right) &.\ x_1x_3x_4 \\
+\left(\sum_{i=0}^{7} f(\alpha_i)\right).\ x_2x_3x_4 &+\left(\sum_{i=0}^{15} f(\alpha_i)\right).\ x_1x_2x_3x_4\ .
\end{aligned}$$

Thus, we find the algebraic normal form of f which is equal to

$$\tilde{f}(x_1, x_2, x_3, x_4) = x_3 + x_4 + x_1x_4 + x_2x_3 + x_3x_4 + x_1x_2x_3 + x_2x_3x_4.$$

Note that if $B = \emptyset$, then all α in V_n satisfying $Supp(\alpha) \subseteq B$ in (2.11) is α_0 only. Similarly, if $B = \{1, 2, 3\}$, then this set is equal to $\{ \alpha_{2i} \mid i = 0, 1, \ldots, 7\}$.

Remark 2.1.3 *From now on, we will make no distinction between the boolean function f and its algebraic normal form $\tilde{f}$. Both of them we will be denoted by f.*

Recall that any function f can be uniquely represented by its truth table T_f which is a 2^n-bit sequence. In other words, the function

$$\begin{aligned} \Xi : \quad \mathcal{F}_n &\longrightarrow V_{2^n} \\ f &\longmapsto T_f \end{aligned} \tag{2.18}$$

is an isomorphism between the vector spaces $\mathcal{F}_n$ and V_{2^n}. It follows that $|\mathcal{F}_n| = 2^{2^n}$.

Recall that a balanced function f is a function with weight 2^{n-1}. We denote the set of all balanced functions in $\mathcal{F}_n$ by $\mathcal{B}_n$. By a simple counting argument $|\mathcal{B}_n| = \begin{pmatrix} 2^n \\ 2^{n-1} \end{pmatrix}$.

Now, we will mention an important relationship between $GF(2^n)$, the Galois field of order 2^n and V_n.

It is well-known that $GF(2^n)$ is a vector space over $GF(2)$ for all positive integers n. Let $\theta = \{\eta_0, \eta_1, \ldots, \eta_{n-1}\}$ be a basis of $GF(2^n)$ over $GF(2)$ and let $x = (x_1, x_2, \ldots, x_n)$ be in V_n. The following function

$$\begin{aligned} \Phi : \qquad V_n \qquad &\longrightarrow GF(2^n) \\ x = (x_1, x_2, \ldots, x_n) \quad &\longmapsto x_1.\eta_0 + x_2.\eta_1 + \cdots + x_n.\eta_{n-1} \end{aligned} \tag{2.19}$$

is an isomorphism when both V_n and $GF(2^n)$ are regarded as vector spaces over $GF(2)$. Let $\theta^d = \{\delta_0, \delta_1, \ldots, \delta_{n-1}\}$ be the dual basis of θ. That is, θ^d is a basis for $GF(2^n)$ over $GF(2)$ and θ, θ^d have the following property :

$$Tr_{GF(2^n)/GF(2)}(\eta_i.\delta_j) = \begin{cases} 1 & \text{if} \quad i = j, \\ 0 & \text{if} \quad i \neq j. \end{cases}$$

where $Tr_{GF(2^n)/GF(2)}(x) = Tr(x) = x+x^2+\cdots+x^{2^{n-1}}$ is the trace function from $GF(2^n)$ to $GF(2)$. Using the trace function and θ^d, one can easily compute the coordinates of x in $GF(2^n)$ with respect to the basis θ by [24] :

$$x = (x_1, x_2, \ldots, x_n) = (Tr(\delta_0.x), Tr(\delta_1.x), \ldots, Tr(\delta_{n-1}.x)) \ . \tag{2.20}$$

Two cryptographically weak but important classes of functions, linear and affine functions, are commonly defined as follows :

A function f is called linear if

$$f(x+y) = f(x) + f(y), \tag{2.21}$$

$$f(c \cdot x) = c \cdot f(x) \tag{2.22}$$

for any x, y in V_n and for any c in $GF(2)$. It is clear that any function f having $f(\alpha_0) = 0$ satisfies (2.22). In other words, $f(\alpha_0) = 0$ for any linear function f. We denote the set of all linear functions in $\mathcal{F}_n$ by $\mathcal{L}_n$. If f and g are both linear functions, then $f + g$ is also a linear function since

$$(f+g)(x+y) = f(x+y) + g(x+y) = (f+g)(x) + (f+g)(y)$$

for any x, y in V_n.

It follows that $\mathcal{L}_n$ is a vector space over $GF(2)$.

Consider the set of all functions in $\mathcal{F}_n$ of the form

$$f(x_1, x_2, \ldots, x_n) = a_1.x_1 + a_2.x_2 + \cdots + a_n.x_n \tag{2.23}$$

where a_i's are in $GF(2)$ for $i = 1, 2, \ldots, n$. It is clear that f is a linear function. Consider the standard ordered bases $\epsilon = \{e_1, e_2, \ldots, e_n\}$ for V_n over $GF(2)$ and $\kappa = \{1\}$ for $GF(2)$ over $GF(2)$. In fact, f is represented by the $1 \times n$ matrix $A = [a_1, a_2, \ldots, a_n]$ relative to the bases ϵ and κ.

Conversely, every linear function is of this form for some $a_1, a_2, \ldots, a_n$ in $GF(2)$. This follows easily from the following :

Let f be a linear function. Then,

$$\begin{aligned} f(x) = f(x_1, x_2, \ldots, x_n) &= f(x_1.e_1 + x_2.e_2 + \cdots + x_n.e_n) \\ &= f(e_1).x_1 + f(e_2).x_2 + \cdots + f(e_n).x_n \\ &= a_1.x_1 + a_2.x_2 + \cdots + a_n.x_n \end{aligned}$$

where $a_i = f(e_i)$ for $i = 1, 2, \ldots, n$.

Recall that $\mathcal{L}_n$ is a vector space over $GF(2)$. The dimension of this space is given by the following :

Theorem 2.1.4 *([19]) The set of all linear functions is an n-dimensional vector space over $GF(2)$.*

Proof. It is clear that the dimension of V_n over $GF(2)$ is n and the dimension of $GF(2)$ over $GF(2)$ is 1. By linear algebra, the dimension of all linear functions from V_n to $GF(2)$, namely the dual space of V_n, denoted by $V_n{}^\star$, has dimension $n.1 = n$. □

By the above theorem or using the algebraic normal form of a linear function, one obtains that $|\mathcal{L}_n| = 2^n$.

Given a vector $\alpha = (a_1, a_2, \ldots, a_n)$ in V_n. We denote a linear function of the form in (2.23) by f_α. Thus, the set of all linear functions is equal to

$$\mathcal{L}_n = \{\ f_\alpha \mid \alpha \in V_n\}.$$

Let $\alpha = (a_1, a_2, \ldots, a_n),\ \beta = (b_1, b_2, \ldots, b_n)$ be in V_n. The standard inner product $\langle ., . \rangle$ is defined on V_n as follows :

$$\langle \alpha, \beta \rangle = \sum_{i=1}^{n} a_i.b_i \tag{2.24}$$

where the addition and multiplication in (2.24) are the corresponding field operations in $GF(2)$.

It is well-known that for any ordered basis $\mu = \{c_1, c_2, \ldots, c_n\}$ for V_n over $GF(2)$, the standard inner product is completely determined by the values $a_{ij} = \langle c_j, c_i \rangle$ for any c_j, c_i in μ. For example, if $\alpha_1 = \sum_{j=1}^{n} u_j.c_j$ and $\alpha_2 = \sum_{i=1}^{n} v_i.c_i$, then

$$\langle \alpha_1, \alpha_2 \rangle \ = \ \langle \sum_{j=1}^{n} u_j.c_j, \sum_{i=1}^{n} v_i.c_i \rangle$$

$$
\begin{aligned}
&= \sum_{j=1}^{n} u_j \sum_{i=1}^{n} v_i \langle c_j, c_i \rangle \\
&= \sum_{i,j=1}^{n} v_i . a_{ij} . u_j \\
&= V.A.U
\end{aligned} \tag{2.25}
$$

where $A = (a_{ij})$ is the matrix with $a_{ij} = \langle c_j, c_i \rangle$ for $i, j = 1, 2, \ldots, n$ and V, U are the coordinate matrices of α_1 and α_2 in the ordered basis μ, respectively [19].

Let α, β be in V_n. Consider the standard inner product on V_n. We say that α and β are orthogonal if $\langle \alpha, \beta \rangle = 0$. Any subset W of V_n is an orthogonal set provided that all pairs of distinct vectors in W are orthogonal.

Theorem 2.1.5 *([19]) For any linear function f, there exists a unique vector α in V_n such that $f(\beta) = \langle \alpha, \beta \rangle$ for all β in V_n.*

Proof. Let $\epsilon = \{e_1, e_2, \ldots, e_n\}$ denote the standard ordered basis for V_n over $GF(2)$. We know that any linear function f is of the form

$$f(x) = f(x_1, x_2, \ldots, x_n) = a_1.x_1 + a_2.x_2 + \cdots + a_n.x_n$$

for some fixed a_i's in $GF(2)$ for $i = 1, 2, \ldots, n$. Clearly, $f(x) = \langle \alpha, x \rangle$ where $\alpha = (a_1, a_2, \ldots, a_n) = (f(e_1), f(e_2), \ldots, f(e_n))$.

For the uniqueness part, suppose that there exists σ in V_n which satisfies $\langle \alpha, \beta \rangle = \langle \sigma, \beta \rangle$ for all β in V_n. Then, $\langle \alpha + \sigma, \beta \rangle = 0$ for all β in V_n. This means that $\langle \alpha + \sigma, e_i \rangle = 0$ for all $i = 1, 2, \ldots, n$. This is equivalent to saying that all components of $\alpha + \sigma$ are zero giving that $\sigma = \alpha$. $\square$

It becomes more clear why the set of all linear functions are denoted by $\mathcal{L}_n = \{ f_\alpha \mid \alpha \in V_n \}$. By the above result, we can also represent $\mathcal{L}_n$ by

$$\mathcal{L}_n = \{ \langle \alpha, x \rangle \mid \alpha, x \in V_n \text{ and } x = (x_1, x_2, \ldots, x_n) \} . \tag{2.26}$$

A function f is said to be an affine function if it is of the form

$$f(x) = f(x_1, x_2, \ldots, x_n) = a_0 + a_1.x_1 + a_2.x_2 + \cdots + a_n.x_n \tag{2.27}$$

for $a_0, a_1, \ldots, a_n$ in $GF(2)$.

The class of affine functions is denoted by $\mathcal{A}_n$. Clearly if $a_0 = 0$, then f is a linear function. In other words, $\mathcal{L}_n$ is properly contained in $\mathcal{A}_n$. Note that for any affine function f, either $f(x+y) = f(x) + f(y)$ holds for all x, y in V_n or never holds for any x, y in V_n. The first case happens for affine functions which are in $\mathcal{L}_n$, whereas the latter holds for functions in $\mathcal{A}_n \setminus \mathcal{L}_n$.

For any function f, the function denoted by $\bar{f}$ is given by $\bar{f}(x) = (f+1)(x)$ where 1 denotes the all-one constant function $1 : V_n \longrightarrow GF(2)$. $\bar{f}$ is called as the complement function, simply the complement of f. In other words, $\bar{f}(x) = f(x) + 1$ for any f and for any x in V_n. So, if T_f is the truth table of f, then $T_{\bar{f}}$, the truth table of $\bar{f}$, is obtained from T_f by simply writing 0 instead of 1 and 1 instead of 0. For any $\alpha = (a_1, a_2, \ldots, a_n)$ in V_n, the complement of α denoted by $\bar{\alpha}$, is similarly defined to be $\bar{\alpha} = (\bar{a_1}, \bar{a_2}, \ldots, \bar{a_n})$ where $\bar{a_i} = a_i + 1$ for $i = 1, 2, \ldots, n$. With this identification, one gets that $T_{\bar{f}} = \bar{T_f}$.

Note also that $Supp(\bar{f}) = V_n \backslash Supp(f)$ which implies that $w(\bar{f}) = 2^n - w(f)$. It follows that f is balanced if and only if $\bar{f}$ is balanced. Similarly, $d(f, \bar{g}) = 2^n - d(f, g)$.

A well-known fact about affine functions is the following which we prove by using algebra :

Lemma 2.1.6 *For every nonconstant affine function f, $w(f) = 2^{n-1}$. Hence, any nonconstant affine function is balanced.*

Proof. Let f be a nonconstant linear function. The kernel of f, defined as $Ker(f) = \{\ x \in V_n \mid f(x) = 0\}$ is a subspace of V_n. As $V_n/Ker(f) \simeq GF(2)$, one gets that $Ker(f) \simeq V_{n-1}$. In other words, $|Ker(f)| = |V_{n-1}| = 2^{n-1}$. This implies that $|Supp(f)| = 2^{n-1}$ giving that f is balanced.

If f is a nonconstant affine function which is not linear, then f is the complement of some linear function. That is, $f(x) = \bar{f_\alpha}(x) = f_\alpha(x) + 1$ for some nonzero α in V_n. Since $w(f) = 2^n - w(f_\alpha)$ and as f_α is linear, one gets that f is balanced. $\square$

2.2 (1,-1)-Sequences

Recall the definition of the truth table T_f of a function f. In this section, an important structure of f which is related to both the function f and its truth table T_f will be investigated.

For any function f, consider the real-valued function $\hat{f}$ which is defined on V_n as follows :

$$\hat{f}(x) = \begin{cases} 1 & \text{if} \quad f(x) = 0, \\ -1 & \text{if} \quad f(x) = 1. \end{cases}$$

It is easy to see that the function $\hat{f} : V_n \longrightarrow \mathbb{R}$ can be expressed as

$$\hat{f}(x) = (-1)^{f(x)} = 1 - 2f(x)$$

for all x in V_n.

The truth table of the function $\hat{f}$ is called as the sequence of f and denoted by ζ_f [44]. In other words,

$$\zeta_f = T_{\hat{f}} = \left((-1)^{f(\alpha_0)}, (-1)^{f(\alpha_1)}, \ldots, (-1)^{f(\alpha_{2^n-1})}\right).$$

If we denote the i-th component in T_f and ζ_f by $T_f[i]$ and $\zeta_f[i]$, respectively, then T_f and ζ_f are related to each other by the obvious relation $(-1)^{T_f[i]} = \zeta_f[i]$ for all $i = 0, 1, \ldots, 2^n - 1$. Hence, the weight $w(f)$ of f and the distance $d(f, g)$ of two functions f, g can also be stated in terms of their sequences as follows :

The weight $w(f)$ of a function f is the number of -1's in ζ_f and the distance between two functions f, g is the distance between their sequences ζ_f and ζ_g, where the distance between two $(1, -1)$-sequences is defined to be the number of unequal (corresponding) components. A function f is balanced if its sequence ζ_f has an equal number of 1's and -1's.

For any two functions f, g, it is easy to see that $\widehat{(f+g)}(x) = \hat{f}(x).\hat{g}(x)$ for all x in V_n [11]. Thus, the sequence of the function $f + g$ is the product of ζ_f and ζ_g where this product is the componentwise multiplication of real numbers. If this operation is denoted by $*$, then $\zeta_{f+g} = \zeta_f * \zeta_g$ [44]. Similarly, for any function f, the sequence of the complement of f satisfies the relation

$\zeta_{\bar{f}} = -\zeta_f$ [44] where the multiplication of ζ_f by a real number c is defined to be the multiplication of all its components by c.

For any function f its sequence ζ_f is a $(1,-1)$-sequence of length 2^n. We denote the set of all $(1,-1)$-sequences of length 2^n by V^{2^n}. It is well-known that the set of all real vectors of length 2^n, denoted by $\mathbb{R}^{2^n}$ is a vector space with the addition of two vectors being the real addition of components and the scalar multiplication being the multiplication of components with the elements of $\mathbb{R}$. Also, it is clear that V^{2^n} is contained in $\mathbb{R}^{2^n}$ but it is not a subspace.

Given two functions f, g, their inner product in $\mathbb{R}^{2^n}$ is defined to be [11] :

$$\langle \hat{f}, \hat{g} \rangle = \sum_{i=0}^{2^n-1} \hat{f}(\alpha_i).\hat{g}(\alpha_i) \tag{2.28}$$

where the addition and multiplication are the usual operations in $\mathbb{R}$. It is clear that we can write the above equation as

$$\langle \zeta_f, \zeta_g \rangle = \sum_{i=0}^{2^n-1} \zeta_f[i].\zeta_g[i] \ . \tag{2.29}$$

Note that the norm of any function f induced by this inner product is constant. Namely, $\|f\| = \sqrt{\langle \hat{f}, \hat{f} \rangle} = 2^{\frac{n}{2}}$ [11].

An important fact about the set $\{ \ \zeta_{f_\alpha} \mid \alpha \in V_n \}$, where ζ_{f_α} is the sequence of the linear function $f_\alpha(x) = \langle \alpha, x \rangle$, is that this set forms an orthogonal basis for $\mathbb{R}^{2^n}$ over $\mathbb{R}$ with respect to the inner product defined in (2.28) [11]. The fact that this set is a basis can be simply seen by noting that the dimension of $\mathbb{R}^{2^n}$ over $\mathbb{R}$ and the cardinality of this set are both 2^n and this set is a linearly independent set. The orthogonality follows from this observation :

Let α, β be vectors in V_n. Then,

$$\begin{aligned} \langle \zeta_{f_\alpha}, \zeta_{f_\beta} \rangle &= \sum_{x \in V_n} (-1)^{\langle \alpha+\beta, x \rangle} \\ &= \begin{cases} 2^n & \text{if} \quad \alpha = \beta, \\ 0 & \text{if} \quad \alpha \neq \beta. \end{cases} \\ &= 2^n \delta(\alpha + \beta) \end{aligned} \tag{2.30}$$

where $\delta(u)$ is the Kronecker delta function which is equal to one if u is equal to the zero vector and zero otherwise.

Thus, for any function f, its associated real-valued function $\hat{f}$ can be written in terms of the associated real-valued functions to the set of linear functions. In other words,

$$\hat{f}(x) = \sum_{\alpha \in V_n} c_\alpha . \hat{f}_\alpha(x) \tag{2.31}$$

where c_α's are the corresponding real coefficients to $\hat{f}_\alpha$'s, namely to the orthogonal basis of $\mathbb{R}^{2^n}$ over $\mathbb{R}$. In Chapter 4, it will be seen that the coefficients c_α's can easily be determined by an important function, called Walsh transform.

The following is a simple but important lemma which will be frequently used in the coming chapters since it relates the distance between two functions to their sequences.

Lemma 2.2.1 *([44]) Let f, g be functions with sequences ζ_f, ζ_g, respectively. Then, $d(f,g) = 2^{n-1} - \frac{1}{2}\langle \zeta_f, \zeta_g \rangle$.*

Proof. We have

$$\begin{aligned} \langle \zeta_f, \zeta_g \rangle &= \sum_{x \in V_n} (-1)^{(f+g)(x)} \\ &= 2^n - 2\, w(f+g). \end{aligned}$$

Since $w(f+g) = d(f,g)$, the result follows. □

2.3 Nonlinearity

Let $\varphi_0, \varphi_1, \ldots, \varphi_{2^{n+1}-1}$ denote all affine functions so that the first half consists of linear functions ordered according to the relation $\varphi_i = f_{\alpha_i}$ for all $i = 0, 1, \ldots, 2^n - 1$ and the second half consists of the (respective) complements of the functions in the first half. Thus, $\varphi_i = \bar{f}_{\alpha_i}$ for all $i = 2^n, 2^n + 1, \ldots, 2^{n+1} - 1$.

The nonlinearity of a function f is defined as [29] :

$$N_f = \min_{i=0,1,\ldots,2^{n+1}-1} d(f, \varphi_i). \tag{2.32}$$

In other words, the nonlinearity of a function is the distance between the function and the set $\mathcal{A}_n$. High nonlinearity is a crucial criterion for a good cryptographic design since it assures resistance against linear cryptanalysis introduced by Matsui [27]. The concept of nonlinearity was introduced by Pieprzyk and Finkelstein [36].

Note that if f is an affine function, then $f = \varphi_{i_k}$ for some $0 \leq i_k \leq 2^{n+1}-1$. In other words, $d(f, \varphi_{i_k}) = 0$ which implies that $N_f = 0$. If f is not an affine function, then $N_f > 0$ as $d(f, \varphi_i) > 0$ for all $i = 0, 1, \ldots, 2^{n+1} - 1$. Hence, $N_f = 0$ if and only if f is an affine function.

Thus, the nonlinearity criterion simply divides all functions as affine functions and nonaffine functions. By abuse of terminology, the first set is sometimes called as linear functions (affine functions) whereas the second set is called as nonlinear functions (nonaffine functions).

Introducing a new method of cryptanalysis of a specific cryptographic design commonly leads to a new design criterion for the similar cryptographic designs. Linear cryptanalysis and nonlinearity is an example to this situation. Nonlinearity measures the quality of a function via its distance to affine functions. In other words, it measures how well a function under consideration may be linearly approximated. Linear cryptanalysis tries to find the best linear approximation, called "effective" linear expression of an algorithm, by finding good approximations to the nonlinear part of the algorithm and extends these approximations to the round function.

In general, exploring the theoretical facts about a cryptological criterion is not enough to understand why a particular design should or should not satisfy it. It is extremely important to work on concrete examples and to apply cryptanalysis methods to systems which are weak in satisfying that criterion. Although two cryptanalysis methods applied to two different designs may seem totally different, they may be based exactly on the same idea and depend on the same kind of pathology occurring in both designs. Similar to cryptanalysis methods, this may happen for cryptographic criteria also. Two cryptographic criteria may seem totally different, although they may be imposing crypto-

graphically equivalent conditions. By cryptographically equivalent conditions, we mean two criteria such that when a design is tested with respect to one of the criteria, then testing with the other one is redundant.

From the designer's point of view, an example to this situation is the nonlinearity criterion. The definition of nonlinearity in this section is the simplest and the most widely accepted one. It is clear that by using this definition of nonlinearity, it is difficult to make a healthy comparison between two functions one of which is in, say $\mathcal{F}_5$ and the other is in, say $\mathcal{F}_9$. By defining the nonlinearity so that the nonlinearity of a function takes values between 0 and 1 would of course handle this problem. However, it is wise not to consider such a form of nonlinearity as a different cryptological criterion from the one stated in (2.32) unless the new form behaves more sensitive in separating cryptologically strong and weak functions or the new form helps in making an algorithm resistant to a cryptanalysis method different from linear cryptanalysis.

In [29], Meier and Staffelbach investigate some properties of the form of nonlinearity which is defined in (2.32) and two additional forms of nonlinearity. One form is defined as the distance to the set of functions having linear structures, called distance to linear structures and the other form is defined as the degree of the considered function, called nonlinear order. We haven't defined what a linear structure means yet. It will be defined in Section 2.6. However, it is worth to mention that the set of functions having linear structures contain the set of affine functions properly and the nonlinear order takes integer values from 0 to n, as noted in Section 2.1. Properties of the nonlinear order are also investigated by O'Connor and Klapper in [33]. They call this form of nonlinearity as algebraic nonlinearity. Hence, from the explanations in the above paragraph, it is clear that the nonlinear order and the algebraic nonlinearity are in fact the same criterion. However, the distance to affine functions (2.32), the distance to linear structures and the nonlinear order should be treated as different cryptological criteria. Meier and Staffelbach proved the invariance of the nonlinearity criterion under nonsingular affine transformations on the input coordinates [29]. This fact will be proved in Section 3.4.

2.4 Sylvester-Hadamard Matrices

The class of square matrices which will be described in this section are very useful in the subsequent sections. We start by a definition :

An $n \times n$ matrix H with entries 1,-1 is called a Hadamard matrix if

$$H.H^t = n.I_n \tag{2.33}$$

where H^t is the transpose of H and I_n is the $n \times n$ identity matrix. Instead of using "an $n \times n$ matrix", we may sometimes use "a matrix of order n" for square matrices.

Theorem 2.4.1 *[40] If a Hadamard matrix of order n exists, then $n = 1, 2$ or $n \equiv 0 \ (mod\ 4)$.*

Proof. Let H be a Hadamard matrix of order n. It is clear by definition that all distinct rows of H are orthogonal. If we change the sign of every entry in any column of H, i.e. if we multiply any column by -1, then the resulting matrix is also a Hadamard matrix. Hence, any matrix obtained from H by multiplying some columns with -1 is also a Hadamard matrix. By changing the signs of all columns for which the entry in the first row is -1, we can make all entries in the first row 1.

Since every other row is orthogonal to the first row, one gets that all these rows have m entries equal to 1 and m entries equal to -1, where $n = 2m$. Moreover, if $n > 2$, then the first three rows are as follows :

$$\begin{array}{cccc} +1\cdots+1 & +1\cdots+1 & +1\cdots+1 & +1\cdots+1 \\ +1\cdots+1 & +1\cdots+1 & -1\cdots-1 & -1\cdots-1 \\ +1\cdots+1 & -1\cdots-1 & +1\cdots+1 & -1\cdots-1 \end{array}$$

Thus, one gets that $n = 4k$, where k is the length of each subsequence above. □

It is conjectured that there is a Hadamard matrix of every order divisible by 4. This is equivalent to saying that the necessary condition in the above

theorem is also sufficient. The smallest multiple of 4 for which no Hadamard matrix has been constructed is currently 428 [40].

