Project Report

On

STUDY ON ALGEBRAIC GRAPH THEORY

Submitted

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

MATHEMATICS

by

AGNES P VARGHESE

(Register No. SM20MAT001)

(2020-2022)

Under the Supervision of

DHANALAKSHMI O.M.

DEPARTMENT OF MATHEMATICS

ST. TERESA'S COLLEGE (AUTONOMOUS)

ERNAKULAM, KOCHI - 682011

APRIL 2022

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

CERTIFICATE

This is to certify that the dissertation entitled, **STUDY ON ALGEBRAIC GRAPH THEORY** is a bonafide record of the work done by Ms. **AGNES P VARGHESE** under my guidance as partial fulfillment of the award of the degree of **Master of Science in Mathematics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

Date : 26.5.2022
Place : Ernakulam

Mrs. Dhanalakshmi O.M.
Assistant Professor,
Department of Mathematics,
St. Teresa's College(Autonomous),
Ernakulam.

ST. TERESA'S COLLEGE DEPT. OF MATHEMATICS ERNAKULAM

Dr.Ursala Paul
Assistant Professor and HOD,
Department of Mathematics,
St. Teresa's College(Autonomous),
Ernakulam.

External Examiners

1:.DIVYA....MARY...DAISE.S

2: MARY RUNIYA

DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of **MRS. DHANALAKSHMI O.M.**, Assistant Professor, Department of Mathematics, St. Teresa's College(Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Agnes.

Place : Ernakulam

Date : 26.5.2022

AGNES P VARGHESE

SM20MAT001

ACKNOWLEDGEMENTS

It is my privilege to express profound gratitude to my teachers who gave me the opportunity to do this wonderful project on the topic STUDY ON ALGEBRAIC GRAPH THEORY which helped me to understand how simpler it is to use algebraic method in illustrating complicated problems.

I would specially mention the the unending help and support provided by my project guide MRS. DHANALAKSHMI O.M. for her invaluable support and guidance throughout my project.

I am extremely thankful to the head of our department, DR.URSALA PAUL for her valuable suggestions, critical examination of my work during the progress.

I am grateful to the teaching and non-teaching staff of Mathematics department, my parents and friends and all those who has given me the moral support and helped me to complete my project and made this venture a success.

I would like to express my sincere thanks to GOD ALMIGHTY, for his constant love and grace that he has showered upon me .I extend my pleasure to him for providing me with all the convenience and strength to undertake and complete this project.

Agnes.

AGNES P VARGHESE

SM20MAT001

Place : Ernakulam

Date: 26.5.2022

Contents

Chapter 1

INTRODUCTION AND HISTORICAL OUTLINE

1.1 INTRODUCTION

Graph theory is one of the newly emerging field in mathematics that has a notable application. Algebraic graph theory is one of the multidimensional branches of Mathematics enriched with many interesting problems. It is one of the most explored and fastest emerging fields in mathematics, with numerous opportunities for research. In the last several decades, in investigating complex graph structures algebraic techniques are being increasingly used. It is an interesting subject concerned with the interplay between algebra and graph theory. Here algebraic techniques are used in the study and explaining problems about graphs. The main aim is to translate the properties of graphs into algebraic properties and then using the results and methods of algebra, deduce the theorems about graphs. There are many interesting algebraic objects associated with graphs, also one can give interesting and elegant proofs of graph theoretic facts using algebraic tools. The literature of algebraic graph theory itself has developed enormously. The problems about graphs is in contrast to geometric,combinatoric, or algorithmic approaches. There are three main branches of algebraic graph theory.

(i) USING LINEAR ALGEBRA:

This is the first branch of algebraic graph theory that involves the study of graphs in connection with linear algebra. It mainly deals with the study of the spectrum

of the adjacency matrix, or the Laplacian matrix of a graph. This part of algebraic graph theory is also called spectral graph theory.

(ii) USE OF GROUP THEORY:

This is the second branch of algebraic graph theory. It involves the study of graphs in connection to group theory, particularly automorphism groups and geometric group theory. The main focus is placed on the various families of graphs based on its symmetry such as symmetric graphs, vertex transitive graphs, edge transitive graphs, etc.

(iii) STUDY OF GRAPH INVARIANTS:

This is the third branch of algebraic graph theory that concerns with the algebraic properties of invariants of graphs especially the knot invariants and chromatic polynomial.

There are two main connections between Algebra and Graph Theory. These arise from two algebraic objects associated with a graph; its Adjacency matrix and Automorphism group they are related.The automorphism group can be regarded as a collection of all permutation matrices that commute with the adjacency matrix. However the two connections involve different algebraic techniques - Linear algebra and group theory.

