Algebraic Graph Theoretic Applications To Cryptography

by

Sonwabile T Mafunda

December 2015

Submitted in fulfillment of the academic requirements for the degree of Master of Science in Mathematics in the School of Mathematics, Statistics and Computer Sciences, College of Agriculture, Engineering and Science, University of KwaZulu-Natal, RSA.

Supervisors:

Dr. Gareth Amery- University of KwaZulu-Natal (UKZN)

Professor Simon Mukwembi- University of Zimbabwe (UKZN & UZ)

Dr. Christine Swart- University of Cape Town (UCT)

Preface and Declaration

The research described in this dissertation was carried out in the School of Mathematcs, Statistics and Computer Sciences, University of KwaZulu-Natal, under the supervision of Dr Gareth Amery (UKZN), Professor Simon Mukwembi (UKZN and UZ) and Dr Christine Swart (UCT).
These studies represent original work by the author and have not otherwise been submitted in any form for any degree or diploma to any tertiary institution. Where use has been made of the work of others it is duly acknowledged in the text.

Signed .....
Author: Sonwabile Templeton Mafunda
December 2015

As the candidate's supervisor(s), I have approved/ disapproved this dissertation for submission.

Signed: ..... Name: .....

December 2015

Declaration

I, Sonwabile T. Mafunda, declare that

(i) The research reported in this thesis, except where otherwise indicated, is my research.

(ii) This thesis has not been submitted for any degree or examination at any other university.

(iii) This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged or referenced as being sourced from other persons.

(iv) This thesis does not contain any other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

  (a) their words have been rewritten but the general information attributed to them has been referenced;

  (b) where their exact words have been used, then they oath to have been referenced.

(v) Where I have reproduced a publication of which I am author, co-author or editor, I have indicated in detail which part of the publication was actually written by myself alone and have fully referenced such publications.

Signed.....

Dedication

First and foremost, to Our Beloved Father who art in Heaven, Creator of Heaven, Earth and all that dwells in it;

and

To my Beloved Parents, Mr B. V Mafunda and Mrs T. K Mafunda

And my Late Grandmother, Mrs B. E Gumede.

(May Our Lord bless them).

Acknowledgements

Allow me to pass my deepest gratitude and votes of appreciation to the One and only, Our Father who art in Heaven, thy Creator of Heaven, Earth and all that dwells in it. Without His authority and love none of this would be a success, Thank You Father.

It is with great appreciation that I thank my supervisors; Professor Simon Mukwembi, Dr. Gareth Amery and Dr. Christine Swart. Their motivation, hard work and effort has not gone unnoticed. Dr. Amery helped me visualize and imagine problems before attempting them, and has instilled in me the willingness to try and never give up; he has been there for me in every step of this research. Professor Mukwembi has always been there for me from the planning of this journey of my MSc, and whenever Dr. Amery called for his input. I truly appreciate Professor Mukwembi's input and advice throughout the year. Thanks also to Dr. Swart at UCT, for allowing me the opportunity to work under her co-supervision.

To the mother of graph theory in South Africa, Professor Henda Swart, for her support and advice, for paving my way and clarifying unclear concepts, and for being my true outside family; my heartfelt thanks.

These extend also to Professor Bernardo Rodrigues who was also willing to help clarify unclear algebraic concepts, and to Mr. Tendai Shumba and Mr. Shalin Singh; great aspiring; dedicated and capable mathematicians and friends who helped me visualize mathematics, supported and respected my study, engaged in understanding my work, and tackled ideas with me. To the admin staff of the School of Mathematics, Statistics and Computer Science and that of the College of Agriculture, Engineering and Science for

all their help.

To my beloved parents; they are indeed a precious gift. They have always supported and inspired me throughout my journey of life, and have stood by me in any educational decision I took. To my family at large.

I also thank the NRF and Dr Gareth Amery for providing me with financial assistance.

Last but not the least, to everyone I mistakenly forgotten but ought to have mentioned, thank you.

Abstract

This dissertation represents a partial review of the literature pertaining to the relationship between algebraic graph theory and cryptography. This requires a preliminary discussion of elementary graph theory, group theory and cryptography. We then focus on the relevant elements of graph theory, namely Cayley graphs and strongly regular graphs; and of cryptography, namely the Boolean and bent functions, which are, respectively, applicable to pseudo-random generation in stream ciphers, and substitution boxes in block ciphers.

In particular, we construct a Cayley graph associated with a Boolean function, $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}) = (V, E)$, where $(\mathbb{F}_2^n, \oplus)$ is assumed to be the group from which the graph is constructed, $\Omega_{wt(f)}$ the Cayley set contained in $\mathbb{F}_2^n$, and $wt(f)$ the Hamming weight of the Boolean function f. Depending on the value of n, we consider two cases, constructing for each an associated Cayley graph.

If n is not necessarily even, then we consider the resulting Cayley graph, $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}) = (V, E)$, and study its properties, and evaluate some cryptographic properties of the associated Boolean function, and hence of the stream cipher. This is possible because the Boolean function acts as a pseudo-random number generator. Since the security of a stream cipher lies in designing strong pseudo-random number generators, it is important to evaluate its properties. This study investigates the cipher attack resistance ability through studying the associated graph. We find that obtaining the regularity of the associated graph is the same as obtaining the Hamming weight of the pseudo-random number generator. Hence, if we know n, we can easily tell whether the cipher stands a chance of resisting statistical dependence as an attack.

Similarly, if n is even and f attains maximum nonlinearity, so that $nl(f) = 2^{n-1} - 2^{\frac{n}{2}-1}$, then by prescribing our resulting Cayley graph to be strongly regular, we can investigate the properties of this graph and evaluate some cryptographic properties of the bent functions and hence of the block cipher. This is because the set of bent functions acts as a substitution box, $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$ for block ciphers. The spectral information of this graph tells us about the Hamming weight of the bent function.

We conclude with a brief discussion of possible future work.

Contents

List of Tables

List of Figures

Symbols and Notations

$\bar{\mathbb{G}}$	Complement of $\mathbb{G}$;
A^T	Transpose of A;
$\oplus$	XOR (exclusive-OR);
G, $\mathbb{G}$	A Group and a Graph respectively;
$\mathcal{G}$	Cayley Graph;
$\mathcal{G}_f$	Cayley graph associated with a Boolean function f;
$N_{\mathbb{G}}(u)$	Neighborhood of $u \in V(\mathbb{G})$;
$\Omega_{wt(f)}$	Cayley set of Cayley graph associated with a Boolean function;
$V(\mathcal{G}_f)$, $E(\mathcal{G}_f)$	Vertex and Edge sets of G_f respectively;
$\mathbf{b(i)}$	Element in the i^{th} row of the adjacency matrix, $0 \leq i \leq 1$;
B_n	Set of n-variable Boolean functions;
BB_n	Set of n-variable bent Boolean functions ;
$\mathbb{F}_2^n$	The $\{0,1\}$ field of n tuples, where $\mathbb{F}_2 = \{0,1\}, n \in \mathbb{N}$;
ASCII	American Standard Code for Information Interchange;
PRNG	Pseudo-Random Number Generator;
S-Box	Substitution Box;
DES	Data Encryption Standards;
AES	Advanced Encryption Standards;
LFSR	Linear Feedback Shift Register;
SAC	Strict Avalanche Criterion;
PC	Propagation Criteria;
SRCG	Strongly Regular Cayley Graph.

Introduction

Cryptography is a very broad research area, as are algebra and graph theory. The particular focus of this dissertation lies at an interface between algebraic graph theory and cryptography. We shall concern ourselves with the question of what information can be obtained from algebraic graphs about the security of certain symmetric cryptosystems.

Graph theory has became a very useful tool in solving mathematically modeled problems, for example network problems, road traffic and ecosystems. In this dissertation we discuss the impact graph theory has on cryptography by considering Cayley graphs of both general form and strongly regular, and studying their properties to elucidate the possible relationship they may have with Boolean functions of both general and bent forms.

Symmetric (private-key) cryptography is a branch of cryptography regarded as not being as strong as asymmetric (public-key) cryptography in terms of security. However, it is widely employed due to the speed and cost saving associated with it compared to asymmetric cryptosystems. For this reason it has became the concern of cryptographers to employ all possible measures to try to maximize the security of private-key ciphers.

This dissertation focuses on two major private-key ciphers, namely: stream and block ciphers, which rely deeply on pseudo-random number generators and substitution boxes respectively, for their security. These ciphers include, the RC4 (Rivest Cipher 4) known to be the most popular stream cipher in the world. It is used to protect much of SSL (secure sockets layer) traffic today, probably summing up to billions of TLS (transport layer security) connections every day. SSL establishes an encrypted link between a server

and a client, eg: a mail server and a mail client. Block ciphers are more often used than stream ciphers for encrypting Internet communications, stream ciphers are more often used when computational resources are constrained; for example, cellphones. Even though stream ciphers encrypt more efficiently than block ciphers since software-optimized stream ciphers need fewer processor instructions and hardware optimized stream ciphers use fewer gates (or smaller chip area), modern block ciphers such as the advanced encryption standards (AES) are also very efficient.

We bring to the attention of the reader an idea of evaluating the strength of the cipher through the knowledge of algebraic graph theory. The study is initiated by recalling some useful elementary linear and abstract algebra and group theory. This is later used to define and study Cayley graphs which are group constructed graphs. In a similar way basic notions of graph theory and cryptography are introduced. Properties and results from Cayley graphs are discussed in the second chapter. These include, their association with circulant graphs, and vertex transitive graphs. Another category of graphs discussed are strongly regular graphs. Strongly regular graphs tend to possess many spectral properties that become useful in elucidating the link between algebraic graph theory and cryptography, in terms of the relationship between their parameters (n, r, λ, μ) and the associated eigenvalues.

By considering some well known cipher attacks, we study and discuss (in the third chapter) the importance of designing cryptographically strong pseudo-random number generators and substitution boxes. The design of these security providers is aligned to some cryptographical requirements drawn from some well known ciphers. The ability to investigate some of these requirements by only studying Cayley graphs constructed from Boolean functions, is the main objective of this dissertation, and is discussed in the last chapter with some theoretical examples illustrating the results. We consider a Cayley graph and associate it with a Boolean function to construct a graph that brings together properties of Cayley graphs and Boolean functions. It is noticed that this graph gives information about the strength of the designed stream cipher to resist well known attacks. It is noted that during this design process there are major trade offs between properties according to levels of importance. Similarly, strongly regular Cayley graphs are associated to bent

Boolean functions to study the strength of a block cipher to resist against some possible well known attacks.

We conclude with a brief summary and outline of possible avenues for further research.

Chapter 1

Mathematical Prerequisites

We begin by considering very basic definitions, properties and results that shall be useful in this dissertation. Accordingly, we discuss elementary graph theory (Section 1.1), elementary group theory (Section 1.2), and elementary cryptography (Section 1.3).

1.1 Elementary Group Theory

This dissertation will deal with algebraic graph theory, and in particular, graphs described in terms of groups. We therefore begin with a review of elementary group theory. This material is covered by the following references: [14], [24], [34], [41] and [44].

Definition 1.1.1. Let G be a non-empty set of elements. An operation that combines any two elements to form a third element is called a binary operation. Then $(G, \circ)$, (where $\circ$ is a binary operation), is called a ***group*** if the following axioms are met:

G1: for any $a, b \in G$ *closure* is preserved:

$$a \circ b \in G;$$

G2: for any $a, b, c \in G$ the *associative law* holds:

$$(a \circ b) \circ c = a \circ (b \circ c);$$

G3: there exists an *identity element* $e \in G$, such that,

$$a \circ e = a = e \circ a, \quad \text{for all } a \in G;$$

G4: for each element $a \in G$, there exists $a' \in G$, an *inverse element* of a, such that,

$$a \circ a' = e = a' \circ a.$$

Moreover, $(G, \circ)$ is an *Abelian group* if in addition to the group axioms G is commutative:

G5: for any $a, b \in G, a \circ b = b \circ a$.

Definition 1.1.2. Let $(G, \circ)$ and $(H, *)$ be groups. A ***group homomorphism*** $\varphi : G \to H$ is a map such that, $\forall\, g_1, g_2 \in \mathrm{G}, \varphi(g_1 \circ g_2) = \varphi(g_1) * \varphi(g_2)$.

Definition 1.1.3. Let $(G, \circ)$ be a group and $GL(n, \mathbb{C})$ a group of $n \times n$ invertible matrices with entries $x \in \mathbb{C}$. Then a homomorphism $\varphi : G \to GL(n, \mathbb{C})$ is called a ***representation*** of G.

Moreover, if φ is a representation, defined above, then $\chi_\varphi : G \to \mathbb{C}$ is defined by $\chi_\varphi(g) = Tr(\varphi(g))$, where $g \in (G, \circ)$, is said to be the ***character*** of G [24]. Here, $Tr(\varphi(g))$ denotes the ***trace*** of the matrix defined by φ above; the sum of the elements in the main diagonal.

Definition 1.1.4. If, in addition to Definition 1.1.2, φ is a bijection, then φ is a ***group isomorphism*** denoted $(G, \circ) \cong (H, *)$.

Definition 1.1.5. Let $(G, \circ)$ be a group. A group isomorphism from $(G, \circ)$ to $(G, \circ)$, i.e., $\varphi : G \to G$ is called an ***automorphism***.

Moreover, the set of all automorphisms of a group $(G, \circ)$ is itself a group and is called an ***automorphism group*** and is denoted $Aut(G)$.

Definition 1.1.6. Let $(G, \circ)$ be a group and Ω be a non-empty set. If $\exists$ a map $\cdot : G \times \Omega \to \Omega$ defined by $\cdot(g, \omega) = g \cdot \omega = g(\omega)$, for each $(g, \omega) \in G \times \Omega$ then $\cdot$ is called a ***group action*** (in fact a ***left group action***) on Ω relative to $(G, \circ)$ if and only if the following are true:

(i) $\forall\, \omega \in \Omega$, $e_G \cdot \omega = \omega$;

(ii) $\forall g_1, g_2 \in G, \omega \in \Omega,\ g_1 \cdot (g_2 \cdot \omega) = (g_1 \circ g_2) \cdot \omega$.

Definition 1.1.7. Let Ω be a non-empty set. Then a bijection from Ω to Ω is called a ***permutation*** on Ω.

Let $\mathbf{I}$ be an $n \times n$ identity matrix. Then the matrix obtained by permuting the rows or columns of $\mathbf{I}$ is called the $n \times n$ ***permutation matrix***. Each $n \times n$ identity matrix has $n!$ permutation matrices.

Moreover, let $(G, \circ)$ be a group and σ a permutation on a set (Ω, say). Let the elements of $(G, \circ)$ be the permutations σ_i, for any i. Then $(G, \circ)$ is a ***permutation group*** if composition is the group binary operation and all the group axioms are met.

Definition 1.1.8. The ***symmetric group*** is a permutation group. It is a group of bijections of n elements to itself. That is, for a finite set Ω (say of order n) then the symmetric group formed by the set of all permutations of Ω under the binary operation "composition" is denoted by S_n.

Consider the relationship given by the following theorem. The proof of the theorem is given in the referenced material.

Theorem 1.1.1. *[34] [**Cayley's Theorem**] Let $(G, \circ)$ be a group. Then G is isomorphic to a subgroup of the symmetric group. For a finite group of order n, the group is isomorphic to a subgroup of S_n.*

Definition 1.1.9. Let $(G, \circ)$ be a group, Ω be a non-empty set and $\cdot$ a group action on Ω relative to $(G, \circ)$. Then the ***orbit of*** $\omega \in \Omega$ under the action $\cdot$ is the set denoted by $G \cdot \omega = \{g \cdot \omega \mid g \in G\}$.

Moreover the subgroup denoted by $G_\omega = \{g \in G \mid g \cdot \omega = \omega\}$ is called the ***stabilizer*** of ω in G.

Definition 1.1.10. Let $(G, \circ)$ be a group and Ω be a non-empty set and $\cdot$ a group action on Ω relative to $(G, \circ)$. Then $(G, \circ)$ is said to be ***transitive, (act transitively)***, if and only if it has only one orbit; i.e $G \cdot \omega = \Omega$, if and only if $|\{G \cdot \omega \mid \omega \in \Omega\}| = 1, \forall \omega \in \Omega$. Moreover, "$\cdot$" is called a ***transitive left action*** on Ω relative to $(G, \circ)$.

Definition 1.1.11. Let $(G, \circ)$ be a permutation group and Ω be a non-empty set and $\cdot$ a group action on Ω relative to $(G, \circ)$. Then $(G, \circ)$ is called ***semi-regular*** if the the stabilizer of ω in $(G, \circ)$ is only the identity element.

Moreover, $(G, \circ)$ is said to be ***regular*** if the following are true [44]:

(i) $(G, \circ)$ is semi-regular;

(ii) $(G, \circ)$ is transitive.

Remark 1.1.1. *$(G \cdot \omega) \cap (G \cdot \psi) \neq \phi \Rightarrow G \cdot \omega = G \cdot \psi$, for $\omega \in \Omega$ and $\psi \in \Psi$. Also, for all $\omega \in \Omega$, $\omega \in (G \cdot \omega)$.*

1.2 Elementary Graph Theory

In this section we describe the fundamentals of elementary graph theory. This material is drawn from [8], [11], [17], [19], [20], [27], [28], [39] and [41].

Definition 1.2.1. Let $\mathbb{G}$ be a finite non-empty set of elements (objects) called *vertices*, together with a (possibly empty) set of unordered pairs (lines joining two distinct vertices) of distinct vertices, called *edges*. Then $\mathbb{G} = (V, E)$, (where, V denotes the set of vertices and E the set of edges), is called a ***graph***.

Definition 1.2.2. Let $\mathbb{G}$ and $\mathbb{H}$ be graphs. Then for $\mathbb{G} = (V_1, E_1)$ and $\mathbb{H} = (V_2, E_2)$, a ***graph homomorphism*** $\varphi : \mathbb{G} \to \mathbb{H}$ is the map $\varphi : V_1 \to V_2$ such that u, $v \in V_1$ and $uv \in E_1$, implies $\varphi(u)\varphi(u) \in E_2$.

Definition 1.2.3. If, in addition to Definition 1.2.2, φ is an injection (i.e if $\exists$ a one-to-one correspondence between V_1 and V_2 irrespective of the geometric appearance/naming of the vertices and such that every and only edges in graph $\mathbb{G}$ have counterparts in graph $\mathbb{H}$) then φ is a ***graph isomorphism*** denoted $\mathbb{G} \cong \mathbb{H}$.

Moreover, if $V_1 = V_2$ and $E_1 = E_2$ then graphs $\mathbb{G}$ and $\mathbb{H}$ are said to be ***identical*** denoted by $\mathbb{G} = \mathbb{H}$.

Definition 1.2.4. Let $\mathbb{G}$ be a graph, a graph isomorphism from $\mathbb{G}$ to $\mathbb{G}$, i.e $\varphi : \mathbb{G} \to \mathbb{G}$ is called an ***automorphism of a graph***, and clearly $\varphi : V_1 \to V_1$ for V_1 the set of vertices of $\mathbb{G}$. Therefore, each automorphism of a graph is a permutation on the set V_1. However, a permutation on V_1 is not necessarily an automorphism.

The set of all automorphisms of a graph $\mathbb{G}$ is the ***automorphism group of the graph*** $\mathbb{G}$ denoted $Aut(\mathbb{G})$, if composition is the group binary operation and all the group axioms are met.

Definition 1.2.5. Let $\mathbb{G} = (V, E)$ be a graph. $\mathbb{G}$ is called ***vertex transitive*** if and only if, for each pair of vertices u and v belonging to V, $\exists\ \varphi \in Aut(\mathbb{G})$ such that $\varphi(u) = v$.
Similarly $\mathbb{G}$ is called ***edge transitive*** if and only if, for each pair of edges uv and wx belonging to E, $\exists\ \varphi \in Aut(\mathbb{G})$ such that $\varphi(uv) = wx$.

Lemma 1.2.1. *Let $\mathbb{G} = (V, E)$ be a graph. Define the map $\cdot : Aut(\mathbb{G}) \times V \to V$ by $\cdot(\varphi, v) = \varphi(v)$, for each $(\varphi, v) \in Aut(\mathbb{G}) \times V$. Then $\cdot$ is a left group action on V relative to $Aut(\mathbb{G})$.*

Proof. Let $v \in V$ be arbitrary and let e be the group identity of $Aut(\mathbb{G})$. Then e is the function on V:

$$e \cdot v = e(v).$$

Let $x,\ y \in Aut(\mathbb{G})$ and let $u \in V$. Then

$$x \cdot (y \cdot u) = x \cdot (y(u)) = x(y(u)),$$

and

$$(x \circ y) \cdot u = (x \circ y)(u) = x(y(u)).$$

Therefore

$$x \cdot (y \cdot u) = (x \circ y) \cdot u.$$

Hence, $\cdot$ is a left group action on V relative to $Aut(\mathbb{G})$. □

Proposition 1.2.2. *Let $\mathbb{G} = (V, E)$ be a graph. Then the following are equivalent:*

(1) $\mathbb{G}$ is vertex transitive;

(2) The map $\cdot$ defined in Lemma 1.2.1 acts transitively on V relative to $Aut(\mathbb{G})$.

Proof. (1) $\Rightarrow$ (2) Assume that $\mathbb{G}$ is a vertex transitive graph.

Recall that we defined $\cdot : Aut(\mathbb{G}) \times V \to V$ by $\cdot(\varphi, v) = \varphi(v)$, $\forall\ (\varphi, v) \in Aut(\mathbb{G}) \times V$.

Then, from Lemma 1.2.1, $\cdot$ is a left group action on V relative to $Aut(\mathbb{G})$.

Let $u \in V$ be arbitrary. We will now show that

$$\begin{aligned} \{Aut(\mathbb{G}) \cdot v \mid v \in V\} = &\quad \{\varphi \cdot v \mid \varphi \in Aut(\mathbb{G}),\ v \in V\} \\ = &\quad \{\varphi \cdot u \mid \varphi \in Aut(\mathbb{G})\} \\ = &\quad \{Aut(\mathbb{G}) \cdot u\}. \end{aligned}$$

Clearly $\{Aut(\mathbb{G}) \cdot u\} \subseteq \{Aut(\mathbb{G}) \cdot v \mid v \in V\}$.

Let $t \in \{Aut(\mathbb{G}) \cdot v \mid v \in V\}$ be arbitrary. Then $t \in Aut(\mathbb{G}) \cdot v$ for some $v \in V$.

Since $\mathbb{G}$ is vertex transitive, $\exists \varphi \in Aut(\mathbb{G})$ such that $\varphi(v) = u$.

Hence, $\varphi \cdot v = u$, and $\varphi \cdot v \in Aut(\mathbb{G}) \cdot v \Rightarrow u \in Aut(\mathbb{G}) \cdot v$. Also, $u \in Aut(\mathbb{G}) \cdot u \Rightarrow (Aut(\mathbb{G}) \cdot v) \cap (Aut(\mathbb{G}) \cdot u) \neq \phi$ and $Aut(\mathbb{G}) \cdot v = Aut(\mathbb{G}) \cdot u$.

$$\begin{aligned} \text{Hence, } t \in Aut(\mathbb{G}) \cdot u &\quad \Rightarrow t \in \{Aut(\mathbb{G}) \cdot u\} \\ &\quad \Rightarrow \{Aut(\mathbb{G}) \cdot v \mid v \in V\} \subseteq \{Aut(\mathbb{G}) \cdot u\}. \end{aligned}$$

Hence, $\{Aut(\mathbb{G}) \cdot v \mid v \in V\} = \{Aut(\mathbb{G}) \cdot u\}$.

However, $|\{Aut(\mathbb{G}) \cdot u\}| = 1$, so $|\{Aut(\mathbb{G}) \cdot v \mid v \in V\}| = 1$.

(1) $\Leftarrow$ (2) Assume condition 2 is true.

Let $u,\ v \in V$. $\{Aut(\mathbb{G}) \cdot u\}, \{Aut(\mathbb{G}) \cdot v\} \subseteq \{Aut(\mathbb{G}) \cdot y \mid y \in V\}$. Since $\cdot$ acts transitively on V relative to $Aut(\mathbb{G})$,

$$|\{Aut(\mathbb{G}) \cdot y \mid y \in V\}| = 1.$$

Hence $Aut(\mathbb{G}) \cdot u = Aut(\mathbb{G}) \cdot v$. Now $v \in (Aut(\mathbb{G}) \cdot v)$
$\Rightarrow v \in (Aut(\mathbb{G}) \cdot u)$, so $\exists \varphi \in Aut(\mathbb{G})$ such that $\varphi \cdot u = v$. Thus

$$\varphi(u) = v.$$

Hence the two statements are equivalent. □

Definition 1.2.6. Let $\mathbb{G} = (V, E)$ be a graph. Then $\mathbb{G}$ is called a ***directed graph (digraph)*** if the edges of $\mathbb{G}$ are given one way directions. Otherwise $\mathbb{G}$ is ***undirected***, i.e edges are not assigned specific directions.

Definition 1.2.7. Let $\mathbb{G} = (V, E)$ be a graph. Then $\mathbb{G}$ is said to be ***connected*** if, for every $u, v \in V$, there exists a path from u to v.

$\mathbb{G}$ is called a ***complete graph*** denoted K_n of order n if each vertex is adjacent to all the other $n-1$ vertices of K_n [17].

$\bar{\mathbb{G}}$ is called the ***complement*** of $\mathbb{G}$ and is defined to be the graph having the same set V but with E replaced by $\bar{E}$ such that, for any $uv \in E$, $uv \notin \bar{E}$ and vice versa. Hence, $\bar{\mathbb{G}} = (V, \bar{E})$. If, in addition, $\mathbb{G}$ is isomorphic to $\bar{\mathbb{G}}$, we say $\mathbb{G}$ is a ***self complementary graph***.