Note that by the proof of the above theorem, one gets that there are several operations which preserve the Hadamard property [40] :

1. Permuting rows and multiplying any row by -1.

2. Permuting columns and multiplying any column by -1.

3. Transposition. In other words, if H is a Hadamard matrix, then H^t is also a Hadamard matrix. This is obvious from (2.33).

Any Hadamard matrix which has every element of its first row and first column +1 is called normalized [40].

It is well-known that if H is a normalized Hadamard matrix of order $4n$, then every row (column) except the first has $2n$ entries equal to -1 and $2n$ entries equal to +1. Furthermore, n -1's in any row (column) overlap with n -1's in any other row (column) [40].

Definition 2.4.1 *([25]) If $A = (a_{ij})$ is an $m \times n$ matrix and $B = (b_{ij})$ is a $p \times q$ matrix, then the Kronecker product of A and B is the $mp \times nq$ matrix given as follows :*

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}. \tag{2.34}$$

That is, $A \otimes B$ is an $mp \times nq$ matrix made up of $m \times n$ blocks where the (i, j) block is $a_{ij}B$. Here, $a_{ij}B$ denotes the $p \times q$ matrix obtained by multiplying each entry of B with a_{ij}.

Lemma 2.4.2 *([25]) Let $A = (a_{ij}), B = (b_{ij}), C = (c_{ij})$ and $D = (d_{ij})$ be matrices. The Kronecker product satisfies the following :*

(1) *$A \otimes (B \otimes C) = (A \otimes B) \otimes C$ (Associativity).*

(2) $(A+B)\otimes C = A\otimes C + B\otimes C$ *(Distributivity).*

(3) $(A\otimes B)(C\otimes D) = AC\otimes BD$.

Proof. Only the third property will be proved. The first two are easy to prove. Note that for the first property A, B and C may be any matrices, while for the second property the matrices A and B should have the same dimensions whereas C may be of any dimension.

(3) ([51]) First, note that in order to have the operations on both sides to be well-defined, the number of columns of A should be equal to the number of rows of C. This is also true for C and D. For these reasons, without loss of generality let A, B, C and D be $m\times n$, $p\times q$, $n\times t$ and $q\times r$ matrices, respectively. By definition $A\otimes B = (a_{ij}B)$ and $C\otimes D = (c_{ij}D)$. Let $(A\otimes B)(C\otimes D) = (\theta_{ij})$, an $mp\times tr$ matrix. Then,

$$\theta_{ij} = \sum_{k=1}^{nq}(a_{ik}B)\ (c_{kj}D) = \sum_{k=1}^{nq} a_{ik}c_{kj}\ BD.$$

Also, let $AC = (\beta_{ij})$. Then, $\beta_{ij} = \sum_{k=1}^{n} a_{ik}c_{kj}$. Since, $AC\otimes BD = (\beta_{ij}\ BD)$ an $mp\times tr$ matrix, one gets that $(A\otimes B)(C\otimes D) = AC\otimes BD$ □

Some examples of Hadamard matrices are $\begin{bmatrix} 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$,

$$\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}. \tag{2.35}$$

Theorem 2.4.3 *(Hadamard [18])*

Let H_1 and H_2 be Hadamard matrices of orders n_1 and n_2. Then, $H_1\otimes H_2$ is a Hadamard matrix of order n_1n_2.

We have seen that the matrix $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ is a Hadamard matrix. By

Hadamard's theorem, the 4×4 matrix

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

is also a Hadamard matrix. In fact, using Hadamard's theorem repeatedly, the iterated Kronecker product of n copies of the Hadamard matrix $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ is a Hadamard matrix of order 2^n. In literature, this class of Hadamard matrices is called as Sylvester-Hadamard matrices. Thus, we have the following :

Theorem 2.4.4 *(Sylvester [50])*

There is a Hadamard matrix of order 2^n for all nonnegative integers n.

Hence, the recursion generating all Sylvester-Hadamard matrices are given as follows :

$$H_0 = \begin{bmatrix} 1 \end{bmatrix}, \quad H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{2.36}$$

and

$$H_n = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes H_{n-1} = \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}. \tag{2.37}$$

Note that H_n is a matrix of order 2^n. Thus the equation (2.33) turns out to be $H_n.{H_n}^t = 2^n.I_{2^n}$ where I_{2^n} is the identity matrix of order 2^n.

By using (2.37) for $n = 2$, one obtains the last matrix in (2.35) as H_2 and

for $n = 3$ H_3 is found as

$$H_3 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}. \tag{2.38}$$

2.5 Relationships Among Sylvester-Hadamard Matrices and Linear Functions $\mathcal{L}_n$

Sylvester-Hadamard matrices are useful because of their relation with linear functions in $\mathcal{F}_n$.

Lemma 2.5.1 *([41]) Let* $H_n = \begin{bmatrix} l_0 \\ l_1 \\ \vdots \\ l_{2^n-1} \end{bmatrix}$ *denote the Sylvester-Hadamard matrix of order* 2^n *for* $n \geq 0$ *where* l_i *denotes the i-th row of* H_n*. Then,* l_i *is the sequence of the linear function* $f_{\alpha_i}(x) = \langle \alpha_i, x \rangle$ *for any* $i = 0, 1, \ldots, 2^n - 1$ *where* α_i *is in* V_n *.*

Proof. We use induction on n.

For $n = 1$, H_1 is given in (2.36). It is easy to see that l_0 is the sequence of $f_{\alpha_0}(x_1) = \langle \alpha_0, x_1 \rangle = 0$ and l_1 is the sequence of $f_{\alpha_1}(x_1) = \langle \alpha_1, x_1 \rangle = x_1$, where f_{α_0} and f_{α_1} are linear functions $\mathcal{F}_1$.

Suppose that the lemma is true for n. Since $H_{n+1} = H_1 \otimes H_n$ from (2.36) and (2.37), it is easy to see that each row δ_i of H_{n+1} is the Kronecker product of a row $(1, 1)$ or $(1, -1)$ of H_1 and a row of H_n. Thus, any row δ_i of H_{n+1} is either (l_i, l_i) or $(l_i, -l_i)$ for some row l_i of H_n. By induction hypothesis, l_i is the sequence of the linear function $f_{\alpha_i}(x_2, x_3, \ldots, x_{n+1}) = \langle \alpha_i, x \rangle$ where α_i is

in V_n and $x = (x_2, x_3, \ldots, x_{n+1})$. However, observe that (l_i, l_i) and $(l_i, -l_i)$ are the sequences of the linear functions $f_{\gamma_i}(x_1, x_2, \ldots, x_{n+1}) = \langle \gamma_i, x \rangle$ where γ_i is in V_{n+1} for $\gamma_i = (0, \alpha_i)$ in the first case and $\gamma_i = (1, \alpha_i)$ in the second case, α_i is in V_n and $x = (x_1, x_2, \ldots, x_{n+1})$. $\square$

The above theorem tells us that the i-th row l_i of H_n is the sequence ζ_i of the linear function corresponding to the binary representation α_i of the integer i. The main importance of this fact is that, the n-th Sylvester-Hadamard matrix is itself nothing but a complete table of the sequences of all linear functions in $\mathcal{F}_n$. Using the trivial relation between the sequence and the truth table of a function, one has the truth tables of all linear functions. What is meant by "to have the truth tables of all linear functions" is obtaining them without performing the evaluation of any linear function on V_n or on any ordered basis of V_n. This work is reduced to only a simple iterated matrix operation.

Another simple observation is that since H_n is a symmetric matrix, the above lemma is also true for columns of H_n.

Note that if l_i is the sequence of the linear function $f_{\alpha_i}(x) = \langle \alpha_i, x \rangle$ for α_i in V_n, then $-l_i$ is the sequence of the complement of f_{α_i}. Thus, the rows of the matrix $-H_n$ contain the sequences of the complements of all linear functions. Consequently, H_n and $-H_n$ together contain the sequences of all affine functions $\varphi_0, \varphi_1, \ldots, \varphi_{2^{n+1}-1}$. As in Lemma 2.5.1, the ones corresponding to linear functions are denoted by $l_0, l_1, \ldots, l_{2^n-1}$ and the ones corresponding to the complements of linear functions are denoted by $l_{2^n}, l_{2^n+1}, \ldots, l_{2^{n+1}-1}$. Thus, $l_{i+2^n} = -l_i$ for all $i = 0, 1, \ldots, 2^n - 1$. By using Lemma 2.5.1, one can find the nonlinearity of a function with a simple algorithm as follows :

Let $\zeta_f.H_n = \mu$ where $\mu = [a_0, a_1, \ldots, a_{2^n-1}]$, a_i's are integers and f is a function with sequence ζ_f. Then, $\tilde{\mu} = (\tilde{a_i})$ is the 1×2^n matrix which contains the distances of f to all linear functions where $\tilde{a_i} = 2^{n-1} - \frac{1}{2}a_i$. By multiplying ζ_f with $-H_n$, one gets that $-\mu = [-a_0, -a_1, \ldots, -a_{2^n-1}]$ and $\tilde{-\mu} = (\tilde{b_i})$ where $\tilde{b_i} = 2^{n-1} + \frac{1}{2}a_i$. $\tilde{-\mu}$ is the matrix containing the distances of f to the complements of all linear functions. Hence, we have the set which

contains the distances of f to all affine functions. Taking the minimum over this set, one gets the nonlinearity of f.

Indeed, performing the multiplication with $-H_n$ is not necessary when finding the nonlinearity of a function. This slight improvement will be proved in this section. Moreover, this simple algorithm can be significantly improved in terms of decreasing the operations made. This is due to an important property satisfied by Sylvester-Hadamard matrices. This property of Sylvester-Hadamard matrices will be proved and the improved algorithm will be demonstrated for $n = 3$ in the chapter devoted to Walsh transform and its properties.

An important question about nonlinearity is the following :

What is the largest possible nonlinearity that can be attained by a function? The following lemma answers this question.

Lemma 2.5.2 *([41]) For any function f, its nonlinearity N_f satisfies the relation $N_f \leq 2^{n-1} - 2^{\frac{n}{2}-1}$.*

Proof. Let $H_n = \begin{bmatrix} l_0 \\ l_1 \\ \vdots \\ l_{2^n-1} \end{bmatrix}$ denote the Sylvester-Hadamard matrix of order 2^n where l_i denotes the i-th row of H_n for $i = 0, 1, \ldots, 2^n - 1$. Since H_n is a symmetric matrix, one has

$$\zeta_f.H_n = \left(\langle \zeta_f, l_0 \rangle, \langle \zeta_f, l_1 \rangle, \ldots, \langle \zeta_f, l_{2^n-1} \rangle\right). \tag{2.39}$$

Also,

$$\left(\zeta_f.H_n\right)\left(\zeta_f.H_n\right)^t = \zeta_f H_n {H_n}^t {\zeta_f}^t = 2^n.\zeta_f {\zeta_f}^t = 2^{2n}. \tag{2.40}$$

Computing the left hand side of (2.40) using (2.39), one gets that

$$\left(\zeta_f.H_n\right)\left(\zeta_f.H_n\right)^t = \sum_{j=0}^{2^n-1} \langle \zeta_f, l_j \rangle^2.$$

By combining these results, one obtains that

$$\sum_{j=0}^{2^n-1} \langle \zeta_f, l_j \rangle^2 = 2^{2n}. \tag{2.41}$$

From (2.41), there exists a j_k satisfying $0 \leq j_k \leq 2^n - 1$ such that $\langle \zeta_f, l_{j_k} \rangle^2 \geq 2^n$. From this, it follows that $\langle \zeta_f, l_{j_k} \rangle \geq 2^{\frac{n}{2}}$ or $\langle \zeta_f, l_{j_k} \rangle \leq -2^{\frac{n}{2}}$. Now,

- If the first case is true, then by Lemma 2.2.1 $d(f, \varphi_{j_k}) \leq 2^{n-1} - 2^{\frac{n}{2}-1}$.
- If the second case holds, then $\langle \zeta_f, -l_{j_k} \rangle = \langle \zeta_f, l_{j_k+2^n} \rangle \geq 2^{\frac{n}{2}}$ where $l_{j_k+2^n} = -l_{j_k}$. Again, by Lemma 2.2.1 $d(f, \varphi_{j_k+2^n}) = d(f, \bar{\varphi}_{j_k}) \leq 2^{n-1} - 2^{\frac{n}{2}-1}$.

However, by the choice of j_k, either $\langle \zeta_f, l_{j_k} \rangle$ or $\langle \zeta_f, l_{j_k+2^n} \rangle$ is the largest among all j's for $j = 0, 1, \ldots, 2^{n+1} - 1$, giving that either $d(f, \varphi_{j_k})$ or $d(f, \bar{\varphi}_{j_k})$ is the smallest among all affine functions $\varphi_0, \varphi_1, \ldots, \varphi_{2^{n+1}-1}$. Hence, $N_f \leq 2^{n-1} - 2^{\frac{n}{2}-1}$. □

Lemma 2.5.3 *([45]) Let f be a function with sequence ζ_f. The nonlinearity N_f of f can be found by $N_f = 2^{n-1} - \frac{1}{2} \max\limits_{i=0,1,\ldots,2^n-1} \{|\langle \zeta_f, l_i \rangle|\}$ where l_i is the i-th row of H_n.*

Proof. Let φ_i be an arbitrary linear function for $i = 0, 1, \ldots, 2^n - 1$. As in the proof of Lemma 2.5.2, $d(f, \bar{\varphi}_i) = 2^{n-1} - \frac{1}{2}\langle \zeta_f, l_{i+2^n} \rangle$ where l_{i+2^n} is the i-th row of $-H_n$, i.e. $l_{i+2^n} = -l_i$ is the sequence of $\bar{\varphi}_i$. Hence, $d(f, \bar{\varphi}_i) = 2^{n-1} + \frac{1}{2}\langle \zeta_f, l_i \rangle$. It follows that, $N_f = \min\limits_{i=0,1,\ldots,2^{n+1}-1} d(f, \varphi_i) = 2^{n-1} - \dfrac{1}{2} \max\limits_{i=0,1,\ldots,2^n-1} \{|\langle \zeta_f, l_i \rangle|\}$. □

Lemma 2.5.3 shows why it is unnecessary to perform the multiplication with $-H_n$ in order to find the nonlinearity of the function. Lemma 2.5.3 states that it is enough to use $\tilde{\mu}$ where $\tilde{\mu} = (\tilde{a}_i)$ is the 1×2^n matrix for which $\tilde{a}_i = 2^{n-1} - \frac{1}{2}|a_i|$ and a_i's are integers such that $a_i = \langle \zeta_f, l_i \rangle$ for $i = 0, 1, \ldots, 2^n - 1$.

Lemma 2.5.4 *Let α, β be two vectors in V_n with even weight. Then, $d(\alpha, \beta)$ is also even.*

Proof. Let $\alpha = (a_1, a_2, \ldots, a_n)$ and $\beta = (b_1, b_2, \ldots, b_n)$. Since

$$w(\alpha + \beta) = |Supp(\alpha + \beta)| = |Supp(\alpha)| + |Supp(\beta)| - 2|Supp(\alpha) \cap Supp(\beta)|$$

and as the right hand side of the above equation is even, we conclude that $w(\alpha+\beta)$ is also even. Thus, $d(\alpha,\beta)$ is even. □

Remark 2.5.5 *The above lemma can be found in [41] with an additional hypothesis that the length of α and β should be even. Its proof is different than the proof we have made above and uses this assumption. However, the proof above does not use this assumption. This shows that n may be arbitrary. In any case, this extra assumption does not effect the following corollary since the length of the truth table of any function in $\mathcal{F}_n$ is always even.*

Corollary 2.5.6 *[44] Let f be a balanced function for $n \geq 3$. Then,*

$$N_f \leq \begin{cases} 2^{n-1} - 2^{\frac{n}{2}-1} & \text{if } n \text{ is even,} \\ \left\lfloor 2^{n-1} - 2^{\frac{n}{2}-1} \right\rfloor & \text{if } n \text{ is odd.} \end{cases}$$

where $\lfloor x \rfloor$ denotes the largest even integer not exceeding x.

Proof. If n is even, then $2^{n-1} - 2^{\frac{n}{2}-1}$ is an integer and by Lemma 2.5.2, $N_f \leq 2^{n-1} - 2^{\frac{n}{2}-1}$.

Now, let n be odd. Since f and all φ_i's are balanced for $i = 0, 1, \ldots, 2^{n+1}-1$, their weights are even. By Lemma 2.5.4, $d(f, \varphi_i)$ is even for all i. As n is odd, $2^{n-1} - 2^{\frac{n}{2}-1}$ is not an integer implying that $N_f < 2^{n-1} - 2^{\frac{n}{2}-1}$. This gives the desired result. □

2.6 Difference Function, Linear Structures, Auto-correlation of f and Properties

Let f be a function with sequence ζ_f. For any α in V_n, the sequence of the function $h(x) = f(x+\alpha)$ is denoted by $\zeta_f(\alpha)$ [57]. Clearly, $\zeta_f = \zeta_f(\alpha_0)$.

The difference function corresponding to α in V_n is defined as the function $f^\alpha(x) = f(x) + f(x+\alpha)$ in the literature.

The auto-correlation of f with a shift α is defined as in [57] :

$$\Delta_f(\alpha) = \langle \zeta_f, \zeta_f(\alpha) \rangle. \tag{2.42}$$

Lemma 2.6.1 *([57]) Let f be a function. The weight of the function f^α for any α in V_n is equal to $2^{n-1} - \frac{1}{2}\Delta_f(\alpha)$.*

Proof. It is enough to observe that

$$\begin{aligned}\Delta_f(\alpha) &= \langle \zeta_f, \zeta_f(\alpha) \rangle = \sum_{x \in V_n} (-1)^{f(x)+f(x+\alpha)} \\ &= 2^n - 2w(f^\alpha).\end{aligned}$$

□

Note that $\Delta_f(\alpha_0) = 2^n$ for any function f.

Corollary 2.6.2 *For any function f and for any nonzero α in V_n, f^α is balanced if and only if $\Delta_f(\alpha) = 0$.*

It is obvious that for any function f, $|\Delta_f(\alpha)| = 2^n$ if and only if the function f^α is constant. This situation is given a special name in the literature.

Definition 2.6.1 *[6] A vector α is said to be a linear structure of the function f if f^α is a constant function.*

The following well-known fact can be found in [33] without proof. For the sake of completeness, we prove it :

Lemma 2.6.3 *For any function f, the set of all linear structures of f forms a vector space over $GF(2)$.*

Proof. Let α, β be linear structures of f. Thus, the functions $f^\alpha(x) = f(x) + f(x+\alpha)$ and $f^\beta(x) = f(x) + f(x+\beta)$ are constant. It follows that, $g(x) = f^\alpha(x) + f^\beta(x) = f(x+\alpha) + f(x+\beta)$ is constant. Note that $g(x) = f(x) + f(x+\alpha+\beta)$ for any x in V_n. Since g is constant, we get that $\alpha + \beta$ is a linear structure of f. □

We denote the set of all linear structures of a fixed function f by $\mathcal{LS}_f$. By previous lemma, $\mathcal{LS}_f$ is a subspace of V_n. The dimension of $\mathcal{LS}_f$ as a vector space is said to be the linearity dimension of f [46].

Now, $|\Delta_f(\alpha)| = 2^n$ if and only if the function f^α is constant. This can be stated in terms of linear structures as follows :

$|\Delta_f(\alpha)| = 2^n$ if and only if α is a linear structure of f, i.e. $\mathcal{LS}_f$ contains α.

Consider the equation in (2.41). It states that

$$\sum_{j=0}^{2^n-1} \langle \zeta_f, l_j \rangle^2 = 2^{2n}$$

holds for any function f, where ζ_f is the sequence of f and $l_0, l_1, \ldots, l_{2^n-1}$ are the rows of H_n. This equation is called as "Parseval's equation" [25]. In Chapter 4, this equation will be written in terms of the Walsh transform of f. In the literature this one is more common than the one in (2.41).

For any function f, the square matrix M_f of order 2^n given by $M_f = (m_{ij})$ where $m_{ij} = (-1)^{f(\alpha_i + \alpha_j)}$ is called the matrix of f [58]. This matrix will be an important tool in proving the important connection between $\langle \zeta_f, l_j \rangle^2$ in (2.41) and $\Delta_f(\alpha_j) = \langle \zeta_f, \zeta_f(\alpha_j) \rangle$ in (2.42).

First of all, it is obvious that M_f is a symmetric matrix. Since, the first row of M_f is $\zeta_f(\alpha_0)$ and the i-th column is $\zeta_f(\alpha_i)$ for $i = 0, 1, \ldots, 2^n - 1$, one gets that the first row of $M_f.{M_f}^t$ is $(\langle \zeta_f, \zeta_f \rangle, \langle \zeta_f, \zeta_f(\alpha_1) \rangle, \ldots, \langle \zeta_f, \zeta_f(\alpha_{2^n-1}) \rangle)$, which is equal to

$$(\Delta_f(\alpha_0), \Delta_f(\alpha_1), \ldots, \Delta_f(\alpha_{2^n-1}))\,. \tag{2.43}$$

According to a result by McFarland (see [14]), the matrix M_f can be represented as

$$M_f = 2^{-n}\ H_n \begin{bmatrix} \langle \zeta_f, l_0 \rangle & & & \\ & \langle \zeta_f, l_1 \rangle & & \\ & & \ddots & \\ & & & \langle \zeta_f, l_{2^n-1} \rangle \end{bmatrix} H_n$$

where the matrix in the middle is a diagonal matrix of order 2^n having zeros outside the diagonal. This matrix will be denoted by

$$diag\,(\langle \zeta_f, l_0 \rangle, \langle \zeta_f, l_1 \rangle, \ldots, \langle \zeta_f, l_{2^n-1} \rangle)$$

as in [57]. Thus,

$$M_f = 2^{-n}\ H_n.diag\left(\langle\zeta_f, l_0\rangle, \langle\zeta_f, l_1\rangle, \ldots, \langle\zeta_f, l_{2^n-1}\rangle\right).H_n\ . \tag{2.44}$$

By using (2.44), we get that

$$M_f.{M_f}^t = 2^{-n}\ H_n.diag\left(\langle\zeta_f, l_0\rangle^2, \langle\zeta_f, l_1\rangle^2, \ldots, \langle\zeta_f, l_{2^n-1}\rangle^2\right).H_n\ . \tag{2.45}$$

By using (2.45), the first row of $M_f.{M_f}^t$ is equal to

$$2^{-n}\ \left(\langle\zeta_f, l_0\rangle^2, \langle\zeta_f, l_1\rangle^2, \ldots, \langle\zeta_f, l_{2^n-1}\rangle^2\right).H_n\ . \tag{2.46}$$

By writing (2.46) explicitly, the first row of $M_f.{M_f}^t$ is equal to

$$2^{-n}\ \left(\langle\xi, l_0\rangle, \langle\xi, l_1\rangle, \ldots, \langle\xi, l_{2^n-1}\rangle\right) \tag{2.47}$$

where

$$\xi = \left(\langle\zeta_f, l_0\rangle^2, \langle\zeta_f, l_1\rangle^2, \ldots, \langle\zeta_f, l_{2^n-1}\rangle^2\right). \tag{2.48}$$

By combining (2.43) and (2.46) and by writing $\Delta(\alpha_i)$ instead of $\Delta_f(\alpha_i)$ and ζ instead of ζ_f, we get that

$$(\Delta(\alpha_0), \Delta(\alpha_1), \ldots, \Delta(\alpha_{2^n-1})) = 2^{-n}(\langle\zeta, l_0\rangle^2, \langle\zeta, l_1\rangle^2, \ldots, \langle\zeta, l_{2^n-1}\rangle^2)H_n. \tag{2.49}$$

Hence, the following theorem is obtained :

Theorem 2.6.4 *([57]) For any function f, the equality*

$$(\Delta_f(\alpha_0), \Delta_f(\alpha_1), \ldots, \Delta_f(\alpha_{2^n-1}))H_n = (\langle\zeta_f, l_0\rangle^2, \langle\zeta_f, l_1\rangle^2, \ldots, \langle\zeta_f, l_{2^n-1}\rangle^2)$$

holds where ζ_f is the sequence of f, $\Delta_f(\alpha_i)$ is the auto-correlation of f with a shift α_i and H_n is the Sylvester-Hadamard matrix of order 2^n with l_i's as its rows for $i = 0, 1, \ldots, 2^n - 1$.

Theorem 2.6.4 is a special form of the Wiener-Khintchine theorem [4].

CHAPTER 3

Two Upper and Two Lower Bounds On Nonlinearity

3.1 Motivation

In the previous sections, we have seen that the nonlinearity N_f of a function f is bounded above by $2^{n-1} - 2^{\frac{n}{2}-1}$. Since the nonlinearity of a function is an integer, it is clear that this upper bound may be achieved only when n is even. If n is odd, the nonlinearity of any function is strictly less than this bound. Additionally, we also know that the nonlinearity of a balanced function should be an even number, as proved in Corollary 2.5.6.

Note that we haven't mentioned whether there exists functions having nonlinearity $2^{n-1} - 2^{\frac{n}{2}-1}$ or not, when n is even. In Chapter 6, we will see that such functions do exist. Bent functions have many applications in digital communications, coding theory and cryptography [3, 1, 13, 22, 23, 29, 25, 31, 34] as stated in [42].

In addition to the known results, there are some important questions about nonlinearity which have not been answered yet. One of them is the upper bound on the nonlinearity of balanced functions for some odd values of n. This question has also a lot in common with another important question of cryptography, which we simply may state as the construction of cryptographically important boolean functions. In Chapter 7, some cryptographically important constructions of boolean functions will be presented [44].