Linear algebra is a branch of mathematics concerning linear equations and linear functions and their representations through matrices and vector spaces. Linear algebra fundamental in modern presentations of geometry.

In mathematics, group theory learns the algebraic structures known as groups. The method of group theory have influenced many parts of algebra. Linear algebraic groups and lie groups are two branches of group theory that have experienced advances.

Graph theory is the study of graph which are mathematical structures used to model pair wise relations between objects. A graph in this context is made up of vertices, nodes or points which are connected by edges, arcs or lines. Graphs are one of the prime objects.This project is divided into five chapters:

CHAPTER 1

This chapter introduces us with a brief note on what is algebraic graph theory and history of graph theory.

CHAPTER 2

This chapter includes the basic definitions, terminology and notations of graph theory and linear algebra which are needed for the subsequent chapters. Also some pre–requisite about Protein structure and Thalassemia as they are needed.

CHAPTER 3

In this chapter we discuss the use of linear algebra in graph theory for the sake of convenience and standardization. This is the first branch of algebraic graph theory that concerns with the study of graphs in connection with linear algebra. Especially it studies the spectrum of adjacency matrix or Laplacian matrix of a graph.

CHAPTER 4

This chapter deals with the Laplacian matrix of a graph and its properties. Also an alternative approach is explored.

CHAPTER 5

In this chapter we discuss about the applications of algebraic graph theory.And some of the applications of algebraic graph theory are briefly described and highlighted.

1.2 HISTORY

The story begins in the 18^{th} century in the city of Konigsberg which was located on the Pregal River in Prussia. The river divided the city into four land masses including the island of Kneiphopf. And these four regions were linked by seven bridges. These seven bridges are termed as Blacksmith's bridge, Connecting bridge, Green bridge, Merchant's bridge, Wooden bridge, High bridge and Honey bridge. During Sunday afternoons, the citizens of Konigsberg used to walk around their beautiful city. While walking around their beautiful city, they decided to create a game for

themselves. The goal was starting from one land mass they have to cross each of the seven bridges exactly once and come back to starting place. None of the citizens of Konigsberg were able to find out a route that would allow them to cross each of the seven bridges exactly once.

Several mathematical studies were done on this problem. Then a Swiss mathematician Leonhard Euler thought about this problem. And on 26th August 1735, Euler presented a paper containing the solution to the Konigsberg bridge problem. He addressed this specific problem, as well as a general solution was given i.e. with any number of land masses and any number of bridges a solution is possible and concluded that there is no solution to Konigsberg Bridge Problem. But we could only solve it if either all the vertices in the graph were even, or if only two of the vertices were odd. This paper was later published in 1741. The method he used to solve this problem is considered to be the birth of graph theory.

Then for a few years no further studies were carried out. Later a study of electrical networks was done. The theory of graphs and their associated matrices shared a long rich synergy and joint development. Starting from the foundational classical work by G. Kirchhoff's modelling and analysis of electric circuits, motivated the birth and the further development of a wide range of graph-theoretical concepts and a certain classes of matrices.

And in 1857, Cayley discovered in the very natural setting of organic chemistry an important class of graphs called trees. He was engaged in enumerating the isomers of the saturated hydrocarbons C_nH_{2n+2}, with a given number n of carbon atoms.

Then the Four Colour Theorem.This is the most famous problem in graph theory and perhaps in all of mathematics is the celebrated four color conjecture. The problem is any map on a plane or the surface of a sphere can be colored with only four colors so that no two adjacent countries have the same color. Each country must consist of a single connected region and adjacent countries are those having a boundary line common. This is a remarkable problem that can be explained by any mathematician to the so called man in the street in five minutes. But at the end of the explanation, both will understand the problem, but neither will be able to solve it.

Chapter 2

PRELIMINARIES

2.1 PRE–REQUISITE ABOUT GRAPH THEORY

2.1.1 Basic Definitions

GRAPH

A Graph is an ordered triple G = (V(G), E(G), I_G), where V(G) is a non empty set , E(G) is a set disjoint from V(G) and I_G is an incidence relation that associates with each element of E(G) an unordered pair of elements of V(G).
Elements of V(G) are called the vertices (or nodes or points) of G and elements of E(G) are called the edges (or lines) of G. V(G) and E(G) are the vertex set and edge set of G respectively.n(G) and m(G) are the number of vertices and edges of the graph G respectively. The number n(G) is called the order of G and m(G) is the size of G.

LOOP

An edge having same end points is called a loop.