Definition 1.2.8. A graph $\mathbb{G} = (V, E)$ is called ***bipartite*** if the set V can be partitioned into two non-empty subsets V_1 and V_2 such that $uv \in E$ if and only if vertices u and v belong to distinct subsets or partite sets of V.

Moreover, if each vertex of V_1 is joined to every vertex of V_2, then G is called a ***complete bipartite graph*** and is denoted $K_{n,m}$, where $n = |V_1|$ and $m = |V_2|$ or vice versa [17].

Definition 1.2.9. Let $\mathbb{G} = (V, E)$ be a graph. Then $\mathbb{G}$ is called ***regular*** if all the vertices of $\mathbb{G}$ have the same degree, more specifically, $\mathbb{G}$ is said to be r-regular, or regular of degree r, where r refers to the degree value.

Moreover $\mathbb{G}$ is called ***strongly regular*** if, in addition to regularity, $\mathbb{G}$ has these two properties:

1 : every pair of adjacent vertices has exactly λ common neighbours;

2 : every pair of non- adjacent vertices has exactly μ common neighbours.

Definition 1.2.10. Let $\mathbb{G} = (V, E)$ be a graph. Then $\mathbb{G}$ is said to be a ***symmetric graph*** if $Aut(\mathbb{G})$ acts transitively on a set of ordered pairs of adjacent vertices. Symmetric graphs are graphs that are both vertex and edge transitive. Symmetric graphs are also called ***arc transitive graphs***.

Definition 1.2.11. Let $\mathbb{G} = (V, E)$ be a graph. Then $\mathbb{G}$ is said to be a ***semi- symmetric graph*** if $\mathbb{G}$ is undirected, edge transitive and regular but lacks the property of vertex transitivity.

Theorem 1.2.3. *Every vertex transitive graph is regular.*

Proof. Let $\mathbb{G} = (V, E)$ be a vertex transitive graph. Let $u, v \in V$. Then there exists $\varphi \in Aut(V, E)$ such that $\varphi(u) = v$. Let $w_1, w_2, \cdots, w_{n-1}, w_n$ be distinct neighbours of u in (V, E). Then

$$\{u, w_1\}, \{u, w_2\}, \cdots, \{u, w_{n-1}\}, \{u, w_n\}$$

are edges in (V, E) and $deg(u) = n$. This implies that

$$\{\varphi(u), \varphi(w_1)\}, \{\varphi(u), \varphi(w_2)\}, \cdots, \{\varphi(u), \varphi(w_{n-1})\}, \{\varphi(u), \varphi(w_n)\}$$

are edges in (V, E)

$\Rightarrow \{v, \varphi(w_1)\}, \{v, \varphi(w_2)\}, \cdots, \{v, \varphi(w_{n-1})\}, \{v, \varphi(w_n)\}$ are edges in (V, E)

$\Rightarrow \{\varphi(w_1)\}, \{\varphi(w_2)\}, \cdots, \{\varphi(w_{n-1})\}, \{\varphi(w_n)\}$ are neighbours of v.

Also, $\{\varphi(w_1)\}, \{\varphi(w_2)\}, \cdots, \{\varphi(w_{n-1})\}, \{\varphi(w_n)\}$ are distinct, since $w_1, w_2, \cdots, w_{n-1}, w_n$ are distinct and φ is injective.

Hence $deg(v) \geq n = deg(u)$, and clearly $deg(v) \geq deg(u)$.

In a similar way we show that $deg(u) \geq deg(v)$.

Now since $deg(v) \geq deg(u)$ and $deg(u) \geq deg(v)$, we have that

$deg(u) = deg(v)$. Therefore, G is regular. □

Remark 1.2.1. *It is not necessarily true that regularity implies vertex transitivity; for example, a semi-symmetric graph is regular and edge transitive but it is not vertex transitive.*

A good example of a type of graph we will soon discuss, (strongly regular graphs), is the Petersen graph, which we define below:

Definition 1.2.12. Let $\mathbb{G} = (V, E)$ be a graph. Then $\mathbb{G}$ is called the ***Petersen graph*** if $\mathbb{G}$ is strongly regular with 10 vertices, 15 edges, degree 3, for adjacent vertices 0 common neighbours, and for non-adjacent vertices 1 common neighbour.

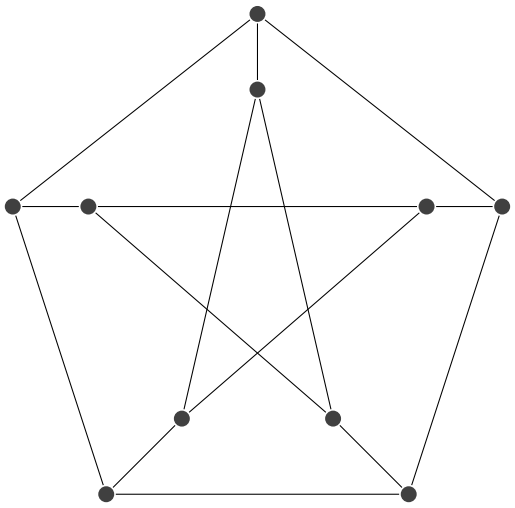


Figure 1.1: The Petersen Graph

Definition 1.2.13. Let $A = (a_{ij})$ be an $n \times n$ matrix. Then A is said to be a ***symmetric matrix*** if $A^T = A$ that is for all i, j, we have $a_{ij} = a_{ji}$.

Definition 1.2.14. Let $\mathbb{G} = (V, E)$ be a graph with vertex set $V = \{v_1, v_2, \cdots, v_n\}$. Then the ***adjacency matrix of*** $\mathbb{G}$ is the matrix $A = (a_{ij})$ where

$$a_{ij} = \begin{cases} 1 \text{ if } v_i v_j \in E \\ 0 \text{ if } v_i v_j \notin E. \end{cases}$$

Remark 1.2.2. *Moreover, note that for digraphs a_{ij} is not necessarily equal to a_{ji} unless it is a multi-graph with the same number of multi-edges on both directions for every vertex. For graphs with loops, a_{ii} is not necessarily* 0*, i.e the diagonal does not necessarily consists of only zeros. (Notice that a_{ij} represents the number of edges v_iv_j between the two vertices and for directed graphs this differs according to the direction of the edge). For weighted graphs (graphs where each edge is assigned a positive real number called the weight of the edge) we should also mention the idea of a weight matrix. Here a_{ij} represents the weight of the v_iv_j if $v_iv_j \in E$ and $a_{ij} = \infty$ if $v_iv_j \notin E$. The notions of weighted directed and simple graphs follow similarly.*

Definition 1.2.15. Let $n \in \mathbb{N}$, and A be an $n \times n$ matrix. Then A is called ***circulant*** if and only if the following is true: let $t \in \{1, 2, \ldots, n-1\}$. If $(a_1, a_2, \ldots, a_n)$ is the t^{th} row then $(a_n, a_1, \ldots, a_{n-1})$ is the $(t+1)^{th}$ row.
Moreover a ***circulant graph*** $\mathbb{G} = (V, E)$ is a graph with circulant adjacency matrix.

Definition 1.2.16. Let A be an $n \times n$ matrix, and I be an $n \times n$ identity matrix. Then we define $det(\lambda I - A)$ to be the ***characteristic polynomial*** of A; $det(\lambda I - A) = 0$ to be its ***characteristic equation***; and the roots of this equation to be the ***eigenvalues*** of A.
Moreover we define the ***spectrum*** of A, denoted $spec(A)$ to be the set of all eigenvalues of A. The ***spectrum of a graph*** $\mathbb{G}$, $spec(\mathbb{G})$, is the spectrum of its adjacency matrix. In addition $\mathbb{G}$ and $\mathbb{H} = (V_2, E_2)$ are said to be ***isospectral*** if $spec(\mathbb{G}) = spec(\mathbb{H})$.

We further state (in addition to the above definition) the following proposition without proof.

Proposition 1.2.4. *[27] Let $\mathbb{G} = (V_1, E_1)$ and $\mathbb{H} = (V_2, E_2)$ be isomorphic graphs. Then $spec(\mathbb{G}) = spec(\mathbb{H})$.*

Remark 1.2.3. *It is not necessarily true that if two graphs are isospectral then they are isomorphic;*

Example 1.2.1. Consider the isospectral digraphs below which are non-isomorphic:

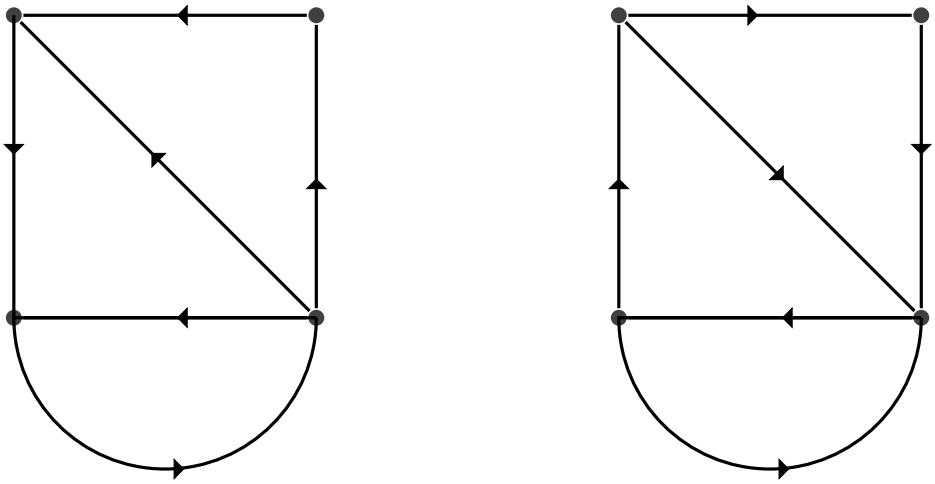


Figure 1.2: Isospectral non- isomorphic digraphs [27]

Let u, v be two vertices of a graph. The ***distance***, $d(u, v)$, between u and v is the length of a shortest path joining u and v. The ***eccentricity***, $e(u)$, of a vertex u is defined as the distance between u and a vertex furthest away from u. The ***radius*** of a graph is the minimum eccentricity of the graph, and the ***diameter*** of the graph is the maximum eccentricity, over all vertices. If the graph contains a cycle we define the ***girth*** of a graph to be the length of the shortest cycle, and the ***circumference*** to be the length of a longest cycle [8], [19].

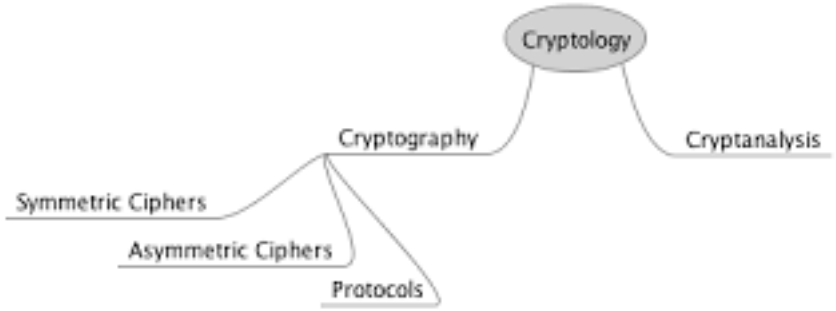
We state without proof the following propositions and refer the reader to the referenced material for proofs:

Proposition 1.2.5. *[11] Let $\mathbb{G} = (V, E)$ be a connected graph with diameter d. Then spec$(\mathbb{G}) \geq d + 1$.*

Proposition 1.2.6. *[16] Let $\mathbb{G} = (V, E)$ be a graph with no odd cycles, such that the number of mutually nonadjacent vertices for any subgraph of $\mathbb{G}$ is at least half the total number of vertices of that subgraph. Then $\lambda \in spec(\mathbb{G})$ implies $-\lambda \in spec(\mathbb{G})$ if and only if $\mathbb{G}$ is bipartite.*

1.3 Elementary Cryptography

Cryptography, which will be more formally introduced in Chapter 3, is a branch of Cryptology. In this section, we use material drawn from [6], [18], [19], [21], [25], [29], [30], [31] and [35] to describe this relationship, and several key properties of symmetric stream and block ciphers.



The term *cryptology* was derived from the Ancient Greek word *'kyptos'* meaning *'hidden secret'*. Cryptology is the science dealing with secret communication, and the mathematics that underpins *cryptography* and *cryptanalysis*. Whereas cryptography is the study of mathematical techniques to provide information security, cryptanalysis is the study of mathematical techniques to defeat information security; it is the study of mathematical techniques to crack the encryption algorithm and obtain the information without knowledge of the cryptographic keys. Cryptography and cryptanalysis have fought an ongoing war against each other since ancient times [35].

Egyptian hieroglyphs is a symbolic language that was used by ancient Egyptians to express their communication. This incorporated alphabets and logo-graphs. The first known evidence of the practice of cryptography uses

non-standard hieroglyphic symbols, although the writing is not necessarily an application of pure cryptography but rather some sort of encoding that aims not to hide information but to change the way it appears. This practice of cryptography dates to 1900 BC [25].

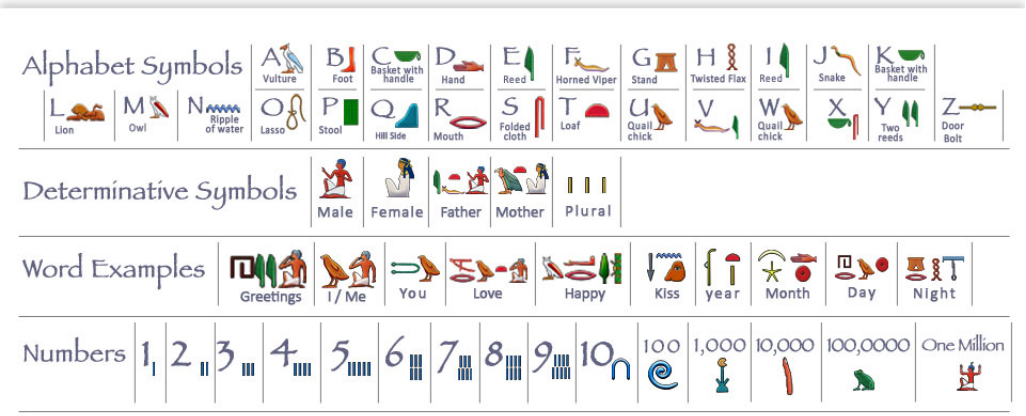


Figure 1.3: Egyptian Standard Hieroglyphic Symbols and Translations

Considerable progress has occurred since, due to the large increase of literate personnel, the invention of pen and paper, the discovery of computers, and so on. Crucial to these developments was also the development of mathematical sophistication. For example, consider the earliest known commonly used transposition cipher where characters of a word are just shuffled in order to hide the meaning, the well known *Caesar's Cipher* by Julius Caesar around 100 BC. Caesar used the *substitution cipher* with the shift 3 to convey secret messages to his army generals. Each letter in the alphabet of his message was shifted 3 units down the English alphabet, such that, **A** would be replaced by **D**, **Z** by **C** and so on. This cipher, although used successfully, is vulnerable to attack, depending entirely on the complete lack of mathematical knowledge on the part of the enemy.

Parallel to the development of cryptography has been cryptanalysis, especially after the development of complex computer based ciphers following the invention of digital computers and electronics after World War II [6].

The type of encryption cipher used in Caesar's cipher and many more built before the 1970's is private-key cryptography. During the 1970's public-

key cryptography was discovered and there has been much work done on both families of encryption ciphers since.

Private-key cryptography (also called symmetric cryptography) is the idea of cryptosystems that make it computationally feasible to compute the decryption from knowing the corresponding encryption key and or vice versa. The concept also includes cryptosystems where the decryption key is exactly the encryption key, in which case the main objective is exchanging the key. This presents many challenges including protecting the key during exchange, and transportation.

Public-key cryptosystems, are asymmetric-key cryptosystems and they aim to make it computationally infeasible to compute either key (encryption or decryption) from the other. The idea of this cryptography was proposed by Diffie and Hellman in their paper [18].

One of the well known public-key cryptosystems is the RSA cryptosystem which was a solution to the problem Diffie and Hellman encounted finding a suitable trapdoor one-way function that would be useful in constructing a public-key cryptosystem. Ronald Rivest, Adi Shamir and Leonard Adleman invented this cryptosystem in 1977, and it is used to this day.

However public-key cryptography can not fully replace private-key cryptography due to numerous challenges, including: speed, resource intensiveness and message size. Hence, the study of private-key cryptosystems continues, and in this dissertation we focus exclusively on private-key cryptography.

We now introduce some basic terminology and mathematics, focusing on private-key cryptography. The main objective of cryptography (irrespective of the type of the cryptosystem) is for secure communication in an unsecured channel. To achieve this, cryptography has had a long history of mathematically achieving its cryptographic goals (confidentiality, data integrity, authentication and non-repudiation).

We shall refer to a *plaintext* or simply a message, say M, as being the original intelligible information fed to the algorithm (maybe in any format or language). The format of M is converted by a process of *encoding* into a particular acceptable format or language for efficient transmission, and this process only requires an algorithm or *cipher* (cypher).

Modern cryptography introduces the concept of a key (encryption key) and is a process of disguising the message so as to hide or protect it from any intruder in an unsecured channel, and this disguised message is then called a *ciphertext* (or cryptogram), whilst the reverse process or the process of reattaining the plaintext (M) from a ciphertext (C) is called *decryption* or deciphering and just like encryption (enciphering), it involves or requires both an algorithm and a key. The presence of a key during encryption and decryption is one that makes it clear that encoding and decoding is not an ideal way of defining cryptography. *Decoding* is simply the reverse process of encoding, at least for symmetric ciphers.

There are two important symmetric-key ciphers that we shall discuss, stream ciphers and block ciphers. *Stream ciphers* act explicitly on each bit of the plaintext by combining it with a generated key. In the case of *block ciphers*, the message M is divided into blocks of fixed length, say d, and encryption is performed separately on each block thus producing a block of the same size for the ciphertext and these are joined together in different special ways that themselves provide better security [6].

Paar and Pelzl, in their text *"Understanding Cryptography"*, [31], give the following figure (Figure 1.4) to illustrate the difference in the mechanism of stream ciphers and block cipher in Symmetric Cryptography. Notice that the block length is d in the illustrations below, but in the first diagram (stream cipher) each bit of the plaintext is encrypted with the encryption key individually, whilst in the second diagram (block cipher), the plaintext is divided into blocks of length d, each of which is then encrypted as a block with the encryption key (K).

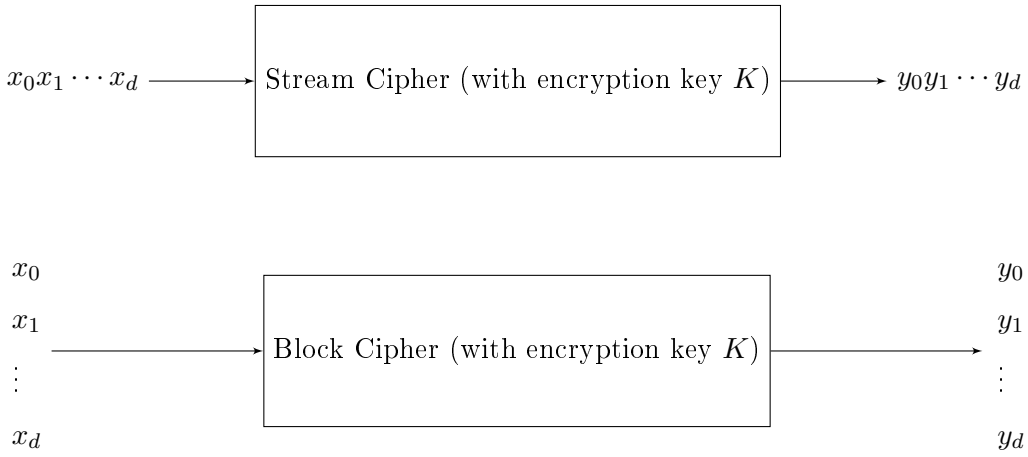


Figure 1.4: The mechanism of stream and block ciphers

The Mechanism of Stream Cipher

Suppose $x_i, y_i, k_i \in \{0, 1\}$, where x_i is a bit of the plaintext, y_i a bit of the ciphertext and k_i a bit of the keystream. Then the encryption is performed as follows:

$$y_i = e_{k_i}(x_i) \equiv x_i + k_i \bmod 2, \quad \text{where } e_{k_i} \text{ denotes using key } k_i \text{ to encrypt.}$$

Next we show that decryption is performed in a similar way, that is:

$$x_i = d_{k_i}(y_i) \equiv y_i + k_i \bmod 2, \quad \text{where } d_{k_i} \text{ denotes using key } k_i \text{ to decrypt.}$$

Proposition 1.3.1. *Let $y_i \equiv x_i + k_i \ mod\, 2$. Then $d_{k_i} = x_i$.*

Proof. Assume that $y_i = e_{k_i}(x_i)$, where $x_i, y_i, k_i \in \{0, 1\}$.

Then

$$\begin{aligned} d_{k_i} &\equiv y_i + k_i \bmod 2 \\ &\equiv (x_i + k_i) + k_i \bmod 2 \\ &\equiv x_i + 2k_i \bmod 2 \\ &\equiv x_i \bmod 2 \end{aligned}$$

□

Remark 1.3.1. *Since calculations are performed in base 2, we must first convert the format to ASC*II.

Example 1.3.1. To encrypt a message such as:

BEWARE THE ARMY!

we look for the corresponding ASCII code and generate the plaintext in the format we require:

01000010 01000101 01010111 01000001 01010010 01000101 00100000 01010100

01001000 01000101 00100000 01000001 01010010 01001101 01011001 00100001.

In order to proceed, we require a keystream to perform encryption, so the following introduces the idea of a keystream in a stream cipher, in order to highlight that the computation of these key bits is in fact the heart of a stream cipher.

The idea of security (cryptographic) keys in stream ciphers, involves an understanding of logic gate truth tables and random number generators (RNG), which are core to this cipher.

From the logic gate truth tables let us single out the 2-input Exclusive-OR (XOR), which we will be using in this cipher. Exclusive-OR is a logic gate function that, for "1:- true and 0:- false", gives as output 1 only when

(only) one of the inputs is 1 else it gives 0. The bits of the plaintext and those of the keystream are "combined" or encrypted using XOR operations.

The security of stream ciphers lies not on the XOR mechanism, but upon the randomness of the keystream generator. Amongst the three RNG-true random number generators (TRNG), pseudo-random number generators (PRNG) and cryptographic secure pseudo-random number generators (CSPRNG)-suitable for stream ciphers, we shall discuss the PRNG and from this example illustrate important attributes common also to the other two. *Pseudo-random number generators* generate a string of numbers which are not in a sense completely random since the sequence is easily reattained unlike in TRNG, where it is almost impossible to reattain the sequence. The randomness of this sequence lies in the initial value called the *seed* (k_0) and the rest of the sequence is the result of the function of the preceding value, that is, given the seed k_0 then $k_1 = f(k_0)$, and

$$k_{i+1} = f(k_i), \quad \text{where } i \in \mathbb{N}_0.$$

On the other hand CSPRNG differs from this in that given some bits, say t bits of the keystream, it is almost impossible to compute the next keystream bit and even the preceding bits using the current t^{th} bit, but possess the pseudo-randomness property to a degree.

Consider Example 1.3.1 and suppose PRNG is being applied to this problem to generate a keystream $k_0, k_1, \cdots$ from a chosen seed value, and a function, (say based on the linear congruential generator) and the generated keystream becomes something like:

$$01110001\cdots\cdots$$

Then, performing encryption gives:

$$\begin{array}{r} 01000010 \\ \oplus 01110001 \\ \hline 00110011 \end{array} \quad .$$

Clearly we notice that with this key the first symbol *"B"* of the plaintext after encryption and decoding by ASCII reads as "3", and also the involvement of the XOR. The rest of the message follows the same pattern.

Proposition 1.3.1 assures us that decryption follows in a similar way with the same key. Hence we will have:

$$\begin{array}{r} 0\,0\,1\,1\,0\,0\,1\,1 \\ \oplus\, 0\,1\,1\,1\,0\,0\,0\,1 \\ \hline 0\,1\,0\,0\,0\,0\,1\,0 \end{array}$$

which returns *"B"* after decoding.

An attack described by Paarl and Pelzl [31] on PRNG motivated the invention of CSPRNG. However, PRNG remains an important cipher.

The Mechanism of Block Cipher

Suppose that x_i is a bit of the plaintext, y_i a bit of the ciphertext and k_i a bit of the keyspace where $x_i, y_i, k_i \in \{0, 1\}$. Suppose M is a n-bit long plaintext. Then M is partitioned in blocks of length d (block-length, where $n \geq d$). Then a block cipher in symmetric cryptography maps each plaintext block of length d to a d-bit ciphertext. As an aside, note that the larger the value of d the higher the security but the slower the operation, and that most of the modern block ciphers have $d \geq 64$-binary bits.

Menezes, Van Oorschot and Vanstone [29], define block cipher encryption function as a bijective map,

$$E : \mathbb{V}_d \times \mathcal{K} \to \mathbb{V}_d,$$

where $\mathcal{K}$ is the keyspace containing subkeys K_i for some i, and the map is denoted $E_K(P)$ or $E(P, K)$, for P the plaintext.

However, the details of the cipher algorithm vary with the type of block cipher, some of which currently in use include the *Data Encryption Standards (DES), Triple-DES, Advanced Encryption Standards*. To motivate the study described in Chapters 3 and 4, we look at the mechanism of this cipher through a non-fully detailed discussion of DES. This should highlight that the security of this cryptosystem relies heavily on the construction of powerful *substitution boxes* (S-boxes), which are functions that take some number of input bits, say r, and transforms them into some number of output bits, say t, where r and t need not be the same. These functions are

presented in a form of tables and are the heart of the encryption in most modern block ciphers. The design of these functions (S-boxes) is a difficult task since one needs to ensure that they possess certain properties to ensure strength against a range of powerful cryptanalysis attacks.