Consider a cryptosystem which is composed of only linear functions or composed of functions which are very close to linear functions. If the functions under consideration are in $\mathcal{F}_n$, then by the results of Section 2.1, it is enough

to know only n outputs which come from n linearly independent vectors in V_n to identify those functions completely. In particular, it is enough to know only the action of these functions on the standard ordered basis of V_n over $GF(2)$. This simple example shows how 2^n operations is reduced to n operations only. What this example suggests as a security criterion is that the functions used in a cryptosystem should be as highly nonlinear as possible. Of course, one should always keep in the mind that, the situation we are talking about can not be generalized blindly. What we mean, when saying "the functions should be highly nonlinear" does not mean that a cryptosystem should not contain any linear function.

The weight of the functions employed in cryptosystems is also important. The use of an unbalanced function repeatedly may result in the outputs of the cryptosystem being biased and hence the cryptosystem can be easily distinguished from a true random source (or a pseudo-random source). This causes a large class of cryptanalysis methods to be applied to the cryptosystem ranging from trivial statistical attacks to much more complex attacks.

Apart from these specific examples, an informal definition for a secure boolean function may be given as follows. A secure boolean function is a function which satisfies the cryptographically important properties in an optimized way. "In an optimized way", is due to the fact that some cryptographically necessary criteria are challenging. That is, strongly satisfying one criterion may cause the function to be the weakest with respect to some other criteria. Keeping these in our mind, the first step is to identify the cryptographic properties which a function should satisfy for being strong. We have already seen some of these criteria, like balance and nonlinearity. In the following chapters, some other cryptographically important criteria will be presented. Note that this thesis does not include all cryptologically important criteria related to boolean functions. Our main interest with these criteria is narrowed by the use of boolean functions in block ciphers .

3.2 Upper Bounds

It is well-known that the bound in Lemma 2.5.2 coincides with the covering radius of the first order Reed-Muller code $RM(1,n)$ of length 2^n [25].

In contrast to this well-known upper bound, less is known about the lower bound on nonlinearity except some progress made in [48] and [56] and some trivial facts as $N_f > 0$ if and only if f is nonlinear.

There are two main questions concerning the nonlinearity. One is how to find the nonlinearity if some additional information is available about the function. The other is that, if the exact value of the nonlinearity can not be easily obtained, how to estimate the nonlinearity using some extra information about the function.

In the following two sections, four formulas will be given in order to estimate the nonlinearity of a function. Two of these bounds are upper bounds while the remaining two are lower bounds. All results in Section 3.2 and Section 3.3 are from the article of Zhang and Zheng [58].

We start with a result which follows easily from Theorem 2.6.4.

Corollary 3.2.1 *([57]) Let f be any function. Then,*

$$\sum_{i=0}^{2^n-1} \Delta_f(\alpha_i)^2 = 2^{-n} \sum_{i=0}^{2^n-1} \langle \zeta_f, l_i \rangle^4.$$

Proof. Let $\xi = (\langle \zeta_f, l_0 \rangle^2, \langle \zeta_f, l_1 \rangle^2, \ldots, \langle \zeta_f, l_{2^n-1} \rangle^2)$ as in the proof of Theorem 2.6.4. Then, $\xi\xi^t = 2^n \sum_{i=0}^{2^n-1} \Delta_f(\alpha_i)^2$ by using Theorem 2.6.4. Since $\xi\xi^t = \sum_{i=0}^{2^n-1} \langle \zeta_f, l_i \rangle^4$, the result follows. □

3.2.1 The First Upper Bound

The first upper bound is a straightforward application of Corollary 3.2.1. Since $2^n \sum_{i=0}^{2^n-1} \Delta_f(\alpha_i)^2 = \sum_{i=0}^{2^n-1} \langle \zeta_f, l_i \rangle^4$, there exists an i_0 satisfying $0 \leq i_0 \leq 2^n - 1$ such

that $\langle \zeta_f, l_{i_0} \rangle^4 \geq \sum_{i=0}^{2^n-1} \Delta_f(\alpha_i)^2$. In other words, $|\langle \zeta_f, l_{i_0} \rangle| \geq \sqrt[4]{\sum_{i=0}^{2^n-1} \Delta_f(\alpha_i)^2}$.

By using Lemma 2.5.3, one obtains that

$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt[4]{2^{2n} + \sum_{i=1}^{2^n-1} \Delta_f(\alpha_i)^2}. \tag{3.1}$$

3.2.2 The Second Upper Bound

It is easy to see that H_n, the n-th Sylvester-Hadamard matrix, satisfies $H_n = H_{n-t} \otimes H_t$ for any integer t such that $0 \leq t \leq n$.

By using this, the equation in Theorem 2.6.4 turns out to be

$$(\Delta(\alpha_0), \Delta(\alpha_1), \ldots, \Delta(\alpha_{2^n-1}))(H_{n-t} \otimes H_t) = (\langle \zeta, l_0 \rangle^2, \langle \zeta, l_1 \rangle^2, \ldots, \langle \zeta, l_{2^n-1} \rangle^2).$$

Set $\sigma_i = \sum_{k=0}^{2^t-1} \langle \zeta_f, l_{2^t.i+k} \rangle^2$ for $i = 0, 1, \ldots, 2^{n-t}-1$. That is, $\sigma_0 = \sum_{k=0}^{2^t-1} \langle \zeta_f, l_k \rangle^2$, $\sigma_1 = \sum_{k=0}^{2^t-1} \langle \zeta_f, l_{2^t+k} \rangle^2$,..., $\sigma_{2^{n-t}-1} = \sum_{k=0}^{2^t-1} \langle \zeta_f, l_{2^n-2^t+k} \rangle^2$.

Let $e = (1, 1, \ldots, 1)$ be the all-one vector of length 2^t and $I_{2^{n-t}}$ be the identity matrix of order 2^{n-t}. Note that $(H_{n-t} \otimes H_t)(I_{2^{n-t}} \otimes e^t) = (H_{n-t} I_{2^{n-t}}) \otimes (H_t e^t) = H_{n-t} \otimes (2^t, 0, \ldots, 0)^t$ by using **(3)** of Lemma 2.4.2, where $(2^t, 0, \ldots, 0)$ is a vector of length 2^t and $(2^t, 0, \ldots, 0)^t$ denotes the transpose of $(2^t, 0, \ldots, 0)$.

If we multiply both sides of the equation

$$(\Delta(\alpha_0), \Delta(\alpha_1), \ldots, \Delta(\alpha_{2^n-1}))(H_{n-t} \otimes H_t) = (\langle \zeta, l_0 \rangle^2, \langle \zeta, l_1 \rangle^2, \ldots, \langle \zeta, l_{2^n-1} \rangle^2)$$

with $(I_{2^{n-t}} \otimes e^t)$, then the left hand side is obtained as

$$(\Delta(\alpha_0), \Delta(\alpha_1), \ldots, \Delta(\alpha_{2^n-1}))(H_{n-t} \otimes H_t)(I_{2^{n-t}} \otimes e^t).$$

The left hand side of this equation is in fact equal to

$$(\Delta(\alpha_0), \Delta(\alpha_1), \ldots, \Delta(\alpha_{2^n-1}))(H_{n-t} \otimes (2^t, 0, \ldots, 0)^t).$$

The right hand side of the above equation after multiplication with $(I_{2^{n-t}} \otimes e^t)$ is equal to

$$(\sigma_0, \sigma_1, \ldots, \sigma_{2^{n-t}-1}).$$

Thus,

$$(\Delta(\alpha_0), \Delta(\alpha_1), \ldots, \Delta(\alpha_{2^n-1}))(H_{n-t} \otimes (2^t, 0, \ldots, 0)^t) = (\sigma_0, \sigma_1, \ldots, \sigma_{2^{n-t}-1}).$$

It is easy to see that the left hand side of the above equation is equal to

$$2^t \left(\Delta(\alpha_0), \Delta(\alpha_{2^t}), \ldots, \Delta(\alpha_{2^t.(2^{n-t}-1)})\right) H_{n-t}.$$

It follows that the equation in Theorem 2.6.4 turns out to be

$$2^t \left(\Delta(\alpha_0), \Delta(\alpha_{2^t}), \ldots, \Delta(\alpha_{2^t.(2^{n-t}-1)})\right) H_{n-t} = (\sigma_0, \sigma_1, \ldots, \sigma_{2^{n-t}-1}). \quad (3.2)$$

Note that the above equation is a generalization of the equation in Theorem 2.6.4. Clearly, two equations become identical when $t = 0$.

By comparing the i-th components of both sides of (3.2), one has

$$2^t \sum_{k=0}^{2^{n-t}-1} h_{i,k} \Delta_f(\alpha_{k.2^t}) = \sigma_i$$

where $l_i = (h_{i,0}, h_{i,1}, \ldots, h_{i,2^{n-t}-1})$ is the i-th row (column) of H_{n-t} for $i = 0, 1, \ldots, 2^{n-t} - 1$. However, since σ_i is defined as $\sigma_i = \sum_{k=0}^{2^t-1} \langle \zeta_f, l_{2^t.i+k} \rangle^2$ for $i = 0, 1, \ldots, 2^{n-t} - 1$, there is a k_0 $0 \leq k_0 \leq 2^t - 1$ such that $|\langle \zeta_f, l_{2^t.i+k_0} \rangle| \geq \sqrt{\sum_{k=0}^{2^{n-t}-1} h_{i,k} \Delta_f(\alpha_{k.2^t})}$ for any fixed i. By using Lemma 2.5.3, one has

$$N_f \leq 2^{n-1} - \frac{1}{2} \sqrt{2^n + \sum_{k=1}^{2^{n-t}-1} h_{i,k} \Delta_f(\alpha_{k.2^t})} \quad (3.3)$$

where t is a fixed integer satisfying $0 \leq t \leq n$, $l_i = (h_{i,0}, h_{i,1}, \ldots, h_{i,2^{n-t}-1})$ is the i-th row (column) of H_{n-t}.

Remark 3.2.2 ***(1)*** *Note that for* $0 \leq t \leq n$, *the set*

$$\Omega = \{\alpha_0, \alpha_{2^t}, \alpha_{2.2^t}, \ldots, \alpha_{(2^{n-t}-1).2^t}\}$$

forms an $n - t$ *dimensional subspace of* V_n *with basis* $\Im = \{\alpha_{2^t}, \alpha_{2^{t+1}}, \ldots, \alpha_{2^{n-t-1}.2^t}\}$.

(2) *The nonlinearity of a function is invariant under a nonsingular affine transformation on the input coordinates. This is easy to prove with the theory introduced up to now. Later, this result will be proved in the section which is totally devoted to invariant properties of a function under such transformations. For the time being, we assume that this is true.*

By using a nonsingular linear transformation on the input coordinates and setting $r = n - t$, the following lemma is obtained :

Lemma 3.2.3 *Let $\beta_1, \beta_2, \ldots, \beta_r$ be r linearly independent vectors in V_n for $0 \leq r \leq n$ and Ω be the subspace of V_n spanned by $\beta_1, \beta_2, \ldots, \beta_r$. In other words, $\Omega = \{\gamma_k \mid k = 0, 1, \ldots, 2^r - 1\}$ where $\gamma_k = a_1.\beta_1 + a_2.\beta_2 + \cdots + a_r.\beta_r$ for some $a_1, a_2, \ldots, a_r$ in $GF(2)$ such that $\psi(a_1, a_2, \ldots, a_r) = k$, ψ being the function in (2.1). Then,*

$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + \sum_{k=1}^{2^r-1} h_{i,k}\Delta_f(\gamma_k)}$$

holds for every row (column) $l_i = (h_{i,0}, h_{i,1}, \ldots, h_{i,2^r-1})$ of H_r where $i = 0, 1, \ldots, 2^r - 1$.

Remark 3.2.4 ***(1)*** *In some situations, it is sufficient to take $r = 1$ in Lemma 3.2.3. This means that for any nonzero vector β in V_n,*

$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n \pm \Delta_f(\beta)} \tag{3.4}$$

holds. Since (3.4) holds for any β in V_n, the following bound is obtained :

For any function f, the nonlinearity N_f of f satisfies

$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + \Delta_{max}} \tag{3.5}$$

where $\Delta_{max} = \max_{\alpha \in V_n, \alpha \neq 0} \{\ |\Delta_f(\alpha)|\ \}$.

(2) *If $r = 2$ is used in Lemma 3.2.3, then a better estimate of nonlinearity than the one above is obtained. Specialization of Lemma 3.2.3 for $r = 2$ is as follows :*

For any function f and for any two different, nonzero vectors β_1, β_2 in V_n, the nonlinearity N_f of f satisfies

$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + \sum_{k=1}^{3} h_{i,k}\Delta_f(\gamma_k)} \tag{3.6}$$

where $l_i = (h_{i,0}, h_{i,1}, h_{i,2}, h_{i,3})$ is the i-th row of H_2 for $i = 0, 1, 2, 3$. By using H_2 in (2.35), the inequality in (3.6) turns to the following four inequalities.

$$\begin{aligned}
N_f &\leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + \Delta_f(\beta_1) + \Delta_f(\beta_2) + \Delta_f(\beta_1 + \beta_2)},\\
N_f &\leq 2^{n-1} - \frac{1}{2}\sqrt{2^n - \Delta_f(\beta_1) + \Delta_f(\beta_2) - \Delta_f(\beta_1 + \beta_2)},\\
N_f &\leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + \Delta_f(\beta_1) - \Delta_f(\beta_2) - \Delta_f(\beta_1 + \beta_2)},\\
N_f &\leq 2^{n-1} - \frac{1}{2}\sqrt{2^n - \Delta_f(\beta_1) - \Delta_f(\beta_2) + \Delta_f(\beta_1 + \beta_2)}.
\end{aligned}$$

By collecting these inequalities, the following bound is obtained.

Corollary 3.2.5 *Let f be any function. Then,*

(1) *For any two different, nonzero vectors β_1, β_2 in V_n, the nonlinearity N_f of f satisfies*

$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + |\Delta_f(\beta_1)| + |\Delta_f(\beta_2)| - |\Delta_f(\beta_1 + \beta_2)|}. \tag{3.7}$$

(2) *Let α, β, γ be three nonzero vectors of V_n with the property that $|\Delta_f(\alpha)| \geq |\Delta_f(\beta)| \geq |\Delta_f(\gamma)| \geq |\Delta_f(\theta)|$ where θ is any nonzero vector in V_n distinct from α, β and γ. Then, the nonlinearity N_f of f satisfies*

$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + |\Delta_f(\alpha))| + |\Delta_f(\beta))| - |\Delta_f(\gamma))|}. \tag{3.8}$$

3.3 Lower Bounds

3.3.1 The First Lower Bound

Let f be a function with sequence $\zeta_f = (a_0, a_1, \ldots, a_{2^n-1})$. Set $\hat{a}_i = (a_{2i}, a_{2i+1})$ for all $i = 0, 1, \ldots, 2^{n-1} - 1$. It is obvious that $\zeta_f = (\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_{2^{n-1}-1})$, i.e. ζ_f is the concatenation of $\hat{a}_i$'s for $i = 0, 1, \ldots, 2^{n-1} - 1$.

Each $\hat{a}_i$ is referred to as a basis. If $\hat{a}_i = (1,1)$ or $(-1,-1)$, then $\hat{a}_i$ is called a $(++)$-basis and if $\hat{a}_i = (+1,-1)$ or $(-1,+1)$, then $\hat{a}_i$ is called a $(+-)$-basis.

Given a function f, its sequence ζ_f may be written as a concatenation of $(++)$ and $(+-)$-bases. Denote the number of $(++)$ and $(+-)$-bases in the sequence of a function f by ${n_f}^+$ and ${n_f}^-$, respectively.

Lemma 3.3.1 *Let f be a function with sequence ζ_f. Then, ${n_f}^+ = 2^{n-2} + \frac{1}{4}\Delta_f(\alpha_1)$ and ${n_f}^- = 2^{n-2} - \frac{1}{4}\Delta_f(\alpha_1)$.*

Proof. With the notation used previously, the sequence of the function $h(x) = f(x+\alpha)$ for any α in V_n is $\zeta_f(\alpha)$. Write ζ_f as the concatenation of the $(++)$ and $(+-)$-bases $\hat{a}_i$'s where $\hat{a}_i = (a_{2i}, a_{2i+1})$. Then, $\zeta_f(\alpha_1)$ with respect to $(++)$ and $(+-)$-bases is $\hat{b}_i = (a_{2i+1}, a_{2i})$. From this, it is easy to see that $\Delta_f(\alpha_1) = \langle \zeta_f, \zeta_f(\alpha_1) \rangle = 2({n_f}^+ - {n_f}^-)$. The result follows since ${n_f}^+ + {n_f}^- = 2^{n-1}$ holds always. □

Lemma 3.3.2 *The nonlinearity N_f of any function f satisfies $N_f \geq 2^{n-2} - \frac{1}{4}|\Delta_f(\alpha_1)|$.*

Proof. Since $w(f) \geq {n_f}^-$, we get that $w(f) \geq 2^{n-2} - \frac{1}{4}\Delta_f(\alpha_1)$ by Lemma 3.3.1. Set $g_j(x) = f(x) + \varphi_j(x)$ where φ_j is a linear function for $j = 0, 1, \ldots, 2^n - 1$. It is easy to see that

$$\Delta_{g_j}(\alpha_1) = \begin{cases} \Delta_f(\alpha_1) & \text{if} \quad \varphi_j(\alpha_1) = 0, \\ -\Delta_f(\alpha_1) & \text{if} \quad \varphi_j(\alpha_1) = 1. \end{cases} \tag{3.9}$$

By Lemma 3.3.1, $w(g_j) \geq 2^{n-2} - \frac{1}{4}\Delta_{g_j}(\alpha_1)$. Since $w(g_j) = d(f, \varphi_j)$, we get that

$$d(f, \varphi_j) \geq \begin{cases} 2^{n-2} - \frac{1}{4}\Delta_f(\alpha_1) & \text{if} \quad \varphi_j(\alpha_1) = 0, \\ 2^{n-2} + \frac{1}{4}\Delta_f(\alpha_1) & \text{if} \quad \varphi_j(\alpha_1) = 1. \end{cases}$$

where φ_j denotes all linear functions.

Now, set $\bar{g}_j(x) = f(x) + \bar{\varphi}_j(x) = f(x) + \varphi_j(x) + 1$ where φ_j is a linear function for $j = 0, 1, \ldots, 2^n - 1$. It is easy to show that (3.9) also holds for $\bar{g}_j$

for $j = 0, 1, \ldots, 2^n - 1$. It follows that

$$d(f, \varphi_j) \geq \begin{cases} 2^{n-2} - \frac{1}{4}\Delta_f(\alpha_1) & \text{if} \quad \varphi_j(\alpha_1) = 0, \\ 2^{n-2} + \frac{1}{4}\Delta_f(\alpha_1) & \text{if} \quad \varphi_j(\alpha_1) = 1. \end{cases}$$

holds for any affine function φ_j. The definition of nonlinearity gives the result. □

Theorem 3.3.3 *For any function f, the nonlinearity N_f of f satisfies*

$$N_f \geq 2^{n-2} - \frac{1}{4}\Delta_{min} \tag{3.10}$$

where $\Delta_{min} = \min_{\alpha \in V_n, \alpha \neq 0} \{ \, |\Delta_f(\alpha)| \, \}$.

Proof. Choose a nonsingular matrix A of order n which satisfies $\alpha_1 A = \alpha_k$ for any fixed k satisfying $0 \leq k \leq 2^n - 1$. In fact, this is equivalent to finding $n-1$ vectors $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}$ in V_n for which the set $\Omega = \{\sigma_1, \sigma_2, \ldots, \sigma_{n-1}, \alpha_k\}$ becomes a linearly independent set. Then, A may be any matrix containing α_k as its last row where the first $n-1$ rows of A are any permutation of σ_i's for $i = 1, 2, \ldots, n-1$.

Set $g(x) = f(xA)$. Then, $g^{\alpha_1}(x) = g(x) + g(x + \alpha_1) = f(xA) + f(xA + \alpha_1 A) = f(u) + f(u + \alpha_k)$ where $u = xA$. Since A is nonsingular, for any x in V_n, there exists a unique u in V_n satisfying $u = xA$ and conversely. This yields that $\Delta_g(\alpha_1) = \Delta_f(\alpha_k)$.

By Lemma 3.3.2, $N_g \geq 2^{n-2} - \frac{1}{4}|\Delta_g(\alpha_1)| = 2^{n-2} - \frac{1}{4}|\Delta_f(\alpha_k)|$. Since A is nonsingular, we have $N_g = N_f$ by **(2)** of Remark 3.2.2. By combining these, $N_f \geq 2^{n-2} - \frac{1}{4}|\Delta_f(\alpha_k)|$ for any arbitrary but specific k satisfying $0 \leq k \leq 2^n - 1$. Hence, $N_f \geq 2^{n-2} - \frac{1}{4}\Delta_{min}$ where $\Delta_{min} = \min_{\alpha \in V_n, \alpha \neq 0} \{ \, |\Delta_f(\alpha)| \, \}$. □

3.3.2 The Second Lower Bound

In [5] it was pointed out that, for any function f, if the difference functions f^α's are balanced with respect to all but a subset $\Re_f$ of vectors in V_n, then the

nonlinearity N_f of f satisfies

$$N_f \geq 2^{n-1} - 2^{\frac{n}{2}-1}|\Re_f|^{\frac{1}{2}}. \tag{3.11}$$

Another improvement which has been made in [48] (See Theorem 11) is

$$N_f \geq 2^{n-1} - 2^{n-\frac{1}{2}\rho-1} \tag{3.12}$$

where ρ is the maximum dimension of the subspaces of $(V_n \setminus \Re_f) \cup \{\alpha_0\}$.

An important shortcoming of (3.11) and (3.12) is that when $|\Re_f|$ is large, the estimates provided by them are far from the real value of N_f.

Let f be a function for which f^α is balanced with respect to all but a subset $\Re_f$ of vectors in V_n. Recall from **(1)** of Remark 3.2.2 that the set

$$\Omega = \{\alpha_0, \alpha_{2^t}, \alpha_{2.2^t}, \ldots, \alpha_{(2^{n-t}-1).2^t}\}$$

is an $n-t$ dimensional subspace of V_n with basis $\Im = \{\alpha_{2^t}, \alpha_{2^{t+1}}, \ldots, \alpha_{2^{n-t-1}.2^t}\}$ for any integer t satisfying $0 \leq t \leq n$. Also, recall that $2^t \sum_{k=0}^{2^{n-t}-1} h_{i,k}\Delta_f(\alpha_{k.2^t}) = \sigma_i$ where $l_i = (h_{i,0}, h_{i,1}, \ldots, h_{i,2^{n-t}-1})$ is the i-th row (column) of H_{n-t} for $i = 0, 1, \ldots, 2^{n-t}-1$ and $\sigma_i = \sum_{k=0}^{2^t-1} \langle \zeta_f, l_{2^t.i+k} \rangle^2$ for $i = 0, 1, \ldots, 2^{n-t}-1$.

By using these, one obtains that

$$\begin{aligned} \sigma_i &= 2^t \sum_{k=0}^{2^{n-t}-1} h_{i,k}\Delta_f(\alpha_{k.2^t}) \\ &\leq 2^t(\Delta_f(\alpha_0) + (|\Re_f \cap \Omega| - 1)\Delta_{max}). \end{aligned} \tag{3.13}$$

This inequality is clear since for any α in $V_n \setminus \Re_f$, $\Delta_f(\alpha) = 0$ as f^α is balanced. Thus, in $\Omega \setminus \{\alpha_0\}$, $\Delta_f(\alpha_{k.2^t})$ may be nonzero at most for $|\Re_f \cap \Omega| - 1$ values of k's where $k = 1, 2, \ldots, 2^{n-t}-1$. Also, note that $\Delta_f(\alpha) \leq \Delta_{max}$ for all α in V_n.

It follows that $\langle \zeta_f, l_{2^t.i+k} \rangle^2 \leq 2^t(2^n + (|\Re_f \cap \Omega| - 1)\Delta_{max})$ for any $i = 0, 1, \ldots, 2^{n-t}-1$ and $k = 0, 1, \ldots, 2^t - 1$, since $\Delta_f(\alpha_0) = 2^n$ and by using the definition of σ_i and (3.13). By using Lemma (2.5.3), one has

$$N_f \geq 2^{n-1} - \frac{1}{2}\sqrt{2^t(2^n + (|\Re_f \cap \Omega| - 1)\Delta_{max})}.$$

By simplifying this expression, one concludes that

$$N_f \geq 2^{n-1} - 2^{\frac{t}{2}-1}\sqrt{2^n + (|\Re_f \cap \Omega| - 1)\Delta_{max}}. \tag{3.14}$$

By setting $r = n - t$ and using a nonsingular linear transformation on the input variables, the following theorem is obtained :

Theorem 3.3.4 *Let f be a function for which f^α is balanced with respect to all but a subset $\Re_f$ of vectors in V_n and let Ω be any r dimensional subspace of V_n for $r = 0, 1, \ldots, n$. Then, the nonlinearity N_f of f satisfies*

$$N_f \geq 2^{n-1} - 2^{\frac{1}{2}(n-r)-1}\sqrt{2^n + (|\Re_f \cap \Omega| - 1)\Delta_{max}}$$

where $\Delta_{max} = \max_{\alpha \in V_n, \alpha \neq 0} \{\ |\Delta_f(\alpha)|\ \}$.