MULTIPLE OR PARALLEL EDGES

A set of two or more edges of a graph G is called a set if multiple or parallel edges if they have the same pair of distinct ends.

ADJACENT EDGE

The vertices u and v are adjacent to each other in G if and only if there is an edge of G with u and v as its ends.

ADJACENT VERTEX

Two distinct edges e and f are said to be adjacent if and only if they have a common end vertex.

SIMPLE GRAPH

A graph G is called simple if it has no loops and no multiple edges.

FINTITE GRAPH

A graph is called finite if both V(G) and E(G) are finite.

INFINITE GRAPH

A graph that is not finite is called an infinite graph.

DEGREE OF VERTEX (also called Valency)

Let G be a graph and v $\in$ V. The number of edges incident at v in G is called the degree (or valency) of the vertex v in G and is denoted by d_G (v) or simply d(v).

ISOMORPHISM OF GRAPHS

Let G = (V(G) , E(G) , I_G) and H = (V(H) , E(H) , I_H) be two graphs.
A graph isomorphism from G to H is a pair (ϕ , θ) where ϕ : V(G) $\rightarrow$ V(H) and θ : E(G) $\rightarrow$ E(H) are bijections with the property that I_G(e) = {u, v} if and only if $I_H(\theta(e))$ = { ϕ(u) , ϕ(v) } .
If (ϕ , θ) is a graph isomorphism, the pair of inverse mappings (ϕ^{-1} , θ^{-1}) is also a graph isomorphism.

BIPARTITE GRAPH

If the vertex set of a graph G can be partitioned into two non-empty subsets X and Y such that each edge has one end in X and other end in Y. The pair (X,Y) is called a bipartition of the bipartite graph.
The most important property of bipartite graphs is that they are the graphs that contain no cycles of odd length. The complete bipartite graph Kp,q is the p x q bipartite graph in which each vertex is adjacent to all those in the other partite set.

COMPLETE GRAPH

A graph G is called complete if every pair of distinct vertices of G are adjacent in G and the complete graph on n vertices is denoted by Kn.

A complete bipartite graph of the form $K_{1,q}$ is called a star.

DEGREE SEQUENCE

A sequence formed by the degrees of the vertices of G is called a degree sequence of G.

DIRECTED GRAPH

A directed graph is a graph in which the edges ares are directed by arrows.

DEGREE SEQUENCE

A sequence of non-negative integers d = $(d_1, d_2, ..., d_n)$ is called graphical if there exists a simple graph whose degree sequence is d.

WALK

A walk in a graph G is an alternating sequences of vertices and edges beginning and ending with vertices in which v_{i-1} and v_i are the ends of e_i.

PATH

A walk is called a path if all the vertices in the walk are distinct.

CYCLE

A cycle is a closed trail in which the vertices are all distinct.

TREE

A tree is a connected graph that has no cycles. They have been characterized in many ways, a few of which are listed below.For a graph G of order n:

- G is connected and has no cycles.
- G is connected and has n-1 edges.
- G has no cycles and has n-1 edges.
- Any graph without cycles is a forest

NOTES:

- If G is a graph with maximum degree Δ that is neither an odd cycle nor a complete graph, then $\chi(G) \leq \Delta(G)$.
- A graph G is bipartite if and only if it contains no odd cycles.
- The sum of the degrees of the vertices of a graph is equal to twice the number of its edges.

2.2 PRE –REQUISITE ABOUT LINEAR ALGEBRA

SQUARE MATRIX

An m x n matrix consist of mn real numbers arranged in m rows and n columns. An m x n matrix is called a square matrix if m = n.Let A be a square matrix of order n.The entries a_{11}, a_{22} ..... a_m are said to constitute the diagonal of A.The trace of A is defined as, Trace A = a_{11} + a_{22} + ... + a_m

$\Longrightarrow$ A zero matrix 0 is a matrix in which each entry is 0.

$\Longrightarrow$ A square matrix A is symmetric if A^T = A

MINOR OF A MATRIX

A minor of a matrix A is the determinant of some smaller square matrix, cut down from A by removing one or more of its rows or columns.

A minor of A of order k is principal if it is obtained by deleting n-k rows and n-k columns with the same numbers.

DEFINITION 2.2.1

If A is a square matrix then the determinant det (A - λI) is a polynomial in the variable λ of degree n and is called the characteristic polynomial of A. The equation det (A -λI) = 0 is called the characteristic equation of A. By the fundamental theorem of algebra the equation has n complex roots and these roots are called the eigenvalues of A.