Mohamed et al in the paper *Study of S-box Properties in Block Cipher*, [30], draw attention to a number of properties S-boxes ought to possess in order to survive powerful cryptanalysis, and furthermore claims that the process of creating new powerful S-boxes never ends, with various methods being applied to make them strong. Some of these properties include non-linearity, balanceness, the strict avalanche criterion, algebraic complexity, differential uniformity, robustness amongst others.

Example 1.3.2. Consider the message

BEWARE THE ARMY!

from Example 1.3.1 which was encoded (ASCII) as:

$M =$ 01000010 01000101 01010111 01000001 01010010 01000101 00100000 01010

10001001000 01000101 00100000 01000001 01010010 01001101 01011001 00100001.

Suppose we apply a DES block cipher with the keyspace:

$\mathcal{K} = 0111000001100111001010100010010101010010011110100010000001011111.$

Although DES partitions the plaintext to block of length $d = 64$-bits and chooses a key of size 64-bits as well, it then ignores the 8^{th}-bit of every byte of that 64-bit key such that:

$$\mathcal{K}^* = 01110000110011001010100100100101001011110100100000101111.$$

Hence, a 64-bits plaintext M is operated using a 56-bits key $\mathcal{K}^*$.

What follows is the desription of the algorithms of DES according to the DES steps. Note that this process involves many given permutations. Since DES uses a Feistel cipher, it operates its algorithms in rounds (r). For DES, $r = 16$-rounds. Hence, we are expected to generate 16-subkeys (K_i, for $1 \leq i \leq 16$).

First, by applying permutation on $\mathcal{K}$, according to table PC-1 given in [29], and ignoring those that do not appear on the table because they do not fall on $\mathcal{K}^*$, we get

$$\text{PC-1}\ (\mathcal{K} \setminus \{8^{th} - \text{bit of every byte of that 64-bit key}\}) =$$

$$0000000010110011011011111011101111101000001010100100001.$$

We then split the result of PC-1 into two equal halves, labelled C_0 = left half and D_0 =right half. Hence each half has 28-bits and is:

$$C_0 = 0000000010110011011011111011,$$

$$D_0 = 1011111010000010101001000 01.$$

To obtain C_i and D_i for all $1 \leq i \leq 16$, perform a left shift according to the left shift schedule of C_{i-1} and D_{i-1} for all 16 iterations, and that will give 17 pairs of C_i and D_i inclusive of C_0 and D_0 to C_{16} and D_{16}.

Next concatenate each pair using table PC-2 given in [29], to form 16-subkeys K_i, such that,

$$\text{PC-2}\ [C_3D_3].$$

That is,

$$K_i = \ \text{PC-2}\ [C_iD_i]\ \text{ for all } 1 \leq i \leq 16.$$

Notice that by performing this operation, each subkey reduces in size from 56-bits to 48-bits as results of PC-2. Now that we have 16-subkeys of 48-bits long each, the next step focuses on algorithms performed on the plaintext itself and is followed by the use of S-boxes, the core of block cipher.

The first process in this step uses an initial permutation function (IP) of 64 characters given in [29], to give the new arrangement of the bits of plaintext. Recall that we have partitioned the plaintext into block of $d = 64$-bits. Hence, we encrypt each 64-bit block individually by applying IP to each partitioned of M, say M_i, for some integer i. The $IP(M_i)$ is also 64-bits long as M_i so we then split $IP(M_i)$ also into two halves of size 32-bits each and label the left half L_0, the right half R_0.

From L_0 we build $L_1 = R_0$ and define a rule that any $L_i = R_{i-1}$, for some i, whilst on the other hand to generate R_i for $i > 0$, we define the rule

$R_i = L_{i-1} \oplus f_{K_i}(R_{i-1},)$. Here, f_{K_i} is the round function. This makes use of the E bit-selection table given in [29], to increase the size of R_{i-1} from 32-bits to 48-bits by repetition. This is done so that the size of each R_{i-1} is the same as the size of a subkeys which is 48-bits. Then $E(R_{i-1})$ which is 48-bits, is combined with the subkey K_i by use of the XOR explained earlier. That is, we calculate:

$$E(R_{i-1}) \oplus K_i,$$

and get a 48-bit, answer. This is then divided into groups of 6 bits named B1-B8, (that is $E(R_{i-1}) \oplus K_i = B1B2B3B4B5B6B7B8$). These 6-bit strings are used as coordinates to locate positions in the respective S-box, which are mathematically constructed tables which are defined by functions of special properties to provide security.

Each 6-bit number gives an idea of a row in an S-box by combining the first and last bit, for example if the number is 111010 and is named B5, then one has 1 and 0, giving rise to 10 in $\mathbb{Z}_2$ and that converts to 2 in $\mathbb{Z}$, and hence row 2 of the S-box; the rest of the 4-bits are 1101 which converts to 13 in $\mathbb{Z}$, and hence column 13 in the S-box. Row 2, column 13 in S_5 has entry 3, that is, $S5(B5) = S5(111010) = 3_{10} = 0011_2$. [Note that the columns and rows are labeled 0-15 and 0-3 respectively].

Iteratively one can construct:

$$S1(B1)S2(B2)S3(B3)S4(B4)S5(B5)S6(B6)S7(B7)S8(B8).$$

The next step involves assigning the value of the function f which is 32-bits long:

$$f = P[S1(B1)S2(B2)S3(B3)S4(B4)S5(B5)S6(B6)S7(B7)S8(B8)]$$

where P is a permutation table and that is XORed with L_i to get $R_i = L_{i-1} \oplus f_{K_i}(R_{i-1})$.

So clearly we managed to get L_i and R_i for all $1 \leq i \leq 16$, so we selected the last iteration L_{16} and R_{16} and swap their positions to get $R_{16}L_{16}$ then permute the 64-bit long $R_{16}L_{16}$ by the IP^{-1} table given in [29], to yield the ciphertext.

It is worth stating that all ciphers are classified or rated according to the

two definitions that follow:

Definition 1.3.1. A cipher is said to possess ***unconditional security*** if it is not defeatable even with infinite computational resources.

Definition 1.3.2. A cipher is said to possess ***computational security*** if the only way it could be defeated is by the application of a particular algorithm for at least n operations.

According to these definitions all currently known practical ciphers both private-key and public-key are NOT classified as unconditionally secure, and at best possess computational security. However, even that is still an issue since identifying a suitable particular algorithm maybe almost impossible for a particular cipher [31].

Summary

The idea of this dissertation is to highlight the relationship between graph theory and cryptography. Hence, this chapter introduced the basics in both fields. We considered the basics of group theory and algebra so as to bring to the readers attention the foundations of group and important properties that link them with certain graphs. This lead us to the next chapter where we look at Cayley graphs and strongly regular graphs which form part of the family of algebraic graphs.

Having considered basic definitions of concepts used in graph theory so as to discuss Cayley graphs and strongly regular graphs, in Section 1.3, we introduced cryptography (both public-key and private-key) with the aim of elucidating the difference. We then focused on private-key cryptosystems thus introducing stream ciphers and block ciphers, their mechanisms and mathematics.

In the chapters that follow we make use of these preliminaries to define in Chapter 2, certain graphs (Cayley graphs and strongly regular graphs) and in Chapter 3, certain cryptographic functions (Boolean functions and bent functions) useful for the design of stream ciphers and block ciphers. Chapter 4 considers the manner in which relationships may be established between the objects of study in the preceding chapters.

Chapter 2

Algebraic Graph Theory

In this chapter, we study the definitions and properties of Cayley and strongly regular graphs which are sub-families of the family of algebraic graphs. We will look at some of the results drawn from the properties of these graphs, although we will limit our study to results that will help us investigate the link between these two families and the cryptographic functions that will be studied in the next chapter. This review will include concepts like eigenvalues of these graphs, circulant graphs defined from Cayley graphs, the spectrum of graphs and partial difference sets. This material is drawn from [3], [4], [8], [10], [23], [26], [25] and [41].

2.1 Introduction

The study of Graph Theory begins with the discoveries of a Swiss Mathematician Leonhard Euler (April 15, 1707- September 18, 1783) in his work on the problem of the *Seven Bridges of Konigsberg* in 1735/6. At this stage Euler had just introduced the idea without naming it and later the idea was used in the *Knight tour problem*. Frequent reference to the idea and technique to solve problems triggered the introduction of the terminology of a ***graph*** in 1838.

Later, more problems were investigated using the idea of graphs. Graph theory as a field grew and questions rose about graphs. Strategies to answer these questions were put into practice, one of which was an extension to

graph theory which looked at addressing problems in graphs by means of algebraic methods. The idea of ***algebraic graph theory*** was introduced and studied, where ideas from algebra and group theory were put into good use.

2.2 Cayley Graphs

Arthur Cayley, a British male mathematician (August 16, 1821 - January 26, 1895), first introduced the idea of group based graphs by consideration of what we today call Cayley's Theorem. He named this group constructed graph the colour group and later it was renamed the Cayley graph or Cayley colour graph.

In this section we take a closer look at Cayley Graphs as they will be used in the application to cryptography. Cayley graphs are graphs constructed via groups. We therefore make use of elementary group and graph theory to elucidate the connection between group theory and graph theory.

Definition 2.2.1. Let $(G, \circ)$ be a group, and let Ω be a non-empty set such that $\Omega \subset G$, and $\forall \omega \in \Omega$ we have $\omega^{-1} \in \Omega$, i.e Ω is symmetric, but $e_G \notin \Omega$ for e_G the identity element of G. This shall henceforth be refereed to as an "inverse stable, identity free set" relative to G or a "Cayley set". Then the ***Cayley graph*** $\mathcal{G}(G, \Omega) = (V, E)$ is the graph with the following properties:

(i) $V = \{g \mid g \in G\}$;

(ii) $E = \{gk \mid k = g \circ \omega \text{ for } \omega \in \Omega, g \in G\}$.

Cayley digraphs differ from Cayley graphs in that they have directions. In this dissertation we focus exclusively on Cayley graphs, that is those where edges have no direction.

Example 2.2.1. Consider the group $(\mathbb{Z}_8, \oplus)$ and $S \subset \mathbb{Z}_8$ such that $S = \{1 + 8\mathbb{Z}, 7 + 8\mathbb{Z}, 3 + 8\mathbb{Z}, 5 + 8\mathbb{Z}\}$. Now since

$$\mathbb{Z}_8 = \{8\mathbb{Z}, 1 + 8\mathbb{Z}, 2 + 8\mathbb{Z}, 3 + 8\mathbb{Z}, 4 + 8\mathbb{Z}, 5 + 8\mathbb{Z}, 6 + 8\mathbb{Z}, 7 + 8\mathbb{Z}\},$$

notice that the identity $e_{\mathbb{Z}_8} \notin S$ but for all $x \in S$ we have $x^{-1} \in S$; all the properties to be met in the construction of Cayley graphs are satisfied. We can now write the Cayley graph:

$$\mathcal{G}((\mathbb{Z}_8, \oplus), S) = (\mathbb{Z}_8, \{\{x, y\} \mid x, y \in \mathbb{Z}_8, \exists k \in S \text{ such that } y = x \oplus k\}).$$

Hence, $V(\mathcal{G}(\mathbb{Z}_8, S)) = \mathbb{Z}_8$, and

$$\begin{aligned} E(\mathcal{G}(\mathbb{Z}_8, S)) =& \{\{8\mathbb{Z}, 1 + 8\mathbb{Z}\}, \{8\mathbb{Z}, 7 + 8\mathbb{Z}\}, \{8\mathbb{Z}, 3 + 8\mathbb{Z}\}, \{8\mathbb{Z}, 5 + 8\mathbb{Z}\}, \\ & \{1 + 8\mathbb{Z}, 2 + 8\mathbb{Z}\}, \{1 + 8\mathbb{Z}, 8\mathbb{Z}\}, \{1 + 8\mathbb{Z}, 4 + 8\mathbb{Z}\}, \{1 + 8\mathbb{Z}, 6 + 8\mathbb{Z}\}, \\ & \{2 + 8\mathbb{Z}, 3 + 8\mathbb{Z}\}, \{2 + 8\mathbb{Z}, 1 + 8\mathbb{Z}\}, \{2 + 8\mathbb{Z}, 5 + 8\mathbb{Z}\}, \{2 + 8\mathbb{Z}, 7 + 8\mathbb{Z}\}, \\ & \{3 + 8\mathbb{Z}, 4 + 8\mathbb{Z}\}, \{3 + 8\mathbb{Z}, 2 + 8\mathbb{Z}\}, \{3 + 8\mathbb{Z}, 6 + 8\mathbb{Z}\}, \{3 + 8\mathbb{Z}, 8\mathbb{Z}\}, \\ & \{4 + 8\mathbb{Z}, 5 + 8\mathbb{Z}\}, \{4 + 8\mathbb{Z}, 3 + 8\mathbb{Z}\}, \{4 + 8\mathbb{Z}, 7 + 8\mathbb{Z}\}, \{4 + 8\mathbb{Z}, 1 + 8\mathbb{Z}\}, \\ & \{5 + 8\mathbb{Z}, 6 + 8\mathbb{Z}\}, \{5 + 8\mathbb{Z}, 4 + 8\mathbb{Z}\}, \{5 + 8\mathbb{Z}, 8\mathbb{Z}\}, \{5 + 8\mathbb{Z}, 2 + 8\mathbb{Z}\}, \\ & \{6 + 8\mathbb{Z}, 7 + 8\mathbb{Z}\}, \{6 + 8\mathbb{Z}, 5 + 8\mathbb{Z}\}, \{6 + 8\mathbb{Z}, 1 + 8\mathbb{Z}\}, \{6 + 8\mathbb{Z}, 3 + 8\mathbb{Z}\}, \\ & \{7 + 8\mathbb{Z}, 8\mathbb{Z}\}, \{7 + 8\mathbb{Z}, 6 + 8\mathbb{Z}\}, \{7 + 8\mathbb{Z}, 2 + 8\mathbb{Z}\}, \{7 + 8\mathbb{Z}, 4 + 8\mathbb{Z}\}\} \end{aligned}$$

which gives the Cayley graph:

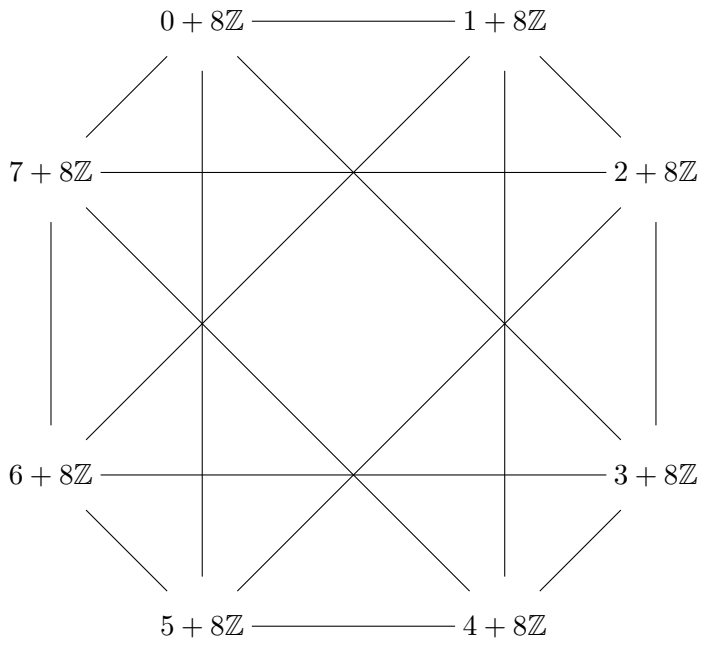


Figure 2.1: Cayley Graph on $(\mathbb{Z}_8)$ and $S \subset \mathbb{Z}_8$

Definition 2.2.2. A ***generating set*** S of $(G,\cdot)$ is defined as a set S for which $G = \langle S\rangle$, where

$$\langle S\rangle = \{s = g_1{\cdot}g_2{\cdot}\ldots{\cdot}g_n \mid \forall\, i \in \{1,\cdots,n\}, g_i \in G, \text{ such that, } g_i \in S \text{ or } g_i^{-1} \in S\}.$$

Theorem 2.2.1. *A Cayley graph $\mathcal{G} = ((G,\cdot),\Omega)$ is connected if and only if Ω is a generating set of the group $(G,\cdot)$.*

Proof. " $\Longrightarrow$ " Suppose $\mathcal{G}$ is connected. Then for any $g_i, g_j \in V(\mathcal{G}((G,\cdot),\Omega))$ there exists a $\{g_i - g_j\}$ path. We want to show that

$$\langle\Omega\rangle = G \Leftrightarrow \{g_1\cdots g_n \mid g_i \in \Omega \text{ or } g_i^{-1} \in \Omega\} = G.$$

Clearly $\{g_1\cdots g_n \mid g_i \in \Omega \text{ or } g_i^{-1} \in \Omega\} \subseteq G$. Let $g \in G$ and e_G the identity element of $(G,\cdot)$ under the binary operation. Since $\mathcal{G} = ((G,\cdot),\Omega)$ is connected, there is a path from e_G to g described as $P = e_G, x_1, x_2, \cdots, x_{n-1}, x_n, g$.

However $\mathcal{G} = ((G,\cdot),\Omega)$ is the Cayley graph of $(G,\cdot)$. Hence, the following is true for any $i \in \mathbb{N}, \omega_i \in \Omega$:

$$\begin{aligned}
x_1 &= e_G \cdot \omega_1,\\
x_2 &= x_1 \cdot \omega_2,\\
&\vdots\\
x_n &= x_{n-1} \cdot \omega_n,\\
g &= x_n \cdot \omega_{n+1},\\
&= x_{n-1} \cdot \omega_n \cdot \omega_{n+1}\\
&= x_{n-2} \cdot \omega_{n-1} \cdot \omega_n \cdot \omega_{n+1}\\
&\vdots\\
&= x_1 \cdot \omega_2 \cdot \omega_3 \cdot \omega_4 \cdot \ldots \cdot \omega_{n+1}\\
&= \omega_1 \cdot \omega_2 \cdot \omega_3 \cdot \omega_4 \cdot \ldots \cdot \omega_{n+1}.
\end{aligned}$$

Then $G \subseteq \{g_1\cdots g_n \mid g_i \in \Omega \text{ or } g_i^{-1} \in \Omega\}$. Therefore $\langle\Omega\rangle = G$.

" $\Longleftarrow$ " Now suppose Ω is a generating set of the group $(G, \cdot)$, that is $\langle \Omega \rangle = G$. Let $g, h \in G \Rightarrow g, h \in \langle \Omega \rangle$. We want to show that there is a path from g to h.

Since $g, h \in \langle \Omega \rangle$, then:

$$\begin{aligned} g &= x_1 \cdot \ldots \cdot x_n \cdot e_G & x_i \in \Omega \quad \text{or} \quad x_i^{-1} \in \Omega \\ h &= e_G \cdot y_1 \cdot \ldots \cdot y_n & y_i \in \Omega \quad \text{or} \quad y_i^{-1} \in \Omega. \end{aligned}$$

Here, $\forall i$, $x_i, y_i \in \Omega$.

Claim: $x_1 \cdots x_n, x_1 \cdots x_{n-1}, x_1 \cdots x_{n-2}, x_1 \cdots x_{n-3}, \cdots, x_1x_2x_3, x_1x_2, x_1, e_G$, $-y_1, y_1y_2, y_1y_2y_3, \cdots, y_1 \cdots y_{m-1}, y_1 \cdots y_m$ is a walk in $\mathcal{G}((G, \cdot), \Omega)$.

Now for $\{x_1 \cdots x_n, x_1 \cdots x_{n-1}\}$ to be an edge in $\mathcal{G}((G, \cdot), \Omega)$:

$$\begin{aligned} x_1 \cdots x_{n-1} &= (x_1 \cdots x_n) \cdot k_1 & \text{for some } k_1 \in \Omega \\ &= (x_1 \cdots x_n) \cdot x_n^{-1} & x_i \in \Omega \Rightarrow x_i^{-1} \in \Omega, \text{ for all } i. \end{aligned}$$

Therefore $\{x_1 \cdots x_n, x_1 \cdots x_{n-1}\}$ is an edge in $\mathcal{G}((G, \cdot), \Omega)$.
Similarly for $\{x_1 \cdots x_{n-1}, x_1 \cdots x_{n-2}\}$ to be an edge in $\mathcal{G}((G, \cdot), \Omega)$:

$$\begin{aligned} x_1 \cdots x_{n-2} &= (x_1 \cdots x_{n-1}) \cdot k_2 & \text{for some } k_2 \in \Omega \\ &= (x_1 \cdots x_{n-1}) \cdot x_{n-1}^{-1} & x_i \in \Omega \Rightarrow x_i^{-1} \in \Omega, \text{ for all } i. \end{aligned}$$

Therefore $\{x_1 \cdots x_{n-1}, x_1 \cdots x_{n-2}\}$ is an edge in $\mathcal{G}((G, \cdot), \Omega)$.

Continuing in this way, we see that

$$x_1 \cdots x_n, x_1 \cdots x_{n-1}, x_1 \cdots x_{n-2}, x_1 \cdots x_{n-3}, \cdots, x_1x_2x_3, x_1x_2, x_1, e_G$$

is a walk.

Next we show that $\{e_G, y_1\}$ is an edge in $\mathcal{G}((G, \cdot), \Omega)$:

$$y_1 = e_G \cdot y_1.$$

Therefore, $\{e_G, y_1\}$ is an edge in $\mathcal{G}((G, \cdot), \Omega)$.

Similarly $\{y_1, y_1 \cdot y_2\}$ is an edge in $\mathcal{G}((G,\cdot),\Omega)$:

$$y_1 \cdot y_2 = y_1 \cdot y_2.$$

Continuing in this way, we see that

$$e_G, y_1, y_1y_2, y_1y_2y_3, \cdots, y_1 \cdots y_{m-1}, y_1 \cdots y_m$$

is a walk, and our claim is substantiated.

Since $g = x_1 \cdots x_n \cdot e_G$ and $h = e_G \cdot y_1 \cdots y_n$, there is a walk between g and h. Therefore, since for all $g, h \in \mathcal{G}((G,\cdot), S)$ there is a path from g to h; $\mathcal{G}((G,\cdot),\Omega)$ is connected.

□

Although the following theorem is not proved in this dissertation, it is listed without proof because it gives a clear relationship between circulant graphs and Cayley graphs for the purpose of classifying Cayley graphs. The referenced material provides the proof.

Theorem 2.2.2. *[4] Circulant graphs are Cayley graphs if and only if they are connected.*

Lemma 2.2.3. *Let $(G,\circ)$ be a group. Let Ω be a non-empty, "inverse stable" (every element has an inverse in Ω), identity free set relative to $(G,\circ)$ and $g \in G$ arbitrary.*

Define $\varphi : G \to G$ by $\varphi(x) = g \circ x$ for each $x \in G$. Then $\varphi \in Aut(\mathcal{G}((G,\circ),\Omega))$, for any $\mathcal{G}((G,\circ),\Omega)$ a Cayley graph.

Proof. First we show that φ is a bijection from G to G. Take $\alpha, \beta \in G$. Assume that $\varphi(\alpha) = \varphi(\beta)$. Then

$$\begin{aligned} & g \circ \alpha = g \circ \beta \\ \Rightarrow \quad & g^{-1} \circ g \circ \alpha = g^{-1} \circ g \circ \beta \\ \Rightarrow \quad & e_G \circ \alpha = e_G \circ \beta \\ \Rightarrow \quad & \alpha = \beta, \qquad \therefore \ \varphi \text{ is injective.} \end{aligned}$$

Now take $\beta \in G$. We want to show that $\beta = \varphi(\alpha)$. That is, we must show

$$\begin{aligned} & \beta = g \circ \alpha \\ \Leftrightarrow\ & g^{-1} \circ \beta = g^{-1} \circ g \circ \alpha \\ \Leftrightarrow\ & g^{-1} \circ \beta = e_G \circ \alpha \\ \Leftrightarrow\ & g^{-1} \circ \beta = \alpha. \end{aligned}$$

Note that $\alpha = g^{-1} \circ \beta \in G$ by closure. Moreover, by construction,

$$\begin{aligned} \varphi(\alpha) &= \varphi(g^{-1} \circ \beta) \\ &= g \circ (g^{-1} \circ \beta) \\ &= (g \circ g^{-1}) \circ \beta \\ &= e_G \circ \beta \\ &= \beta \qquad \therefore\ \varphi \text{ is surjective.} \end{aligned}$$

This shows that φ is bijective.

Let a and b be arbitrary objects. Suppose $\{a, b\}$ is an edge of $\mathcal{G}((G, \circ), \Omega)$. Then there exists some $k \in \Omega$ such that $b = a \circ k$. Hence

$$\begin{aligned} g \circ b &= g \circ (a \circ k) \\ &= (g \circ a) \circ k \\ \Rightarrow \varphi(b) &= \varphi(a) \circ k \end{aligned}$$

$\Rightarrow \{\varphi(a), \varphi(b)\}$ is an edge in $\mathcal{G}((G, \circ), \Omega)$
$\Rightarrow \varphi \in Aut(\mathcal{G}((G, \circ), \Omega))$. □

Theorem 2.2.4. *Every Cayley graph is vertex transitive.*

Proof. Let $\mathcal{G}$ be a Cayley graph. Let x, y be arbitrary vertices of $\mathcal{G}$.