Since $|\Delta_f(\alpha)| \leq 2^n$ for any α in V_n, it is clear that $\Delta_{max} \leq 2^n$. If one substitutes 2^n for Δ_{max} in Theorem 3.3.4, the following corollary is obtained :

Corollary 3.3.5 *Let f be a function for which f^α is balanced with respect to all but a subset $\Re_f$ of vectors in V_n. Let Ω be any r dimensional subspace of V_n for $r = 0, 1, \ldots, n$. Then, the nonlinearity N_f of f satisfies*

$$N_f \geq 2^{n-1} - 2^{n-\frac{1}{2}r-1}\sqrt{|\Re_f \cap \Omega|}.$$

Theorem 3.3.4 is more general and gives a better estimate of lower bound than the bound in (3.11) because of the following :

Let $\Omega = V_n$, i.e. $r = n$. As $\Delta_{max} \leq 2^n$, we have $(|\Re_f| - 1)\Delta_{max} \leq 2^{\frac{n}{2}}|\Re_f|^{\frac{1}{2}}$. Thus, $N_f \geq 2^{n-1} - \frac{1}{2}\sqrt{2^n + (|\Re_f| - 1)\Delta_{max}} \geq 2^{n-1} - 2^{\frac{n}{2}-1}|\Re_f|^{\frac{1}{2}}$ giving the result.

Theorem 3.3.4 gives also a better estimate of lower bound than the bound in (3.12) because of the following :

Let Ω be such that $\Re_f \cap \Omega = \{\alpha_0\}$ only. Then, by Corollary 3.3.5, $N_f \geq 2^{n-1} - 2^{n-\frac{1}{2}r-1}$ which is equal to (3.12).

Given a function f, if there exists an integer r, $0 \leq r \leq n$ and an integer $p > 0$ such that $N_f \leq 2^{n-1} - 2^{n-\frac{1}{2}r-1}p$, then as $N_f \geq 2^{n-1} - 2^{n-\frac{1}{2}r-1}\sqrt{|\Re_f \cap \Omega|}$

holds for any r dimensional subspace Ω of V_n, one concludes that there is an r dimensional subspace of V_n such that $|\Re_f \cap \Omega| \geq p^2$. This in turn shows that the nonlinearity of f is not only related to $\Re_f$ but also to the distribution of $\Re_f$.

3.4 Relations With Nonsingular Affine Transformations

Recall that in **(2)** of Remark 3.2.2, it was mentioned but not proved that the nonlinearity of a function is invariant under a nonsingular affine transformation on the input coordinates. In this section, the proof of this fact will be given. In fact, this section is completely devoted to the properties of the functions which remain invariant under nonsingular affine transformations on the input coordinates.

Let f be a function, A be a nonsingular matrix of order n with entries from $GF(2)$ and α be a vector in V_n. We denote the composition of two functions f and θ by $(f \circ \theta)(x) = f(\theta(x))$ where $\theta(x) = xA + \alpha$ denotes a nonsingular affine transformation on V_n. In particular if α is the zero vector, then θ is called a linear transformation. Note that θ is a bijection from V_n to V_n and $f \circ \theta$ is in $\mathcal{F}_n$.

Consider the degree of f, which is the degree of the highest degree term appearing in the algebraic normal form of f. The degree of f takes values from 0 to n. It is obvious that the degree of f is equal to the degree of $f \circ \theta$ for any nonsingular affine transformation θ [29].

Now, consider the weight of f. By definition, it is the weight of $T_f = (f(\alpha_0), f(\alpha_1), \ldots, f(\alpha_{2^n-1}))$. The weight of $f \circ \theta$ is the weight of $T_{f\circ\theta} = \big(f(\alpha_{i_0}), f(\alpha_{i_1}), \ldots, f(\alpha_{i_{2^n-1}})\big)$ where $\{i_0, i_1, \ldots, i_{2^n-1}\}$ is a permutation of $\{0, 1, \ldots, 2^n - 1\}$ since θ is a bijection. This yields that $w(f) = w(f \circ \theta)$. In other words, the balance of a function is preserved under any nonsingular affine transformation on the input coordinates [44].

In order to show that the nonlinearity N_f of the function f is invariant under any nonsingular affine transformation on the input coordinates, we need

the following result.

Lemma 3.4.1 *Let f be a function, θ be a nonsingular affine transformation corresponding to a nonsingular matrix A of order n and to a vector α in V_n and let φ be any affine function. Then, $d(f, \varphi) = d(f \circ \theta, \varphi \circ \theta)$.*

Proof.

Let $T_f = (f(\alpha_0), f(\alpha_1), \ldots, f(\alpha_{2^n-1}))$, $T_\varphi = (\varphi(\alpha_0), \varphi(\alpha_1), \ldots, \varphi(\alpha_{2^n-1}))$. Assume that $T_{f\circ\theta} = \big(f(\alpha_{i_0}), f(\alpha_{i_1}), \ldots, f(\alpha_{i_{2^n-1}})\big)$ where $\{i_0, i_1, \ldots, i_{2^n-1}\}$ is a permutation of $\{0, 1, \ldots, 2^n - 1\}$. Then, $T_{\varphi\circ\theta} = \big(\varphi(\alpha_{i_0}), \varphi(\alpha_{i_1}), \ldots, \varphi(\alpha_{i_{2^n-1}})\big)$. Thus, $f(\alpha_k) = \varphi(\alpha_k)$ if and only if $f(\alpha_{i_k}) = \varphi(\alpha_{i_k})$, giving that $d(f, \varphi) = d(f \circ \theta, \varphi \circ \theta)$. $\square$

Given a function f, its nonlinearity is defined as $N_f = \min_{\varphi_i \in \mathcal{A}_n} d(f, \varphi_i)$ where $\mathcal{A}_n = \{\ \varphi_i \mid i = 0, 1, \ldots, 2^{n+1} - 1\ \}$ denotes the set of all affine functions. Consider the set $\mathcal{A}_n\prime = \{\ \varphi_i \circ \theta \mid i = 0, 1, \ldots, 2^{n+1} - 1\ \}$ where θ is defined as above. Since θ is nonsingular, $\mathcal{A}_n = \mathcal{A}_n\prime$ as sets. Thus,

$$N_f = \min_{\varphi_i \in \mathcal{A}_n} d(f, \varphi_i) = \min_{(\varphi_i \circ \theta) \in \mathcal{A}_n\prime} d(f \circ \theta, \varphi_i \circ \theta) = N_{f\circ\theta}.$$

This gives that the nonlinearity N_f of a function f is invariant under any nonsingular affine transformation on the input coordinates [29, 44].

Now, consider the set of all α's in V_n such that the difference function f^α is balanced. Recall that the set of all α's for which the difference function f^α is not balanced is the set $\Re_f$. Thus, the set under consideration is $V_n \setminus \Re_f$. We claim that the number of elements of this set is invariant under any nonsingular affine transformation on the input coordinates. That is, $|V_n \setminus \Re_f| = |V_n \setminus \Re_{f\circ\theta}|$ [44] :

Let β be a nonzero vector in V_n, $\theta(x) = xA + \alpha$ where A is a nonsingular matrix of order n and α is any vector in V_n. The function $(f \circ \theta)^\beta$ is balanced if and only if

$$(f \circ \theta)^\beta(x) \quad = \quad f(xA + \alpha) + f((x + \beta)A + \alpha)$$

$$
\begin{aligned}
&= f(u) + f(u+v) \\
&= f^v(u) \qquad (3.15)
\end{aligned}
$$

is balanced where $u = xA + \alpha$ and $v = \beta A$. Since A is nonsingular, if x runs through all vectors in V_n, then so does u. Also, v is nonzero since β is nonzero. From these, we get the desired result.

In particular, any nonzero β in V_n is a linear structure of $(f \circ \theta)$ if and only if the nonzero vector $v = \beta A$ is a linear structure of f. This means that, the number of linear structures of f and $f \circ \theta$ and hence the linearity dimension of f and $f \circ \theta$ are the same.

By summarizing what is proved in this section, we get that

Theorem 3.4.2 *([44, 29]) For any function f, the degree, the weight, the nonlinearity, the linearity dimension and the number of α's for which f^α is balanced are invariant under a nonsingular affine transformation on the input coordinates.*

In [29], the two additional forms of nonlinearity (the distance to linear structures and the nonlinear order) mentioned in Section 2.3 are also shown to be invariant under nonsingular affine transformations on the input coordinates. Hence, they also serve as useful nonlinearity criteria.

CHAPTER 4

Walsh Transform and Properties

Recall that in Section 2.5, a simple algorithm to calculate the nonlinearity of a function is presented by using the sequence ζ_f of f and the n-th Sylvester-Hadamard matrix H_n. By using Lemma 2.5.3 in that section, a slight improvement of this algorithm is also mentioned.

The main purpose of this chapter is to introduce one of the most important functions in cryptology, the Walsh transform (or Hadamard or discrete Fourier transform) of a function. After examining the Walsh transform with its various properties, a fast method of computing the nonlinearity will be presented at the end of the chapter.

4.1 Walsh Transform

If ζ is a $(1,-1)$-sequence in V^{2^n}, then its Walsh transform is defined as [25] :

$$\hat{\zeta} = \zeta H_n. \tag{4.1}$$

Let f be a function. The Walsh transform of f is commonly defined in the literature as

$$W_f(\alpha) = \sum_{x \in V_n} f(x)(-1)^{\langle \alpha, x \rangle} \tag{4.2}$$

where α is in V_n.

Although this form of Walsh transform is sometimes used in the literature, the form which will be used in this thesis differs from the one in (4.2) and is also commonly used in the literature. After presenting this form, the simple

relation between these two forms of Walsh transform will be proved in this chapter. Any property of Walsh transform enjoyed by one of these forms can be rewritten in terms of the other form.

Recall from Section 2.2 that the real-valued function $\hat{f}$ associated to a function f is defined as $\hat{f}(x) = (-1)^{f(x)} = 1 - 2f(x)$ for all x in V_n. The function $\hat{f}$ is called as the "sign function" of f. In some places, this function is represented by χ_f, but in order not to make a confusion with the usage of the characteristic function χ_α in Section 2.1, we prefer this notation.

The Walsh transform of the sign function of a function f is defined as

$$W_{\hat{f}}(\alpha) = \sum_{x \in V_n} \hat{f}(x)(-1)^{\langle \alpha, x \rangle} = \sum_{x \in V_n} (-1)^{f(x)+\langle \alpha, x \rangle} \tag{4.3}$$

where α is in V_n. Note that the Walsh transform of the sign function of f takes integer values in $[-2^n, 2^n]$. If we denote by $T_{W_{\hat{f}}}$ the ordered values of the Walsh transform of the sign function of f as

$$T_{W_{\hat{f}}} = \Big(W_{\hat{f}}(\alpha_0), W_{\hat{f}}(\alpha_1), \ldots, W_{\hat{f}}(\alpha_{2^n-1}) \Big)$$

and by $T_{\hat{f}} = \zeta_f$ the truth table of the sign function of f as usual, then we get that

$$T_{W_{\hat{f}}} = \zeta_f H_n. \tag{4.4}$$

(4.1) and (4.4) show that the Walsh transform of the sign function of f is equal to obtaining the Walsh transform of ζ_f via multiplying by H_n. By multiplying (4.4) with H_n, one obtains that $\zeta_f = \frac{1}{2^n} T_{W_{\hat{f}}} H_n$. Thus,

$$\hat{f}(x) = \frac{1}{2^n} \sum_{\alpha \in V_n} W_{\hat{f}}(\alpha)(-1)^{\langle \alpha, x \rangle} \tag{4.5}$$

where x is in V_n.

The equation in (4.5) is called as the inverse Walsh transform or the inversion formula for (4.3). Note that the equation in (4.5) and the equation

$$\hat{f}(x) = \sum_{\alpha \in V_n} c_\alpha . \hat{f}_\alpha(x)$$

in (2.31) are the same except the appearance of the constant $\frac{1}{2^n}$ since $\hat{f}_\alpha(x) = (-1)^{\langle \alpha, x \rangle}$. Thus, for any function f and for any α, the value $W_{\hat{f}}(\alpha)$ appearing in $T_{W_{\hat{f}}}$ is a constant multiple (2^n) of the coefficient c_α in the representation of the function f with respect to the orthogonal basis $\{ \zeta_{f_\alpha} \mid \alpha \in V_n \}$ where ζ_{f_α} is the sequence of the linear function $f_\alpha(x) = \langle \alpha, x \rangle$. This fact gives a method for the representation of a function with respect to the orthogonal basis $\{ \zeta_{f_\alpha} \mid \alpha \in V_n \}$ by using the Walsh transform, as mentioned in Section 2.2. The Walsh transform is sometimes called as the spectral distribution or the spectrum of f in the literature.

Now, consider the Walsh transform of the sign function of f given by $W_{\hat{f}}(\alpha) = \sum_{x \in V_n} (-1)^{f(x)+\langle \alpha, x \rangle}$. It is easy to see from this equation that $W_{\hat{f}}(\alpha)$ is equal to the number of 0's minus the number of 1's of the function $f + f_\alpha$. Thus, $W_{\hat{f}}(\alpha) = 2^n - 2w(f + f_\alpha) = 2^n - 2d(f, f_\alpha)$. In particular, if $\alpha = \alpha_0$ is the zero vector, then $W_{\hat{f}}(\alpha_0) = 2^n - 2w(f)$.

It follows that for a function f and a fixed linear function $f_\alpha(x) = \langle \alpha, x \rangle$, we have the following equality :

$$d(f, f_\alpha) = \frac{1}{2}(2^n - W_{\hat{f}}(\alpha)). \tag{4.6}$$

Since $d(f, \bar{g}) = 2^n - d(f, g)$ for any two functions f, g, we also have the following equality :

$$d(f, \bar{f}_\alpha) = \frac{1}{2}(2^n + W_{\hat{f}}(\alpha)). \tag{4.7}$$

The equations (4.6) and (4.7) imply that the nearest affine function $\varphi_\alpha(x) = a_0 + \langle \alpha, x \rangle$, $a_0 \in GF(2)$, to f in the sense of the Hamming distance is the function for which $|W_{\hat{f}}(\alpha)|$ is the largest. We give an example to demonstrate these facts about the Walsh transform.

Example 4.1.1

Let $n = 3$. Consider the function $f(x_1, x_2, x_3) = 1 + x_1 + x_2 + x_2x_3 + x_1x_2x_3$ with $T_f = (1, 1, 0, 1, 0, 0, 1, 1)$ and $T_{\hat{f}} = \zeta_f = (-1, -1, +1, -1, +1, +1, -1, -1)$.

By using (4.3), $W_{\hat{f}}(\alpha_0) = -2$, $W_{\hat{f}}(\alpha_1) = +2$, $W_{\hat{f}}(\alpha_2) = +2$, $W_{\hat{f}}(\alpha_3) = -2$, $W_{\hat{f}}(\alpha_4) = -2$, $W_{\hat{f}}(\alpha_5) = +2$, $W_{\hat{f}}(\alpha_6) = -6$ and $W_{\hat{f}}(\alpha_7) = -2$. Thus,

$$T_{W_{\hat{f}}} = (-2, +2, +2, -2, -2, +2, -6, -2)\,.$$

We can verify the above computations by using (4.5). From these calculations, we obtain that $\hat{f}(\alpha_0) = -1$, $\hat{f}(\alpha_1) = -1$, $\hat{f}(\alpha_2) = 1$, $\hat{f}(\alpha_3) = -1$, $\hat{f}(\alpha_4) = 1$, $\hat{f}(\alpha_5) = 1$, $\hat{f}(\alpha_6) = -1$ and $\hat{f}(\alpha_7) = -1$ which are compatible with ζ_f. Since $|W_{\hat{f}}(\alpha)|$ is the largest when $\alpha = \alpha_6$, we conclude that the function $\varphi_{\alpha_6}(x) = 1 + x_1 + x_2$ is the nearest function to f. Note that since $W_{\hat{f}}(\alpha_6) = 2^3 - 2d(f, f_{\alpha_6}) = -6$, we get that $d(f, f_{\alpha_6}) = 7$ giving that f and φ_{α_6} agree on 7 points out of 8.

By using (4.2) and (4.3), we can now prove the relationship between the two forms of Walsh transforms. This result can be found in [17, 38] without proof. Since $W_f(\alpha) = \sum_{x \in V_n} f(x)(-1)^{\langle \alpha, x \rangle}$ and $W_{\hat{f}}(\alpha) = \sum_{x \in V_n} (-1)^{f(x) + \langle \alpha, x \rangle}$, it is easy to see that $W_f(\alpha) = \sum_{x \in Supp(f)} (-1)^{\langle \alpha, x \rangle}$ and

$$W_{\hat{f}}(\alpha) = \sum_{x \in V_n \setminus Supp(f)} (-1)^{\langle \alpha, x \rangle} - \sum_{x \in Supp(f)} (-1)^{\langle \alpha, x \rangle}.$$

However, it is known that

$$\sum_{x \in V_n \setminus Supp(f)} (-1)^{\langle \alpha, x \rangle} + \sum_{x \in Supp(f)} (-1)^{\langle \alpha, x \rangle} = 0$$

for any nonzero α in V_n. So, $\sum_{x \in V_n \setminus Supp(f)} (-1)^{\langle \alpha, x \rangle} = -\sum_{x \in Supp(f)} (-1)^{\langle \alpha, x \rangle}$ implying that $W_{\hat{f}}(\alpha) = -2W_f(\alpha)$ for any nonzero α in V_n. Note that if $\alpha = \alpha_0$ is the zero vector, then $W_f(\alpha_0) = w(f)$ and $W_{\hat{f}}(\alpha_0) = 2^n - 2w(f)$. It follows that $W_{\hat{f}}(\alpha_0) = 2^n - 2W_f(\alpha_0)$.

4.2 Cross-Correlation of f

Now, pausing for some time about the Walsh transform, another important function which operates on two functions f, g will be investigated. This function is a generalization of (2.42) on two functions f, g instead of one and also

generalizes the correlation concept given in [29] and [11]. Its outputs are between 0 and 1. The definition is as follows :

Let f, g be two functions. Their cross-correlation with a shift α or simply α-correlation is defined as

$$C(f,g)(\alpha) = \frac{1}{2^n} \sum_{x \in V_n} (-1)^{f(x)+g(x+\alpha)}. \tag{4.8}$$

Note that $C(f,g)(\alpha) = \frac{1}{2^n}\langle \zeta_f, \zeta_g(\alpha) \rangle$ where ζ_f is the sequence of f and $\zeta_g(\alpha)$ is the sequence of $g(x+\alpha)$. The auto-correlation of f with a shift α, defined as $\Delta_f(\alpha) = \langle \zeta_f, \zeta_f(\alpha) \rangle$ in (2.42), is actually $2^n.C(f,f)(\alpha)$.

What $C(f,g)(\alpha)$ measures can be seen from the observation below :

$$\langle \zeta_f, \zeta_g(\alpha) \rangle = 2|\{\ x \in V_n \mid f(x) = g(x+\alpha)\ \}| - 2^n. \tag{4.9}$$

It follows that $C(f,g)(\alpha) = \frac{1}{2^n}\langle \zeta_f, \zeta_g(\alpha) \rangle = 2.P\{f(x) = g(x+\alpha)\} - 1$ where $P\{A\}$ denotes the probability of an event A.

We call the α_0-correlation between a function f and a linear function $f_\alpha(x) = \langle \alpha, x \rangle$ for α in V_n as the 0-correlation and we denote it by $C(f, f_\alpha)(0)$. Thus, $C(f,f_\alpha)(0) = \frac{1}{2^n} \sum_{x \in V_n} (-1)^{f(x)+f_\alpha(x)} = \frac{1}{2^n} W_{\hat{f}}(\alpha)$. Using this, $C(f,f_\alpha)(\alpha_i)$ $= \frac{1}{2^n}(-1)^{\langle \alpha, \alpha_i \rangle} W_{\hat{f}}(\alpha)$ for all $i = 0, 1, \ldots, 2^n - 1$. Moreover, $C(f, f_\alpha + a_0)(\alpha_i) = \frac{1}{2^n}(-1)^{\langle \alpha, \alpha_i \rangle + a_0} W_{\hat{f}}(\alpha)$ for all $i = 0, 1, \ldots, 2^n - 1$ and for any a_0 in $GF(2)$.

Thus, by using (4.5) and the above results, one obtains that

$$\begin{aligned} \hat{f}(x) &= \frac{1}{2^n} \sum_{\alpha \in V_n} W_{\hat{f}}(\alpha)(-1)^{\langle \alpha, x \rangle} \\ &= \sum_{\alpha \in V_n} C(f, f_\alpha)(0)\ (-1)^{f_\alpha(x)} \\ &= \sum_{\alpha \in V_n} C(f, f_\alpha)\ (-1)^{f_\alpha(x)} \end{aligned} \tag{4.10}$$

where $C(f,f_\alpha)(0)$ is denoted by $C(f,f_\alpha)$. This is meaningful since $C(f,f_\alpha)(0)$ coincides with the definition of the correlation between f and f_α denoted by $C(f,f_\alpha)$ in [29] and [11].

Now, we turn back to Example 4.1.1. By using (4.10), we get that $\hat{f}(x) = -\frac{1}{4}\hat{f}_{\alpha_0}(x) + \frac{1}{4}\hat{f}_{\alpha_1}(x) + \frac{1}{4}\hat{f}_{\alpha_2}(x) - \frac{1}{4}\hat{f}_{\alpha_3}(x) - \frac{1}{4}\hat{f}_{\alpha_4}(x) + \frac{1}{4}\hat{f}_{\alpha_5}(x) - \frac{3}{4}\hat{f}_{\alpha_6}(x) - \frac{1}{4}\hat{f}_{\alpha_7}(x)$.

4.3 Properties of The Walsh Transform

Lemma 4.3.1 *([25])* $\sum_{\alpha \in V_n} W_{\hat{f}}(\alpha)\, W_{\hat{f}}(\alpha+\beta) = \begin{cases} 2^{2n} & \text{if } \beta = 0, \\ 0 & \text{if } \beta \neq 0. \end{cases}$

Proof. By writing the left hand side of the above equation explicitly, one gets that

$$\begin{aligned} LHS &= \sum_{\alpha \in V_n} \sum_{x \in V_n} \hat{f}(x)(-1)^{\langle \alpha, x \rangle} \sum_{y \in V_n} \hat{f}(y)(-1)^{\langle \alpha+\beta, y \rangle} \\ &= \sum_{x \in V_n} \sum_{y \in V_n} (-1)^{\langle \beta, y \rangle} \hat{f}(x)\hat{f}(y) \sum_{\alpha \in V_n} (-1)^{\langle \alpha, x+y \rangle}. \end{aligned}$$

Since $\sum_{\alpha \in V_n} (-1)^{\langle \alpha, x+y \rangle} = 2^n \delta(x+y)$ where $\delta(x)$ is the Kronecker delta, the above expression turns out to be

$$\begin{aligned} LHS &= 2^n \sum_{x \in V_n} (-1)^{\langle \beta, x \rangle} \hat{f}(x)^2 \\ &= 2^{2n} \delta(\beta) \end{aligned}$$

which is the desired result. □

Corollary 4.3.2 *([25], Parseval's equation)*

$$\sum_{\alpha \in V_n} W_{\hat{f}}(\alpha)^2 = 2^{2n}.$$

Note that this result was already proved in Lemma 2.5.3. There, the equation was in the form $\sum_{j=0}^{2^n-1} \langle \zeta_f, l_j \rangle^2 = 2^{2n}$ where ζ_f is the sequence of f and l_j's are the rows of H_n for $j = 0, 1, \ldots, 2^n - 1$. The equivalence of these two fact is due to $W_{\hat{f}}(\alpha_j) = \langle \zeta_f, l_j \rangle$ where l_j is the sequence of the linear function f_{α_j}.

Theorem 4.3.3 *([11],[39],[25]) Let f, g be functions and $\hat{f}, \hat{g}$ denote their sign functions, respectively. Let $z = (x, y)$ be in V_{n+m} such that x is in V_n and y is in V_m. Define the functions $r(x) = \bar{f}(x)$, $h(x) = (f+g)(x)$, $k(x) = f(x)g(x)$, $t(x) = f(x) + f_\beta(x)$ where $f_\beta(x) = \langle \beta, x \rangle$ is the linear function corresponding to β in V_n and $s(z) = f(x) + g(y)$. Thus, r, h, k, t are functions in $\mathcal{F}_n$ and s*

is a function in $\mathcal{F}_{n+m}$. Then, the following hold :

(a) $W_{\hat{h}}(\alpha) = \frac{1}{2^n} \sum_{x \in V_n} W_{\hat{f}}(x)\ W_{\hat{g}}(x+\alpha)$.

(b) $W_{\hat{r}}(\alpha) = -W_{\hat{f}}(\alpha)$.

(c) $W_{\hat{t}}(\alpha) = W_{\hat{f}}(\alpha+\beta)$.

(d) $W_{\hat{k}}(\alpha) = \frac{1}{2}\left(2^n\delta(\alpha) + W_{\hat{f}}(\alpha) + W_{\hat{g}}(\alpha) - W_{\hat{h}}(\alpha)\right)$ *where $\delta(\alpha)$ is the Kronecker delta.*

(e) $W_{\hat{s}}(\alpha) = W_{\hat{f}}(\beta).W_{\hat{g}}(\gamma)$ *for any $\alpha = (\beta, \gamma)$ in V_{n+m} where β is in V_n and γ is in V_m.*

(f) $W_{\hat{f}}(\alpha_0) = 0$ *if and only if f is balanced.*

(g) $\sum_{\alpha \in V_n} C(f, f_\alpha)^2 = 1$ *where f_α is the linear function corresponding to $\langle \alpha, x \rangle$ for any α in V_n.*

Proof. Only **(a)** and **(d)** are proved. The rest follow from the definitions.