The set of eigenvalues is the spectrum of A. The eigenvalues might not all be distinct. The number of times an eigenvalue occurs as a root of the characteristic equation is called the algebraic multiplicity of the eigenvalue.

If 1, 2 ..... n are the eigenvalues of A, then det A = λ_1, λ_2 ..... λ_n while trace A = λ_1 + λ_2 + ... + λ_n

A principal submatrix of a square matrix is a submatrix formed by a set of rows and the corresponding set of columns.
A principal minor of A is the determinant of a principal submatrix. The sum of the products of the eigenvalues of A, taken k at a time, equals the sum of the k x k principal minors of A. When k = I this reduces to the familiar fact that the sum of the eigenvalues equals the trace. The Cayley Hamilton theorem states that every matrix satisfies its characteristic equation.

DEFINITION 2.2.2
An n x n matrix A is said to be positive definite if it is symmetric and if for any nonzero vector x, $x^T Ax \geq 0$. The identity matrix is clearly positive definite, then the following conditions is equivalent to A being positive definite :
1) The eigenvalues of A are positive.
2) All principal minors of A are positive.

2.3 PRE –REQUISITE OF PROTEIN STRUCTURE

Proteins are macro molecules found in the cells that play many critical roles in the body. They do most of the work in cells and are required for all metabolic activities in an organism. Proteins are made up of smaller sub units called amino acids, which are attached to one another by peptide bond to form long polypeptide chains. There are about 20 types of amino acids that can be combined to make a protein. The sequence of amino acids determines the structure of proteins. Protein exist in four different structures such as primary, secondary, tertiary and quaternary.

2.4 PRE –REQUISITE ABOUT THALASSEMIA

Thalassemia is an inherited blood disorder in which the body makes an abnormal form of haemoglobin and have less haemoglobin than normal. Haemoglobin is the protein molecule in red blood cells that carries oxygen. A person with Thalassemia can become anaemic and fatigued due to the excessive destruction of red blood cells. Anemia is a condition in which your body doesn't have enough normal, healthy red blood cells.

Thalassemia is inherited, in the sense that at least one of your parents must

be a carrier of the disorder. It's caused by either a genetic mutation or a deletion of certain key gene fragments. A person with mild Thalassemia might not need treatment since it is a less serious form of the disorder.

There are two main forms of Thalassemia that are more serious. They are:

(i) Alpha Thalassemia:

In Alpha Thalassemia, at least one of the alpha globin genes has a mutation or abnormality.

(ii) Beta Thalassemia:

In Beta Thalassemia, the beta globin genes are affected. Each of these forms of Thalassemia has different subtypes.

But the severe forms might require bone marrow transplant or regular blood transfusions or continuous medications and supplements. But this is not a permanent cure. These are only just precautions to suppress the disease.

Chapter 3

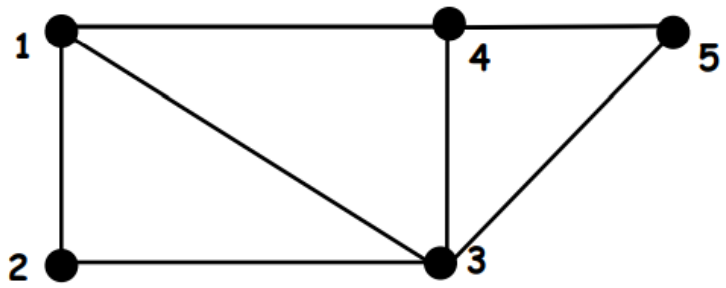
LINEAR ALGEBRA IN GRAPH THEORY

DEFINITION 3.1 : *ADJACENCY MATRIX*

An adjacency matrix is a means of representing which vertices (nodes) of a graph are adjacent to which other vertices.

Let G be a graph with V(G) = $\{v_1, v_2, \ldots.., v_n\}$ and E(G) = $\{ e_1, e_2, \ldots\ldots., e_n\}$. The adjacency matrix of G, denoted by A(G), is the n x n matrix, whose entries a_{ij} are given by,

$$a_{ij} = \begin{cases} 1 \ , & \textit{if } v_i \textit{ and } v_j \textit{ are adjacent} \\ 0 \ , & \textit{otherwise} \end{cases}$$



$$A(G) = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

Here we can see that the diagonal of an adjacency matrix of the graph G contains only zeros because there are no self-loops. Remember that our graphs have no multiple edges or loops. This causes the trace of the adjacency matrix, denoted by tr(A), the sum of its main diagonal, to be zero. Also, when A(G) represents a graph, it is symmetric matrix and all of the elements are non-negative. In other words, $a_{ij} = a_{ji}$.