Since $\mathcal{G}$ is Cayley, there exists a group $(G, \circ)$ and Ω, a non-empty inverse stable and identity free set relative to $(G, \circ)$, such that $\mathcal{G} = \mathcal{G}((G, \circ), \Omega)$.

Now $y \circ x^{-1} \in (G, \circ)$. Define $\varphi : G \to G$ by $\varphi(g) = (y \circ x^{-1}) \circ g$ for all $g \in (G, \circ)$.

We have $\varphi \in Aut(\mathcal{G}((G, \circ), \Omega))$ by Lemma 2.2.7. Hence,

$$\begin{aligned}\varphi(x) &= (y \circ x^{-1}) \circ x \\ &= y \circ (x^{-1} \circ x) \\ &= y \circ e_G \\ &= y \quad \text{which implies that} \quad \mathcal{G}((G, \circ), \Omega) \quad \text{is vertex transitive.}\end{aligned}$$

□

Remark 2.2.1. *It is not necessarily true that vertex transitivity implies Cayley, (the converse of Theorem 2.2.4). The fundamental theorem for recognizing Cayley graphs (given below) helps us to identify vertex transitive graphs that are not Cayley.*

In an effort to identify Cayley graphs, Gert Sabidussi presented the fundamental theorem below. We state without proof Sabidussi's theorem and refer the reader to the referenced paper for the proof.

Theorem 2.2.5. *[8], [37] [**Sabidussi's theorem**] A graph $\mathcal{G} = (V, E)$ is a Cayley graph if and only if $Aut(\mathcal{G})$ contains a regular subgroup.*

Example 2.2.2. The Petersen graph $\mathbb{P}$ is one good example of a vertex transitive graph such that $\nexists$ r-regular H subgroup of $Aut(\mathbb{P})$ for some r. This is a difficult and lengthy result to prove, and the reader is referred to [8] for details.

We also show that the cyclic graph C_6 is vertex transitive:

Example 2.2.3. Consider the group $\mathbb{Z}_6 = \{6\mathbb{Z}, 1 + 6\mathbb{Z}, 2 + 6\mathbb{Z}, 3 + 6\mathbb{Z}, 4 + 6\mathbb{Z}, 5 + 6\mathbb{Z}\}$. If $S \subset \mathbb{Z}_6$ such that $S = \{1 + 6\mathbb{Z}, 5 + 6\mathbb{Z}\}$, then we note that $e_{\mathbb{Z}_6} \notin S$ and for all $x \in S$ there is $x^{-1} \in S$. Hence, we may construct

$$\mathcal{G}((\mathbb{Z}_6, \oplus), S) = (\mathbb{Z}_6, \{\{x, y\} \mid x, y \in \mathbb{Z}_6, \exists k \in S \text{ such that } y = x \oplus k\})$$

where $V(\mathcal{G}(\mathbb{Z}_6, S)) = \mathbb{Z}_6$ and

$$\begin{aligned} E(\mathcal{G}(\mathbb{Z}_6, S)) =& \{\{6\mathbb{Z}, 1+6\mathbb{Z}\}; \{6\mathbb{Z}, 5+6\mathbb{Z}\}; \\ & \{1+6\mathbb{Z}, 2+6\mathbb{Z}\}; \{1+6\mathbb{Z}, 6\mathbb{Z}\}; \\ & \{2+6\mathbb{Z}, 3+6\mathbb{Z}\}; \{2+6\mathbb{Z}, 1+6\mathbb{Z}\}; \\ & \{3+6\mathbb{Z}, 4+6\mathbb{Z}\}; \{3+6\mathbb{Z}, 2+6\mathbb{Z}\}; \\ & \{4+6\mathbb{Z}, 5+6\mathbb{Z}\}; \{4+6\mathbb{Z}, 3+6\mathbb{Z}\}; \\ & \{5+6\mathbb{Z}, 6\mathbb{Z}\}; \{5+6\mathbb{Z}, 4+6\mathbb{Z}\}\}. \end{aligned}$$

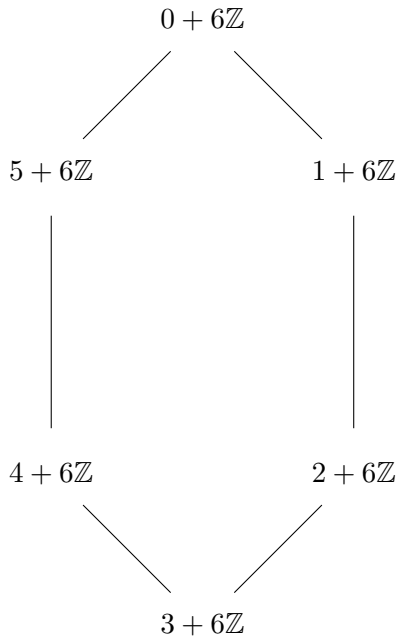


Figure 2.2: Cayley Graph on $(\mathbb{Z}_6)$ and $S = \{1+6\mathbb{Z}, 5+6\mathbb{Z}\}$

Proposition 2.2.6. *Let $\mathcal{G}(G, \Omega)$ be a Cayley graph. Then $\mathcal{G}(G, \Omega)$ is $|\Omega|$-regular.*

Proof. Suppose $\mathcal{G}(G, \Omega)$ is a Cayley graph. Then $\Omega \subset (G, \circ)$ and for all $g \in V(\mathcal{G}(G, \Omega)), g \in G$. Also, define the neighbors of g as being:

$$\{k \mid gk \in E(\mathcal{G}(G, \Omega)) \text{ if and only if } k = g \circ \omega \text{ for } \omega \in \Omega, g \in G\}.$$

Hence, the degree of $g \in G$ would be the number of all $\omega \in \Omega$. Therefore $\mathcal{G}(G, \Omega)$ is $|\Omega|$-regular. $\square$

Next, we recall our earlier definition of the adjacency matrix and the eigenvalues of a graph, and use them to define the notion of an adjacency operator of any graph on any given eigenfunction. We are, of course, particularly interested in the case of a Cayley graph.

Definition 2.2.3. Let $\mathcal{G} = (V, E)$ be a graph with $A = (a_{ij})$ as its adjacency matrix, where v_i, v_j label elements of V. Define f to be an ***eigenfunction*** of A. Then we define an ***adjacency operator*** A ***of*** $\mathcal{G}$ ***on an eigenfunction*** f, by $(Af)(v_i) = \sum_{j \in V} a_{ij} f(v_j)$.

Let $v_1, v_2, \cdots, v_n \in V$. Then generally in matrix form, [26],

$$Af = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} f(v_1) \\ f(v_2) \\ \vdots \\ f(v_n) \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} a_{1k} f(v_k) \\ \sum_{k=1}^{n} a_{2k} f(v_k) \\ \vdots \\ \sum_{k=1}^{n} a_{nk} f(v_k) \end{pmatrix}.$$

For the special case of Cayley graphs $\mathcal{G}(G, \Omega)$, the adjacency operator on an eigenfunction can be simplified to be $(Af)(g) = \sum_{\omega \in \Omega} f(g \circ \omega)$, where $g \in G$ [26].

Example 2.2.4. Let us consider the Cayley graph defined in Example 2.2.3 and as an example show that the special rule for obtaining the adjacency operator of Cayley graphs on an eigenfunction gives the same answer as the general method for all graphs for obtaining the adjacency operator on an eigenfunction. The given Cayley graph has the adjacency matrix:

$$A_{\mathcal{G}(\mathbb{Z}_6, S))} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

$$Af = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} f(6\mathbb{Z}) \\ f(1+6\mathbb{Z}) \\ f(2+6\mathbb{Z}) \\ f(3+6\mathbb{Z}) \\ f(4+6\mathbb{Z}) \\ f(5+6\mathbb{Z}) \end{pmatrix}$$

From this we notice that $(Af)(4+6\mathbb{Z}) = f(3+6\mathbb{Z}) + f(5+6\mathbb{Z})$.

Considering the method defined for the special case of Cayley graphs for obtaining the adjacency operator on an eigenfunction we evaluate $(Af)(4+6\mathbb{Z})$: In Example 2.2.3 we are given $S = \{1+6\mathbb{Z}, 5+8\mathbb{Z}\}$, so

$$\begin{aligned}(Af)(4+6\mathbb{Z}) &= \sum_{s\in S} f((4+6\mathbb{Z}) \circ s) \\ &= f((4+6\mathbb{Z}) + (1+6\mathbb{Z})) + f((4+6\mathbb{Z}) + (5+6\mathbb{Z})) \\ &= f((5+6\mathbb{Z}) + f((3+6\mathbb{Z});\end{aligned}$$

and we notice that the two methods agree.

Consider the following lemma that displays the relationship between the spectral information of a Cayley graph and characters of the Abelian group used in constructing the graph. The proof for this lemma is given in the referenced material.

Lemma 2.2.7. *[42] Let $(G, \circ)$ be an Abelian group, $\chi_\varphi : G \to \mathbb{C}$ the character of $(G, \circ)$ and Ω the Cayley set. Let $\mathcal{G}(G, \Omega) = (V, E)$ be a Cayley graph, and $A_\mathcal{G}$ its adjacency matrix. Then*

$$\frac{1}{|\Omega|} \sum_{\omega\in\Omega} \chi_\varphi(\omega), \quad \textit{gives the eigenvalue of } \mathcal{G}(G, \Omega),$$

associated with χ_φ and the characters of $(G, \circ)$ are the corresponding eigenvectors of $\mathcal{G}(G, \Omega)$.

Definition 1.2.8 introduced the idea of bipartite graphs. The example below will help us explore bipartite Cayley graphs:

Example 2.2.5 (**Bipartite Cayley Graph**). Suppose that

$$\mathcal{G}((D_4,\cdot),S) = (V,E) = (D_4, \{\{x,y\} \mid x,y \in D_4, \exists k \in S \text{ such that } y = x{\cdot}k\})$$

is a Cayley graph of the Dihedral group on two generators defined by the vertex set below, $\alpha, \beta \in (G,\cdot)$ and $\alpha \neq \beta$, where $(G,\cdot)$ is a group.

If $S = \{\alpha, \alpha^3, \beta\}$, then:

$$\begin{aligned} V(\mathcal{G}(D_4,S)) &= \langle \alpha, \beta \mid \alpha^4 = \beta^2 = (\alpha\beta)^2 = e_{D_4} \rangle \\ &= \{e_{D_4}, \alpha, \alpha^2, \alpha^3, \beta, \alpha\beta, \alpha^2\beta, \alpha^3\beta\}, \end{aligned}$$

and

$$\begin{aligned} E(\mathcal{G}(D_4,S)) = &\{\{e_{D_4},\alpha\},\{e_{D_4},\beta\},\{e_{D_4},\alpha^3\}, \\ &\{\alpha,\alpha^2\},\{\alpha,\alpha\beta\},\{\alpha,e_{D_4}\}, \\ &\{\alpha^2,\alpha^3\},\{\alpha^2,\alpha^2\beta\},\{\alpha^2,\alpha\}, \\ &\{\alpha^3,e_{D_4}\},\{\alpha^3,\alpha^3\beta\},\{\alpha^3,\alpha^2\}, \\ &\{\beta,\alpha\beta\},\{\beta,e_{D_4}\},\{\beta,\alpha^3\beta\}, \\ &\{\alpha\beta,\alpha^2\beta\},\{\alpha\beta,\alpha\},\{\alpha\beta,\beta\}, \\ &\{\alpha^2\beta,\alpha^3\beta\},\{\alpha^2\beta,\alpha^2\},\{\alpha^2\beta,\alpha\beta\}, \\ &\{\alpha^3\beta,\beta\},\{\alpha^3\beta,\alpha^3\},\{\alpha^3\beta,\alpha^2\beta\}\}. \end{aligned}$$

Now let $V(\mathcal{G}(D_4,S))$ be partitioned into two partite sets, $V_1, V_2 \in V(\mathcal{G}(D_4,S))$ such that:

$$\begin{aligned} V_1 &= \{e_{D_4}, \alpha^2, \alpha\beta, \alpha^3\beta\} \\ V_2 &= \{\alpha, \beta, \alpha^3, \alpha^2\beta\}. \end{aligned}$$

This yields the defined bipartite Cayley graph:

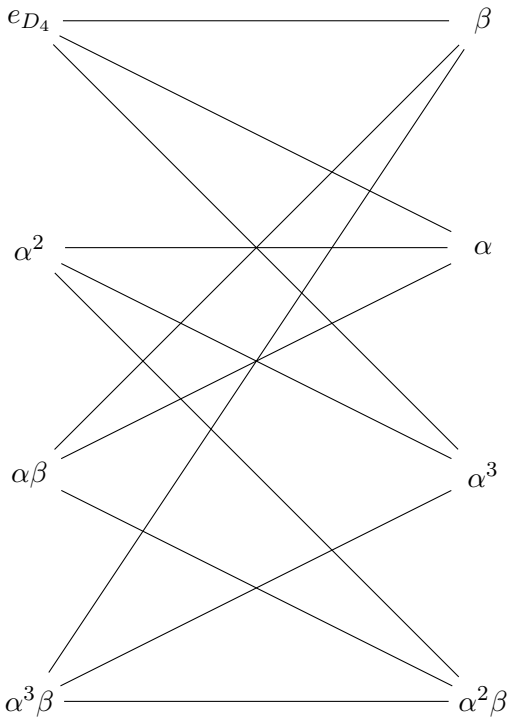


Figure 2.3: Bipartite Cayley Graph

2.3 Strongly Regular Graphs

Strongly regular graph, were first discovered and studied in 1963 by Raj Chandra Bose when he used the ideas of graph theory to solve problems in his work in design theory and in the theory of error-correcting codes [23]. (To be specific, with the aim to survey association schemes of partially balanced incomplete block designs [10]). These graphs quickly became core and used in many other studies inclusive of that by Donald Gordon Higman in his work on representation theory. Later they were seen to play a major role in the study of cryptography.

In this section we take a closer look at strongly regular graphs so as to prepare for the chapters that follow where they will be used in application to cryptography. The results reviewed below include both those which are group theory based and those combinatorial; the reader is referred to the following sources for further details: [3], [28].

Definition 2.3.1. Let $\mathbb{G} = (V, E)$ be a graph. Then $\mathbb{G}$ is called a ***strongly regular graph*** if and only if the following is true:

1 : $\mathbb{G}$ is a regular graph;

2 : There exists $\lambda, \mu \in \mathbb{N}_0$ such that:

i : For all $u, v \in V(\mathbb{G})$ if $u \neq v$ and $\{u, v\} \in E$, then

$$|N_{\mathbb{G}}(u) \cap N_{\mathbb{G}}(v)| = \lambda;$$

ii : For all $u, v \in V(\mathbb{G})$ if $u \neq v$ and $\{u, v\} \notin E$, then

$$|N_{\mathbb{G}}(u) \cap N_{\mathbb{G}}(v)| = \mu,$$

where $N_{\mathbb{G}}(u)$ is defined as the set of vertices that are neighbours of vertex u.

Remark 2.3.1. *A strongly regular graph has the following parameters:*

- $n :=$ *number of vertices of the graph;*

- $r :=$ *uniform degree per vertex;*
- λ *as defined in Definition 2.3.1 [number of common neighbors for adjacent vertices];*
- μ *as defined in Definition 2.3.1 [number of common neighbors for non-adjacent vertices],*

and these would generally be given in the form (n, r, λ, μ) *for any strongly regular graph. Hence, we shall follow this convention in this text.*

Example 2.3.1. A widely used example of a strongly regular graph is the Petersen graph discussed in Definition 1.2.12. Notice that the parameters of the Petersen graph as given in the definition satisfy the conditions of strongly regular graphs.

There are many other examples of strongly regular graphs. For example, Paley graphs, are graphs constructed from the ring $\mathbb{Z}/p\mathbb{Z}$ and the identity free, inverse stable set $\Omega = \{x^2 | x \in \mathbb{Z}/p\mathbb{Z}\}$ and they are strongly regular with parameters (n, r, λ, μ) as $\left(p, \frac{(p-1)}{2}, \frac{(p-5)}{4}, \frac{(p-1)}{4}\right)$. A particular case is described below.

Example 2.3.2. Let $\mathcal{G} = (V, E)$ be a Paley graph with $p = 13$. Then $V(\mathcal{G}(\mathbb{Z}_{13}, S)) = \mathbb{Z}_{13}$, and since $e_{\mathbb{Z}_{13}} \notin S$,

$$S = \{1 + 13\mathbb{Z}, 3 + 13\mathbb{Z}, 4 + 13\mathbb{Z}, 9 + 13\mathbb{Z}, 10 + 13\mathbb{Z}, 12 + 13\mathbb{Z}\}.$$

To avoid having a messy graph we give a twisted drawing with none of the mathematics changed and with every vertex given to mod 13. Note that the ordering of the vertices does not affect the graph mathematically. Hence any mathematically correct drawing is acceptable. The key to understanding this diagram is sticking to the name of each vertex and noting that repeated vertices are to be regarded as just one vertex.

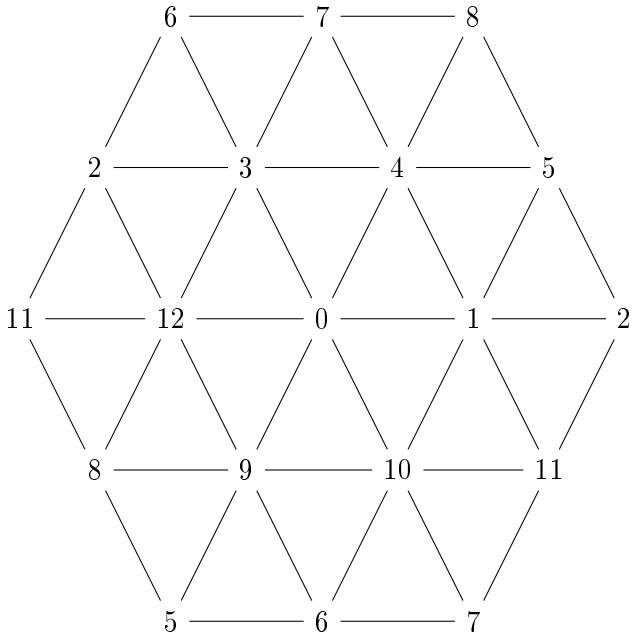


Figure 2.4: Paley Graph of $p = 13$

Notice that we have $(13, 6, 2, 3) = \left(p, \frac{(p-1)}{2}, \frac{(p-5)}{4}, \frac{(p-1)}{4}\right)$ for $p = 13$, as it should.

Proposition 2.3.1. *Every bipartite strongly regular graph has* $\lambda = 0$.

Proof. Let $\mathbb{G}$ be a bipartite graph. Pick vertices $u \in V(\mathbb{G}_1)$ and $v \in V(\mathbb{G}_2)$, where $V(\mathbb{G}_1)$ and $V(\mathbb{G}_2)$ are the partite sets of $V(\mathbb{G})$.

Then, since $\mathbb{G}$ is bipartite the only possible neighbours of $u \in V(\mathbb{G}_1)$ are in $V(\mathbb{G}_2)$, and; similarly, the only possible neighbours of $v \in V(\mathbb{G}_2)$ are in $V(\mathbb{G}_1)$. Hence, there does not exist w such that w is a neighbour of both u and v.
Therefore, every bipartite strongly regular graph has $\lambda = 0$. □

Remark 2.3.2. *Let* $\mathbb{G}$ *be a strongly regular graph with parameters* (n, r, λ, μ). *It is not necessarily true that if* $\lambda = 0$ *then* $\mathbb{G}$ *is a bipartite strongly regular graph.*

Example 2.3.3. Consider the Petersen graph, as defined and discussed in this text. This has $\lambda = 0$, but is not bipartite.

Proposition 2.3.2. *[3] Let $\mathbb{G} = (V, E)$ be a strongly regular graph. Then the complement, $\bar{\mathbb{G}}$, of $\mathbb{G}$ is also strongly regular and has parameters $(n, \bar{r}, \bar{\lambda}, \bar{\mu})$ where,*

$$\begin{aligned}\bar{r} &= n - r - 1,\\ \bar{\lambda} &= n - 2 - 2r + \mu,\\ \bar{\mu} &= n - 2r + \lambda.\end{aligned}$$

Proof. By the definition of $\bar{\mathbb{G}}$ we have that $|\bar{\mathbb{G}}| = |\mathbb{G}| = n$.

(a) <u>To show $\bar{r} = n - r - 1$:</u>

Let $x \in V(\bar{\mathbb{G}})$ be arbitrary. Since $deg_{\mathbb{G}}(x) = r$, let $y_1, y_2, \cdots, y_r \in V(\mathbb{G})$ be all the neighbors of x in G listed with no repetitions.

Claim: $\{y \in V(\bar{\mathbb{G}}) \mid xy \in E(\bar{\mathbb{G}})\} = V(\mathbb{G})\backslash\{x, y_1, y_2, \cdots, y_r\}$. We prove this by demonstrating that the $LHS \subseteq RHS$ and conversely.

<u>To show $LHS \subseteq RHS$</u>: Pick $t \in \{y \in V(\bar{\mathbb{G}}) \mid xy \in E(\bar{\mathbb{G}})\}$

$$\Rightarrow t \in V(\bar{\mathbb{G}}),$$

$$\Rightarrow t \in V(\mathbb{G}).$$

Suppose $t = x$ then $xx \in E(\bar{\mathbb{G}})$, which is a contradiction because that would be a loop, therefore $t \neq x$.
Therefore, suppose $t \in \{y_1, y_2, \cdots, y_r\}$

$$\Rightarrow xy_i \in E(\bar{\mathbb{G}}) \text{ for } i \in \{1, 2, \cdots, r\}.$$

However, $xy_i \in E(\mathbb{G})$, leading to a contradiction, therefore
$t \notin \{y_1, y_2, \cdots, y_r\}$. Hence $t \notin \{x, y_1, y_2, \cdots, y_r\}$. Therefore $t \in V(\mathbb{G})\backslash\{x, y_1, y_2, \cdots, y_r\} \Rightarrow LHS \subseteq RHS$.

<u>To show $RHS \subseteq LHS$</u>: Pick $t \in V(\mathbb{G})\backslash\{x, y_1, y_2, \cdots, y_r\}$

$$\Rightarrow t \in V(\mathbb{G}),$$

$$\Rightarrow t \in V(\bar{\mathbb{G}}), \text{ and also } t \notin \{y_1, y_2, \cdots, y_r\}.$$

Therefore, t is not a neighbor of x in G:

$$xt \notin E(\mathbb{G}),$$

$$\Rightarrow xt \in E(\bar{\mathbb{G}}).$$

Therefore, $t \in \{y \in V(\bar{\mathbb{G}}) \mid xy \in E(\bar{\mathbb{G}})\} \Rightarrow RHS \subseteq LHS$.

$$\therefore\ RHS = LHS.$$

This establishes that $\{y \in V(\bar{\mathbb{G}}) \mid xy \in E(\bar{\mathbb{G}})\} = V(\mathbb{G})\backslash\{x, y_1, \cdots, y_r\}$.

$$\begin{aligned}\text{Moreover, } deg_{\bar{\mathbb{G}}}(x) &= |\{y \in V(\bar{\mathbb{G}}) \mid xy \in E(\bar{\mathbb{G}})\}| \\ &= |V(\mathbb{G})\backslash\{x, y_1, y_2, \cdots, y_r\}|.\end{aligned}$$

However, since $\{x, y_1, y_2, \cdots, y_r\} \subseteq V(\mathbb{G})$

$$|V(\mathbb{G})\backslash\{x, y_1, y_2, \cdots, y_r\}| = |V(\mathbb{G})| - |\{x, y_1, y_2, \cdots, y_r\}| \qquad (2.1)$$

The $y_1, y_2, \cdots, y_r$ are distinct, and none of $y_1, y_2, \cdots, y_r$ are equal to x since they are neighbors of x and the graph does not have loops. Therefore $x, y_1, y_2, \cdots, y_r$ are distinct. It follows that:

$$|\{x, y_1, y_2, \cdots, y_r\}| = r + 1.$$

Therefore, from (2.1) we obtain:

$$\begin{aligned}deg_{\bar{\mathbb{G}}}(x) = \bar{r} &= |V(\mathbb{G})| - (r+1) \\ &= n - (r+1) \\ &= n - r - 1.\end{aligned}$$

(b) **To show $\bar{\lambda} = n - 2 - 2r - \mu$:**

Pick $x, y \in V(\bar{\mathbb{G}})$ such that $xy \in E(\bar{\mathbb{G}}) \Rightarrow xy \notin E(\mathbb{G})$. Therefore $|N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)| = \mu$.

Let $c_1, c_2, \cdots, c_\mu$ be the distinct listing of all elements of $N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)$, and let $v_1, v_2, \cdots, v_t$ be the distinct listing of all elements of $N_{\mathbb{G}}(x)\backslash N_{\mathbb{G}}(y)$, and let $u_1, u_2, \cdots, u_k$ be the distinct listing of all elements of $N_{\mathbb{G}}(y)\backslash N_{\mathbb{G}}(x)$.