(a) Write (4.10) for h :

$$\begin{aligned}
\hat{h}(x) &= \hat{f}(x)\hat{g}(x) \\
&= \left(\sum_{\alpha \in V_n} C(f, f_\alpha)(-1)^{f_\alpha(x)}\right)\left(\sum_{\beta \in V_n} C(g, f_\beta)(-1)^{f_\beta(x)}\right) \\
&= \sum_{\beta \in V_n}\sum_{\alpha \in V_n} C(f, f_\alpha)C(g, f_\beta)(-1)^{f_{\alpha+\beta}(x)} \\
&= \sum_{\sigma \in V_n}\left(\sum_{\beta \in V_n} C(f, f_{\beta+\sigma})C(g, f_\beta)\right)(-1)^{f_\sigma(x)}.
\end{aligned}$$

Thus, $C(h, f_\sigma) = \sum_{\beta \in V_n} C(f, f_{\beta+\sigma}).C(g, f_\beta)$ by (4.10) and the above equations. Equivalently, $C(h, f_\sigma) = \frac{1}{2^n} W_{\hat{h}}(\sigma) = \frac{1}{2^{2n}} \sum_{\beta \in V_n} W_{\hat{f}}(\beta+\sigma) W_{\hat{g}}(\beta)$. Hence, $W_{\hat{h}}(\sigma) = \frac{1}{2^n} \sum_{\beta \in V_n} W_{\hat{f}}(\beta+\sigma) W_{\hat{g}}(\beta)$ giving the result.

(d) First of all note that $\hat{k}(x) = \frac{1}{2}(1 + \hat{f}(x) + \hat{g}(x) - \hat{f}(x).\hat{g}(x))$. Use this in $W_{\hat{k}}(\alpha) = \sum_{x \in V_n} \hat{k}(x)(-1)^{f_\alpha(x)}$ and note that $\sum_{x \in V_n} (-1)^{f_\alpha(x)} = 2^n\delta(\alpha)$. Since $h(x) = (f+g)(x)$ and by using part **(a)**, the result follows. □

4.4 Fast Walsh Transform

The calculation of the Walsh transform of the sign function of a function, denoted by $W_{\hat{f}}(\alpha)$, would require about $2^n \times 2^n = 2^{2n}$ additions and subtractions. However, there is a faster way to obtain $T_{W_{\hat{f}}}$ which is called the Fast Walsh Transform. It is a discrete version of the so-called Fast Fourier Transform. This faster algorithm is the one which we mentioned in Section 2.5. The efficiency of this algorithm comes from the fact that H_n can be written as the product of n matrices of order 2^n where these matrices have only two nonzero elements per column. In other words, by writing H_n as a product of n sparse matrices, it is enough to perform only $n2^n$ additions and subtractions to compute $T_{W_{\hat{f}}}$.

The following lemma gives the method to write H_n in the form mentioned above. This result is used in (4.4) for $n = 3$ to demonstrate the faster algorithm [25].

Lemma 4.4.1 *([25]) Let H_n be the Sylvester-Hadamard matrix of order 2^n. H_n can be written as the product $H_n = H_n^{(1)}.H_n^{(2)} \dots H_n^{(n)}$ of n matrices $H_n^{(1)}$, $H_n^{(2)}, \dots, H_n^{(n)}$ each containing only two nonzero elements per column where $H_n^{(i)} = I_{2^{n-i}} \otimes H_2 \otimes I_{2^{i-1}}$ for $i = 1, 2, \dots, n$ and I_{2^i} is the identity matrix of order 2^i.*

Proof. Use induction on n. For $n = 1$, the result is obvious. Assume that the result is true for n. Then,

$$\begin{aligned}
H_{n+1} &= H_2 \otimes H_n \\
&= H_2 \otimes (H_n^{(1)}.H_n^{(2)} \dots H_n^{(n)}) \\
&= (I_2 \otimes H_n^{(1)})(I_2 \otimes H_n^{(2)}) \dots (I_2 \otimes H_n^{(n)})(H_2 \otimes I_{2^n}) \\
&= (I_2 \otimes (I_{2^{n-1}} \otimes H_2 \otimes I_1))(I_2 \otimes (I_{2^{n-2}} \otimes H_2 \otimes I_2)) \dots \\
&\quad (I_2 \otimes (I_0 \otimes H_2 \otimes I_{2^{n-1}}))(I_0 \otimes H_2 \otimes I_{2^{n+1-1}}) \\
&= H_{n+1}^{(1)}.H_{n+1}^{(2)} \dots H_{n+1}^{(n)}.H_{n+1}^{(n+1)}.
\end{aligned}$$

Note that the third equality is due to **(3)** of Lemma 2.4.2. □

Example 4.4.1 *([25])*

(a) Let $n = 2$. H_2 is given in (2.35). Now, $H_2^{(1)}$ and $H_2^{(2)}$ are as follows :

$$H_2^{(1)} = I_2 \otimes H_2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix},$$

$$H_2^{(2)} = H_2 \otimes I_2 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}.$$

It is easy to see that $H_2 = H_2^{(1)}.H_2^{(2)}$.

(b) Let $n = 3$. H_3 is given in (2.38). $H_3^{(1)}$, $H_3^{(2)}$ and $H_3^{(3)}$ are as follows :

$$H_3^{(1)} = I_{2^2} \otimes H_2 \otimes I_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix},$$

$$H_3^{(2)} = I_2 \otimes H_2 \otimes I_2 = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix},$$

$$H_3^{(3)} = I_1 \otimes H_2 \otimes I_{2^2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

By some tedious work, one can easily verify that indeed $H_3 = H_3^{(1)}.H_3^{(2)}.H_3^{(3)}$.

Now, let $n = 3$ and f be in $\mathcal{F}_3$ with $\zeta_f = (a_0, a_1, a_2, a_3, a_4, a_5, a_6, a_7)$. Since $T_{W_{\hat{f}}} = \zeta_f H_3$, by using $H_3 = H_3^{(1)}.H_3^{(2)}.H_3^{(3)}$, i.e. by multiplying ζ_f first with $H_3^{(1)}$, second with $H_3^{(2)}$ and finally with $H_3^{(3)}$, one obtains that

$$\begin{aligned} {\zeta_f}^{(1)} &= \zeta_f H_3^{(1)} &&= (x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7), \\ {\zeta_f}^{(2)} &= {\zeta_f}^{(1)} H_3^{(2)} &&= (y_0, y_1, y_2, y_3, y_4, y_5, y_6, y_7), \\ {\zeta_f}^{(3)} &= {\zeta_f}^{(2)} H_3^{(3)} &&= (z_0, z_1, z_2, z_3, z_4, z_5, z_6, z_7), \end{aligned}$$

where

$$
\begin{aligned}
x_0 &= a_0 + a_1 \\
x_1 &= a_0 - a_1 \\
x_2 &= a_2 + a_3 \\
x_3 &= a_2 - a_3 \\
x_4 &= a_4 + a_5 \\
x_5 &= a_4 - a_5 \\
x_6 &= a_6 + a_7 \\
x_7 &= a_6 - a_7
\end{aligned}
$$

$$
\begin{aligned}
y_0 &= x_0 + x_2 \\
y_1 &= x_1 + x_3 \\
y_2 &= x_0 - x_2 \\
y_3 &= x_1 - x_3 \\
y_4 &= x_4 + x_6 \\
y_5 &= x_5 + x_7 \\
y_6 &= x_4 - x_6 \\
y_7 &= x_5 - x_7
\end{aligned}
$$

$$
\begin{aligned}
z_0 &= y_0 + y_4 \\
z_1 &= y_1 + y_5 \\
z_2 &= y_2 + y_6 \\
z_3 &= y_3 + y_7 \\
z_4 &= y_0 - y_4 \\
z_5 &= y_1 - y_5 \\
z_6 &= y_2 - y_6 \\
z_7 &= y_3 - y_7.
\end{aligned}
$$

Thus, starting from a_i's for $i = 1, 2, \ldots, 7$, after three steps one obtains z_i's for $i = 1, 2, \ldots, 7$. It is clear that these three matrix multiplications take 3.2^3 operations instead of 2^6 operations if direct multiplication of ζ_f and H_n is performed.

By writing y_i's and z_i's in terms of a_i's for $i = 1, 2, \ldots, 7$, after each mul-

tiplication we have

$$
\begin{array}{lllll}
a_0 & \rightarrow & a_0 + a_1 & \rightarrow & a_0 + a_1 + a_2 + a_3 \quad \rightarrow \quad \ldots \\
a_1 & \rightarrow & a_0 - a_1 & \rightarrow & a_0 - a_1 + a_2 - a_3 \quad \rightarrow \quad \ldots \\
a_2 & \rightarrow & a_2 + a_3 & \rightarrow & a_0 + a_1 - a_2 - a_3 \quad \rightarrow \quad \ldots \\
a_3 & \rightarrow & a_2 - a_3 & \rightarrow & a_0 - a_1 - a_2 + a_3 \quad \rightarrow \quad \ldots \\
a_4 & \rightarrow & a_4 + a_5 & \rightarrow & a_4 + a_5 + a_6 + a_7 \quad \rightarrow \quad \ldots \\
a_5 & \rightarrow & a_4 - a_5 & \rightarrow & a_4 - a_5 + a_6 - a_7 \quad \rightarrow \quad \ldots \\
a_6 & \rightarrow & a_6 + a_7 & \rightarrow & a_4 + a_5 - a_6 - a_7 \quad \rightarrow \quad \ldots \\
a_7 & \rightarrow & a_6 - a_7 & \rightarrow & a_4 - a_5 - a_6 + a_7 \quad \rightarrow \quad \ldots
\end{array}
$$

$$
\begin{array}{ll}
\ldots & \rightarrow \quad a_0 + a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 \\
\ldots & \rightarrow \quad a_0 - a_1 + a_2 - a_3 + a_4 - a_5 + a_6 - a_7 \\
\ldots & \rightarrow \quad a_0 + a_1 - a_2 - a_3 + a_4 + a_5 - a_6 - a_7 \\
\ldots & \rightarrow \quad a_0 - a_1 - a_2 + a_3 + a_4 - a_5 - a_6 + a_7 \\
\ldots & \rightarrow \quad a_0 + a_1 + a_2 + a_3 - a_4 - a_5 - a_6 - a_7 \\
\ldots & \rightarrow \quad a_0 - a_1 + a_2 - a_3 - a_4 + a_5 - a_6 + a_7 \\
\ldots & \rightarrow \quad a_0 + a_1 - a_2 - a_3 - a_4 - a_5 + a_6 + a_7 \\
\ldots & \rightarrow \quad a_0 - a_1 - a_2 + a_3 - a_4 + a_5 + a_6 - a_7.
\end{array}
$$

showing each obtained intermediate result which occurs as a result of the three matrix multiplications.

CHAPTER 5

More Cryptological Properties

5.1 Strict Avalanche Criterion and Propagation Criterion Of Degree k

The strict avalanche criterion is introduced by Webster and Tavares [55]. A function f is said to satisfy the strict avalanche criterion if complementing any single input coordinate results in the output of f changing with probability exactly one half. In other words, the difference function f^α is a balanced function for any α in V_n with $w(\alpha) = 1$. Hence, strict avalanche criterion or in short SAC characterizes the output of the function when there is a single bit change on the input.

An important generalization of the strict avalanche criterion which was introduced in [2] and [38] is the following :

Definition 5.1.1 *A function f is said to satisfy*
(a) *the propagation criterion with respect to a nonzero vector α in V_n if the difference function f^α is balanced.*
(b) *the propagation criterion of degree k if it satisfies the propagation criterion with respect to all α in V_n with $1 \leq w(\alpha) \leq k$. In this case, f is said to satisfy $PC(k)$ and f is said to be a $PC(k)$ function.*

Note that SAC is equivalent to the propagation criterion of degree 1 which is denoted by $PC(1)$.

Recall from Section 3.4 that the number of vectors for which f^α is balanced,

i.e. the number of vectors for which f satisfies the propagation criterion is invariant under any affine transformation on the input coordinates. This is not true for SAC. In other words, the strict avalanche criterion is not invariant under affine transformations on the input coordinates. Thus, by an affine transformation on coordinates, one can construct a SAC fulfilling function from a function which does not satisfy SAC and conversely. This is an indicator that SAC itself is not a strong measurement of propagation criterion. A function may very well be a $PC(1)$ function whereas it may not satisfy the propagation criterion for many vectors with weight greater than one. Moreover, it may even have nonzero linear structures with weight greater than one. It is obvious that having nonzero linear structures is the worst case with respect to propagation criterion. The following result shows how to obtain a SAC fulfilling function from any function by an affine transformation of input coordinates.

Theorem 5.1.1 *([43]) Let f be a function and A be a nonsingular matrix of order n with entries from $GF(2)$. If f satisfies the propagation criterion with respect to each row of A when a row of A is considered as a vector of V_n, then $\psi(x) = f(xA)$ satisfies SAC.*

Proof. Let $\{e_0, e_1, \ldots, e_n\}$ be the standard ordered basis of V_n. Then, $\psi(x) + \psi(x + e_i) = f(xA) + f((x + e_i)A) = f(xA) + f(xA + \gamma_i)$ where $\gamma_i = e_iA$ is the i-th row of A. Hence, $\psi(x) + \psi(x + e_i) = f(u) + f(u + \gamma_i)$ where $u = xA$. Since A is nonsingular, u runs through V_n when x runs through V_n. By hypothesis, f^{γ_i} is balanced for any row γ_i of A which implies that ψ satisfies SAC. □

An important remark is that a function satisfying SAC need not necessarily be balanced. By using the definition, a function f satisfies SAC if and only if $\sum_{x \in V_n} (f(x) + f(x + e_i)) = 2^{n-1}$ for all $i = 1, 2, \ldots, n$ where e_i is the vector in V_n with all entries except the i-th are zero. Equivalently, f satisfies SAC if and only if $\sum_{x \in V_n} (\hat{f}(x).\hat{f}(x + e_i)) = 0$ for all $i = 1, 2, \ldots, n$.

Example 5.1.1 *([17]) Let f be in $\mathcal{F}_3$ with $T_f = (0,0,0,1,1,0,0,0)$. Note that $\sum_{x \in V_n} (f(x) + f(x+e_i)) = 4$ for $i = 1,2,3$. Thus, f satisfies SAC but f is not balanced.*

In [32], $S(n,k)$ denotes the number of functions for which the output changes with probability exactly one half if any of the input variables $x_1, x_2, \ldots, x_k$ among $x = (x_1, x_2, \ldots, x_n)$ is complemented and $S(n,n)$ denotes the number of functions satisfying SAC. We use S_n for the number of functions satisfying SAC as used in [8].

In [32], explicit formulas for $S(n,1)$ and $S(n,2)$ are given both of which are in fact upper bounds for the number of functions satisfying SAC. In [8] and [53], asymptotics for the sizes of $S(n,1)$ and $S(n,2)$ are given, quantifying the number of functions satisfying SAC.

Lemma 5.1.2 *([8]) $S(n,1) \sim 2\pi^{-\frac{1}{2}} 2^{2^n - \frac{n}{2}}$.*

Proof. From Lemma 1 of [32], $S(n,1) = \begin{pmatrix} 2^{n-1} \\ 2^{n-2} \end{pmatrix}$. By applying Stirling's formula $n! = (2\pi n)^{\frac{1}{2}} (\frac{n}{e})^n$ to the binomial coefficient, the result follows. □

Lemma 5.1.3 *([8]) For $n \geq 2$, $S(n,2) > 2^{2^n - n}$.*

The following theorem gives a lower bound for S_n.

Theorem 5.1.4 *([8]) One can explicitly construct $2^{2^{n-2}}$ functions which satify SAC.*

By construction, all functions in Theorem 5.1.4 are balanced.

In [8], t_n is defined as $t_n = \frac{\log_2 S_n}{2^n}$. By Theorem 5.1.4, we get that $t_n \geq \frac{1}{4}$. In [8], a stronger result is proposed as Conjecture 4 and it is proved in [9].

Conjecture 4 (of [8]): Given any choice of the values $f(\alpha_i)$ for $i = 0, 1, \ldots, 2^{n-1} - 1$, there exists a choice of $f(\alpha_i)$, for $2^{n-1} \leq i \leq 2^n - 1$ such that the resulting function f satisfies SAC.

The proof of Conjecture 4 implies that there are at least $2^{2^{n-1}}$ functions which satisfy SAC and improve the bound from $t_n \geq \frac{1}{4}$ to $t_n \geq \frac{1}{2}$. This inequality was proved independently in [54] by using a different method. Later, Daniel Biss has given a much more complicated argument that shows $t_n = 1$, thereby disproving the Conjecture 1 of [8].

CHAPTER 6

Bent Functions and Properties

6.1 Bent Functions

In this chapter, a non-exhaustive survey is given about the properties of one of the most important classes of functions in $\mathcal{F}_n$, so called bent functions. Almost all results of this chapter are from the article of Rothaus [39] in which bent functions are introduced and from MacWilliams and Sloane [25]. In this chapter, the characterization of bent functions by using the Walsh transform and by other cryptological means will be presented.

A function f in $\mathcal{F}_n$ is called bent if all the Walsh transform coefficients $W_{\hat{f}}(\alpha)$ given in (4.3) have the same absolute value, i.e. $|W_{\hat{f}}(\alpha)|$ is constant for all α in V_n. By using Parseval's equation in Corollary 4.3.2, f is a bent function if and only if $|W_{\hat{f}}(\alpha)| = 2^{n/2}$ for all α in V_n. Since $W_{\hat{f}}(\alpha)$ is an integer for all α in V_n, if f is a bent function, then n must be even. In this chapter, unless otherwise stated explicitly, we assume that n is even and $n \geq 2$.

Lemma 6.1.1 *([39, 14]) Let f be a function.*

(a) *f is bent if and only if $\bar{f}$, the complement of f, is bent.*

(b) *Being bent is invariant under nonsingular affine transformations on the input coordinates. In other words, f is bent if and only if the function $h = f \circ \theta$ is bent where $\theta(x) = xA + \alpha$, A is a nonsingular matrix of order n and α is any vector in V_n.*

(c) *f is bent if and only if the function $f + \varphi$ is bent where φ as an affine function.*

(d) *f is bent if and only if $\langle \zeta_f, \zeta_\varphi \rangle = \pm 2^{n/2}$ where ζ_φ is the sequence of an*

affine function φ. The sequence of an affine function is called an affine sequence.

(e) *f is bent if and only if the function $h = f + f_\alpha$, where $f_\alpha(x) = \langle \alpha, x \rangle$, has weight $2^{n-1} \pm 2^{\frac{n}{2}-1}$.*

Proof. Only the proof of **(b)** is given. The others can be deduced easily from the definition of a bent function.

(b) f is bent if and only if $|W_{\hat{f}}(\alpha)| = 2^{n/2}$ for all α in V_n. Let $\theta(x) = xA+\alpha$ where A is nonsingular and α in V_n. By Lemma 3.4.1, $d(f, f_\alpha) = d(f \circ \theta, f_\alpha \circ \theta)$ where f_α is the linear function corresponding to α. Note that $|W_{\hat{f}}(\alpha)| = 2^{n/2} = |2^n - 2d(f, f_\alpha)| = |2^n - 2d(f \circ \theta, f_\alpha \circ \theta)| = |2^n - 2d(h, f_\beta)| = |W_{\hat{h}}(\beta)|$ where $h = f \circ \theta$ and $f_\beta = \langle \beta, x \rangle + a_0 = f_\alpha \circ \theta$ for some a_0 in $GF(2)$. Since A is nonsingular, as θ runs through all nonsingular affine transformations f_β runs through all linear functions giving that $h = f \circ \theta$ is also bent. $\square$

One of the most important characterizations of bent functions is the following :

Theorem 6.1.2 *([25]) Let f be a function. Then, f is bent if and only if $d(f, \mathcal{A}_n) = N_{max}$ where $N_{max} = 2^{n-1} - 2^{\frac{n}{2}-1}$ is the largest value of nonlinearity as proved in Lemma 2.5.2 and $\mathcal{A}_n$ is the set of all affine functions. In other words, f is the furthest function away from the set of all affine functions with respect to the Hamming distance.*

Proof. Let f be bent. Then, $|W_{\hat{f}}(\alpha)| = 2^{n/2}$ for all α in V_n. Thus, $d(f, f_\alpha) = \frac{1}{2}(2^n - W_{\hat{f}}(\alpha)) = 2^{n-1} \pm 2^{n/2-1}$ and $d(f, \bar{f}_\alpha) = \frac{1}{2}(2^n + W_{\hat{f}}(\alpha)) = 2^{n-1} \mp 2^{n/2-1}$ implies that $d(f, \varphi) = 2^{n-1} \pm 2^{n/2-1}$ for any affine function φ in $\mathcal{A}_n$. It is obvious that $N_f = \min_{\varphi \in \mathcal{A}_n} \{d(f, \varphi)\} = \min\{2^{n-1} + 2^{n/2-1}, 2^{n-1} - 2^{n/2-1}\} = 2^{n-1} - 2^{n/2-1} = N_{max}$.

For the converse, suppose that f is a not bent. Then, $|W_{\hat{f}}(\alpha)| \neq 2^{n/2}$ for all α in V_n. By Parseval's equation there exists α in V_n such that $|W_{\hat{f}}(\alpha)| \geq 2^{n/2}$. Since $d(f, f_\alpha) = \frac{1}{2}(2^n - W_{\hat{f}}(\alpha))$ and $d(f, \bar{f}_\alpha) = \frac{1}{2}(2^n + W_{\hat{f}}(\alpha))$, either

$d(f, f_\alpha) < N_{max}$ or $d(f, \bar{f}_\alpha) < N_{max}$. Thus, $N_f < N_{max}$. $\square$

The upper bound $N_{max} = 2^{n-1} - 2^{\frac{n}{2}-1}$ in Lemma 2.5.2 is an integer, hence it is attainable if n is even. However, even if n is even, this does not imply that there exist functions having this nonlinearity. The above theorem guarantees that there exists a certain class of functions attaining the largest nonlinearity. Furthermore, this class may be solely described in terms of its property of having the largest nonlinearity. Although they have the largest nonlinearity among all functions, bent functions have an important drawback to be used in cryptography directly, as shown in the following lemma.

Lemma 6.1.3 *([39]) Let f be a bent function. Then, $w(f) = 2^{n-1} \pm 2^{\frac{n}{2}-1}$.*

Proof. Since $W_{\hat{f}}(\alpha_0) = 2^n - 2w(f)$, $|W_{\hat{f}}(\alpha_0)| = |2^n - 2w(f)| = 2^{n/2}$. This implies that $w(f) = 2^{n-1} \pm 2^{n/2-1}$. $\square$

Thus, by Lemma 6.1.3, bent functions are not balanced. So, a necessary caution should be taken when using a bent function in a cryptosystem. Another important fact obtained from Theorem 6.1.2 and Lemma 6.1.3 is that balanced functions can not attain the largest nonlinearity. In other words, balance and largest nonlinearity can not be simultaneously satisfied.

Although bent functions are not balanced and can not be used directly, they are used as building blocks of many cryptologically important constructions such as the construction of highly nonlinear balanced functions with good propagation characteristics [44], the construction of cryptographically robust S-boxes [47] and many more.

Since the introduction of bent functions [39], although a significant amount of work has been spent on them, still very few distinct classes of bent functions are known. An important problem in this theory is to construct new classes of bent functions either by using the previously known classes or by different methods. From the cryptological point of view, the nonlinearity and the propagation characteristics of bent functions are very attractive. The propagation

characteristics of bent functions will also be investigated in this chapter. The result is that bent functions are also the best among all functions with respect to the propagation criterion. In other words, bent functions are $PC(n)$. As mentioned in the above paragraph, bent functions are a good source to construct cryptological functions or mappings. In the following chapter, methods of constructing highly nonlinear balanced functions with good propagation characteristics will be presented [44]. Most of the methods presented in the following chapter will be based on concatenating, splitting and modifying known bent functions.

Bent functions are also useful to observe the relations between cryptologically important properties. As noted in the relation of balance and largest nonlinearity, not all these properties can be simultaneously satisfied. This is also true if one considers balance and nonlinearity with propagation criterion and correlation immunity (or resiliency) [49]. It is well-known that bent functions are not correlation immune. These facts show that there are important trade offs between cryptologically important properties. Hence, if one looks for a function which satisfies a list of properties (possibly some of them are conflicting), then the best he can do is to seek a function in some special subsets of $\mathcal{F}_n$ by exhaustive search or to construct a function explicitly which optimizes these properties.

We continue to investigate the properties of bent functions. Let f be a bent function. Define the function g by setting $(-1)^{g(\alpha)} = \frac{W_{\hat{f}}(\alpha)}{2^{n/2}}$ for all α in V_n where $W_{\hat{f}}(\alpha)$ denotes the Walsh transform of the sign function of f. Since f is bent, it is clear that g is in $\mathcal{F}_n$. Note that for any α in V_n the Walsh transform coefficients of the sign function of g are also $\pm 2^{n/2}$ since :

$$\begin{aligned} \hat{f}(x) = \frac{1}{2^n} \sum_{\alpha \in V_n} W_{\hat{f}}(\alpha)(-1)^{\langle \alpha, x \rangle} &= \frac{1}{2^{n/2}} \sum_{\alpha \in V_n} (-1)^{g(\alpha)+\langle \alpha, x \rangle} \\ &= \frac{1}{2^{n/2}} W_{\hat{g}}(x). \end{aligned}$$

Thus, g defined as above is also a bent function. This fact can be stated equivalently in terms of the sequence of f as follows :

Let f be a bent function with sequence ζ_f. Then, the $(1,-1)$-sequence $2^{-n/2}\zeta_f H_n$ is the sequence of a bent function. The sequence of a bent function is commonly called as a bent sequence in the literature.

Consider now a bent function f and the function g obtained via f as above. Define a function h by using g exactly in the same way as g is defined by using f as $(-1)^{h(\alpha)} = \frac{W_{\hat{g}}(\alpha)}{2^{n/2}}$. It is obvious that $h = f$. This shows that there is a natural pairing $f \longleftrightarrow g$ of bent functions [39].

Lemma 6.1.4 *([39],[25]) A function f is bent if and only if the matrix $A = (a_{ij})$ of order 2^n where $a_{ij} = \frac{1}{2^{n/2}} W_{\hat{f}}(\alpha_i + \alpha_j)$ for $0 \leq i, j \leq 2^n - 1$ is a Hadamard matrix.*

Proof. Let $AA^t = (x_{ij})$ where $x_{ij} = \frac{1}{2^n} \sum_{k=0}^{2^n-1} W_{\hat{f}}(\alpha_i + \alpha_k) W_{\hat{f}}(\alpha_j + \alpha_k)$. By using Lemma 4.3.1, we get that

$$x_{ij} = \begin{cases} 2^n & \text{if } \ i = j, \\ 0 & \text{if } \ i \neq j. \end{cases}$$

Thus, A is a Hadamard matrix. The converse is trivial. □

By using Theorem 2.6.4 and the definition of a bent function, the following important property of bent functions is obtained.

Theorem 6.1.5 *([14]) A function f is bent if and only if f^α is balanced for any nonzero α in V_n. Equivalently, f is bent if and only if f satisfies $PC(n)$.*

In the literature, the difference function f^α corresponding to α in V_n is sometimes called as the directional derivative of f in the direction of α. By Theorem 6.1.5, another characterization of bent functions is obtained via their propagation characteristics. A simple but worth to mention fact is that a balanced function can not be $PC(n)$. This shows that balance and $PC(n)$ can not be simultaneously satisfied.

Lemma 6.1.6 *([39]) A function f is bent if and only if the matrix of f, $M_f = (m_{ij})$ where $m_{ij} = (-1)^{f(\alpha_i + \alpha_j)}$ for $0 \leq i, j \leq 2^n - 1$, is a Hadamard matrix.*

Proof. Let $M_f {M_f}^t = (y_{ij})$ where $y_{ij} = \sum_{\alpha_k \in V_n} (-1)^{f(\alpha_i+\alpha_k)+f(\alpha_j+\alpha_k)}$. By a simple change of variable in the bound of this summation, one can obtain that $y_{ij} = \sum_{\theta \in V_n} (-1)^{f(\theta)+f(\alpha_i+\alpha_j+\theta)}$. Since the function inside the summation is the directional derivative of f in the direction of $\alpha_i + \alpha_j$, Theorem 6.1.5 gives that

$$y_{ij} = \begin{cases} 2^n & \text{if } \ i = j, \\ 0 & \text{if } \ i \neq j. \end{cases}$$

Thus, M_f is a Hadamard matrix. The converse is also true again by using Theorem 6.1.5. □

An important fact about the degree of bent functions is that if a function has degree strictly greater than $n/2$, then that function can not be bent.

Proposition 6.1.7 *([39]) If f is a bent function and $n > 2$, then the degree of f is less than or equal to $n/2$, i.e. $deg(f) \leq n/2$.*

Proof. Since f is bent, let $n = 2k$ where $k > 1$. Let r be an integer satisfying $1 < k < r \leq n$. Define $f(x_1, x_2, \ldots, x_k, \ldots, x_r, 0, 0, \ldots, 0) = g(x_1, x_2, \ldots, x_k, \ldots, x_r)$. By (4.5),

$$(-1)^{g(x_1,x_2,\ldots,x_r)} = \frac{1}{2^r} \sum_{\alpha_1,\alpha_2,\ldots,\alpha_r \in GF(2)} W_{\hat{g}}(\alpha_1, \alpha_2, \ldots, \alpha_r)(-1)^{\alpha_1 x_1 + \cdots + \alpha_r x_r}$$

and

$$(-1)^{f(x_1,x_2,\ldots,x_r,0,\ldots,0)} = \frac{1}{2^n} \sum_{\alpha_1,\alpha_2,\ldots,\alpha_n \in GF(2)} W_{\hat{f}}(\alpha_1, \alpha_2, \ldots, \alpha_n)(-1)^{\alpha_1 x_1 + \cdots + \alpha_r x_r}.$$

Since the left hand sides of these two equations are equal, so are the right hand sides. Equating them and by using the uniqueness of the expansion in (4.5), one obtains that

$$W_{\hat{g}}(\alpha_1, \alpha_2, \ldots, \alpha_r) = \frac{1}{2^{n-r}} \sum_{\alpha_{r+1},\alpha_{r+2},\ldots,\alpha_n \in GF(2)} W_{\hat{f}}(\alpha_1, \alpha_2, \ldots, \alpha_r; \alpha_{r+1}, \ldots, \alpha_n).$$

Since $W_{\hat{g}}(\alpha_0) = 2^r - 2w(g)$ where α_0 is in V_r, the number of zeros of the function $g(x_1, x_2, \ldots, x_r)$ is equal to $2^{r-1} + \frac{1}{2} W_{\hat{g}}(0, 0, \ldots, 0)$. By using the

above equation, the number of zeros of g is equal to

$$2^{r-1} + \frac{1}{2^{n-r+1}} \sum_{\alpha_{r+1},\alpha_{r+2},\ldots,\alpha_n \in GF(2)} W_{\hat{f}}(0,0,\ldots,0;\alpha_{r+1},\ldots,\alpha_n).$$

Thus, the number of zeros of g is equal to

$$2^{r-1} + 2^{r-n-1} \sum_{\alpha_{r+1},\alpha_{r+2},\ldots,\alpha_n \in GF(2)} W_{\hat{f}}(0,0,\ldots,0;\alpha_{r+1},\ldots,\alpha_n).$$

There are 2^{n-r} summands in the above summation and since f is bent,

$$|W_{\hat{f}}(0,0,\ldots,0;\alpha_{r+1},\ldots,\alpha_n)| = 2^{\frac{n}{2}}$$

for any $\alpha_{r+1},\ldots,\alpha_n \in GF(2)$. Hence, the number of zeros of g is even. However, $w(g) \equiv a_{12\ldots r} \pmod 2$ by (2.11) where $a_{12\ldots r}$ is the coefficient of the term $x_1x_2\ldots x_r$ in g. Thus, $a_{12\ldots r} = 0$. In other words, the degree of g is strictly less than r. Since r is arbitrary in the range $1 < k < r \leq n$, $deg(f) \leq k$. □

Proposition 6.1.8 *([39]) Let the function h be defined as $h(z) = f(x) + g(y)$ for $z = (x, y)$ in V_{n+m}, x in V_n and y in V_m. Then, h is bent if and only if f and g are bent.*

Proof. ([25]) Let $\alpha = (\beta, \gamma)$ be in V_{n+m} for β in V_n and γ in V_m. By **(e)** of Theorem 4.3.3, $|W_{\hat{h}}(\alpha)| = |W_{\hat{f}}(\beta)|.|W_{\hat{g}}(\gamma)|$.

If f and g are both bent then, $|W_{\hat{f}}(\beta)| = 2^{\frac{n}{2}}$ and $|W_{\hat{g}}(\gamma)| = 2^{\frac{m}{2}}$ for any β in V_n and for any γ in V_m. Thus, $|W_{\hat{h}}(\alpha)| = 2^{\frac{n+m}{2}}$ for any α in V_{n+m} giving that h is bent.

Conversely, assume that h is bent but f is not bent. Thus, there exists some β in V_n such that $|W_{\hat{f}}(\beta)| > 2^{\frac{n}{2}}$. Then, for any $\alpha = (\beta, \gamma)$ in V_{n+m} $2^{\frac{n+m}{2}} = |W_{\hat{f}}(\beta)|.|W_{\hat{g}}(\gamma)|$. It follows that $|W_{\hat{g}}(\gamma)| < 2^{\frac{m}{2}}$ for all γ in V_m. However, this is impossible due to Corollary 4.3.2. □

The functions of type h used in Proposition 6.1.8 are given a special name in [25] as follows :

Definition 6.1.1 *A function h is called decomposable if there is a linear transformation on the input coordinates such that h can be written as a sum of functions on disjoint variables as in Proposition 6.1.8.*

In other words, h in $\mathcal{F}_n$ is decomposable if there exists a binary matrix of order n such that $h(zA) = f(x) + g(y)$ where $z = (x, y)$ is in V_n for x in V_k and y in V_t satisfying $k + t = n$. If there exists no such matrix, then h is said to be indecomposable.

If h is a decomposable bent function for $n = 2k$, then by Proposition 6.1.7 the degree of each function f and g is necessarily strictly less than k, except in the case when $k = 2$. This gives the following :

Proposition 6.1.9 *([39]) If f is a bent function in $\mathcal{F}_n$ where $n = 2k$ for $k \geq 3$, then f is indecomposable.*

After having investigated many properties of bent functions, it is time to see some examples of bent functions.

Theorem 6.1.10 *([39]) Let $n = 2k$ and g be any function in $\mathcal{F}_k$. The function, $f(z) = f(x, y) = \langle x, y\rangle + g(x) = x_1y_1 + x_2y_2 + \cdots + x_ky_k + g(x)$ is bent where $z = (x, y)$ in V_n, x, y are in V_k with $x = (x_1, x_2, \ldots, x_k)$ and $y = (y_1, y_2, \ldots, y_k)$.*

Proof. Let f be in $\mathcal{F}_n$ of the form described in the statement of the theorem and let f_α be the linear function corresponding to $\langle \alpha, z\rangle$ for any α in V_n. It is enough to show that $(f + f_\alpha)(z)$ has $2^{2k-1} \pm 2^{k-1}$ zeros. Part **(e)** of Lemma 6.1.1 yields the result.

Write $\alpha = (\beta, \gamma)$ where β, γ are in V_k. Then, $f_\alpha(z) = f_\beta(x) + f_\gamma(y)$. Set $h(z) = (f + f_\alpha)(z)$. Then, $h(z) = h(x, y) = g(x) + \langle \beta, x\rangle + \langle x + \gamma, y\rangle$.

For the values of x where $x + \gamma = \alpha_0$, $h(z) = h(\gamma, y) = g(\gamma) + \langle \beta, \gamma\rangle$ and is a constant for any y, i.e. independent of $y_1, y_2, \ldots, y_k$. If $g(\gamma) + \langle \beta, \gamma\rangle = 0$, then h has 2^k zeros and no zeros otherwise.

For the values of x where $x + \gamma \neq \alpha_0$, the function $h(z)$ is a nonconstant linear function in the variable y and hence has 2^{k-1} zeros. Since, there are $2^k - 1$ choices for x, h has $2^{k-1}.(2^k - 1) = 2^{2k-1} - 2^{k-1}$ zeros.

Thus, h has in total $2^{2k-1} \pm 2^{k-1}$ zeros. This gives the result. □

As a corollary of Theorem 6.1.10, the function defined as $f(x_1, x_2, \ldots, x_n) = x_1x_2 + x_3x_4 + \cdots + x_{n-1}x_n$ is a bent function when n is even [25].

Theorem 6.1.11 *([25]) Let f be the function defined as $f(x_1, x_2, \ldots, x_n) = x_1x_2 + x_1x_3 + \ldots + x_{n-1}x_n$ where n is even. Then, f is bent.*

Definition 6.1.2 *([25]) Two functions f, g are said to be equivalent if the difference $h(x) = f(x) + g(x)$ is an affine function. If f, g are equivalent, then they are denoted by $f \sim g$.*

Theorem 6.1.12 *([25]) Let $n = 2k$ and f be the function defined as $f(z) = f(x, y) = \langle x, y \rangle$ where $z = (x, y)$ in V_n, x, y are in V_k with $x = (x_1, x_2, \ldots, x_k)$ and $y = (y_1, y_2, \ldots, y_k)$, i.e. take $g = 0$ in Theorem 6.1.10. Then, the functions $f, f + x_1x_2x_3, f + x_1x_2x_3x_4, \ldots, f + x_1x_2 \ldots x_k$ are $k - 1$ inequivalent bent functions of degrees $2, 3, \ldots, k$.*

Theorem 6.1.13 *([39]) Let $n = 2k$ and f, g, h be in $\mathcal{F}_n$ such that $f + g + h$ is bent. Define the function θ in $\mathcal{F}_{n+2}$ as $\theta(z) = \theta(x_1, x_2, \ldots, x_n, u, v) = f(x).g(x) + g(x).h(x) + h(x).f(x) + (f(x) + g(x)).u + (f(x) + h(x)).v + u.v$ where $z = (x, u, v)$ is in V_{n+2} for $x = (x_1, x_2, \ldots, x_n)$ in V_n and u, v in $GF(2)$. Then, θ is also bent.*

Proof. The proof proceeds exactly as the proof of Theorem 6.1.10 and can be found in [39]. □

Remark 6.1.14 *The class of bent functions in Theorem 6.1.13 give the most general polynomial of the form $\theta(x, u, v) = u.v + f(x).v + g(x).u + h(x)$.*

The class of bent functions in Theorem 6.1.13 contains the class in Theorem 6.1.10.

Now, an important construction of bent functions is given without proof. We state the form as in [25].

Theorem 6.1.15 *(Maiorana-McFarland) Let g be any function in $\mathcal{F}_n$ and φ be a bijective transformation of V_n given by*

$$\varphi(x) = (\varphi_1(x_1, x_2, \ldots, x_n), \varphi_2(x_1, x_2, \ldots, x_n), \ldots, \varphi_n(x_1, x_2, \ldots, x_n))$$

where $x = (x_1, x_2, \ldots, x_n)$ is in V_n. Then, $f(x, y) = \langle \varphi(x), y \rangle + g(y) = \varphi_1(x).y_1 + \varphi_2(x).y_2 + \ldots + \varphi_n(x).y_n + g(y)$ is a bent function in $\mathcal{F}_n$ where x, y in V_n with $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$.

The following paragraph is from [31] :

The general case when $q = p^n$, a prime power and g, φ over $\mathbb{Z}_q^n$ instead of V_n is due to [21]. When $q = 2$, it was proved by Maiorana (unpublished, see [15]) generalizing the construction method of Rothaus [39]. An equivalent method is given by McFarland [28]. A third equivalent way of looking at this construction when $q = 2$ is to make use of Hadamard matrices as in [22]. The constructions given in [3] and [52] are special cases of Theorem 6.1.15.

As mentioned in [30], different choices for φ and g in Theorem 6.1.15 yield different bent functions. It follows that, the number of bent functions is lower bounded by $2^{2^{\frac{n}{2}}}.(2^{\frac{n}{2}}!)$.

Definition 6.1.3 *([30]) Let G be an additive abelian group of order v. A subset D of G is called a (v, k, λ)-difference set if the order of D is k and if every element $a \in D$ can be expressed in λ different ways as a difference $a = b - c$ where b, c are in D.*

The following theorem is stated as it is in [30].

Theorem 6.1.16 *([15]) A function f in $\mathcal{F}_n$ is bent if and only if it is a characteristic function of a difference set in V_n where the parameters of a difference set in V_n are $(2^n, 2^{n-1} \pm 2^{\frac{n}{2}-1}, 2^{n-2} \pm 2^{\frac{n}{2}-1})$ due to [26].*

The main result of [15] is that for $k > 3$ there exist bent functions in $\mathcal{F}_{2k}$ which are not equivalent to any functions in Theorem 6.1.15 [30].

CHAPTER 7

Constructions

7.1 Constructing Highly Nonlinear Functions

All constructions in this chapter are from the article of Seberry, Zhang and Zheng [44].

Let f_1, f_2 be in $\mathcal{F}_n$. Consider g in $\mathcal{F}_{n+1}$ defined as

$$g(z) = (y+1)f_1(x_1, x_2, \ldots, x_n) + yf_2(x_1, x_2, \ldots, x_n) \tag{7.1}$$

where $z = (y, x_1, x_2, \ldots, x_n)$ is in V_{n+1}. Siegenthaler [49] showed that if f_1 and f_2 are m-th order correlation immune functions in $\mathcal{F}_n$, then g is an m-th order correlation immune function in $\mathcal{F}_{n+1}$.

It is clear from (7.1) that $g(0, \alpha) = f_1(\alpha)$ for any $\beta = (0, \alpha)$ in V_{n+1} and $g(1, \alpha) = f_2(\alpha)$ for any $\beta = (1, \alpha)$ in V_{n+1} where α is arbitrary in V_n. This gives that the truth table of g is the concatenation of the truth tables of f_1 and f_2. In such a case, g is said to be the concatenation of f_1, f_2 and the truth table of g is denoted by $T_g = (T_{f_1}, T_{f_2})$. The following lemma gives a lower bound on the nonlinearity of a function obtained by the concatenation of two special functions.

Lemma 7.1.1 *Let f_1, f_2 be in $\mathcal{F}_n$ and g be a function in $\mathcal{F}_{n+1}$ obtained as in (7.1). Suppose that $\langle \zeta_{f_i}, l \rangle \leq P_i$ holds where ζ_{f_i} is the sequence of f_i, l is the sequence of any affine function and P_i is a positive integer for $i = 1, 2$. Then, the nonlinearity of g satisfies $N_g \geq 2^n - \frac{1}{2}(P_1 + P_2)$.*

Proof. By construction, the sequence ζ_g of g is equal to $\zeta_g = (\zeta_{f_1}, \zeta_{f_2})$. Let φ be an arbitrary affine function in $\mathcal{A}_{n+1}$ with sequence ζ_φ. By Lemma 2.5.1,

$\zeta_\varphi = (l_i, \pm l_i)$ where l_i is the sequence of an affine function in $\mathcal{F}_n$. Thus, $\langle \zeta_g, \zeta_\varphi \rangle = \langle \zeta_{f_1}, l_i \rangle \pm \langle \zeta_{f_2}, l_i \rangle$ which implies that $|\langle \zeta_g, \zeta_\varphi \rangle| \leq P_1 + P_2$. By Lemma 2.2.1, $d(g, \varphi) = 2^n - \frac{1}{2}\langle \zeta_g, \zeta_\varphi \rangle \geq 2^n - \frac{1}{2}(P_1 + P_2)$. Since φ is arbitrary, the result follows. □

The construction which was introduced by Meier and Staffelbach [29] as a special case of Lemma 7.1.1 shows that highly nonlinear functions may be obtained by concatenating bent sequences.

Corollary 7.1.2 *Let $n = 2k$ and f_1, f_2 be bent functions in $\mathcal{F}_n$. Then, g constructed as in Lemma 7.1.1 has nonlinearity $N_g \geq 2^{2k} - 2^k$.*

One can also get similar results by concatenating four functions instead of two :

Lemma 7.1.3 *Let f_i be in $\mathcal{F}_n$ with sequence ζ_{f_i} and suppose that $\langle \zeta_{f_i}, l \rangle \leq P_i$ holds for any affine sequence l of length 2^n where P_i's are positive integers for $i = 0, 1, 2, 3$. Let g be in $\mathcal{F}_{n+2}$ obtained by the concatenation of f_i's. In other words, $g(z) = \sum_{i=0}^{3} \chi_{\alpha_i}(y).f_i(x)$ where $z = (y, x)$ is in V_{n+2} for $y = (y_1, y_2)$ in V_2, $x = (x_1, x_2, \ldots, x_n)$ in V_n and χ_{α_i} is the characteristic function of α_i in V_2. Then, $N_g \geq 2^{n+1} - \frac{1}{2}(P_0 + P_1 + P_2 + P_3)$. As in Lemma 7.1.1, if n is even and f_i's are bent functions for $i = 0, 1, 2, 3$, then $N_g \geq 2^{n+1} - 2^{\frac{n}{2}+1}$.*

Proof. The proof is exactly the same with the proof of Lemma 7.1.1. Just note that $H_{n+2} = H_2 \otimes H_n$. It follows that the sequence ζ_φ of any affine function φ is equal to (l_i, l_i, l_i, l_i) or $(l_i, -l_i, l_i, -l_i)$ or $(l_i, l_i, -l_i, -l_i)$ or $(l_i, -l_i, -l_i, l_i)$ where l_i is the sequence of some affine function in $\mathcal{F}_n$. The rest is the same. □

Remark 7.1.4 ***(a)*** *Lemma 7.1.3 can easily be generalized to the case where 2^t functions are concatenated.*

(b) *In Lemma 7.1.1 and Lemma 7.1.3, one can obtain balanced functions by using suitable functions. In Lemma 7.1.1, if f_0 and f_1 are both balanced or*

more generally if f_0, f_1 satisfy $w(f_0) + w(f_1) = 2^n$, then g is balanced. These are also true for the function in Lemma 7.1.3.

Another way to obtain highly nonlinear balanced functions apart from concatenating bent functions is by splitting bent sequences.

Lemma 7.1.5 *Let $n = 2k$ and $f(x_1, x_2, \ldots, x_n)$ be a bent function. Define two functions g_0, g_1 in $\mathcal{F}_{n-1}$ by $g_0(x_2, x_3, \ldots, x_n) = f(0, x_2, x_3, \ldots, x_n)$ and $g_1(x_2, x_3, \ldots, x_n) = f(1, x_2, x_3, \ldots, x_n)$ with sequences ζ_{g_0}, ζ_{g_1} respectively. For any affine sequence l of length 2^{n-1} and for $i = 0, 1$, the following holds :*

$$-2^k \leq \langle \zeta_{g_i}, l \rangle \leq 2^k.$$

Proof. Note that $f(x) = (x_1 + 1)g_0(x_2, x_3, \ldots, x_n) + x_1 g_1(x_2, x_3, \ldots, x_n)$. Let L be an affine sequence of length 2^n. Since $H_n = H_1 \otimes H_{n-1}$, $L = (l, l)$ or $L = (l, -l)$ for some affine sequence l of length 2^{n-1}.

Assume that $-2^k \leq \langle \zeta_{g_0}, l \rangle \leq 2^k$ is not true. Without loss of generality, let $\langle \zeta_{g_0}, l \rangle > 2^k$. We have, $\langle \zeta_f, L \rangle = \langle \zeta_{g_0}, l \rangle \pm \langle \zeta_{g_1}, l \rangle$ according to $L = (l, l)$ or $L = (l, -l)$.

If $\langle \zeta_{g_1}, l \rangle > 0$, then $\langle \zeta_f, L \rangle > 2^k$ for $L = (l, l)$ and if $\langle \zeta_{g_1}, l \rangle < 0$, then $\langle \zeta_f, L \rangle > 2^k$ for $L = (l, -l)$, both of which contradict the fact that $\langle \zeta_f, L \rangle = \pm 2^k$. Thus, $-2^k \leq \langle \zeta_{g_0}, l \rangle \leq 2^k$ holds for any affine sequence l of length 2^{n-1}. The fact $-2^k \leq \langle \zeta_{g_1}, l \rangle \leq 2^k$ is proved exactly in the same way. □

Let $n = 2k$ and f be a bent function. Then, by using Lemma 7.1.5 one concludes that $N_{g_i} \geq 2^{2k-2} - 2^{k-1}$ for $i = 0, 1$.

Note that the concatenation of two bent functions in $\mathcal{F}_{2k-2}$ by using Corollary 7.1.2 yields a function g with $N_g \geq 2^{2k-2} - 2^{k-1}$. Thus, concatenating two bent functions in $\mathcal{F}_{2k-2}$ and splitting a bent function in $\mathcal{F}_{2k}$ both result in functions having nonlinearities bounded below by the same value.

We have seen that concatenating two bent functions properly yields a balanced function. Similarly, one can also obtain a balanced function by splitting a bent function. For this, the result obtained by Adams and Tavares [2] will

be used. It states that the concatenation of the rows $l_0, l_1, \ldots, l_{2^k-1}$ of H_k is a bent sequence of length 2^{2k}. Denote the resulting function by $f(x_1, x_2, \ldots, x_{2k})$. Since $f(x) = (x_1 + 1)g_0(x_2, x_3, \ldots, x_{2k}) + x_1 g_1(x_2, x_3, \ldots, x_{2k})$, where g_0 and g_1 are as in Lemma 7.1.5, the second half of the sequence of f is the sequence of $g_1(x_2, x_3, \ldots, x_{2k})$. This sequence is equal to $\zeta_{g_1} = (l_{2^{k-1}}, l_{2^{k-1}+1}, \ldots, l_{2^k-1})$. Since all rows of H_k except the all-one sequence l_0 is balanced, g_1 is a balanced function with nonlinearity $N_{g_1} \geq 2^{2k-2} - 2^{k-1}$.