Claim 1:

$N_{\bar{\mathbb{G}}}(x) \cap N_{\bar{\mathbb{G}}}(y) = (V(\mathbb{G}) \backslash \{x, y\}) \backslash \{c_1, c_2, \cdots, c_\mu, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}$

To show $LHS \subseteq RHS$: Pick $t \in N_{\bar{\mathbb{G}}}(x) \cap N_{\bar{\mathbb{G}}}(y)$. Then

$$t \in N_{\bar{\mathbb{G}}}(x),$$

$$\Rightarrow t \in V(\bar{\mathbb{G}}),$$

$$\Rightarrow t \in V(\mathbb{G}).$$

Suppose $t = x$. Then $xx \in E(\bar{\mathbb{G}})$, which is a contradiction because that would be a loop, therefore $t \neq x$.
Similarly, suppose $t = y$. Then $yy \in E(\bar{\mathbb{G}})$, which is a contradiction because that would be a loop. Therefore $t \neq y$, and

$$\Rightarrow t \notin \{x, y\}$$

$$\Rightarrow t \in (V(\mathbb{G}) \backslash \{x, y\}).$$

If $t \in \{c_1, c_2, \cdots, c_\mu\}$, then $t \in (N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y))$,

$$\Rightarrow t \in N_{\mathbb{G}}(x).$$

However, $t \in N_{\bar{\mathbb{G}}}(x)$, yields a contradiction. Therefore $t \notin \{c_1, c_2, \cdots, c_\mu\}$.
If $t \in \{v_1, v_2, \cdots, v_t\}$, then $t \in (N_{\mathbb{G}}(x) \backslash N_{\mathbb{G}}(y))$,

$$\Rightarrow t \in N_{\mathbb{G}}(x).$$

However, $t \in N_{\bar{\mathbb{G}}}(x)$ yields a contradiction. Therefore $t \notin \{v_1, v_2, \cdots, v_t\}$.
Finally, choosing $t \in \{u_1, u_2, \cdots, u_k\} \Rightarrow t \in (N_{\mathbb{G}}(y) \backslash N_{\mathbb{G}}(x))$,

$$\Rightarrow t \in N_{\mathbb{G}}(y).$$

However, $t \in N_{\bar{\mathbb{G}}}(y)$ yields a contradiction. Therefore $t \notin \{u_1, u_2, \cdots, u_k\}$,

$$\Rightarrow t \notin \{c_1, c_2, \cdots, c_\mu, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}.$$

Therefore $t \in ((V(\mathbb{G})\backslash\{x,y\})\backslash\{c_1, c_2, \cdots, c_\mu, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}))$

$$\Rightarrow LHS \subseteq RHS.$$

To show $RHS \subseteq LHS$:
Pick $t \in ((V(\mathbb{G})\backslash\{x,y\})\backslash\{c_1, c_2, \cdots, c_\mu, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\})$

$$\Rightarrow t \neq x.$$

Suppose $xt \in E(\mathbb{G})$. Then $t \in N_{\mathbb{G}}(x)$. However, $N_{\mathbb{G}}(x) = (N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)) \cup (N_{\mathbb{G}}(x)\backslash N_{\mathbb{G}}(y))$,

$$\Rightarrow t \in (N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)) \text{ or } t \in (N_{\mathbb{G}}(x)\backslash N_{\mathbb{G}}(y)),$$

$$\Rightarrow t \in \{c_1, c_2, \cdots, c_\mu\} \text{ or } t \in \{v_1, v_2, \cdots, v_t\},$$

which is a contradiction. Hence $xt \notin E(\mathbb{G})xt \in E(\bar{\mathbb{G}})$, and hence

$$t \in N_{\bar{\mathbb{G}}}(x).$$

Similarly, from $t \neq y$ we may show that

$$t \in N_{\bar{\mathbb{G}}}(y).$$

Thus, $t \in (N_{\bar{\mathbb{G}}}(x) \cap N_{\bar{\mathbb{G}}}(y)), \Rightarrow RHS \subseteq LHS$.

Hence, we have established Claim 1:

$$N_{\bar{\mathbb{G}}}(x) \cap N_{\bar{\mathbb{G}}}(y) = (V(\mathbb{G})\backslash\{x,y\})\backslash\{c_1, c_2, \cdots, c_\mu, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}.$$

Claim 2: $\{c_1, c_2, \cdots, c_\mu, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\} \subseteq (V(\mathbb{G})\backslash\{x,y\})$.

Pick $t \in \{c_1, c_2, \cdots, c_\mu, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}$. Then, to show $t \in (V(\mathbb{G})\backslash\{x,y\})$:

(1) If $t = c_i$, then $t \in N_{\mathbb{G}}(x)$ and $t \in N_{\mathbb{G}}(y)$,

$$\Rightarrow t \neq x \text{ and } t \neq y.$$

(2) If $t = v_i$, then $t \in N_{\mathbb{G}}(x) \Rightarrow t \neq x$, (and, by definition of v_i,)

$$\Rightarrow t \notin N_{\mathbb{G}}(y).$$

If $t = y$

$$\Rightarrow tx \in E(\mathbb{G}) \text{ that is } yx \in E(\mathbb{G})$$

which is a contradiction. Hence $t \neq y$.

(3) If $t = u_i$, then $t \in N_{\mathbb{G}}(y) \Rightarrow t \neq y$, (and, by definition of u_i,)

$$\Rightarrow t \notin N_{\mathbb{G}}(x).$$

If $t = x$

$$\Rightarrow ty \in E(\mathbb{G}), \text{ that is } xy \in E(\mathbb{G})$$

which is a contradiction. Hence $t \neq x$.

Therefore, we have established Claim 2:

$$\{c_1, c_2, \cdots, c_\mu, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\} \subseteq (V(\mathbb{G})\backslash\{x, y\}).$$

It follows that:

$$\begin{aligned} |N_{\bar{\mathbb{G}}}(x) \cap N_{\bar{\mathbb{G}}}(y)| &= |(V(\mathbb{G})\backslash\{x, y\})\backslash\{c_1, c_2, \cdots, c_\mu, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}| \\ &= |(V(\mathbb{G})\backslash\{x, y\})| - |\{c_1, c_2, \cdots, c_\mu, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}| \end{aligned} \tag{2.2}$$

Also since $c_1, c_2, \cdots, c_\mu, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k$ are distinct and non-repetitive:

$$c_i \neq c_j, v_i \neq v_j, u_i \neq u_j, \text{ for } i \neq j \text{ and } c_i \neq v_i \neq u_i \text{ for any } i,$$

we have that,

$$\begin{aligned} N_{\mathbb{G}}(x) &= (N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)) \cup (N_{\mathbb{G}}(x) \backslash N_{\mathbb{G}}(y)) \text{ disjoint} \\ |N_{\mathbb{G}}(x)| &= |(N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y))| + |(N_{\mathbb{G}}(x) \backslash N_{\mathbb{G}}(y))| \\ r &= \mu + |(N_{\mathbb{G}}(x) \backslash N_{\mathbb{G}}(y))| \\ r - \mu &= |(N_{\mathbb{G}}(x) \backslash N_{\mathbb{G}}(y))| \\ &= t. \end{aligned}$$

Similarly

$$\begin{aligned} N_{\mathbb{G}}(y) &= (N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)) \cup (N_{\mathbb{G}}(y) \backslash N_{\mathbb{G}}(x)) \text{ disjoint} \\ |N_{\mathbb{G}}(y)| &= |(N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y))| + |(N_{\mathbb{G}}(y) \backslash N_{\mathbb{G}}(x))| \\ r &= \mu + |(N_{\mathbb{G}}(y) \backslash N_{\mathbb{G}}(x))| \\ r - \mu &= |(N_{\mathbb{G}}(y) \backslash N_{\mathbb{G}}(x))| \\ &= k. \end{aligned}$$

Therefore from (2.2) we get:

$$\begin{aligned} |N_{\bar{\mathbb{G}}}(x) \cap N_{\bar{\mathbb{G}}}(y)| = \bar{\lambda} &= |V(\mathbb{G})| - |\{x, y\}| - |c_1, c_2, \cdots, c_\mu, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k| \\ &= n - 2 - (|c_1, c_2, \cdots, c_\mu| + |v_1, v_2, \cdots, v_t| + |u_1, u_2, \cdots, u_k|) \\ &= n - 2 - (\mu + r - \mu + r - \mu) \\ &= n - 2 - (2r - \mu) \\ &= n - 2 - 2r + \mu. \end{aligned}$$

(c) To show $\bar{\mu} = n - 2r + \lambda$:

Pick $x, y \in V(\bar{\mathbb{G}})$ such that $xy \notin E(\bar{\mathbb{G}})$. Then $xy \in E(\mathbb{G})$. Therefore $|N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)| = \lambda$.

Let $c_1, c_2, \cdots, c_\lambda$ be the distinct listing of all elements of $N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)$, and let $v_1, v_2, \cdots, v_t$ be the distinct listing of all elements of $N_{\mathbb{G}}(x) \backslash N_{\mathbb{G}}(y)$, and let $u_1, u_2, \cdots, u_k$ be the distinct listing of all elements of $N_{\mathbb{G}}(y) \backslash N_{\mathbb{G}}(x)$.

Claim 1:

$$N_{\bar{\mathbb{G}}}(x) \cap N_{\bar{\mathbb{G}}}(y) = (V(\mathbb{G}) \backslash \{x, y\}) \backslash \{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}$$

To show $LHS \subseteq RHS$: Pick $t \in N_{\bar{\mathbb{G}}}(x) \cap N_{\bar{\mathbb{G}}}(y) \Rightarrow t \in N_{\bar{\mathbb{G}}}(x)$ and $t \in N_{\bar{\mathbb{G}}}(y)$,

$$\Rightarrow t \in V(\bar{\mathbb{G}}),$$

$$\Rightarrow t \in V(\mathbb{G}.$$

If $t = x$, then $xy \in E(\bar{\mathbb{G}})$ which is a contradiction, therefore $t \neq x$. Similarly, $t = y \Rightarrow xy \in E(\bar{\mathbb{G}})$ which is a contradiction. Therefore $t \neq y$,

$$\Rightarrow t \notin \{x, y\},$$

$$\Rightarrow t \in (V(\mathbb{G})\backslash\{x, y\}).$$

If $t \in \{c_1, c_2, \cdots, c_\lambda\}$, then $t \in (N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y))$,

$$\Rightarrow t \in N_{\mathbb{G}}(x).$$

However, $t \in N_{\bar{\mathbb{G}}}(x)$, which yields a contradiction. Therefore $t \notin \{c_1, c_2, \cdots, c_\lambda\}$.

If $t \in \{v_1, v_2, \cdots, v_t\}$, then $t \in (N_{\mathbb{G}}(x)\backslash N_{\mathbb{G}}(y))$,

$$\Rightarrow t \in N_{\mathbb{G}}(x).$$

However, $t \in N_{\bar{\mathbb{G}}}(x)$, which yields a contradiction. Therefore $t \notin \{v_1, v_2, \cdots, v_t\}$.

Finally, choosing $t \in \{u_1, u_2, \cdots, u_k\}$ then $t \in (N_{\mathbb{G}}(y)\backslash N_{\mathbb{G}}(x))$,

$$\Rightarrow t \in N_{\mathbb{G}}(y).$$

However, $t \in N_{\bar{\mathbb{G}}}(y)$ which yields a contradiction. Therefore $t \notin \{u_1, u_2, \cdots, u_k\}$,

$$\Rightarrow t \notin \{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}.$$

Therefore, $t \in ((V(\mathbb{G})\backslash\{x, y\})\backslash\{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}))$,

$$\Rightarrow LHS \subseteq RHS.$$

To show $RHS \subseteq LHS$:
Pick $t \in ((V(\mathbb{G})\backslash\{x,y\})\backslash\{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\})$,

$$\Rightarrow t \neq x.$$

If $xt \in E(\mathbb{G})$, then $t \in N_{\mathbb{G}}(x)$. However,

$$N_{\mathbb{G}}(x) = (N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)) \cup (N_{\mathbb{G}}(x)\backslash N_{\mathbb{G}}(y)),$$

$$\Rightarrow t \in (N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)) \text{ or } t \in (N_{\mathbb{G}}(x)\backslash N_{\mathbb{G}}(y)),$$

$$\Rightarrow t \in \{c_1, c_2, \cdots, c_\lambda\} \text{ or } t \in \{v_1, v_2, \cdots, v_t\},$$

which yields a contradiction. Hence $xt \notin E(\mathbb{G})$,

$$\Rightarrow xt \in E(\bar{\mathbb{G}}),$$

$$\Rightarrow t \in N_{\bar{\mathbb{G}}}(x).$$

Similarly $t \neq y \Rightarrow t \in N_{\bar{\mathbb{G}}}(y)$. So $t \in (N_{\bar{\mathbb{G}}}(x) \cap N_{\bar{\mathbb{G}}}(y))$. $\Rightarrow RHS \subseteq LHS$.

Hence we have established Claim1:
$N_{\bar{\mathbb{G}}}(x) \cap N_{\bar{\mathbb{G}}}(y) = (V(\mathbb{G})\backslash\{x,y\})\backslash\{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}$.
Moreover,

$$(V(\mathbb{G})\backslash\{x,y\})\backslash\{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}$$

$$= V(\mathbb{G})\backslash\{x, y, c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}.$$

Claim 2: $\{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\} \subseteq V(\mathbb{G})$.

To show $LHS \subseteq RHS$: Pick $t \in \{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}$.
To show $t \in (V(\mathbb{G})\backslash\{x,y\})$:

(1) If $t = c_i$, then $t \in N_{\mathbb{G}}(x)$ and $t \in N_{\mathbb{G}}(y)$,

$$\Rightarrow t \neq x \text{ and } t \neq y.$$

(2) If $t = v_i$, then $t \in N_{\mathbb{G}}(x) \Rightarrow t \neq x$,

$$\Rightarrow t \notin N_{\mathbb{G}}(y).$$

If $t = y$

$$\Rightarrow tx \in E(\mathbb{G}), \text{ that is } yx \in E(\mathbb{G}),$$

which is possible. Hence $t = y$.

(3) If $t = u_i$, then $t \in N_{\mathbb{G}}(y) \Rightarrow t \neq y$,

$$\Rightarrow t \notin N_{\mathbb{G}}(x).$$

If $t = x$

$$\Rightarrow ty \in E(\mathbb{G}) \text{ that is } xy \in E(\mathbb{G}),$$

which is possible. Hence $t = x$.

Hence $x = u_i$ and $y = v_j$.
Therefore $\{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\} \subseteq V(\mathbb{G})$, and
$\{x, y, c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}$
$= \{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}$

It follows that:

$$\begin{aligned} |N_{\bar{\mathbb{G}}}(x) \cap N_{\bar{\mathbb{G}}}(y)| &= |(V(\mathbb{G})\backslash\{x, y\})\backslash\{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}| \\ &= |(V(\mathbb{G})\backslash\{x, y, c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}| \\ &= |(V(\mathbb{G})\backslash\{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}| \\ &= |V(\mathbb{G})| - |\{c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k\}|. \end{aligned} \tag{2.3}$$

Also, since $c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k$ are distinct and non- repetitive, we have that

$$c_i \neq c_j, v_i \neq v_j, u_i \neq u_j, \text{ for } i \neq j \text{ and } c_i \neq v_i \neq u_i \text{ for any } i.$$

Moreover,

$$\begin{aligned} N_{\mathbb{G}}(x) &= (N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)) \cup (N_{\mathbb{G}}(x)\backslash N_{\mathbb{G}}(y)) \text{ disjoint} \\ |N_{\mathbb{G}}(x)| &= |(N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y))| + |(N_{\mathbb{G}}(x)\backslash N_{\mathbb{G}}(y))| \\ r &= \lambda + |(N_{\mathbb{G}}(x)\backslash N_{\mathbb{G}}(y))| \\ r - \lambda &= |(N_{\mathbb{G}}(x)\backslash N_{\mathbb{G}}(y))| \\ &= t. \end{aligned}$$

Similarly

$$\begin{aligned} N_{\mathbb{G}}(y) &= (N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)) \cup (N_{\mathbb{G}}(y) \backslash N_{\mathbb{G}}(x)) \text{ disjoint} \\ |N_{\mathbb{G}}(y)| &= |(N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y))| + |(N_{\mathbb{G}}(y) \backslash N_{\mathbb{G}}(x))| \\ r &= \lambda + |(N_{\mathbb{G}}(y) \backslash N_{\mathbb{G}}(x))| \\ r - \lambda &= |(N_{\mathbb{G}}(y) \backslash N_{\mathbb{G}}(x))| \\ &= k. \end{aligned}$$

Therefore from (2.3) we obtain:

$$\begin{aligned} |N_{\bar{\mathbb{G}}}(x) \cap N_{\bar{\mathbb{G}}}(y)| = \bar{\mu} &= |V(\mathbb{G})| - |c_1, c_2, \cdots, c_\lambda, v_1, v_2, \cdots, v_t, u_1, u_2, \cdots, u_k| \\ &= n - (|c_1, c_2, \cdots, c_\lambda| + |v_1, v_2, \cdots, v_t| + |u_1, u_2, \cdots, u_k|) \\ &= n - (\lambda + r - \lambda + r - \lambda) \\ &= n - (2r - \lambda) \\ &= n - 2r + \lambda. \end{aligned}$$

□

It should be stressed out that not all given sequences of parameters generate a strongly regular graph. For example, Bahman Ahmadi (2009) shows that $(21, 10, 4, 5)$ are not valid parameters for a strongly regular graph. However, the following theorem assures us that should there exist a strongly regular graph of some parameters, then one can complete a sequence given incomplete parameters, via a relationship between them.

Theorem 2.3.3. *Let $\mathbb{G} = (V, E)$ be a strongly regular graph. Then from the parameters discussed in Remark 2.3.1*

$$r(r - \lambda - 1) = (n - r - 1)\mu.$$

Proof. Let $\mathbb{G}$ be a strongly regular graph. Pick an arbitrary fixed vertex, $u \in V(\mathbb{G})$, and let Υ be the set of all vertices in $V(\mathbb{G})$ adjacent to u. Then $|\Upsilon| = r$, since r is the degree of each vertex.

It clearly follows from Theorem 2.3.2 above that the order of all the other

vertices outside this set is:

$$|V(\mathbb{G}) \setminus \Upsilon| = \left|\bar{\Upsilon}\right| = \bar{r} = n - r - 1. \tag{2.4}$$

Let S denote the set of all edges of connecting Υ and $\bar{\Upsilon}$. Then $|S|$ can be calculated in at least two ways:

1. by considering the number of vertices each vertex of Υ is adjacent to; that is, each $v \in \Upsilon$ is adjacent to u and λ other vertices in Υ, since λ vertices are adjacent to both u and v, i.e:

$$\begin{aligned} |S| &= r[r - (\lambda + 1)] \\ &= r(r - \lambda - 1); \end{aligned} \tag{2.5}$$

2. or by picking a vertex in $\bar{\Upsilon}$ and considering the number of vertices it is adjacent to in Υ; that is, each $w \in \bar{\Upsilon}$ is adjacent to μ vertices in Υ, but from Equation (2.4), there are $n - r - 1$ elements in $\bar{\Upsilon}$, so:

$$|S| = (n - r - 1)\mu. \tag{2.6}$$

Therefore from (2.5) and (2.6), we obtain:

$$r(r - \lambda - 1) = (n - r - 1)\mu.$$

□

Example 2.3.4. Consider the parameters $(100, 20, 10, 5)$, and the equation given in Theorem 2.3.3. We can easily verify that there does not exist a strongly regular graph with the given parameters, by establishing that the $LHS \neq RHS$:

$$LHS = r(r - \lambda - 1) = 20(100 - 10 - 1) = 1780$$

which is not equal to:

$$RHS = (n - r - 1)\mu = (100 - 20 - 1)5 = 395 \quad .$$

Hence invalid parameters.

Lemma 2.3.4. *[3] Let $\mathbb{G} = (V, E)$ be a strongly regular graph. Then the following are equivalent:*

i : $\mathbb{G}$ *is not connected;*

ii : $\mu = 0$*;*

iii : $\lambda = r - 1$*;*

iv : *Each component of* $\mathbb{G}$ *is isomorphic to the complete* $(r+1)$*-regular graph.*

Proof. $(i) \Rightarrow (ii)$ Suppose $\mathbb{G}$ is a disconnected graph. Then $\mathbb{G}$ has at least two components $\mathbb{G}_1$ and $\mathbb{G}_2$. Let $x \in V(\mathbb{G}_1)$ and $y \in V(\mathbb{G}_2)$. Then there is no path from x to $y \Rightarrow xy \notin E(\mathbb{G})$. If

$$N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y) \neq \emptyset,$$

then $t \in N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)$ which implies xty is a path in G; a contradiction. Hence, $N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y) = \emptyset$

$$\Rightarrow |N_{\mathbb{G}}(x) \cap N_{\mathbb{G}}(y)| = \mu = 0.$$

$(ii) \Rightarrow (iii)$ Let $\mu = 0$, and assume there exists $u, v, w \in V(\mathbb{G})$, where u and w are neighbors of v. Then u must be adjacent to w, since $\mu = 0$ which implies that each vertex must be adjacent to $r - 1$ other vertices; that is, $\lambda = r - 1$.
$(iii) \Rightarrow (iv)$ Let $\lambda = r - 1$. Then any component of $\mathbb{G}$ is complete, and since degree of $\mathbb{G}$ is r then each component is the complete graph K_{r+1}.
$(iv) \Rightarrow (i)$ Let each component of $\mathbb{G}$ be isomorphic to the complete $(r+1)$-regular graph.

Then all vertices of a component must have same degree, which implies that the components are not connected; that is $\mathbb{G}$ is not connected. □

The following stated results help in proving most of the results that will soon follow in this dissertation, the reader is advised to consult the referenced material for the proofs.

Theorem 2.3.5. *[3] Let* $\mathbb{G}$ *be a strongly regular graph. Then the following expression is true about the adjacency matrix,* A*, of* $\mathbb{G}$*:*

$$A^2 = (\lambda - \mu)A + (r - \mu)I + \mu J,$$

where I and J are the identity and the matrix consisting of all entries equal to one respectively.

Lemma 2.3.6. *[3] Let $\mathbb{G} = (V, E)$ be a strongly regular graph. Then $\mathbb{G}$ has at most three distinct eigenvalues.*

Theorem 2.3.7. *[9] Let $\mathbb{G} = (V, E)$ be a connected r-regular graph. Then $\mathbb{G}$ is strongly regular if and only if it has exactly three distinct eigenvalues, r, s, t.*

Corollary 2.3.8. *[9] Let $\mathbb{G}$ be a strongly regular graph. Then,*

$$\lambda = r + s \cdot t + s + t,$$
$$\mu = r + s \cdot t.$$

The following definition is used to establish relationships between SRG and groups as will be discussed in Chapters 3 and 4.

Definition 2.3.2. Let $(G, \circ)$ be a group of order n. A r-subset Ω of G is called a (n, r, λ, μ)-***Partial Difference Set*** in G if, for any $g, h \in \Omega$ and $g \neq h$, the mathematical expression $g \circ h^{-1}$ represents a non-identity element in Ω exactly λ times and represents a non-identity element in $G \setminus \Omega$ exactly μ times.

Example 2.3.5. Consider $\mathbb{Z}_4$

$$(\mathbb{Z}_4, \oplus) = \{4\mathbb{Z}, 1 + 4\mathbb{Z}, 2 + 4\mathbb{Z}, 3 + 4\mathbb{Z}\}$$

and $S \subset \mathbb{Z}_4$ such that $S = \{1 + 4\mathbb{Z}, 3 + 4\mathbb{Z}\}$. We notice from Figure 2.5 below that this Cayley graph is strongly regular with parameters $(4, 2, 0, 2)$.

Next we show that S is a $(4, 2, 0, 2)$-partial difference set. We observe that: $|Z_4| = 4$ and $|S| = 2$.

Since $S = \{1 + 4\mathbb{Z}, 3 + 4\mathbb{Z}\}$, pick any $s, \omega \in S$ with $s \neq \omega$, compute $s \oplus \omega^{-1} = \{((1 + 4\mathbb{Z}) \oplus (1 + 4\mathbb{Z})), ((3 + 4\mathbb{Z}) \oplus (3 + 4\mathbb{Z}))\} = \{2 + 4\mathbb{Z}, 2 + 4\mathbb{Z}\}$, and notice that none of the non-identity elements in this set also lie in S. Hence $\lambda = 0$. Also, for $s \oplus \omega^{-1} = \{2 + 4\mathbb{Z}, 2 + 4\mathbb{Z}\} = \{2 + 4\mathbb{Z}\}$, each non-identity member of the complement of S in $\mathbb{Z}_4$ appears exactly twice and so $\mu = 2$. Altogether this indicates that S is a $(4, 2, 0, 2)$-partial difference set as per the definition.

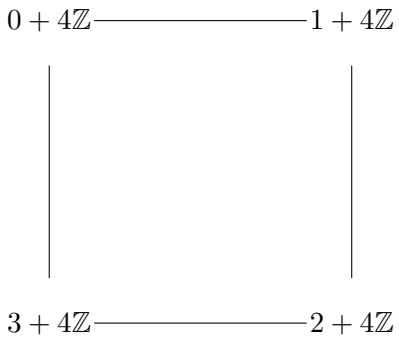


Figure 2.5: Cayley Graph on $(\mathbb{Z}_4)$ and $S \subset \mathbb{Z}_4$

Summary

In this chapter we introduced the concepts of Cayley graphs and that of SRG. We discussed how to identify Cayley graphs and SRG from different sort of graphs by considering Sabiddusi's Theorem for Cayley graphs. We studied some properties that set Cayley graphs aside from other graphs, and looked at their construction. Cayley graphs play a major role in developing the relationship between graph theory and cryptography. The next chapter introduces some notions in cryptography that link well with Cayley graphs and SRG. The discussion of this relationship is the topic of Chapter 4. There we will notice the role Cayley graphs play in stream ciphers, and further extend the notion of Cayley graphs to that of strongly regular Cayley graphs in order to study their relationship with block ciphers.

Chapter 3

Cryptographic Functions

In Chapter 1, we observed that the security of stream ciphers and block ciphers rests upon the randomness of the keystream generators and the design of cryptographically strong s-boxes respectively.

This chapter introduces the properties used to quantify such cryptographic strength. We do so by studying two mathematical functions (Boolean and bent functions) whose properties are suitable for the design of strong pseudo-random number generators and s-boxes.