By permuting the l_i's appearing in the sequence of g_1 for $2^{k-1} \leq i \leq 2^k - 1$, one obtains a different balanced sequence $\zeta^\star_{g_1} = (l_{i_{2^{k-1}}}, l_{i_{2^{k-1}+1}}, \ldots, l_{i_{2^k-1}})$ where $\{i_{2^{k-1}}, i_{2^{k-1}+1}, \ldots, i_{2^k-1}\}$ is any permutation of $\{2^{k-1}, 2^{k-1}+1, \ldots, 2^k-1\}$. The function corresponding to $\zeta^\star_{g_1}$ has also the same nonlinearity as the function corresponding to ζ_{g_1}. Thus, $\zeta^\star = (a_{2^{k-1}}.l_{i_{2^{k-1}}}, a_{2^{k-1}+1}.l_{i_{2^{k-1}+1}}, \ldots, a_{2^k-1}.l_{i_{2^k-1}})$ are balanced sequences with the same nonlinearity as ζ_{g_1} where $a_i \in \{+1, -1\}$ for $2^{k-1} \leq i \leq 2^k - 1$. Hence, there are $2^{2^{k-1}}.(2^{k-1}!)$ different balanced sequences with this nonlinearity which are obtained by permuting l_i's and changing the signs of a_i's for $i = 2^{k-1}, 2^{k-1} + 1, \ldots, 2^k - 1$.

Now, bent sequences of length 2^{2k} obtained by concatenating the rows of the Sylvester-Hadamard matrices will be modified so that the resulting functions on V_{2k} are balanced and have a much higher nonlinearity than those which are obtained by concatenating four bent sequences. This result with the sequences in [35] will lead to the construction of balanced functions on V_{2k+1} for $k \geq 14$. These functions have higher nonlinearities than those which are obtained by concatenating or splitting bent sequences . These results bring significant improvements to the previously known construction methods. Even and odd dimensional cases will be considered separately.

(a) <u>On V_{2k}</u> :

Lemma 7.1.6 *For any integer $t \geq 1$, there exists*
(i) *a balanced function f on V_{4t} such that $N_f \geq 2^{4t-1} - 2^{2t-1} - 2^t$.*
(ii) *a balanced function f on V_{4t+2} such that $N_f \geq 2^{4t+1} - 2^{2t} - 2^t$.*

Proof. First of all, note that an even number $n \geq 4$ can be expressed as

$n = 4t$ or $n = 4t + 2$ where $t \geq 1$.

(i) The concatenation of all l_i's is a bent sequence of length 2^{4t} where l_i is the i-th row of H_{2t} for $i = 0, 1, \ldots, 2^{2t} - 1$. Since all rows of H_{2t} except l_0 are balanced, replacing l_0 with a balanced sequence of length 2^{2t} makes $\zeta = (l_0, l_1, \ldots, l_{2^{2t}-1})$ a balanced sequence. Thus, the crucial point here is a replacement which makes the function balanced while making the function as nonlinear as possible.

Denote the rows of H_t by $e_0, e_1, \ldots, e_{2^t-1}$. Set $l_0^{'} = (e_1, e_1, e_2, e_3, \ldots, e_{2^t-1})$ which is a balanced sequence of length 2^{2t}. Hence, $\zeta^{'} = (l_0^{'}, l_1, \ldots, l_{2^{2t}-1})$ is a balanced sequence. Now, a lower bound for the nonlinearity of the function f with sequence $\zeta^{'}$ will be given.

Let φ be an affine function in $\mathcal{F}_{4t}$ with sequence ζ_φ which is a row of $\pm H_{4t}$. Since $H_{4t} = H_{2t} \otimes H_{2t}$, we get that $\zeta_\varphi = \pm l_i \otimes l_j$ where l_i and l_j are rows of H_{2t}. Denote $l_i = (h_{i,1}, h_{i,2}, \ldots, h_{i,2^{2t}})$. Then, $\zeta_\varphi = \pm(h_{i,1}.l_j, h_{i,2}.l_j, \ldots, h_{i,2^{2t}}.l_j)$. It follows that $|\langle \zeta^{'}, \zeta_\varphi \rangle| \leq |\langle l_0^{'}, l_j \rangle| + |\langle l_j, l_j \rangle| = |\langle l_0^{'}, l_j \rangle| + 2^{2t}$ since any two distinct rows of H_{2t} are orthogonal.

As $H_{2t} = H_t \otimes H_t$, $l_j = e_k \otimes e_l$ where l_j is a row of H_{2t} and e_k, e_l are rows of H_t. Denote e_k by $(a_1, a_2, \ldots, a_{2^t})$. Then, $l_j = (a_1.e_l, a_2.e_l, \ldots, a_{2^t}.e_l)$. Hence, one obtains that

$$\begin{aligned} |\langle l_0^{'}, l_j \rangle| &\leq |\langle e_1, e_l \rangle| + |\langle e_1, e_l \rangle| + |\langle e_2, e_l \rangle| + \cdots + |\langle e_{2^t-1}, e_l \rangle| \\ &= \begin{cases} 2^{t+1} & \text{if } \ l = 1, \\ 2^t & \text{if } \ l = 2, 3, \ldots, 2^t - 1, \\ 0 & \text{if } \ l = 0. \end{cases} \end{aligned}$$

Thus, $|\langle \zeta^{'}, \zeta_\varphi \rangle| \leq |\langle l_0^{'}, l_j \rangle| + 2^{2t} \leq 2^{t+1} + 2^{2t}$. By using Lemma 2.2.1, one can conclude that $d(f, \varphi) \geq 2^{4t-1} - \frac{1}{2}\langle \zeta^{'}, \zeta_\varphi \rangle \geq 2^{4t-1} - 2^{2t-1} - 2^t$. Since φ is arbitrary, $N_f \geq 2^{4t-1} - 2^{2t-1} - 2^t$.

(ii) Consider this time H_{2t+1}. The concatenation $\zeta = (l_0, l_1, \ldots, l_{2^{2t+1}-1})$ of the rows of H_{2t+1} is a bent sequence. As in part **(i)**, replace l_0 by the balanced sequence $l_0^{'} = (e_{2^t}, e_{2^t+1}, \ldots, e_{2^{t+1}-1})$ where e_i's are the rows of H_{t+1} for $2^t \leq i \leq 2^{t+1} - 1$ with length 2^{t+1}. Also, let $\zeta^{'} = (l_0^{'}, l_1, \ldots, l_{2^{t+1}-1})$.

Let φ be an affine function in $\mathcal{F}_{4t+2}$ with sequence ζ_φ. It is equal to $\zeta_\varphi = \pm l_i \otimes l_j$ where l_i, l_j are rows of H_{2t+1}. Thus, $|\langle \zeta^{'}, \zeta_\varphi \rangle| \leq |\langle l_0^{'}, l_j \rangle| + 2^{2t+1}$.

Since $l_0^{'} = (e_{2^t}, e_{2^t+1}, \ldots, e_{2^{t+1}-1})$ is the sequence of the function

$$g_1(x_2, x_3, \ldots, x_{2t+2})$$

and obtained from the bent sequence $(e_0, e_1, \ldots, e_{2^{t+1}-1})$ by splitting as in Lemma 7.1.5, $|\langle l_0^{'}, l_j \rangle| \leq 2^{t+1}$. Hence, $|\langle \zeta^{'}, \zeta_\varphi \rangle| \leq 2^{t+1} + 2^{2t+1}$ which yields that $N_f \geq 2^{4t+1} - 2^{2t} - 2^t$. $\square$

By using Lemma 7.1.6 and applying it iteratively, the nonlinearity of a balanced function can be further improved as the following theorem suggests.

Theorem 7.1.7 *For any even number $n \geq 4$, there exists a balanced function f with nonlinearity N_f satisfying*

$$N_f \geq \begin{cases} 2^{2^m-1} - \frac{1}{2}(2^{2^{m-1}} + 2^{2^{m-2}} + \cdots + 2^{2^2} + 2^{2^1}) & \text{if} \quad n = 2^m, \\ 2^{2^s(2t+1)-1} - \frac{1}{2}(2^{2^{s-1}(2t+1)} + 2^{2^{s-2}(2t+1)} + \cdots + & \text{if} \quad n = 2^s(2t+1). \\ 2^{2(2t+1)} + 2^{(2t+1)} + 2^{t+1}) & \end{cases}$$

Proof. First of all, observe that an even number $n \geq 4$ can be expressed as $n = 2^m$ for $m \geq 2$ or $n = 2^s(2t+1)$ for $s \geq 1$ and $t \geq 1$.

Case 1 : $n = 2^m$ for $m \geq 2$.

Consider $H_{2^{m-1}}$. The concatenation of the rows of $H_{2^{m-1}}$ is a bent sequence which contains $2^{2^{m-1}}$ sequences of length $2^{2^{m-1}}$. Replace the first all-one sequence with a bent sequence of the same length $2^{2^{m-1}}$. The bent sequence of length $2^{2^{m-1}}$ needed is obtained through the concatenation of the rows of $H_{2^{m-2}}$ which are of length $2^{2^{m-2}}$. The first all-one row of $H_{2^{m-2}}$ appears now in the new sequence. We replace this all-one sequence by a bent sequence of the same length. We continue this process until the length of the all-one leading sequence becomes $2^2 = 4$. Finally, the all-one sequence of length four is replaced with the sequence $(+1, -1, +1, -1)$. By means of all these replacements, the resulting sequence turns out to be a balanced sequence. It can be

proved by induction that the nonlinearity of the final function satisfies

$$N_f \geq 2^{2^{m-1}} - \frac{1}{2}(2^{2^{m-1}} + 2^{2^{m-2}} + \cdots + 2^{2^2} + 2^{2^1}).$$

Case 2 : $n = 2^s(2t+1)$ for $s \geq 1$ and $t \geq 1$.

In this case, the replacing process continues until the length of the leading all-one sequence is 2^{2t+1}. The final leading all-one sequence is replaced by $l_0{}' = (e_{2^t}, e_{2^t+1}, \ldots, e_{2^{t+1}-1})$ where e_i's are the rows of H_{t+1} for $i = 2^t, 2^t + 1, \ldots, 2^{t+1} - 1$. Note that $l_0{}'$ is the second half of the sequence of the bent function obtained by the concatenation of the rows of H_{t+1}. It can be shown by induction that the nonlinearity of the function obtained satisfies

$$N_f \geq 2^{2^s(2t+1)-1} - \frac{1}{2}(2^{2^{s-1}(2t+1)} + 2^{2^{s-2}(2t+1)} + \cdots + 2^{2(2t+1)} + 2^{(2t+1)} + 2^{t+1}).$$

□

Let $\zeta = (\zeta_0, \zeta_1, \ldots, \zeta_{2^k-1})$ be a sequence of length 2^{2k} which is obtained from a bent sequence by modifying the leading all-one sequence as in Theorem 7.1.7. By permuting ζ_i's and by changing their signs for $i = 0, 1, \ldots, 2^k - 1$ as it is done in splitting a bent sequence, one obtains $2^{2^k}.(2^k!)$ different balanced sequences of length 2^{2k} all of which have the same nonlinearity. In fact, note that the final leading sequence ζ_0 has the same structure as the large sequence ζ. Thus, permuting and changing signs can also be applied to ζ_0.

In the following table, entries in the first row are the upper bounds on the nonlinearities of balanced functions in $\mathcal{F}_n$ given by the bound in Corollary 2.5.6 where $n = 4, 6, 8, 10, 12, 14$. Entries in the second row are the lower bounds obtained by Theorem 7.1.7 on the nonlinearities of balanced functions. The third row contains the lower bounds on the nonlinearities of balanced functions due to Lemma 7.1.3.

Vector Spaces	V_4	V_6	V_8	V_{10}	V_{12}	V_{14}
Lemma 2.5.6, $N_f \leq$	4	26	118	494	2014	8126
Theorem 7.1.7, $N_f \geq$	4	26	116	492	2010	8120
Lemma 7.1.3, $N_f \geq$	4	24	112	480	1984	8064

(b) On V_{2k+1} :

The proof of the following lemma can be made easily.

Lemma 7.1.8 *Let f be defined in $\mathcal{F}_{n+m}$ as $f(z) = f_1(x) + f_2(y)$ where $z = (x, y)$ is in V_{n+m} for x in V_n, y in V_m, $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_m)$. Then, the following hold :*

(i) *f is balanced if and only if f_1 or f_2 is balanced.*

(ii) *The sequence ζ_f of f is equal to $\zeta_f = \zeta_{f_1} \otimes \zeta_{f_2}$ where ζ_{f_i} is the sequence of f_i for $i = 1, 2$.*

Lemma 7.1.9 *Let f in $\mathcal{F}_{n+m}$ be defined as in Lemma 7.1.8 from f_1 in $\mathcal{F}_n$ and f_2 in $\mathcal{F}_m$. Assume that $\langle \zeta_{f_1}, l_1 \rangle \leq P_1$ and $\langle \zeta_{f_2}, l_2 \rangle \leq P_2$ hold for any affine sequence l_1, l_2 of length 2^n and 2^m, respectively where P_1, P_2 are positive integers. Then, $N_f \geq 2^{n+m-1} - \frac{1}{2} P_1 P_2$.*

Proof. Let φ be an affine function in $\mathcal{F}_{n+m}$ with sequence ζ_φ. Then, $\zeta_\varphi = \pm l_1 \otimes l_2$ where l_1 is a row of H_n and l_2 is a row of H_m. It follows that, $\langle \zeta_f, \zeta_\varphi \rangle = \langle \zeta_{f_1} \otimes \zeta_{f_2}, \pm l_1 \otimes l_2 \rangle = \pm \langle \zeta_{f_1}, l_1 \rangle \langle \zeta_{f_2}, l_2 \rangle$ by **(3)** of Lemma 2.4.2, giving that $|\langle \zeta_f, \zeta_\varphi \rangle| = |\langle \zeta_{f_1}, l_1 \rangle| . |\langle \zeta_{f_2}, l_2 \rangle| \leq P_1 P_2$. Since φ is arbitrary, Lemma 2.2.1 implies that $N_f \geq 2^{n+m-1} - \frac{1}{2} P_1 P_2$. $\square$

By using Lemma 7.1.9 and a result of [35], one can obtain a function in $\mathcal{F}_{2k+15}$ with nonlinearity greater than all those obtained by concatenating or splitting bent sequences for all $k \geq 7$ as follows :

Let ζ_1 be a balanced sequence of length 2^{2k} which is obtained by using the method in the proof of Theorem 7.1.7 for $k \geq 2$. Let ζ_2 be a sequence of length 2^{15} obtained by [35]. The nonlinearity of the function with sequence ζ_2 is 16276 and there are 13021 such sequences. Denote the functions corresponding to

ζ_1, ζ_2 by f_1, f_2 respectively. Using these two functions in Lemma 7.1.8, one gets a function f in $\mathcal{F}_{2k+15}$ defined as $f(x_1, x_2, \ldots, x_{2k}, x_{2k+1}, \ldots, x_{2k+15}) = f_1(x_1, x_2, \ldots, x_{2k}) + f_2(x_{2k+1}, x_{2k+2}, \ldots, x_{2k+15})$. By Theorem 7.1.7 and Lemma 2.2.1, it is known that

$$\langle \zeta_1, l_1 \rangle \leq \begin{cases} 2^{2^{m-1}} + 2^{2^{m-2}} + \cdots + 2^{2^2} + 2^{2^1} & \text{if } \ 2k = 2^m, \\ 2^{2^{s-1}(2t+1)} + 2^{2^{s-2}(2t+1)} + \ldots + & \text{if } \ 2k = 2^s(2t+1). \\ 2^{2(2t+1)} + 2^{2t+1} + 2^{t+1} & \end{cases}$$

and $\langle \zeta_2, l_2 \rangle \leq 2(2^{14} - 16276) = 216$ where l_1 is an affine sequence of length 2^{2k} and l_2 is an affine sequence of length 2^{15}. Then, the following theorem is obtained by using Lemma 7.1.9.

Theorem 7.1.10 *Let f be a function in $\mathcal{F}_{2k+15}$ defined by f_1 and f_2 as above for $k \geq 2$. Then, the nonlinearity N_f of f satisfies*

$$N_f \geq \begin{cases} 2^{2^m+14} - 108(2^{2^{m-1}} + 2^{2^{m-2}} + \ldots + 2^{2^2} + 2^{2^1}) & \textit{if } \ 2k = 2^m, \\ 2^{2^s(2t+1)+14} - 108(2^{2^{s-1}(2t+1)} + 2^{2^{s-2}(2t+1)} + \cdots + & \textit{if } \ 2k = 2^s(2t+1). \\ 2^{2(2t+1)} + 2^{(2t+1)} + 2^{t+1}) & \end{cases}$$

7.2 Constructing Highly Nonlinear Balanced Functions Satisfying Strict Avalanche Criterion

In this section, methods for constructing highly nonlinear balanced functions satisfying strict avalanche criterion will be presented.

(a) On V_{2k+1} :

Let $n = 2k$ for $k \geq 1$, f be a bent function and g be a nonconstant affine function both of which are in $\mathcal{F}_n$. By **(c)** of Lemma 6.1.1, it is clear that the function $f(x) + g(x)$ is also bent. Thus, the weight of $f + g$ is $2^{n-1} \pm 2^{\frac{n}{2}-1}$ by **(e)** of the same lemma. Since f is bent, the weight of f is also $2^{n-1} \pm 2^{\frac{n}{2}-1}$. It follows that after fixing a bent function f, one can always find an affine function g such that the weight of f is different from the weight of $f + g$. In other words, if $w(f) = w(f + g)$ for any bent function f and for any affine function g, then $\bar{g}$, the complement of g, is the desired function. By using

those properly chosen f and g, define the function h in $\mathcal{F}_{n+1}$ as follows :

$$h(z) = (y+1)f(x_1, x_2, \ldots, x_{2k}) + y(f(x_1, x_2, \ldots, x_{2k}) + g(x_1, x_2, \ldots, x_{2k}))$$

where $z = (y, x)$ is in V_{n+1} for y in $GF(2)$ and $x = (x_1, x_2, \ldots, x_{2k})$ in V_n. Thus, h is the concatenation of the bent functions f and $f + g$. It is easy to see that the function h is in fact equal to

$$h(y, x_1, x_2, \ldots, x_{2k}) = f(x_1, x_2, \ldots, x_{2k}) + yg(x_1, x_2, \ldots, x_{2k}).$$

The properties of this construction are as follows :

Theorem 7.2.1 *The function h in $\mathcal{F}_{n+1}$ defined as above is a balanced function with nonlinearity $N_h \geq 2^{2k} - 2^k$ satisfying strict avalanche criterion. Moreover, the degree of h is equal to the degree of the bent function used in the construction of h. Additionally, the number of vectors in V_{n+1} such that h satisfies the propagation criterion is $2^{2k} + 2^{2k-1}$. In other words, h satisfies the propagation criterion with respect to 75% of all vectors in V_{n+1}.*

Proof. h is balanced since $w(h) = w(f) + w(f + g) = 2^{2k}$.

By using Corollary 7.1.2, $N_h \geq 2^{2k} - 2^k$. h satisfies the strict avalanche criterion since :

Let $\alpha = (u, v_1, v_2, \ldots, v_{2k})$ be a vector in V_{n+1} with $w(\alpha) = 1$. It is enough to show that the directional derivative of h in the direction of α is a balanced function. There are two cases :

(i) $u = 0$. Since $w(\alpha) = 1$, this implies that $w(\beta) = 1$ where $\beta = (v_1, v_2, \ldots, v_{2k})$ in V_n. Then, $h^\alpha(z) = h(z) + h(z + \alpha) = f(x) + f(x + \beta) + y(g(x) + g(x + \beta))$. Since g is affine, $g(x) + g(x + \beta)$ is constant, say equal to θ in $GF(2)$. Thus, $h^\alpha(z) = f^\beta(x) + \theta y$. As f is bent, f^β is balanced by Theorem 6.1.5 since β is nonzero in V_n. By using Lemma 7.1.8, h^α is balanced.

(ii) $u = 1$. Since $w(\alpha) = 1$, this implies that $w(\beta) = 0$, i.e. v_i is zero for all $i = 1, 2, \ldots, 2k$. Then, $h^\alpha(z) = h(z) + h(z + \alpha) = g(x)$. Since g is a nonconstant affine function, by Lemma 2.1.6, g is balanced implying that h^α is balanced.

Thus, h satisfies the strict avalanche criterion.

Since g is a nonconstant affine function, the degree of g is one. By Proposition 6.1.7, the degree of f satisfies $1 \leq deg(f) \leq k$. From the construction of h, the degree of h is equal to the degree of f.

In order to prove the propagation characteristics of h, the number of $\alpha = (u, v_1, v_2, \ldots, v_{2k})$'s in V_{n+1} for which h satisfies the propagation criterion will be counted. Let $\beta = (v_1, v_2, \ldots, v_{2k})$. Now,

$$\begin{aligned} h^\alpha(z) &= h(z) + h(z+\alpha) \\ &= f(x) + f(x+\beta) + y(g(x) + g(x+\beta)) + ug(x+\beta). \end{aligned} \tag{7.2}$$

There are three cases in (7.2). These are **(i)** $u = 0$, $\beta \neq 0$; **(ii)** $u \neq 0$, $\beta = 0$ and **(iii)** $u \neq 0$, $\beta \neq 0$.

(i) $u = 0$, $\beta \neq 0$. Then, $h^\alpha(z) = f^\beta(x) + y\theta$ where $g^\beta(x)$ is constant, say equal to θ in $GF(2)$, since g is affine. As f is bent and $\beta \neq 0$, f^β is balanced. The number of vectors $\alpha = (0, \beta)$ in V_{n+1} where $\beta \neq 0$ and h satisfies the propagation criterion is $2^{2k} - 1$.

(ii) $u \neq 0$, $\beta = 0$. Then, $h^\alpha(z) = g(x)$ giving that $w(h^\alpha) = 2.2^{2k-1}$ since $w(g) = 2^{2k-1}$. Hence, h^α is balanced. In other words, for $\alpha = (1, 0, 0, \ldots, 0)$, the function h^α is balanced.

(iii) $u \neq 0$, $\beta \neq 0$. Now,

$$h^\alpha(z) = f(x) + f(x+\beta) + y(g(x) + g(x+\beta)) + g(x+\beta)$$

where $\alpha = (1, \beta)$ is in V_{n+1} and $\beta = (v_1, v_2, \ldots, v_{2k})$ is in V_n. Since g is affine, $g^\beta(x) = g(x) + g(x+\beta)$ is constant, say equal to θ in $GF(2)$. Hence, $h^\alpha(z) = f(x) + f(x+\beta) + g(x+\beta) + \theta y$. There are two cases :

(1) $\underline{\theta = 1}$: We have $h^\alpha(z) = f(x) + f(x+\beta) + g(x+\beta) + y$. By Lemma 7.1.8, $h^\alpha(z)$ is balanced. There are 2^{2k-1} vectors β in V_n satisfying $g^\beta(x) = 1$.

(2) $\underline{\theta = 0}$: We have $h^\alpha(z) = f^\beta(x) + g(x)$ since $\theta = g^\beta(x) = 0$ implies that $g(x+\beta) = g(x)$. By using (2.4), we get that

$$w(h^\alpha) = |Supp(f^\beta)| + |Supp(g)| - 2|Supp(f^\beta) \cap Supp(g)|.$$

Since f is bent and $\beta \neq 0$, we get that f^β is a balanced function. Hence, $|Supp(f^\beta)| = 2^{2k-1}$. Also, since g is a nonconstant affine function, $|Supp(g)| = 2^{2k-1}$. Thus, $w(h^\alpha) = 2^{2k} - 2|Supp(f^\beta) \cap Supp(g)|$. In order for $h^\alpha(z)$ to be balanced, $Supp(f^\beta) \cap Supp(g)$ must be the empty set. However, this is impossible implying that $h^\alpha(z)$ is not balanced.

Thus, the number of vectors for which h satisfies the propagation criterion is $(2^{2k} - 1) + 1 + 2^{2k-1} = 2^{2k} + 2^{2k-1}$. $\square$

(b) On V_{2k} :

Let $k \geq 2$, $n = 2k-2$ and f be a bent function in $\mathcal{F}_n$. Also let g_1, g_2 and g_3 be three nonconstant affine functions in $\mathcal{F}_n$ such that $g_i + g_j$ is nonconstant for any $1 \leq i < j \leq 3$. It is clear that for $k \geq 2$, such affine functions exist in $\mathcal{F}_n$. It is possible to choose g_1, g_2 and g_3 in such a way that $w(f) = w(f + g_1) = 2^{2k-3} + 2^{k-2}$ and $w(f + g_2) = w(f + g_3) = 2^{2k-3} - 2^{k-2}$ since f and $f + g_i$'s are all bent functions for $i = 1, 2, 3$. By using these functions, define the function h in $\mathcal{F}_{n+2}$ as

$$h(z) = h(y, x) = \sum_{i=0}^{3} \chi_{\alpha_i}(y) h_i(x)$$

where $z = (y, x)$ is in V_{n+2} for $y = (y_1, y_2)$ in V_2 and $x = (x_1, x_2, \ldots, x_{2k-2})$ in V_n. The function χ_{α_i} is the characteristic function of α_i in V_2 and the functions h_i's are defined as $h_0(x) = f(x)$, $h_1(x) = (f + g_1)(x)$, $h_2(x) = (f + g_2)(x)$ and $h_3(x) = (f + g_3)(x)$. Thus, it is clear that the function h is the concatenation of four bent functions each of which differs from another by a suitably chosen affine function in $\mathcal{F}_n$. By using the definitions of h_i's in the algebraic normal form of h, one obtains that h is in fact equal to

$$h(z) = h(y, x) = f(x) + y_2 g_1(x) + y_1 g_2(x) + y_1 y_2 (g_1(x) + g_2(x) + g_3(x))$$

where $z = (y, x)$, $y = (y_1, y_2)$ and $x = (x_1, x_2, \ldots, x_{2k-2})$.

The properties of this construction are as follows :

Theorem 7.2.2 *The function h in $\mathcal{F}_{n+2}$ defined as above is a balanced function with nonlinearity $N_h \geq 2^{2k-1} - 2^k$ satisfying strict avalanche criterion.*

Moreover, the degree of h is equal to the degree of the bent function f used in the construction and the number of vectors in V_{n+2} for which h satisfies the propagation criterion is at least $2^{2k-2}+1$. In other words, h satisfies the propagation criterion with respect to at least 25% of all vectors in V_{n+2}.

Proof. Since h is the concatenation of h_i's for $i = 0, 1, 2, 3$, it follows that $w(h) = \sum_{i=0}^{3} w(h_i)$. However, the functions g_1, g_2, g_3 are chosen according to f in such a way that $w(h_0) = w(h_1) = 2^{2k-3} + 2^{k-2}$ and $w(h_2) = w(h_3) = 2^{2k-3} - 2^{k-2}$ where h_0, h_1, h_2 and h_3 are defined as in the construction above. Hence, $w(f) = 2^{2k-1}$ implying that h is balanced.

By using Lemma 7.1.3, the nonlinearity N_h of h satisfies $N_h \geq 2^{2k-1} - 2^k$. h satisfies the strict avalanche criterion since :

Let $\alpha = (u, t, v_1, v_2, \ldots, v_{2k-2})$ be a vector in V_{n+2} with $w(\alpha) = 1$. There are three cases :

(i) $u = 1$. Then, $t = 0$ and $w(\beta) = 0$ for $\beta = (v_1, v_2, \ldots, v_{2k-2})$ in V_n. Now, $h^\alpha(z) = g_2(x) + y_2(g_1(x) + g_2(x) + g_3(x))$. Equivalently, $h^\alpha(z) = y_2 g_1(x) + (y_2 + 1)g_2(x) + y_2 g_3(x)$. If $y_2 = 0$, then $h^\alpha(z) = g_2(x)$ and if $y_2 = 1$, then $h^\alpha(z) = g_1(x) + g_3(x)$. Hence, $w(h^\alpha) = 2w(g_2) + 2w(g_1 + g_3) = 2^{2k-1}$ since g_2 and $g_1 + g_3$ are nonconstant affine functions by the choices of g_1, g_2 and g_3. This gives that h^α is balanced.

(ii) $t = 1$. Then, $u = 0$ and $w(\beta) = 0$. Now, $h^\alpha(z) = (y_1 + 1)g_1(x) + y_1 g_2(x) + y_1 g_3(x)$. If $y_1 = 0$, then $h^\alpha(z) = g_1(x)$ and if $y_1 = 1$, then $h^\alpha(z) = g_2(x) + g_3(x)$. Similar to part **(i)**, one gets that $w(h^\alpha) = 2^{2k-1}$ giving that h^α is balanced.

(iii) $u = 0, t = 0$. Then, $w(\beta) = 1$. Now, $h^\alpha(z) = h(y_1, y_2, x) = f^\beta(x) + a_1 y_2 + a_2 y_1 + (a_1 + a_2 + a_3) y_1 y_2$ where $a_i = g_i(x) + g_i(x + \beta)$ is a constant in $GF(2)$ since g_i is an affine function for $i = 1, 2, 3$. As $\beta \neq 0$ and f is bent, f^β is balanced. By Lemma 7.1.8, h^α is balanced.

Thus, h satisfies the strict avalanche criterion.

As proved in Theorem 7.2.1 the degree of h is equal to the degree of f.

The proof showing that h satisfies the propagation criterion for at least

$2^{2k-2}+1$ vectors in V_{n+2} is similar to the proof done in detail in Theorem 7.2.1. □

Remark 7.2.3 *An important note is that by using bent functions which have degrees* $2,3,\ldots,k$ *and Theorem 7.2.1, one can obtain new functions having degrees* $2,3,\ldots,k$. *Similarly, using bent functions with degrees* $2,3,\ldots,k$ *in Theorem 7.2.2, one obtains functions with degrees* $2,3,\ldots,k-1$. *Recall that a simple way to obtain bent functions with all possible degrees (*$2,3,\ldots,k$*) is to use Theorem 6.1.15.*

7.3 Constructing Highly Nonlinear Balanced Functions With Good Propagation Characteristics

In this section, methods of constructing highly nonlinear balanced functions with good propagation characteristics will be given. Recall from Theorem 6.1.5 that bent functions are the only class of functions in $\mathcal{F}_{2k}$ to satisfy $PC(2k)$. However, bent functions are not balanced. Thus, if a function is to meet several cryptological properties including balance, nonlinearity and the propagation criterion, then it is clear that no function can satisfy most of these properties completely. Thus, to construct highly nonlinear, balanced functions which do not satisfy $PC(n)$ but satisfy the propagation criterion for almost all vectors in V_n is an important problem in cryptology.

Moreover, it will be shown in this section that there are some functions which satisfy only $PC(0)$ although they satisfy the propagation criterion for almost all vectors in V_n. This is due to the fact that those functions satisfy the propagation criterion for almost all vectors except for some vectors α with weight one. However, recall from Theorem 3.4.2 that for any function f, the balance, the degree, the nonlinearity and the number of vectors for which f satisfies the propagation criterion are invariant under nonsingular affine transformations on the input coordinates. Hence, by a suitable affine transformation, the vectors for which f does not satisfy the propagation crite-

rion can be transformed into vectors with larger weights. This new obtained function satisfies the propagation criterion of a higher degree having the same degree, weight and nonlinearity with the starting function. This is the main technique employed in this section. The two cases are considered separately :

(a) On V_{2k+1} :

Let $n = 2k$ and f be a bent function in $\mathcal{F}_n$. Define the function g in $\mathcal{F}_{n+1}$ by using f as follows :

$$\begin{aligned} g(z) = g(y, x) &= (y+1)f(x) + y\bar{f}(x) \\ &= y + f(x_1, x_2, \ldots, x_{2k}) \end{aligned}$$

where $z = (y, x)$ is in V_{n+1} for y in $GF(2)$ and $x = (x_1, x_2, \ldots, x_{2k})$ in V_n. In other words, g is the concatenation of f and $\bar{f}$. Rewriting the variables $\{y, x_1, x_2, \ldots, x_{2k}\}$ as $\{x_1, x_2, \ldots, x_{2k+1}\}$, g is in fact equal to

$$g(x_1, x_2, \ldots, x_{2k+1}) = x_1 + f(x_2, x_3, \ldots, x_{2k+1}).$$

Theorem 7.3.1 *The function g in $\mathcal{F}_{n+1}$ defined as above is a balanced function with $N_g \geq 2^{2k} - 2^k$ satisfying the propagation criterion with respect to all nonzero vectors α in V_{n+1} except for $\alpha = (1, 0, 0, \ldots, 0)$. Furthermore, by a linear transformation on the input coordinates, the function $g^\star(x) = g^\star(x_1, x_2, \ldots, x_{2k+1}) = g(xA)$ is a balanced function with nonlinearity $N_{g^\star} \geq 2^{2k} - 2^k$ and it satisfies the propagation criterion with respect to all nonzero vectors α in V_{n+1} except for $\alpha = (1, 1, \ldots, 1)$. In other words, $g^\star$ satisfies $PC(n)$.*

Proof. It is easy to see that g is balanced and $N_g \geq 2^{2k} - 2^k$. Moreover, as in the proof of Theorem 7.2.1, it can be shown that g satisfies the propagation criterion with respect to all nonzero vectors α in V_{n+1} except for $\alpha = (1, 1, \ldots, 1)$.

Now, the last part of the theorem will be proved. Let β_1, β_2 be subsets of V_{n+1} such that $\beta_1 = \{\alpha, e_2, \ldots, e_{2k+1}\}$ and $\beta_1 = \{e_1, e_2, \ldots, e_{2k+1}\}$ where e_i is the vector in V_{n+1} whose all coordinates are 0 except the i-th one. It is clear that β_1 and β_2 are bases of V_{n+1} over $GF(2)$. From linear algebra, there exists

a unique linear transformation $\theta : V_{n+1} \longrightarrow V_{n+1}$ such that $\theta(\alpha) = e_1$ and $\theta(e_i) = e_i$ for $i = 2, 3, \ldots, 2k+1$. This linear transformation can be written as $\theta(x) = xA$ where A is nonsingular matrix of order $n+1$ given by

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Now, consider the function $g^\star$ in $\mathcal{F}_{n+1}$ defined as

$$g^\star(x) = g^\star(x_1, x_2, \ldots, x_{2k+1}) = g(xA).$$

$g^\star$ is also a balanced function with $N_{g^\star} \geq 2^{2k} - 2^k$. Explicitly, $g^\star$ is equal to

$$\begin{aligned} g^\star(x_1, x_2, \ldots, x_{2k+1}) &= g(x_1, x_1 + x_2, \ldots, x_1 + x_{2k+1}) \\ &= x_1 + f(x_1 + x_2, x_1 + x_3, \ldots, x_1 + x_{2k+1}). \end{aligned}$$

It is easy to see that $g^\star$ satisfies the propagation criterion with respect to all nonzero vectors α in V_{n+1} except for $\alpha = (1, 1, \ldots, 1)$. In other words, $g^\star$ satisfies $PC(n)$. $\square$

(b) On V_{2k} :

Let $n = 2k - 2$ and f be a bent function in $\mathcal{F}_n$. The function g in $\mathcal{F}_{n+2}$ is defined as

$$g(z) = g(y, x) = \sum_{i=0}^{3} \chi_{\alpha_i}(y) f_i(x)$$

where $z = (y, x)$ is in V_{n+2} for $y = (y_1, y_2)$ in V_2 and $x = (x_1, x_2, \ldots, x_{2k-2})$ in V_n and χ_{α_i} is the characteristic function of α_i in V_2. The functions f_i's are defined as $f_0 = f$, $f_1 = \bar{f}$, $f_2 = \bar{f}$ and $f_3 = f$. Hence, g is the concatenation of four bent functions. Rewriting the variables $\{y_1, y_2, x_1, x_2, \ldots, x_{2k-2}\}$ as $\{x_1, x_2, \ldots, x_{2k}\}$ and simplifying the above summation, g is in fact equal to

$$g(x_1, x_2, \ldots, x_{2k}) = x_1 + x_2 + f(x_3, x_4, \ldots, x_{2k}).$$

Theorem 7.3.2 *The function g in $\mathcal{F}_{n+2}$ defined as above is a balanced function with $N_g \geq 2^{2k-1} - 2^k$ satisfying the propagation criterion with respect to all nonzero vectors α in V_{n+2} except for $\alpha = (1, 0, 0, \ldots, 0)$, $\beta = (0, 1, \ldots, 0)$ and $\alpha + \beta = (1, 1, 0, \ldots, 0)$. Furthermore, by a linear transformation on the input coordinates, the function $g^\star(x) = g^\star(x_1, x_2, \ldots, x_{2k}) = g(xA)$ is a balanced function with nonlinearity $N_{g^\star} \geq 2^{2k-1} - 2^k$ and it satisfies the propagation criterion of degree at most $\frac{4}{3}k$.*

Proof. It is easy to see that g is balanced and $N_g \geq 2^{2k-1} - 2^k$. The propagation characteristics of g can be shown as in the proof of Theorem 7.2.1.

Let β_1, β_2 be two bases of V_{n+2} such that $\beta_1 = \{{\alpha_1}^\star, {\alpha_2}^\star, \delta_3, \ldots, \delta_{2k}\}$ and $\beta_2 = \{e_1, e_2, \gamma_3, \ldots, \gamma_{2k}\}$ where ${\alpha_1}^\star, {\alpha_2}^\star$ are two nonzero, distinct (hence linearly independent) vectors in V_{n+2}, e_i's are as in Theorem 7.3.1 for $i = 1, 2$ and δ_i, γ_i's are arbitrary vectors which make β_1, β_2 bases of V_{n+2} for $i = 3, 4, \ldots, 2k$. Let A denote the matrix of the linear transformation sending β_1 to β_2 in that order. Then, the function $g^\star(x) = g^\star(x_1, x_2, \ldots, x_{2k}) = g(xA)$ satisfies the propagation criterion with respect to all but the vectors ${\alpha_1}^\star, {\alpha_2}^\star$ and ${\alpha_1}^\star + {\alpha_2}^\star$. Choosing ${\alpha_1}^\star$ and ${\alpha_2}^\star$ properly, it will be shown that $g^\star$ satisfies the propagation criterion of degree at most $\frac{4}{3}k$.

Note that $2k$ can be written as $3t + c$ where t is an integer and $c = 0, 1$ or 2. Let ${\alpha_1}^\star = (v_1, v_2, \ldots, v_{3t+c})$ and ${\alpha_2}^\star = (u_1, u_2, \ldots, u_{3t+c})$ where

$$v_i = \begin{cases} 1 & \text{for } i = 1, 2, \ldots, 2t + c_1, \\ 0 & \text{for } i = 2t + c_1 + 1, 2t + c_1 + 2, \ldots, 3t + c. \end{cases}$$

and

$$u_i = \begin{cases} 0 & \text{for } i = 1, 2, \ldots, t + c_1, \\ 1 & \text{for } i = t + c_1 + 1, t + c_1 + 2, \ldots, 3t + c. \end{cases}$$

$$\text{for } c_1 = \begin{cases} 0 & \text{if } c = 1, \\ \frac{c}{2} & \text{otherwise.} \end{cases} \quad \text{and } c_2 = \begin{cases} 1 & \text{if } c = 1, \\ \frac{c}{2} & \text{otherwise.} \end{cases}.$$

Note that $w({\alpha_1}^\star) = 2t + c_1$, $w({\alpha_2}^\star) = 2t + c_2$ and $w({\alpha_1}^\star + {\alpha_2}^\star) = 2t + c$. Since ${\alpha_1}^\star$ has the minimum weight among ${\alpha_1}^\star, {\alpha_2}^\star$ and ${\alpha_1}^\star + {\alpha_2}^\star$, for any nonzero

α in V_{n+2} with $w(\alpha) \leq 2t + c_1 - 1$ it is clear that $\alpha \neq \alpha_1{}^\star, \alpha_2{}^\star, \alpha_1{}^\star + \alpha_2{}^\star$. Hence, $g^\star$ satisfies the propagation criterion of degree $2t + c_1 - 1$. By using the definition of c_1, if $c = 0$ or 1, then $g^\star$ satisfies the propagation criterion of degree $2t - 1$ and if $c = 2$, then $g^\star$ satisfies the propagation criterion of degree $2t$. $\square$

Remark 7.3.3 *Note that the constructions in Theorems 7.2.1 and 7.3.1 differ only in the selection of the affine functions used. Theorem 7.2.1 uses a nonconstant affine function while Theorem 7.3.1 uses the constant function 1. This is also true for the constructions in Theorems 7.2.2 and 7.3.2.*

REFERENCES

[1] C. M. Adams. On immunity against Biham and Shamir's "differential cryptanalysis". *Information Processing Letters* **41** (1992) 77-80.

[2] C. M. Adams and S. E. Tavares. Generating and Counting binary bent sequences. *IEEE Transactions on Information Theory* **IT-36 No.5** (1990) 1170-1173.

[3] C. M. Adams and S. E. Tavares. The use of bent sequences to achieve higher-order strict avalanche criterion. (Technical Report, TR 90-013, Department of Electrical Engineering, Queen's University) (1990).

[4] K. G. Beauchamp. *Applications of Walsh and Related Functions with an Introduction to Sequency Functions.* (Microelectronics and Signal Processing Academic. Academic Press, London, New York, Tokyo, 1984).

[5] C. Carlet. Partially-bent functions. *Designs, Codes and Cryptography* **3** (193) 135-145.

[6] D. Chaum and J. H. Evertse. Cryptanalysis of DES with a reduced number of rounds. *Advances in Cryptology - CRYPTO'85* ed. H. C. Williams (Lecture Notes in Computer Science. Springer-Verlag, 1986) **218** 192-211.

[7] G. D. Cohen, M. G. Karpovsky, Jr. H. F. Mattson, and J. R. Schatz.

Covering Radius-survey and recent results. *IEEE Transactions on Information Theory* **IT-31(3)** (1985) 328-343.

[8] T. W. Cusick. Bounds on the number of functions satisfying the strict avalanche ariterion. *Information Processing Letters* **57** (1996) 261-263.

[9] T. W. Cusick and P. Stanica. Bounds on the number of functions satisfying the strict avalanche criterion. *Information Processing Letters* **60(4)** (1996) 215-219.

[10] J. Daemen, R. Govaerts and J. Vandewalle. Correlation Matrices. *Fast Software Encryption, Second International Workshop* ed. B. Preneel (Springer-Verlag, 1994) 275-285.

[11] J. Daemen, V. Rijmen. *The Design of Rijndael. AES The Advanced Encryption Standard* (Springer-Verlag, 2002).

[12] M. H. Dawson and S. E. Tavares. An expanded set of s-box design criteria based on information theory and its relation to differential-like attacks. *Advances in Cryptology - EUROCRYPT'91*

[13] J. Detombe and S. Tavares. Constructing large cryptographically strong S-boxes. *AUSCRYPT'92* (1992).

[14] J. F. Dillon. A survey of bent functions. *The NSA Technical Journal*, unclassified, (1972) 191-215.

[15] J. F. Dillon. Elementary Hadamard difference sets. *Proceedings of the Sixth Southeastern Conference on Combinatorics, Graph Theory and*

Computing, Boca Raton, Florida (1975) 237-249; *Congressus Numerantium No. XIV, Utilitas Math., Winnipeg, Manitoba* (1975).

[16] H. Feistel. Cryptography and Computer Privacy. *Scientific American* **Vol 228 No 5, 15** (1973).

[17] R. Forré. The strict avalanche criterion: Spectral properties of boolean functions and extended definition. *Advances in Cryptology - CRYPTO'88* (Lecture Notes in Computer Science. Springer-Verlag, Berlin, Heidelberg, New York, 1989) **403** 450-468.

[18] J. Hadamard. Resolution d'une question relative aux determinants. *Bull. des Sci. Math.* **17** (1893) 240-246.

[19] K. Hoffman and R. Kunze. 1971. *Linear Algebra*. Prentice Hall, New Jersey.

[20] J. B. Kam, G. I. Davida. Structured design of substitution permutation encryption networks. *IEEE Transactions on Computers* **Vol 28 No 10, 747** (1979).

[21] P. V. Kumar, R. A. Scholtz and L. R. Welch. Generalized bent functions and their properties. *J. Combinatorial Theory* **A 40** (1985) 90-107.

[22] A. Lempel and M. Cohn. Maximal families of bent sequences. *IEEE Trans. Inform. Theory* **IT-28 No.6** (1982) 865-868.

[23] V. V. Losev. Decoding of sequences of bent functions by means of a fast Hadamard transform. *Radiotechnika i elektronika* **7** (1987) 1479-1492.

[24] R. Lidl and H. Niederreiter. 1983. *Finite Fields in Encyclopedia of Mathematics*. **Vol 20** Cambridge University Press: Cambridge.

[25] F. J. MacWilliams and N. J. A. Sloane. 1978. *The Theory of Error-Correcting Codes*. Amsterdam, New York, Oxford: North-Holland.

[26] H. B. Mann. 1965. *Addition theorems*. John Wiley and Sons, New York.

[27] M. Matsui. Linear cryptanalysis method for DES cipher. *Advances in Cryptology - EUROCRYPT'93* (Lecture Notes In Computer Science. Springer-Verlag, Berlin, Heidelberg, New York, 1991) **765** 386-397.

[28] R. L. McFarland. A family of difference sets in non-cyclic groups. *J. Combinatorial Theory* **A 15** (1973) 1-10.

[29] W. Meier and O. Staffelbach. Nonlinearity criteria for cryptographic functions. *Advances in Cryptology - EUROCRYPT'89* (Lecture Notes in Computer Science, Springer-Verlag, 1990) **434** 549-562.

[30] K. Nyberg. Constructions of bent functions and difference sets. *Proceedings of Eurocrypt'90* (Springer-Verlag, 1991) 151-160.

[31] K. Nyberg. Perfect nonlinear S-boxes. *Advances in Cryptology - EUROCRYPT'91* (Lecture Notes in Computer Science, Springer-Verlag,1991) **547** 378-386.

[32] L. O'Connor. An upper bound on the number of functions satisfying the strict avalanche criterion. *Information Processing Letters* **52** (1994) 325-327.

[33] L. O'Connor, A. Klapper. Algebraic nonlinearity and its applications to cryptography. *Journal of Cryptology* **7(4)** (1994) 213-228.

[34] J. D. Olsen, R. A. Scholtz, and L. R. Welch. Bent-function sequences. *IEEE Trans. Inform. Theory* **IT-28 No.6** (1982) 858-864.

[35] N. J. Patterson, D. H. Wiedemann. The covering radius of the $(2^{15}, 16)$ Reed-Muller code is at least 16276. *IEEE Transactions on Information Theory* **IT-28 No. 6** (1983) 858-864.

[36] J. Pieprzyk and G. Finkelstein. Towards effective nonlinear cryptosystem design. *IEE Proceedings (Part E)* **135** (1988) 325-335.

[37] B. Preneel, R. Govaerts, and J.Vandewalle. Boolean functions satisfying higher order propagation criteria. *Advances in Cryptology - EUROCRYPT'91* (Lecture Notes in Computer Science. Springer-Verlag, Berlin, Heidelberg, New York, 1991) **547** 141-152.

[38] B. Preneel, W. V. Leekwijck, L. V. Linden, R.Govaerts, and J.Vandewalle. Propagation characteristics of boolean functions. *In Advances in Cryptology - EUROCRYPT'90* (Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, 1991) **437** 155-165.

[39] O. S. Rothaus. On "bent" functions. *Journal of Combinatorial Theory* **Ser. A, 20** (1976) 300-305.

[40] J. Seberry and M. Yamada. Hadamard Matrices, sequences and block

designs. *Contemporary Design Theory : A Collection of Surveys* ed. J. H. Dinitz and D. R. Stinson (Wiley, New York, 1992) 431-560.

[41] J. Seberry, X. M. Zhang. Highly nonlinear 0-1 balanced boolean functions satisfying strict avalanche criterion. *Advances in Cryptology - AUSCRYPT'92* (Lecture Notes in Computer Science. Springer-Verlag, Berlin, Heidelberg, New York, 1993) **718** 145-155.

[42] J. Seberry, X. M. Zhang. On group of bent functions. *Australasian Journal of Combinatorics* (1993).

[43] J. Seberry, X. M. Zhang, and Y. Zheng. Improving the strict avalanche characteristics of cryptographic functions. *Information Processing Letters* **50** (1994) 37-41.

[44] J. Seberry, X. M. Zhang, and Y. Zheng. Nonlinearity and propagation characteristics of balanced boolean functions. *Information and Computation* **119(1)** (1995) 1-13.

[45] J. Seberry, X. M. Zhang, and Y. Zheng. Nonlinearly balanced boolean functions and their propagation characteristics. *Advances in Cryptology - CRYPTO'93* (Lecture Notes in Computer Science. Springer-Verlag, Berlin, Heidelberg, New York, 1994) **773** 49-60.

[46] J. Seberry, X. M. Zhang, and Y. Zheng. Relationships among nonlinearity criteria. *Advances in Cryptology - EUROCRYPT'94* (Lecture Notes in

Computer Science. Springer-Verlag, Berlin, Heidelberg, New York, 1995) **950** 376-388.

[47] J. Seberry, X. M. Zhang, and Y. Zheng. Systematic generation of cryptographically robust s-boxes. *Proceedings of the First ACM Conference on Computer and Communications Security* (The Association for Computing Machinery, New York, 1993) 172-182.

[48] J. Seberry, X. M. Zhang, and Y. Zheng. The relationship between propagation characteristics and nonlinearity of cryptographic functions. *Journal of Universal Computer Science* **1(2)** (1995) 136-150.

[49] T. Siegenthaler. Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information Theory* **IT-30 No. 5** (1984) 776-779.

[50] J. J. Sylvester. Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tesselated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers. *Phil. Mag.* **34** (1867) 461-475.

[51] S. A. Vanstone, P. C. Van Oorschot. 1989. *An Introduction to Error Correcting Codes with Applications.* Kluwer International Series in Engineering and Computer Science, 71.

[52] R. Yarlagadda and J. E. Hershey. Analysis and synthesis of bent sequences. *IEE Proceeding (Part E)* **136** (1989) 112-123.

[53] A. M. Youssef, T. W. Cusick, P. Stanica and S. E. Tavares. New bounds on the number of functions satisfying the strict avalanche criterion. *Selected Areas in Cryptography'96* (Kingston, Ontario, Canada) 49-56.

[54] A. M. Youssef and S. E. Tavares. Comment on "Bounds on the number of functions satisfying the Strict Avalanche Criterion". *Information Processing Letters* **60(5)** (1996) 271-275.

[55] A. F. Webster and S. E. Tavares. On the design of S-boxes. *Advances in Cryptology - CRYPTO'85* ed. H.C. Williams (Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, New York, 1986) **218** 523-534.

[56] 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(1/2)** (1996) 111-134.

[57] X. M. Zhang and Y. Zheng. GAC- the criterion for global avalanche characteristics of cryptographic functions. *Journal of Universal Computer science* **1(5)** (1995) 316-333.

[58] X. M. Zhang and Y. Zheng. New Bounds on the nonlinearity of boolean functions. *Advances in Cryptography - EUROCRYPT'96* (Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, 1996) **1070** 294-306.