We will discuss known properties that classify keystream generators as being random enough to provide cryptographic security and s-boxes as being cryptographically strong. We introduce and discuss some known results, and the properties of Boolean and bent functions that make them suitable to the cryptographic needs of pseudo-random number generators and s-boxes respectively. This will complete our background on the mathematical design of keystream generators and s-boxes. This leads us to the next chapter where we will link these functions to suitable algebraic graphs with the required properties for cryptography. This material is drawn from [5], [12], [13], [15], [22], [43].

3.1 Introduction

The study of mathematical techniques to defeat information security, (cryptanalysis), is an ongoing process. Hence, many different attack methods have

been successfully studied and implemented. Cryptography on the other hand responds by investigating these attacks and creating cryptosystems that are less vulnerable to them, and the cycle continues.

Stream ciphers make use of Boolean functions to achieve standard security in pseudo-random number generators because of the properties these functions possess. The Boolean functions focused on here need to be balanced, non-linear and have a high algebraic degree in order to resist a number of known attacks by cryptanalysis.

Block ciphers on the other hand make use of bent functions to achieve security in substitution-boxes. Bent functions are also Boolean functions but they achieve maximum non-linearity and, in order to support the objective of constructing suitable cryptographic substitution-boxes, they need to additonally satisfy the strict avalanche criterion; the bit independence criterion; and also be bijective. Unlike the requirement for stream ciphers, these functions should not be balanced. Additional properties and their relations (such as the Hamming weight of these bent functions) are also considered.

In order to understand the role played by the properties desired for a "nice" cryptographic function, we need to review several cryptanalytic attacks.

The linear approximation attack takes advantage of the linearity of the expression that involves plaintext bits, ciphertext bits and subkey bits [22]. Another old and famous attack, the differential cryptanalysis attack considers the XOR difference between plaintexts and its propagations through nonlinear and linear transformations of a primitive. The correlation attack focuses on the choice of the Boolean function used: it uses this function to regenerate the keystream by combining the outputs of the linear feedback shift registers (LFSR - to be defined later on this chapter). The algebraic attack considers algebraic methods to break the cipher. It expresses the cipher operations as systems of equations and substitutes known information for certain known variables, then it attempts to solve for the key. So the choice of the Boolean function is important.

We introduce Boolean and bent functions and discuss their properties, thus

hinting at techniques to defeat these attacks.

3.2 Boolean Functions

The study of Boolean functions (named after George Boole) is widely discussed in the field of algebraic logic. Boolean functions occur in the study of the mathematical formulation of logical problems. The language of Boolean functions has lately become fundamental to the applications of discrete mathematics, including the analysis and construction of cryptosystems [5], [16].

In this section we explore Boolean functions for cryptographic use in stream ciphers. We explore their mathematical properties and align them with the requirements of cryptography. They will later (in Chapter 4) be compared to the properties of Cayley graphs.

Remark 3.2.1. *Let $\mathbb{F}_2$ denote the finite field of two elements. Then $\mathbb{F}_2$ is closed under addition and multiplication modulo 2. In this context the elements of $\mathbb{F}_2$ are bits and the addition is XOR ($\oplus$).*

Also, let $\mathbb{F}_2^n$ be a $\{0,1\}$ vector space of n tuples, $n \in \mathbb{N}$, such that $X \in \mathbb{F}_2^n$ if and only if $X = (x_1, \cdots, x_n)$, where $x_i \in \{0,1\}$ for all $1 \leq i \leq n$. It is the set of all n-dimensional bit-strings.

*We will therefore refer to $\mathbb{F}_2$ as being the set of all **Boolean values**.*

Definition 3.2.1. Let f be a map that takes the vector $X \in \mathbb{F}_2^n$ and maps it to some $x_i \in \{0,1\}$,

$$f : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2,$$
$$f : X \longmapsto x_i, \quad \text{where } X \in \mathbb{F}_2^n,\ x_i \in \mathbb{F}_2.$$

Then f is called a ***Boolean Function***. We denote B_n to be the set of n-variable Boolean functions, that is, $f \in B_n$.

Remark 3.2.2. *1. $|\mathbb{F}_2^n| = 2^n$, since it is simply n-tuples of $\{0,1\}$, and $|\{0,1\}| = 2$.*

2. $|B_n| = 2^{2^n}$.

Example 3.2.1. Consider the Boolean function

$$f(X) = x_1 \oplus x_2x_3 \oplus x_4, \text{ where } X = (x_1, x_2, x_3, x_4),\ x_i \in \mathbb{F}_2.$$

Then $X \in \mathbb{F}_2^4$.
By Remark 3.2.2

$$|\mathbb{F}_2^4| = 2^4 = 16.$$

Hence in the truth table representation we have 16 rows of 4 columns of inputs and 16 rows of 1 column of output.

Input				**Output**
x_1	x_2	x_3	x_4	$f(x_1, x_2, x_3, x_4)$
0	0	0	0	0
0	0	0	1	1
0	0	1	0	0
0	1	0	0	0
1	0	0	0	1
0	0	1	1	1
0	1	0	1	1
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
0	1	1	1	0
1	0	1	1	0
1	1	0	1	0
1	1	1	0	0
1	1	1	1	1

Table 3.1: Truth table of the 4-variable Boolean Function f

Definition 3.2.2. Let $f \in B_n$. Then f can be expressed in the ***algebraic normal form (ANF)***,

$$\begin{aligned} f(x_1, \cdots, x_n) &= \bigoplus a_t \left(\prod_{i=1}^{n} x_i^{t_i} \right), \\ &= \bigoplus a_t X^t, \\ &= a_0 \oplus a_1 x_1 \oplus a_2 x_2 \oplus \cdots \oplus a_n x_n \oplus a_{1,2} x_1 x_2 \oplus a_{2,3} x_2 x_3 \oplus \\ &\quad \cdots \oplus a_{n-1,n} x_{n-1} x_n \oplus \cdots \oplus a_{1,\ldots,n} x_1 \cdots x_n. \end{aligned}$$

Here, $x_i, t_i, a_t \in \mathbb{F}_2$ and $X, t \in \mathbb{F}_2^n$. Moreover the ***algebraic degree*** of the ANF of f, denoted $deg(f)$, is the number of variables in the highest order term with non-zero coefficient.

Definition 3.2.3. For the same f defined in Definition 3.2.2 above, we define the number of vectors $X \in \mathbb{F}_2^n$, for which $f(X) = 1$, to be the ***Hamming weight*** of f, and we denote that by $wt(f)$:

$$wt(f) = \sum_{X \in \mathbb{F}_2^n} f(X).$$

We can also define/calculate the algebraic degree of an n-variable Boolean function from its Hamming weight:

$$deg(f) = \max \{wt(f) | a_t \neq 0, t \in \mathbb{F}_2^n\}.$$

Moreover, if $wt(f) = wt(f \oplus 1)$ then we call f a ***balanced*** n-variable Boolean function.

We state without proof the following propositions to explain deductions made later such as the link between the definition of a balanced n-variable Boolean function and Proposition 3.2.1.

Proposition 3.2.1. *Let f be a n-variable Boolean function and $wt(f)$ denote the Hamming weight of f. Then $wt(f \oplus 1) = 2^{n-1}$.*

Proposition 3.2.2. *Let f be a n-variable Boolean function. Then $wt(f)$ is odd if and only if $deg(f) = n$.*

Definition 3.2.4. Let $f, g \in B_n$. Then the ***Hamming distance*** between f and g in B_n, (denoted $d(f,g)$), is the number of instances in which corresponding values of the functions differ, that is the number of values of $(x_1, \cdots, x_n)$ for which $f(x_1 \cdots, x_n)$ and $g(x_1, \cdots x_n)$ differ. Thus

$$\begin{aligned} d(f,g) &= wt(f \oplus g) \\ &= |\{X \in \mathbb{F}_2^n | f(X) \oplus g(X) = 1\}|. \end{aligned}$$

From the above definition the following results are clear, proofs to all these results is provided by the referenced material. We state these results to provide clarity to the results that conclude this study, in Chapter 4.

Proposition 3.2.3. *Let $f, g \in B_n$. Then $d(f,g) = 2^{n-1} - \frac{1}{2}$.*

Proposition 3.2.4. *Let d be the Hamming distance of pairs of functions in B_n. Then d is a metric on B_n.*

Proposition 3.2.5. *Let $d(f,g)$ be Hamming distance between f and g. If $\bar{g} = g + 1$ is the negation of g, then*

$$d(f, \bar{g}) = 2^n - d(f,g).$$

Definition 3.2.5. Let $f(X)$ be a n-variable Boolean function such that $X \in \mathbb{F}_2^n$, $f(X) \in \mathbb{F}_2$. Then we define the ***sign function of*** f to be the integer valued function

$$sgn(f(X)) = (-1)^{f(X)}.$$

Moreover, let $Y \in \mathbb{F}_2^n$, such that $Y = (y_1 \cdots y_n)$ and $X \cdot Y = x_1 y_1 \oplus \cdots \oplus x_n y_n$. Then the integer valued function

$$W_f(Y) = \sum_{X \in \mathbb{F}_2^n} (-1)^{f(X) + X \cdot Y}$$

is called the ***Walsh transform*** of a Boolean function f at Y.

Moreover, in general the ***discrete Fourier transform,***

$$W_f^*(Y) = f(X)(-1)^{X \cdot Y},$$

is sometimes used in place of the Walsh transform as they are closely related as follows:

$$W_f^*(Y) = -\frac{1}{2}W_f(Y) + 2^{n-1}\delta(Y),$$

where $\delta(Y)$ is the Kronecker delta function defined as:

$$\delta(Y) = \begin{cases} 1 & \text{if } Y = 0 \\ 0 & \text{if } Y \neq 0. \end{cases}$$

Remark 3.2.3. *Clearly the Walsh transform of a balanced Boolean function f on a 0-vector is given as $W_f(0) = 0$.*

Proposition 3.2.6. *Let $f \in B_n$, $X, Y \in \mathbb{F}_2^n$, such that $k = Y \cdot X$. Then the Walsh transform of f at Y can be given as*

$$W_f(Y) = 2^n - 2wt(f \oplus k).$$

Remark 3.2.4.

If $Y, Z \in \mathbb{F}_2^n$ then $\bar{Y}$ is the complement of Y and to say, $Z \leq \bar{Y}$ means each $z_i \leq \bar{y}_i$ for any i.

Corollary 3.2.7. *[16] Let $f \in B_n$, $X, Y, Z \in \mathbb{F}_2^n$. Then*

$$\sum_{Z \leq Y} W_f^*(Y) = 2^{wt(Y)} \sum_{Z \leq \bar{Y}} f(Y).$$

Definition 3.2.6. Let $g \in B_n$. Then g is said to be ***affine*** (or an ***affine function***) if and only if $deg(g(X)) \leq 1$, where $X \in \mathbb{F}_2^n$.

Moreover, let $f \in B_n$ and A_{B_n} denote the set of all n-variable affine Boolean functions. Then

$$nl(f) = \min_{g \in A_{B_n}} d(f, g)$$

is called the ***nonlinearity of*** f.

Proposition 3.2.8. *[5] Let $f \in B_n$, $X, Y \in \mathbb{F}_2^n$. Then*

$$nl(f) = 2^{n-1} - \frac{1}{2} max_{\, Y \in \mathbb{F}_2^n} \left| W_f(Y) \right|.$$

Definition 3.2.7. Let $f, g \in B_n$, and $g \neq 0$. Then $g \in B_n$ is called a ***annihilator of*** $f \in B_n$ if $f \cdot g = 0$ (where $\cdot$ is defined to be scalar multiplication). The set of all annihilators of $f \in B_n$ is given by

$$AN(g) = \{g \in B_n \mid f \cdot g = 0\}.$$

Moreover, we define ***algebraic immunity of*** f, where $g \neq 0$, to be

$$Al(f) = \min_{g \in \mathbb{F}_2^n}\{deg(g) \mid f \cdot g = 0 \text{ or } (f \oplus 1) \cdot g = 0\}.$$

Definition 3.2.8. Let $f \in B_n$. If for every $Y \in \mathbb{F}_2^n$ and $1 \leq wt(Y) \leq m$, $W_f(Y) = 0$, then f is called m^{th}***-order correlation immune,*** $(cl(m))$. If, moreover, f is balanced then f is called ***m-resilent***.

Proposition 3.2.9. *[16] Let $f \in B_n$ be m-resilent and $0 \leq m \leq n-1$. Then*

$$deg(f) \leq n - m - 1.$$

Proposition 3.2.10. *[16] Let $f \in B_n$ be $(n-1)$-resilent. Then f is affine.*

Definition 3.2.9. Let $f, g \in B_n$, $X, Y \in \mathbb{F}_2^n$, such that $Y \neq 0$. Then the ***autocorrelation function*** of f with respect to Y is given by

$$AC_f(Y) = \sum_X f(X) \cdot f(X \oplus Y).$$

Moreover the ***autocorrelation value of*** f ***with respect to*** Y is given by

$$|AC_f^*(Y)| = \max_{Y \in \mathbb{F}_2^n} \left| \sum_X f(X) \cdot f(X \oplus Y) \right|.$$

The properties described above have been associated with resistance to cryptanalysis in various manners which we now review. The property of ***balancedness*** allows one to distribute the output uniformly and avoid attacks by statistical dependence between plaintext and ciphertex. Hence, the function used for PRNG must be balanced.

Moreover, resistance to correlation attacks on PRNG requires ***correlation immunity of order m,*** $cl(m)$. If $f(X)$ is not $cl(m)$ then an exhaustive

initiation search as an attack reveals that there is a correlation between the output and almost m-bits of the input. Furthermore, if m is relatively small, then the cipher stands the risk of correlation attack (divide and conquer attack).

Low ***algebraic immunity of*** f is always desired for an algebraic attack resistance of the cipher.

High ***nonlinearity*** and high level of ***algebraic degree*** of f is generally a requirement for cryptographic functions so as to resist attack by linear and differential cryptanalysis. The ***Hamming distance***, for all $f, g \in B_n$ and $g \in A_{B_n}$ is desired to be kept high.

The design of cryptographically strong Boolean functions for stream ciphers involves taking into account of all of the above properties as part of the requirements to overcome well researched attacks and possibly new ones. On the other hand there are trade-offs between these properties according to the specific requirements of the cipher.

In stream ciphers, linear feedback shift registers are used in generating the key-stream (pseudo-random sequence) from the key. A ***Linear Feedback Shift Register, (LFSR)***, is a shift register of key bits, a linear function taking the key bits and performing XOR's on them to yield the next bit in the shift register. The output of the LFSR then becomes the input of the (typically nonlinear) Boolean function used to produce the key-stream. Although the methodology would differ depending on the type of generator (combination or filter), the focus here is that, regardless of the type of generator, the output of the LFSR is the input of the Boolean function.

The idea of maximal possible level is an emphasis that the trade offs between properties during the design of a strong Boolean functions is necessary. Methods of designing cryptographically strong Boolean functions include *random generation, algebraic and heuristic* techniques and many others.

Having introduced the concept of Boolean functions we shall investigate to what extent their required properties align with those of Cayley graphs and hence deduce whether cryptographically useful Boolean functions can be use-

fully described in terms of Cayley graphs.

Boolean functions with additional properties (such as maximizing certain properties etc) are generally grouped and classified. In the next section we review some properties of a special type of Boolean functions, the bent functions, and discuss their use in cryptography.

3.3 Bent Functions

Bent functions, like many other mathematical discoveries, do not have a solid recorded beginning. However, results by Rothaus (1976) and Eliseev are some of the earliest mentions of the notion. Since then the study of bent functions has intensified as their properties lend themselves to employment in cryptography, amongst other uses.

In Chapter 2, we explored algebraic graphs. One of the families of graphs we reviewed was the family of strongly regular graphs. In Chapter 4 we shall study the cryptographic strength of block ciphers via the properties of these graphs. In this section we extend the material of Section 3.2 to define and understand bent functions. We explore their nature for cryptographic use in block ciphers, thus distinguishing them from the normal Boolean functions discussed in Section 3.2. Without studying the details of the design of these functions we review their application to the construction of strong substitution boxes for a block cipher.

Proposition 3.3.1 is the basis from which one of the properties of bent functions is drawn; the proof is explained in the reference:

Proposition 3.3.1. *[5] Let $f : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2$ be an unbalanced Boolean function with $n = 2k$, $k \in \mathbb{Z}$. Then the upper bound for nonlinearity is*

$$nl(f) \leq 2^{n-1} - 2^{\frac{n}{2}-1}.$$

Remark 3.3.1. *[40] If f is a n-variable bent functions, and g any affine function then $f \oplus g$ is also a bent function. It then follows that the hamming weight of any bent function is given as $wt(f) = 2^{n-1} \pm 2^{\frac{n}{2}-1}$.*

Definition 3.3.1. Let $f : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2$ be an n-variable unbalanced Boolean function for even n. Then f is said to be a ***bent function*** if and only if the Hamming distance,

$$d(f, g) = 2^{n-1} - 2^{\frac{n}{2}-1}, \text{ for all } g \in A_{B_n},$$

where A_{B_n} denotes the set of all n-variable affine Boolean functions. We denote by BB_n the set of n-variable bent functions.

Remark 3.3.2. *1. If $f : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2$ is a bent function, then n is even.*

2. If $f \in BB_n, Y \in \mathbb{F}_2^n$ and $1 \leq wt(Y) \leq n$, then $f(X) \oplus f(X \oplus Y)$ is balanced, where $X \in \mathbb{F}_2^n$.

Example 3.3.1. Consider the Boolean function

$$f(X) = x_1 \cdot x_2 \oplus x_3 \cdot x_4, \text{ where } X = (x_1, x_2, x_3, x_4),\ x_i \in \mathbb{F}_2,$$

to be bent, then

$$\begin{aligned} nl(f) \leq d(f, g) &= 2^{4-1} - 2^{\frac{4}{2}-1} \\ &= 6, \end{aligned}$$

If we let $Y = 1011 \in \mathbb{F}_2^4$ then $1 \leq wt(Y) = 3 \leq 4$.

Next we consider the truth table representation of $f(X)$, $f(X \oplus Y)$ and $f(X) \oplus f(X \oplus Y)$:

Input				**Output**
x_1	x_2	x_3	x_4	$f(X)$
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	1	0	0	0
1	0	0	0	0
0	0	1	1	1
0	1	0	1	0
0	1	1	0	0
1	0	0	1	0
1	0	1	0	0
1	1	0	0	1
0	1	1	1	1
1	0	1	1	1
1	1	0	1	1
1	1	1	0	1
1	1	1	1	0

Table 3.2: Truth Table of $f(X) = x_1 \cdot x_2 \oplus x_3 \cdot x_4$

.

Input				**Output**
$x_1 \oplus y_1$	$x_2 \oplus y_2$	$x_3 \oplus y_3$	$x_4 \oplus y_4$	$f(X \oplus Y)$
1	0	1	1	1
1	0	1	0	0
1	0	0	1	0
1	1	1	1	0
0	0	1	1	1
1	0	0	0	0
1	1	1	0	1
1	1	0	1	1
0	0	1	0	0
0	0	0	1	0
0	1	1	1	1
1	1	0	0	1
0	0	0	0	0
0	1	1	0	0
0	1	0	1	0
0	1	0	0	0

Table 3.3: Truth Table of $f(X \oplus Y) = (x_1 \oplus y_1) \cdot (x_2 \oplus y_2) \oplus (x_3 \oplus y_3) \cdot (x_4 \oplus y_4)$

.

$f(X)$	$f(X \oplus Y)$	$f(X) \oplus f(X \oplus Y)$
0	1	1
0	0	0
0	0	0
0	0	0
0	1	1
1	0	1
0	1	1
0	1	1
0	0	0
0	0	0
1	1	0
1	1	0
1	0	1
1	0	1
1	0	1
0	0	0

Table 3.4: Truth Table of $f(X) \oplus f(X \oplus Y)$

.

Remark 3.3.2 claims that if f is bent then $f(X) \oplus f(X \oplus Y)$ is balanced, which in this case is true since

$$\begin{aligned} wt(f(X) \oplus f(X \oplus Y)) &= \sum_{X \in \mathbb{F}_2^4} (f(X) \oplus f(X \oplus Y)) \\ &= 8, \end{aligned}$$

which coincides with Proposition 3.2.1 which says $f(X) \oplus f(X \oplus Y)$ is balanced if

$$\begin{aligned} wt(f(X) \oplus f(X \oplus Y)) &= 2^{n-1} \\ &= 2^{4-1} \\ &= 8. \end{aligned}$$

Definition 3.3.2. Let $f \in BB_n$. Then the ***Walsh spectrum*** of f at Y is defined to be:

$$|W_f(Y)| = 2^{\frac{n}{2}},$$

where $W_f(Y) = \pm 2^{\frac{n}{2}}$, for all Y, is the Walsh transform of f at $Y \in \mathbb{F}_2^n$.

Definition 3.3.3. Let $f_i \in BB_n$, where $i = 1, \ldots, m$. Then an ***S-box*** is defined to be:

$$f : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^m,$$

such that each $f_i : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2$ forms a column of the s-box, where the input bits gives the position and the entry gives the output.

Remark 3.3.3. *1. An s-box is a collection of m highly nonlinear Boolean functions, (bent functions).*

2. Positions of an entry in a s-box starts from row 0, column 0.

Example 3.3.2. Consider the bent function defined in Example 3.3.1. Let $f : \mathbb{F}_2^4 \longrightarrow \mathbb{F}_2^4$ be an s-box. Then $f_1(X) = x_1 \cdot x_2 \oplus x_3 \cdot x_4$ forms the first column of f while f_2, f_3, f_4 occupy columns $2, 3, 4$ respectively as follows. Suppose we choose another bent function, $f_2(X) = 1 \oplus x_1 \cdot x_2 \oplus x_1 \cdot x_3 \oplus x_1 \cdot x_4 \oplus x_2 \cdot x_3 \oplus x_2 \cdot x_4 \oplus x_3 \cdot x_4$, and some bent functions, $f_3(X)$ and $f_4(X)$, where $X = (x_1, x_2, x_3, x_4)$, $x_i \in \mathbb{F}_2$:

Input				Output			
x_1	x_2	x_3	x_4	$f_1(X)$	$f_2(X)$	$f_3(X)$	$f_4(X)$
0	0	0	0	0	·	·	·
0	0	0	1	0	·	·	·
0	0	1	0	0	·	·	·
0	1	0	0	0	1	1	0
1	0	0	0	0	·	·	·
0	0	1	1	1	·	·	·
0	1	0	1	0	·	·	·
0	1	1	0	0	·	·	·
1	0	0	1	0	·	·	·
1	0	1	0	0	·	·	·
1	1	0	0	1	·	·	·
0	1	1	1	1	0	1	0
1	0	1	1	1	·	·	·
1	1	0	1	1	·	·	·
1	1	1	0	1	·	·	·
1	1	1	1	0	·	·	·

Table 3.5: Truth Table of the S-box $f : \mathbb{F}_2^4 \longrightarrow \mathbb{F}_2^4$

.

Assuming that the 4^{th} row of the truth table is as shown above, then the input bits are 0100 which corresponds to, outer elements $00 = 0$ in decimals and gives the row position of the entry, and the middle elements $10 = 2$ in decimals, giving the column position of the entry.

Now the output bits are $6_{10} = 0110_2$, which is the entry.
Hence, labelling this s-box S1:

S1			
·	·	6	·
·	·	·	10
·	·	·	·
·	·	·	·

Table 3.6: S-box 1 $f : \mathbb{F}_2^4 \longrightarrow \mathbb{F}_2^4$

.

Similarly if we assume that the 12^{th} row of the truth table to be as given then: Input bits are 0111, where 01= row 1 of the s-box and 11=column 3 of the s-box. The output bits are $10_{10} = 1010_2$, which is the entry. Similar calculations are performed for all entries of the truth table to construct the entire s-box.

Since nonlinearity is very important in constructing secure s-boxes, special types Boolean functions have been classified as attaining maximum nonlinearity (amongst other cryptographic requirements) and these have been used to build attack resistant block ciphers. Among these other cryptographic requirements we consider the strict avalanche criterion and the propagation criterion, and evaluate their link with bent functions.

Definition 3.3.4. Let $f \in BB_n$. Then f is said to satisfy the ***Strict Avalanche Criterion, (SAC),*** if flipping/ changing a single input bit $x_i \in X \in \mathbb{F}_2^n$ results in the output bits changing exactly half the time.

We state without proof the following lemma and provide reference to the proof.

Lemma 3.3.2. *[16] Let $f \in BB_n$, such that $\sum_X f(X) \oplus f(X \oplus Y) = 2^{n-1}$, for any $X, Y \in \mathbb{F}_2^n$. Then f satisfies the SAC if and only if $wt(Y) = 1$.*

Corollary 3.3.3. *Let $f \in BB_n$, such that $n > 2$, and $deg(f) = n$. Then f does not satisfy the SAC.*

Proof. Let $f \in BB_n$ with $n > 2$ and $deg(f) = n$. Then $n = 2t$ for some integer $t > 1$

$\Rightarrow deg(f) = 2t$ for some integer $t > 1$.

Consider Lemma 3.3.2 and Remark 3.3.2. Since f is bent we have:

$$\sum_X f(X) \oplus f(X \oplus Y) = 2^{n-1}.$$

All that remains to be shown is $wt(f)$ not even, and $wt(f) \neq 1$.

Assume $wt(f)$ is even. Then by Proposition 3.2.2, $deg(f) \neq n$, which is a contradiction.

Hence $wt(f)$ is not even $\Rightarrow wt(f)$ is odd.

Since $wt(f)$ is odd and $\sum_X f(X) \oplus f(X \oplus Y) = 2^{n-1}$,

$$\begin{aligned} wt(f) &= \sum_X f(X) \\ &= \sum_X f(X \oplus Y) \\ &= \frac{1}{2} \sum_X f(X) \oplus f(X \oplus Y), \text{ where } Y \in \mathbb{F}_2^n \\ &= \frac{1}{2}(2^{n-1}), \text{ for some integer } n > 1 \\ &= 2^{n-2} \\ &> 2, \text{ since } n > 1 \text{ and } wt(f) \text{ is odd.} \end{aligned}$$

Hence $wt(f) \neq 1$. Therefore by Lemma 3.3.2 above f does not satisfy the SAC. □

Definition 3.3.5. Let $f \in BB_n$. Then f is said to satisfy the ***Propagation Criterion*** of degree (l), denoted $PC(l)$, if flipping/ changing k input bits $x_i \in X \in \mathbb{F}_2^n$, for $1 \leq k \leq l,\ 1 \leq i \leq n$, results in the output bits changing exactly half the time.

Remark 3.3.4. *The propagation criteria $-PC(l)-$ is a general case of the Strict Avalanche Criterion, $PC(1)$.*

Lemma 3.3.4. *Let $f \in BB_n$, and $X, Y \in \mathbb{F}_2^n$ such that $wt(Y) = l$, where $0 \leq l \leq n$. Then $f(X)$ is $PC(l)$ if and only if*

$$\sum_{Z \leq \bar{Y}} W_f(Z \oplus V)^2 = 2^{wt(\bar{Y}) + wt(Y)}, \qquad \text{where } V, X, Y, Z \in \mathbb{F}_2^n.$$

Proof. Let $f \in BB_n$ and $wt(f) = l$, where $0 \leq l \leq n$. Then

$$\begin{aligned}\sum_{Z\leq\bar{Y}} W_f\left(Z\oplus V\right)^2 &= \sum_{Z\leq\bar{Y}}\left[\sum_X (-1)^{f(X)\oplus(Z\oplus V)\cdot X}\right]^2 \\ &= \sum_{Z\leq\bar{Y}}\left[\left(\sum_X (-1)^{f(X)\oplus(Z\oplus V)\cdot X}\right)\left(\sum_X (-1)^{f(X)\oplus(Z\oplus V)\cdot X}\right)\right] \\ &= \sum_{Z\leq\bar{Y}}\sum_{X,K\in\mathbb{F}_2^n} (-1)^{f(X)\oplus f(K)\oplus(Z\oplus V)\cdot(X\oplus K)} \\ &= \sum_{Z\leq\bar{Y}} (-1)^{Z\cdot(X\oplus K)} \sum_{X,K\in\mathbb{F}_2^n} (-1)^{f(X)\oplus f(K)\oplus V\cdot(X\oplus K)}.\end{aligned} \tag{3.1}$$

Considering Corollary 3.2.7 we have (3.1) as:

$$\begin{aligned}\sum_{Z\leq\bar{Y}} W_f\left(Z\oplus V\right)^2 &= 2^{wt(\bar{Y})} \sum_{X,K\in\mathbb{F}_2^n} (-1)^{f(X)\oplus f(K)\oplus V\cdot(X\oplus K)} \\ &= 2^{wt(\bar{Y})} \sum_{X\oplus K\leq Y} (-1)^{f(X)\oplus f(K)\oplus V\cdot(X\oplus K)} \\ &= 2^{wt(\bar{Y})} \sum_{X\oplus K\leq Y} (-1)^{V\cdot(X\oplus K)} \sum_{X\oplus K\leq Y} (-1)^{f(X)\oplus f(X\oplus(K\oplus X))}.\end{aligned} \tag{3.2}$$

By same Corollary 3.2.7 (3.2) becomes:

$$\begin{aligned}\sum_{Z\leq\bar{Y}} W_f\left(Z\oplus V\right)^2 &= 2^{wt(\bar{Y})}\cdot 2^{wt(Y)} \sum_{X\oplus K\leq Y} (-1)^{f(X)\oplus f(X\oplus K\oplus X))}. \\ &= 2^{wt(\bar{Y})+wt(Y)} \sum_{X\oplus K\leq Y} (-1)^{f(X)\oplus f(X\oplus K\oplus X))}.\end{aligned} \tag{3.3}$$

Now since we are considering a bent Boolean function, by Remark 3.3.2, the number of zero's and one's produced by $f(X)\oplus f(X\oplus(K\oplus X))$ are equal,

$\Rightarrow (-1)^{f(X)\oplus f(X\oplus(K\oplus X))}$ gives equal number of -1's and 1's.

Therefore 3.3 is equal to $2^{wt(\bar{Y})+wt(Y)}$. □

Summary

In this chapter we considered private-key cyptography, by focussing on the cryptographic functions that are used in stream and block ciphers. We defined Boolean functions and discussed the properties that make them cryptographically useful. We further investigated the likelihood of Boolean functions to resist different attacks by considering some cryptographic requirements for cryptographic functions.

We then extended our analysis to a special class of Boolean function (the bent functions), evaluated their strength with respect to certain attacks, and discussed how it achieves the upper bound of one of the discussed cryptographic properties; nonlinearity.

In the next chapter we will consider the relationship between algebraic graphs (the Cayley graphs and strongly regular graphs discussed in the previous chapter) and the cryptographic functions discussed in this chapter to explore the possibilities of interpreting the properties of a stream and/or block cipher through its associated graph.

Chapter 4

Algebraic Graph Theory applied to Cryptographic Functions

The main objective of the study carried out in this dissertation is to investigate and discuss the links between algebraic graphs and symmetric cryptography.

In this chapter we reconsider the properties and results discussed in Chapters 2 and 3, and we use these properties and results to elucidate the connections between cryptography based on Boolean and bent functions on the one hand, and characterizations of these in terms of particular graphs, on the other hand. This allows one to draw conclusions about joint properties. This material is drawn from [7], [9], [33], [38].

4.1 Introduction

In Section 4.2 we consider the Cayley graph associated with a Boolean function, and use its spectral information to investigate the cryptographic properties of the stream cipher. In Section 4.3 a similar investigation is carried out for strongly regular graphs and bent functions as applied to building substitution boxes for block ciphers.

A cipher is said to be cryptographically strong if it can resist almost every known attack. The term "cryptographically strong" is commonly in use even though it is imprecise, in the sense that ciphers are generally rated in comparison to other existing ciphers in their ability to resist a number of attacks that have been investigated in cryptanalysis literature. To this end one seeks to ensure that a cipher satisfies (as a minimum) known math properties, such as balanceness of the Boolean function in use, that is, $wt(f) = wt(f \oplus 1)$, and other properties described in Chapter 3. This chapter describes how one can make some of these cryptographic decisions about a cipher by studying its associated graph.

4.2 Boolean functions characterized by Cayley graphs

The security of stream ciphers relies on the design of cryptographically strong Boolean functions to account for the production of pseudo-random sequences. Stream ciphers were first introduced by Gilbert Sandford Vernam in 1917 and for that reason they are sometimes referred to as the Vernam Ciphers.

In this section we compare properties of Cayley graphs introduced in Section 2.2 with the cryptographic requirements for Boolean functions to be cryptographically strong discussed in Section 3.2. We construct an associated Cayley graph, and from this graph we determine the strength of the cipher by reading off some Boolean function properties.

Recall that Cayley graphs are those graphs constructed via groups, as discussed in Section 2.2. We consider a Boolean function as defined in the preceding chapter, (a map from a vector space of n-tuples with elements from $\mathbb{F}_2$). It can be shown that $\mathbb{F}_2^n$ is a group under XOR, which in this study we use to construct the associated Cayley graph.

The following definition follows directly from the Definition 2.2.1:

Definition 4.2.1. Let $(\mathbb{F}_2^n, \oplus)$ be a group, f a Boolean function, $\Omega_{wt(f)} = \{\omega \in \mathbb{F}_2^n \mid f(\omega) = 1\}$, set of elements making up the Hamming weight of f, such that $\Omega_{wt(f)} \subset \mathbb{F}_2^n$ and $\forall \omega \in \Omega_{wt(f)}$ we have $\omega^{-1} \in \Omega_{wt(f)}$. Then the ***Cayley graph associated with the Boolean function***, $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}) = (V, E)$, is the graph with the following properties:

(i) $V = \{X \mid X \in \mathbb{F}_2^n\}$;

(ii) $E = \{XY \mid Y = X \oplus \omega \text{ for } \omega \in \Omega_{wt(f)},\ X, Y \in \mathbb{F}_2^n\}$
$= \{XY \mid X \oplus Y = X \oplus X \oplus \omega\}$
$= \{XY \mid X \oplus Y \in \Omega_{wt(f)},\ X, Y \in \mathbb{F}_2^n\}$
$= \{XY \mid f(X \oplus Y) = 1, X, Y \in \mathbb{F}_2^n\}$.

Example 4.2.1. Let $f \in B_3$, $f(X) = x_1x_3 \oplus x_2$.

Since $V(\mathcal{G}_f(\mathbb{F}_2^3, \Omega_{wt(f)})) = \mathbb{F}_2^3$, then

$$\left|V(\mathcal{G}_f(\mathbb{F}_2^3, \Omega_{wt(f)}))\right| = \left|\mathbb{F}_2^3\right| = 2^3.$$

Considering the corresponding truth table of the Boolean function f,

Input			**Output**
x_1	x_2	x_3	$f(x_1, x_2, x_3)$
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	0

Table 4.1: Truth Table of $f \in B_3$, $f(X) = x_1x_3 \oplus x_2$

.

we obtain $\Omega_{wt(f)} = \{010, 011, 101, 110\}$. From this we notice that any pair of vertices $X, Y \in \mathbb{F}_2^3$ is adjacent if $X \oplus Y$ is any one of the above elements that give output 1, which follows from the definition that:

$E(\mathcal{G}_f(\mathbb{F}_2^3, \Omega_{wt(f)})) = \{XY \mid f(X \oplus Y) = 1, X, Y \in \mathbb{F}_2^3\}$.

This yields the Cayley graph associated to the Boolean function:

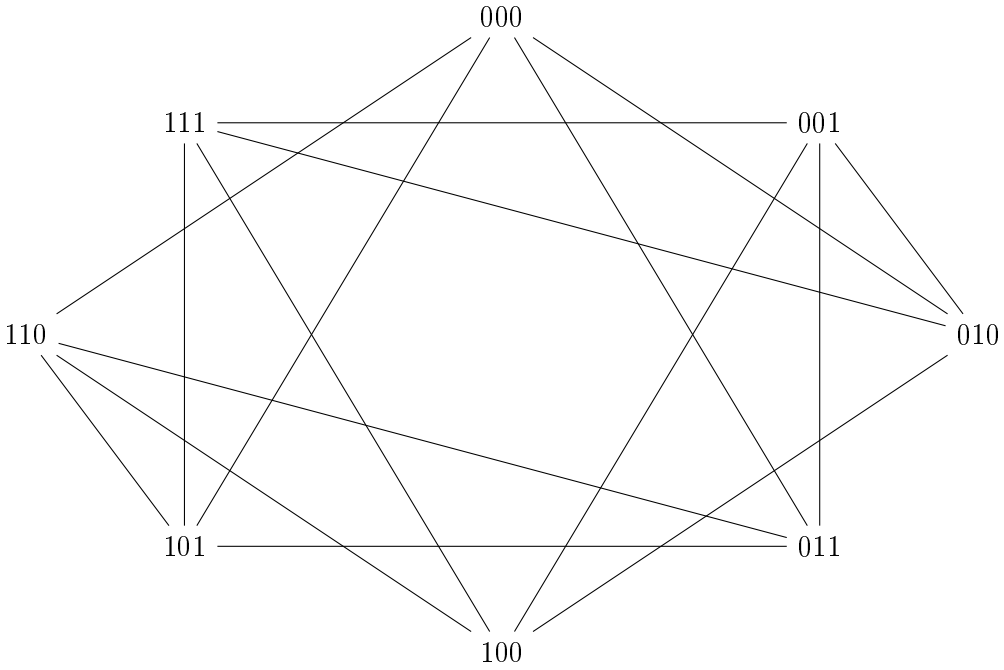


Figure 4.1: Cayley graph associated with the Boolean function $f \in B_3$

In Chapter 1 we defined the notions of an adjacency matrix and the spectrum of a graph $\mathbb{G} = (V, E)$. Hence, we consider the adjacency matrix of Cayley graphs.

Definition 4.2.2. Let $\mathcal{G}_f = (V, E)$ be a Cayley graph associated with a given Boolean function $f \in B_n$, and $\mathbf{b(i)}, \mathbf{b(j)} \in \mathbb{F}_2^n$ being the binary representation of integers, i and j, rows and columns of the corresponding adjacency matrix respectively, such that $0 \leq i, j \leq n-1$. Then, from the definition of a Cayley graph associated with a Boolean function, the adjacency matrix of this graph is easily attained by:

$$[a_{ij}]_{n\times n} = f(\mathbf{b}(i) \ \oplus \ \mathbf{b}(j)).$$

We state without proof the following propositions which are useful in obtaining the adjacency matrix of $\mathcal{G}_f$;

Proposition 4.2.1. *[7] Addition mod-2 of binary representation of numbers has the property:*

$$i \oplus j = (i + 2^n) \oplus (j + 2^n) = j \oplus i$$

for $i, j \in \mathbb{N}_0$ *such that* $0 \leq i, j \leq 2^n - 1$. *Whence the above matrix has the following property:*

$$[a_{ij}]_{n\times n} = \left[a_{i+2^{n-1},j+2^{n-1}}\right]_{n\times n} = \left[a_{j+2^{n-1},i+2^{n-1}}\right]_{n\times n} = [a_{ji}]_{n\times n}.$$

Proposition 4.2.2. *Let* $[a_{ij}]_{n\times n}$ *be the adjacency matrix of the Cayley graph* $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}) = (V, E)$. *Then* $\sum_{i\ fixed}[a_{ij}] = wt(f)$, *where* $wt(f)$ *is the Hamming weight of* $f \in B_n$.

Remark 4.2.1. *A Cayley graph associated with a Boolean function* $f \in B_n$ *is* $wt(f)$*-regular, since* $\left|\Omega_{wt(f)}\right| = wt(f)$ *and* $\Omega_{wt(f)} \subset \mathbb{F}_2^n$, *where for all* $\omega \in \Omega_{wt(f)}$ *there is* $\omega^{-1} \in \Omega_{wt(f)}$.

In what follows we discuss the nature of a strong link between Cayley graphs and cryptographic Boolean functions by presenting a spectral perspective, where the spectral information of a Cayley graph can give necessary but not sufficient results about the strength of the designed Boolean function. We recall the properties to consider when determining the capability of a cipher to withstand some known attacks; these include the balancedness of the Boolean function for resistance against statistical dependence, and others discussed in the preceding chapter. In particular, the Walsh transform of a cryptographic function can be obtained from the eigenvalues of the associated Cayley graph. We also discuss the possibility of investigating the ability of a cipher to resist correlation attack, by examining the spectrum of the Cayley graph associated with the Boolean function and from it concluding whether the function is m^{th}-correlation immune (or resilent) or not.

The next theorem, (Theorem 4.2.3), paves the way for the results that conclude and explain the relationship between algebraic graphs and cryptographic functions. The proof to Theorem 4.2.3 is given in the referenced material. The results that follow are then proved from this theorem.

Theorem 4.2.3. *[9] Let $f \in B_n$, define*

$$\lambda_i = 2^n W_f^*(\boldsymbol{b}(i)) \ \text{ for, } \ 0 \le i, j \le 2^n - 1.$$
$$\text{Then } \ \{\lambda_i\} = Spec\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}).$$

Proposition 4.2.4. *Let $f \in B_n$. Then f is $(cl(m))$ if and only if*

$$\lambda_i \in Spec(\mathcal{G}_f), \ \lambda_i = 0, \ \text{ for all } \ 1 \le wt(b(i)) \le m.$$

*Moreover, f is **m-resilent** if and only if $\lambda_i = 0$, for all $1 \le wt(b(i)) \le m$ and $\lambda_0 = 2^{n-1}$.*

Proof. Let f be an n-variable Boolean function.
" $\Rightarrow$ ": If f is m^{th}-order correlation immune, then

$W_f(b(i)) = 0, \ \text{ for, } \ 0 \le i \le 2^n - 1.$
From Theorem 4.2.3:

$$\begin{aligned} \lambda_i &= 2^n W_f^*(\mathbf{b}(i)), \ \text{ for all } \ 1 \le wt(b(i)) \le m \\ &= 2^n \left(-\frac{1}{2} W_f(b(i)) + 2^{n-1} \delta\left(b(i)\right) \right) \\ &= -2^{n-1} W_f(b(i)) + 2^{2n-1} \delta(b(i)). \end{aligned} \tag{4.1}$$

Recall that,

$$\delta(b(i)) = \begin{cases} 1 & \text{if } b(i)) = 0 \\ 0 & \text{if } b(i)) \neq 0. \end{cases}$$

However, $1 \le wt(b(i)) \le m \Rightarrow \delta(b(i)) = 0$, since for $wt(b(i)) > 0$ we must have $b(i) \neq 0$.

Then (4.1) becomes

$$\lambda_i = -2^{n-1} W_f(b(i)).$$

Also, we have that $W_f(b(i)) = 0$. Hence, $\lambda_i = 0$.

" $\Leftarrow$ ": Now, assume $\lambda_i \in Spec(\mathcal{G}_f), \ \lambda_i = 0, \ \text{ for all } \ 1 \le wt(b(i)) \le m.$

Then following from Theorem 4.2.3

$$0 = \lambda_i = -2^{n-1} W_f(b(i)) + 2^{2n-1}\delta(b(i)).$$

$$\text{Hence, } \; W_f(b(i)) = 2^n \delta(b(i))$$

$$= \begin{cases} 1 & \text{if } b(i)) = 0 \\ 0 & \text{if } b(i)) \neq 0. \end{cases} \tag{4.2}$$

However, $1 \leq wt(b(i)) \leq m \Rightarrow \delta(b(i)) = 0$, since for $wt(b(i)) > 0$ we must have $b(i) \neq 0$.

Then (4.2) becomes

$$W_f(b(i)) = 0.$$

Similarly, to demonstrate m-resilence we proceed as follows:
" $\Rightarrow$ ": If f is m-resilent, then $wt(f) = 2^{n-1}$ and $W_f(b(i)) = 0$, for, $0 \leq i \leq 2^n - 1$.
So, from Theorem 4.2.3,

$$\lambda_i = 2^n W_f^*(\mathbf{b}(i)), \text{ for all } 1 \leq wt(b(i)) \leq m,$$

and it follows (in a similar fashion to the presented above) that $\lambda_i = 0$.

Also, by definition, $\lambda_0 = r = wt(f)$. However, since f is m-resilent, f is balanced. Hence, $wt(f) = 2^{n-1} \Rightarrow \lambda_0 = 2^{n-1}$

" $\Leftarrow$ ": Now, assume $\lambda_i \in Spec(\mathcal{G}_f)$, $\lambda_i = 0$, for all $1 \leq wt(b(i)) \leq m$ and $\lambda_0 = 2^{n-1}$. Then, by definition, $\lambda_0 = wt(f) \Rightarrow wt(f) = 2^{n-1} = wt(f \oplus 1)$. Thus, f is balanced.

Also, by a similar technique as that used above, $W_f(b(i)) = 0$. Hence, since $W_f(b(i)) = 0$ and f is balanced, f must be m-resilent.

□

Theorem 4.2.5. *Let $f \in B_n, \left|Spec(\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}))\right| = 2$, such that $\lambda_0 \neq \lambda_1$, for $\lambda_0, \lambda_1 \in Spec(\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}))$. Then the connected components of $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$ are complete graphs. Moreover, $\Omega_{wt(f)} \cup \{\boldsymbol{b}(0)\}$ is a group, where $\boldsymbol{b}(0) \in \mathbb{F}_2^n$.*

Proof. Let $\mathcal{G}_f$ be a Cayley graph associated to a Boolean function with two

distinct eigenvalues. Then, from Proposition 1.2.5, if $|Spec(\mathcal{G}_f)| = s+1$,

$$diam(\mathcal{G}_f)\,|Spec(\mathcal{G}_f)| - 1 = 1.$$

Hence any connected components of $\mathcal{G}_f$ are complete graphs.

Next we show that $\left(\Omega_{wt(f)} \cup \{\mathbf{b}(0)\}, \oplus\right)$ meets all properties of a group;

(i) Pick any pair of $\omega_i \in \Omega_{wt(f)}$ for any $0 \leq i \leq n-1$, (say ω_1 and ω_2).
Then, since $diam(\mathcal{G}_f) \leq 1$, for any connected component,

$$d(\omega_1, \omega_2) = 1,$$

$\Rightarrow$ any pair of $\omega_i's$ is adjacent,
$\Rightarrow f(\omega_1 \oplus \omega_2) = 1$,
$\Rightarrow \omega_1 \oplus \omega_2 = \Omega_{wt(f)}$,
Hence, $\left(\Omega_{wt(f)} \cup \{\mathbf{b}(0)\}, \oplus\right)$ is closed under $\oplus$.

(ii) Let $\omega_1, \omega_2, \omega_3 \in \left(\Omega_{wt(f)} \cup \{\mathbf{b}(0)\}, \oplus\right)$. Then, since $\oplus$ is associative;

$$(\omega_1 \oplus \omega_2) \oplus \omega_3 = \omega_1 \oplus (\omega_2 \oplus \omega_3)\,.$$

(iii) Let $\omega_i, \mathbf{b}(0) \in \left(\Omega_{wt(f)} \cup \{\mathbf{b}(0)\}, \oplus\right)$. Then, since any n-dimensional vector XORed with the 0-vector returns the same vector, and XOR is symmetric, it therefore, suffices to say there is a 0-vector $\mathbf{b}(0) \in \left(\Omega_{wt(f)} \cup \{\mathbf{b}(0)\}, \oplus\right)$ such that $\omega_i \oplus \mathbf{b}(0) = \omega_i = \mathbf{b}(0) \oplus \omega_i$.

(iv) Let $\omega_i \in \left(\Omega_{wt(f)}\right)$ for any i. Then, by the definition of $\Omega_{wt(f)}$, for all $\omega_i \in \Omega_{wt(f)}$ there exists $\omega_j \in \Omega_{wt(f)}$ such that $\omega_i \oplus \omega_j = \mathbf{b}(0) = \omega_j \oplus \omega_i$ for any i and $j \Rightarrow \omega_j = \omega_i^{-1}$.
Also, since $\mathbf{b}(0)^{-1} = \mathbf{b}(0)$, for every $\omega_i \in \left(\Omega_{wt(f)} \cup \{\mathbf{b}(0)\}\right)$,
there exists $\omega_j \in \left(\Omega_{wt(f)} \cup \{\mathbf{b}(0)\}\right)$, such that,

$$\omega_i \oplus \omega_j = e_{\Omega_{wt(f)} \cup \{\mathbf{b}(0)\}} = \mathbf{b}(0) = \omega_j \oplus \omega_i.$$

Hence, it is clear that $\left(\Omega_{wt(f)} \cup \{\mathbf{b}(0)\}, \oplus\right)$ is a group.

□

Corollary 4.2.6. *Let $f \in B_n, \left|Spec(\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})\right| = 2$, such that $\lambda_0 \neq \lambda_1$, for $\lambda_0, \lambda_1 \in Spec(\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}))$. If $\boldsymbol{b}(0) \in \Omega_{wt(f)}$, then*

$$\lambda_0 = \left|\Omega_{wt(f)}\right| \quad and \quad \lambda_1 = 0,$$

where $\boldsymbol{b}(0) \in \mathbb{F}_2^n$.

Proof. Let $\mathbf{b}(0) \in \Omega_{wt(f)}$. Then, $\Omega_{wt(f)} \cup \{\mathbf{b}(0)\} = \Omega_{wt(f)}$. By definition $\lambda_0 = r = \left|\Omega_{wt(f)}\right|$, so all we are left to show is that $\lambda_1 = 0$.

By Proposition 1.2.5, $diam(\mathcal{G}_f) \leq 1$ which implies that, for each connected component of $\mathcal{G}_f$ we have $d(X, Y) = 1$, for all $X, Y \in \mathbb{F}_2^n$, (since the components are complete via Theorem 4.2.5).

Also, since $\mathbf{b}(0) \in \Omega_{wt(f)}$ and $\mathcal{G}_f$ is complete, the graph has self loops,

$\Rightarrow$ the adjacency matrix, $A_{\mathcal{G}_f}$, of the associated graph is;

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix},$$

from which the second eigenvalue, $\lambda_1 = 0$, may be calculated.

□

Corollary 4.2.7. *Let $f \in B_n, \left|Spec(\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}))\right| = 2$, such that $\lambda_0 \neq \lambda_1$, for $\lambda_0, \lambda_1 \in Spec(\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$. If $\boldsymbol{b}(0) \notin \Omega_{wt(f)}$, then*

$$\lambda_0 = \left|\Omega_{wt(f)}\right| \quad and \quad \lambda_1 = -1,$$

where $\boldsymbol{b}(0) \in \mathbb{F}_2^n$.

Proof. Let $\mathbf{b}(0) \notin \Omega_{wt(f)}$. By definition $\lambda_0 = r = \left|\Omega_{wt(f)}\right|$, so all we are required to show is that $\lambda_1 = -1$.
Similarly as for Corollary 4.2.6, we may construct the adjacency matrix.

However, now $\mathbf{b}(0) \notin \Omega_{wt(f)}$, so the main diagonal of the adjacency matrix, $A_{\mathcal{G}_f}$, of the associated graph has zero's only;

$$\begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 0 \end{pmatrix}.$$

Hence, the second eigenvalue may be calculated as $\lambda_1 = -1$.

□

Theorem 4.2.8. *Let $f \in B_n$. If $\mathcal{G}_f$ is connected and $|Spec(\mathcal{G}_f)| = s + 1$, where $s \leq \frac{n}{2}$ then*

$$n \leq \log_2 \left(wt(f) + \binom{wt(f)}{s} \right).$$

Proof. Let $\mathcal{G}_f$ be a connected Cayley graph associated with a Boolean function. Since $|Spec(\mathcal{G}_f)| = s + 1$, from Proposition 1.2.5;

$diam(\mathcal{G}_f) \leq (s+1) - 1 = s$.

Also, any pair of vertices $X, Y \in \mathcal{G}_f$ are adjacent if $Y = X \oplus \omega_i$, where $\omega_i \in \Omega_{wt}(f)$ for some i. Thus, if Z is adjacent to Y,

$Z = Y \oplus \omega_2 = X \oplus \omega_1 \oplus \omega_2$ for some ω_1, ω_2.

Hence, any $Z \in \left(\mathbb{F}_2^n \setminus \Omega_{wt(f)}\right)$ can be given as

$$Z = \sum_i \omega_i, \text{ where } \omega_i \in \Omega_{wt(f)}.$$

It follows then that $i \leq s$ since $diam(\mathcal{G}_f) \leq s$.

Hence $Z = \sum_j^r c_j \omega_j$, where $r = wt(f)$ and $c_j \in \mathbb{F}_2$. Now,

$$\left|\mathbb{F}_2^n \setminus \Omega_{wt(f)}\right| = 2^n - r \leq \binom{r}{s}$$

since each c_j is either 0 or 1 for any $\omega_i \in \Omega_{wt(f)}$, but $\left|\Omega_{wt(f)}\right| = r$. Hence

each $Z \in \mathbb{F}_2^n$ can be made up of r or less $\omega_i's$ so

$$\begin{aligned} & 2^n - r && \leq && \binom{r}{s} \\ \Rightarrow \quad & 2^n && \leq && r + \binom{r}{s} \end{aligned}$$

Therefore, substituting $r = wt(f)$,

$$n \leq \log_2 \left(wt(f) + \binom{wt(f)}{s} \right).$$

□

To illustrate the relationship between Cayley graphs and the Boolean functions underpining the security of stream ciphers, we consider the following continuation of Example 4.2.1:

Example 4.2.2. It is clear from Figure 4.2.1 that $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$ in Example 4.2.1 is 4-regular, so

$$\begin{aligned} 2^{n-1} &= 2^{3-1} \\ &= 4. \end{aligned}$$

Also Remark 4.2.1 assures us that the regularity of $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$ is the Hamming weight of its associated Boolean function. Hence $wt(f) = 4$, which is as expected. Therefore, amongst other known attacks we are at least certain (to some probability) that the cipher can resist statistical dependence as an attack, since the Boolean function used is balanced by Definition 3.2.3 and Proposition 3.2.1. One can further test for resistance against other attacks.

4.3 Bent functions characterized by Strongly regular graphs

Just as pseudo-random number generators are core to the security of stream ciphers, bent functions possessing necessary cryptographic properties are used for construction of strong s-boxes which are central to the security of block ciphers.
Block ciphers took over as an important shield for ensuring security of electronic data after the US National Bureau of Standards (NBS) called for a strong encryption primitive in 1973. Since then many implementations have been made, including designs of DES, AES.

In this section we build upon Chapters 2 and 3 by comparing the properties and results for algebraic graphs and cryptographic functions. In Chapter 2 we introduced Cayley graphs and SRGs. In Chapter 3 we introduced general Boolean functions and a special case; bent functions.

We consider n to be even and construct a Cayley graph associated to the Boolean function, $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}) = (V, E)$. Then the resulting graph is said to be associated to a bent Boolean function. If, in addition, $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}) = (V, E)$ is a strongly regular graph then we say $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$ is a ***strongly regular Cayley graph associated with the bent Boolean function***, with both the vertex and edge set defined as in Definition 4.2.1.

We show the following powerful relationship between strongly regular Cayley graphs and cryptographic bent Boolean functions. Recall that we consider n to be even when dealing with bent functions.

Remark 4.3.1. *The **spectral coefficients** of $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$ are the eigenvalues of the corresponding adjacency matrix.*

Considering $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$ to be connected we show that there is a link (via the spectral coefficient of $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$) between strongly regular Cayley graphs and cryptographic bent functions. We show that the Hamming weight of a cryptographic function has a lower bound. Furthermore we explore some corresponding properties of these strongly regular Cayley graphs.

Corollary 4.3.1. *Let $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}) = (V, E)$ be a strongly regular Cayley graph associated with a bent function. Then*

$$wt(f) \geq \frac{-1 + \sqrt{2^{n+3} + 1}}{2}.$$

Proof. Let $\mathcal{G}_f$ be a SRCG associated to a bent Boolean function. Since $\mathcal{G}_f$ is connected, by Theorem 2.3.7, $|Spec(\mathcal{G}_f)| = 3$. This implies that the maximum eccentricity of $\mathcal{G}_f$ ($diam(\mathcal{G}_f)$) is not more than 2.

We omit the case where $diam(\mathcal{G}_f) = 0$, because it violates the requirement of SRG **Case I:** $diam(\mathcal{G}_f) = 1 \Rightarrow \mathcal{G}_f$ is complete and $|Spec(\mathcal{G}_f)| = 2$. However, $|Spec(\mathcal{G}_f)| = 3$, which is a contradiction.
Case II: $diam(\mathcal{G}_f) = 2 \Rightarrow$ since a pair of vertices $X, Y \in V(\mathcal{G}_f)$ is adjacent if, for $\omega_i \in \Omega_{wt(f)}$,

$Y = X \oplus \omega_i$.

Similarly, for pair of nonadjacent vertices $X, Z \in V(\mathcal{G}_f)$, sharing vertex Y,

$$\begin{aligned} Z &= Y \oplus \omega_2 \\ &= X \oplus \omega_1 \oplus \omega_2. \end{aligned}$$

i.e any element outside the set $\Omega_{wt(f)}$ can be given by the sum of two elements inside the set $\Omega_{wt(f)}$.

Hence any $Z \in \left(\mathbb{F}_2^n \setminus \Omega_{wt(f)}\right)$ can be given as

$$\begin{aligned} Z &= \sum_i \omega_i, \text{ where } \omega_i \in \Omega_{wt(f)} \\ &= \sum_j^r c_j \omega_j, \text{ where } r = wt(f) \text{ and } c_j \in \mathbb{F}_2 \end{aligned}$$

but the number of c_j that are not equal to zero is 2.

Hence, $\left|\mathbb{F}_2^n \setminus \Omega_{wt(f)}\right| = 2^n - r \leq \binom{r}{2}$

$$
\begin{aligned}
\Rightarrow \quad & 2^n - r && \leq && \frac{r(r-1)}{2} \\
\Rightarrow \quad & r^2 + r - 2^{n+1} && \geq && 0 \\
\Rightarrow \quad & r && \geq && \frac{-1 \pm \sqrt{1 + 2^{n+3}}}{2}
\end{aligned}
$$

However, $r > 0$. Therefore

$$wt(f) \geq \frac{-1 + \sqrt{1 + 2^{n+3}}}{2}.$$

□

Example 4.3.1. The lower bound of the Hamming weight of f in Example (3.3.1) is

$$
\begin{aligned}
wt(f) &\geq \frac{-1 + \sqrt{2^{4+3} + 1}}{2} \\
&= \frac{-1 + \sqrt{129}}{2} \\
&> 5.
\end{aligned}
$$

The next theorem, Theorem 4.3.2 paves the way for the results that conclude and explain the relationship between strongly regular graphs and bent cryptographic functions. The proof to Theorem 4.3.2 is given in the referenced material. The results that follow are then proved from this theorem.

In particular Theorem 4.3.3 demonstrates a special property in the family of strongly regular graphs. This is when $\lambda = \mu$. Strongly regular graphs with the property that $\lambda = \mu$, correlate with symmetric balanced incomplete block designs, also known as the 2-(n, r, λ) designs [11]. This gives rise to a natural question on the possible interplay between Boolean functions and 2-designs or a more general question on the possible interplay between cryptographic functions and symmetric 2-designs [1]. Block designs form part of design theory, a study in combinatorics. The literature (e.g. [2] and [36]) reveals interactions between specific types of block designs and cryptography.

Theorem 4.3.2. *[16] Let $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}) = (V, E)$ be a strongly regular Cayley graph associated to a bent function. Then*

$$Spec(\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})) = \left\{ \left|\Omega_{wt(f)}\right|, \sqrt{\left|\Omega_{wt(f)}\right| - \mu}, -\sqrt{\left|\Omega_{wt(f)}\right| - \mu} \right\}.$$

Theorem 4.3.3. *Let $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)}) = (V, E)$ be a strongly regular graph associated to a bent function. Then $\lambda = \mu$ if (n, r, λ, μ) are the parameters of $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$.*
Moreover, the corresponding adjacency matrix satisfies

$$A^2 = (2^{n-1} \pm 2^{\frac{n}{2}-1} - \mu)I + \mu J.$$

Proof. Let $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$ be a SRCG associated to a bent Boolean function. Then from Theorem 2.3.5 we have that a connected (n, r, λ, μ) strongly regular graph with the property

$$A^2 = (\lambda - \mu)A + (r - \mu)I + \mu J, \tag{4.3}$$

where I and J are the identity and the matrix consisting of all entries equal to 1, respectively.

However, $r = wt(f)$. From Theorem 4.3.2

$$\left\{ \sqrt{wt(f) - \mu}, -\sqrt{wt(f) - \mu} \right\} \subset Spec(\mathcal{G}_f).$$

Then it follows from Corollary 2.3.8 that

$$\begin{aligned} \lambda &= wt(f) + \left[\sqrt{wt(f) - \mu} \cdot \sqrt{wt(f) - \mu} \right] + \sqrt{wt(f) - \mu} - \sqrt{wt(f) - \mu}, \\ &= wt(f) - \sqrt{(wt(f) - \mu)(wt(f) - \mu)} \end{aligned} \tag{4.4}$$

$$\begin{aligned} \mu &= wt(f) + (\sqrt{wt(f) - \mu}) \cdot (-\sqrt{wt(f) - \mu}), \\ &= wt(f) - \sqrt{(wt(f) - \mu)(wt(f) - \mu)} \end{aligned} \tag{4.5}$$

Since (4.4) and (4.5) are equal, it follows that $\lambda = \mu$.

From (4.3) we have that

$$\begin{aligned} A^2 &= (\lambda - \mu)A + (r - \mu)I + \mu J \\ &= 0A + [wt(f) - wt(f) + (wt(f) - \mu)]\, I + \mu J \\ &= (wt(f) - \mu)\, I + \mu J. \end{aligned}$$

Then, from Remark 3.3.1 we have:

$$A^2 = \left(2^{n-1} \pm 2^{\frac{n}{2}-1} - \mu\right) I + \mu J.$$

□

Theorem 4.3.4. *Let $f \in BB_n$. Then $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$ is not a bipartite graph.*

Proof. Let $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$ be a strongly regular Cayley graph associated with a bent Boolean function f. Then, if $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$ is bipartite, we have $Spec(\mathcal{G}_f)$ symmetric with respect to 0 by Proposition 1.2.6. Hence, if $\lambda \in Spec(\mathcal{G}_f)$ then $-\lambda \in Spec(\mathcal{G}_f)$.

From Theorem 4.3.2 above we have that $\left|\Omega_{wt(f)}\right| = wt(f) \in Spec(\mathcal{G}_f)$. Hence, it would follow that $-wt(f) \in Spec(\mathcal{G}_f)$. This is a contradiction, according to the properties of $Spec(\mathcal{G}_f)$ in Theorem 4.3.2. Therefore; $\mathcal{G}_f(\mathbb{F}_2^n, \Omega_{wt(f)})$ is not a bipartite graph.

□

Example 4.3.2. Consider $f \in BB_4$ defined in Example 3.3.1 as

$$f(X) = x_1 \cdot x_2 \oplus x_3 \cdot x_4.$$

Then $V(\mathcal{G}_f(\mathbb{F}_2^4, \Omega_{wt(f)})) = \mathbb{F}_2^4$, so $\left|V(\mathcal{G}_f(\mathbb{F}_2^4, \Omega_{wt(f)}))\right| = 2^4 = 16$.

From Table 3.3.2,

$$\Omega_{wt(f)} = \{0011, 1100, 0111, 1011, 1101, 1110\},$$

so

$$\begin{aligned} E(\mathcal{G}_f(\mathbb{F}_2^3, \Omega_{wt(f)})) &= \{XY \mid f(X \oplus Y) = 1, X, Y \in \mathbb{F}_2^4\} \\ &= \{XY \mid (X \oplus Y) \in \Omega_{wt(f)},\ X, Y \in \mathbb{F}_2^4\}. \end{aligned}$$

Hence, we have $\mathcal{G}_f(\mathbb{F}_2^4, \Omega_{wt(f)})$ as:

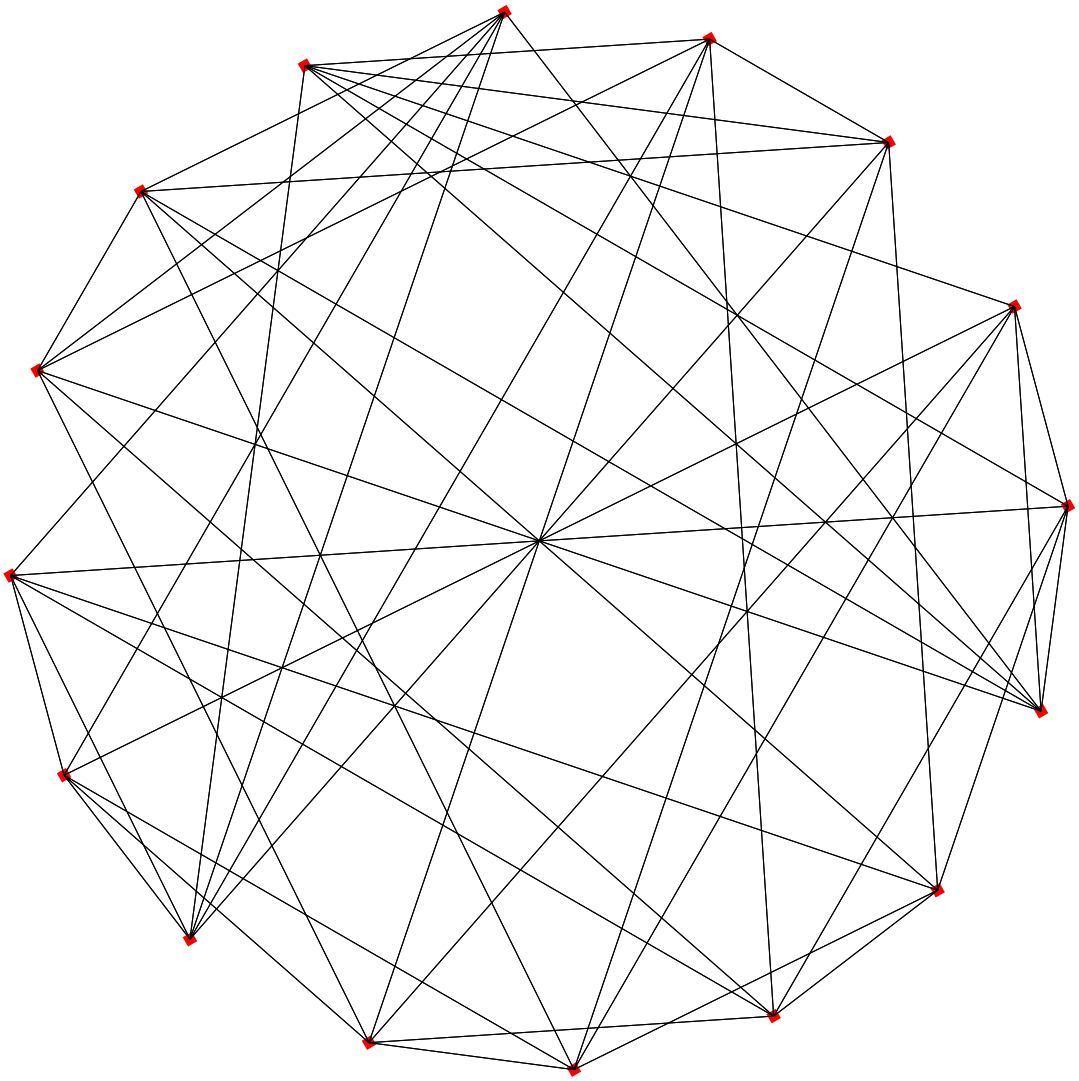


Figure 4.2: Strongly regular Cayley graph associated with the bent Boolean function $f \in BB_4$

Summary

In this chapter we reviewed the application of Cayley and strongly regular graphs to cryptographic use. We did this by considering the vector space

from a Boolean function to be the group from which the Cayley graph is constructed. Hence this bring together both fields (graph theory and cryptography) and the graph we end up with is the Cayley graph associated with a Boolean function. We further noticed that from the Cayley graph associated with a Boolean function we can derive numerous cryptographic properties of a stream cipher.

Further to this we saw how strongly regular Cayley graphs play a role in describing the strength of a block cipher by studying bent Boolean functions. We concluded with examples drawn from the study of the relationship between graph theory and cryptography.

Conclusion

This dissertation discussed the construction of Cayley graphs and the properties they possess with a view to applications in cryptography. Strongly regular graphs were also discussed so as to define strongly regular Cayley graphs and distinguish them from general Cayley graphs. We provided background material on cryptographic functions, and in particular Boolean functions, (which are used in stream ciphers), and a special case, bent functions (which are used in block ciphers). We then presented material discussing the links between Cayley graphs and Boolean functions, as well as those between strongly regular graphs and bent functions.

The key idea being that of constructing and defining a Cayley graph associated with a Boolean function both generally and those in the special case of a strongly regular Cayley graph associated with bent Boolean function. These graphs elucidate the connection between cryptography based on Boolean and bent functions, on the one hand, and the characterization of these in terms of general Cayley and strongly regular Cayley graphs on the other hand.

We showed that the construction of these graphs follows directly from the definition of Cayley and strongly regular graphs, with the group used for the construction of Cayley graphs being $(\mathbb{F}_2^n, \oplus)$. In some cases the Cayley set $\Omega_{wt(t)}$ maybe chosen without regarding the condition, $b(0) \notin \Omega_{wt(f)}$, where $b(0)$ is the identity element of the group under the binary operation XOR.

These algebraic graphs can be used to measure some cryptographic properties of the underlying cipher. The strength of the cipher is measured by considering the cryptographic functions that make up the security part of it. Boolean functions make up the pseudo-random number generator of the

stream cipher, so the design of the Boolean function is the crucial part of the cipher and needs to align with the cryptographic requirements. Similarly the set of bent (Boolean) functions makes up the substitution box of the block cipher, so these bent (Boolean) functions need to be checked against the relevant cryptographic requirements.

These requirements are drawn from understanding currently well researched and implemented attacks by cryptanalysts. Some attacks considered in this dissertation are: statistical dependence between plaintext and ciphertext, (fast) correlation attacks, algebraic attacks, as well as linear and differential cryptanalysis. Some fundamental requirements drawn from analysis of these attacks include: the Boolean function must be balanced, which means, the choice of f must be such that $wt(f) = wt(f \oplus 1) = 2^{n-1}$; and the Boolean function must have high nonlinearity, which is in fact attained to its maximum by the bent function. Other requirement briefly discussed include, SAC, propagation, $cl(m)$, high Hamming distance etc. We noticed that during the attempt to achieve these requirements there are trade-offs that appear, for instance we know that, for block ciphers, we could increase the number of rounds to make it more secure but at the same time that would lead to a disadvantage on the speed requirement of the cipher. Also [32] makes known that correlation immunity and the algebraic degree are conflicting properties and it is not possible to obtain a function with both properties optimal.

We managed to conclude that, from the Cayley graph associated with the Boolean function, one can actually tell whether the designed Boolean function is suitable against statistical dependence as an attack, since the regularity of the graph is equivalent to obtaining the Hamming weight of the function, from which we may decide whether the function is balanced or not. Theorem 4.3.2 from the last chapter shows that the Hamming weight of the bent function can be given in terms of the spectral information of the associated graph.

This dissertation considers PRNG; the author challenges the reader to investigate the possibility of considering CSPRNG for a similar study.

Bibliography

[1] Anonymous Examiner, *MSc Dissertation,* S.T Mafunda, University of KwaZulu-Natal (2015);

[2] Adhikar A, *Design Theory and Visual Cryptographic Schemes*, University of Calcutta, Kolkata (2013);

[3] Ahmadi B, *Strongly Regular Graphs*, Dissertation University of Regina (2009);

[4] Alspach B, *CS E6204 Lecture 6 Cayley Graphs*, Lecture Notes, University of Regina, Canada;

[5] Al-Shehhi M. A, Baek J, Yeun C. Y, *The Use of Boolean Function in Stream Ciphers*, Khalifa University of Science, Technology and Research, (2011);

[6] Al-Vahed A, Sahhavi H, *An overview of modern cryptography*, Mathematic School of Fada, Vol. 1 No. 1 (2011);

[7] Arazi B, *Some Properties of Hadamard Matrices Generated Recursively by Kronecker Products*, National Electrical Engineering Research Institute, SA;

[8] Beineke L. W, Wilson R. J, Cameron P. J, *Topics in Algebraic Graph Theory*, C.U.P (2004);

[9] Bernasconi A, Codenotti B, *Spectral Analysis of Boolean Functions as a Graph Eigenvalue Problem*, Vol. 48 No. 3 (1999);

[10] Bose R. C, *Strongly Regular Graphs, Partial Geometries and Partially Balanced Designs*, Pacific J. Math 13 (1963);

[11] Brouwer A. E, Haemers W. H, *Spectra of Graphs*, Springer (2011);

[12] Burnett L, Millan W, Dawson E, Clark A, *Simpler methods for generating better Boolean Functions with good Cryptographic proprieties* , Queensland University of Technology, (2004);

[13] Carlet C, Mesnager S, *On the supports of the Walsh transforms of Boolean Functions*, (2004);

[14] Cameron P. J, *Permutation Groups*, C.U.P (1999);

[15] Carlet C, *Boolean Functions for Cryptography and Error Correcting Codes*, University of Paris 8, France;

[16] Cusick T. W, Stanica P, *Cryptographic Boolean Functions and Applications*, A.P (2009);

[17] Dlamini G, *Aspect of Distance in Graphs*, Dissertation, University of KwaZulu-Natal (2003);

[18] Diffie W, Hellman M. E *New Directions in Cryptography*, IEEE Transactions on Information Theory Vol. IT-22 No. 6 (1963);

[19] Erwin D. J, Mukwembi S, Swart H. C, Henning M, Course Notes, *Discrete Mathematics with Applications*, University of KwaZulu-Natal;

[20] Farrugia A, *Self-complementary graphs and generalisations: a comprehensive reference manual*, University of Malta (1976);

[21] Grabbe J. O, *The DES Algorithm Illustrated*;

[22] Heys H. M, *A Tutorial on Linear and Differential Cryptanalysis*, Faculty of Engineering and Applied Sciences, Memorial University of Newfoundland;

[23] Hubaut X. L, *Strongly Regular Graphs*, Free University of Brussels (1975);

[24] Kang D, *Group Representations and Character Theory*;

[25] Khan D, *The Codebreakers* (1973);

[26] Krebs M, Shaheen A, *Expander Families and Cayley Graphs, A Beginner's Guide*, O.U.P (2011);

[27] Lazenby F. J, *Circulant Graphs and their Spectra*, Reed College (2008);

[28] Magliaro P. A, Weaver A. D, *Investigates into a possible new family of Partial Difference Sets*, University of Richmond;

[29] Menezes A, Van Oorschot P, Vanstone S, *Handbook of Applied Cryptography*, C.R.C (1996);

[30] Mohamed K, Pauzi M. N. M, Ali F. H. M, Ariffin S, Zulkipli N. H. N, *Study of S-box Properties in Block Cipher*, (2014);

[31] Paar C, Pelzl J, *Understanding Cryptography*, Springer (2010);

[32] Picek S, Carlet C, Jakobovic D, Miller J.F, Batina L, *Correlation Immunity of Boolean Functions: An Evolutionary Algorithms Perspective*, Association for Computing Machinery, (2015);

[33] Pommerening K, *Fourier Analysis of Boolean Maps, A Tutorial*, Fachbereich Mathematik, der Johannes Gutenberg Universitaet, (2005);

[34] Rodrigues B. G, *Notes on Classical Algebra (Further Group Theory)*, Course Notes, University of KwaZulu-Natal (2014);

[35] Rothe J, *Complexity Theory and Cryptology*, Springer;

[36] Roy B, *PBIBD and its application in Cryptology*, Indian Statistical Institute (2012);

[37] Sabidussi G, *On a Class of Fixed-Point-Free Graphs*, Proc. Amer. Math. Soc. **9** (1958) 800-804;

[38] Stanica P, *Graph Eigenvalues and Walsh Spectrum of Boolean Functions*, Naval Postgraduate School, Monterey (2007);

[39] Swart H. C, Swart J. H, *Introduction to the method of Operations Research*, Course Notes, University of KwaZulu-Natal;

[40] Tokareva N, *Bent Functions: Results and Application to Cryptography*, Novosibirsk State University, A.P (2015);

[41] Toomey G, *Algebraic Graph Theory: Automorphism Groups and Cayley graphs*, C.U.P (2014);

[42] Trevisan L, *Graph Partitioning and Expanders* , Stanford University, (2011);

[43] Wei Y, Hu Y, *Maximum Autocorrelation Analysis of Nonlinear Combining Functions in Stream Cipher*, Xidian University, (2007);

[44] Wielandt H, *Finite Permutation Groups*, University of Tubingen, (2007);