Thus we can say A is a symmetric matrix with zeros on the diagonal and that the trace of A is zero.

LEMMA 3.2:

Let G be a connected graph with vertices 1,2,.....n and let A be the adjacency matrix of G. If i,j are the vertices of G with d(i,j) = m, then the matrices I,A,.....A^m are linearly independent.

DEFINITION 3.3 : *SPECTRUM OF A GRAPH*

The spectrum of a graph G is the set of numbers which are eigenvalues of A(G), together with their multiplicities of A(G). If the distinct eigenvalues of A(G) are $\lambda_0 \geq \lambda_1 \ldots\ldots \geq \lambda_{s-1}$ and their multiplicities are $m(\lambda_0)$, $m(\lambda_1)$, ... $m(\lambda_{s-1})$, then we shall write

$$\text{Spec G} = \begin{pmatrix} \lambda_0\ ,\ \lambda_1\ .... & \lambda_{s-1} \\ m(\lambda_0),\ m(\lambda_1)\ .... & m(\lambda_1) \end{pmatrix}$$

DEFINITION 3.4 : *ADJACENCY ALGEBRA*

The adjacency algebra of a graph G is the algebra of polynomials in the adjacency matrix A(G). We shall denote the adjacency algebra of G by A(G). Since every element of the adjacency algebra is a linear combination of powers of A, we can obtain results about A(G) from a study of these powers. We already a walk of length l in G, joining v_1 to v_p to be a finite sequence of vertices of G. $v_i = u_0$, u_1,..., $u_1 = v_p$ such that u_{i-1} and u_1 are adjacent for $l \leq t \leq i$

DEFINITION 3.5 : *CHARACTERISTIC POLYNOMIAL*

If A is a square matrix then the det($A-\lambda$I) is a polynomial in the variable of

degree n and is called the characteristic polynomial of A. The equation det (A - λI) = 0 is called the characteristic equation of A. By the fundamental theorem of algebra the equation has n complex roots and these roots are called the eigenvalues of A. The set of eigenvalues is the spectrum of A. The eigenvalues might not all be distinct. The number of times an eigenvalue occurs as a root of the characteristic equation is called the algebraic multiplicity of the eigenvalue.

DEFINITION 3.6 : *EIGEN VALUES AND WALKS*

The oldest, and perhaps the most fundamental, relationship between the eigenvalues of a graph and its geometric properties concerns walks – particularly, closed walks. Suppose that A is the adjacency matrix of a graph. Then the powers of A enumerate the walks in that graph.

DEFINITION 3.7 : *EIGEN VALUES AND LABELLINGS OF GRAPHS*

The columns of the adjacency matrix A correspond to the vertices of the graph. If Aacts on a vector x, then the entries of that vector can have the same correspondence,so a vector may be used to label the vertices. When we use an eigenvector, the eigenvalues will reflect the geometric structure of the graph.

DEFINITION 3.8 : *COSPECTRAL*

Two non isomorphic graphs are said to be cospectral if they have the same eigenvalues with the same multiplicities. Two methods are used for the construction of cospectral graphs. One uses operations on graphs (Complements, Product etc) to produce new cospectral ones, while the other logically pastes different graphs together.

One easy way to construct cospectral graph is due to Hoffman. Take two non isomorphic graphs G_1 and G_2, and consider the graph formed by taking k copies of G_1 and s-k copies of G_2. Let H_k be the complement of this graph. The graphs H_k (k = 0,1,. . . ..s) are cospectral. Thus we can have arbitrarily large sets of non isomorphic cospectral graphs.

A second method of constructing cospectral graphs is by pasting smaller graphs together. One way is to take two graphs, G and H, designate a special vertex in each of them and let it be u and w. We denote this new graph by G,H. Then find

the characteristics polynomials of G and H. If the characteristic polynomials are equal the two graphs G and H are cospectral.

DEFINITION 3.9 : *ALGEBRAIC MULTIPLICITY*

The algebraic multiplicity of an eigenvalue is the number of times that the value occurs as a root of the characteristic polynomial.

DEFINITION 3.10 : *GEOMETRIC MULTIPLICITY*

The geometric multiplicity is the dimension of the eigenspace, or the subspace spanned by all of the eigenvectors.

THEOREM 3.11: *MATRIX TREE THEOREM*

Let H is a subgraph of G if V (H) $\subset$V (G) and E(H) $\subset E(G)$. A subgraph H of G is a spanning subgraph of G if V (H) = V (G). Hence, a spanning subgraph of G is obtained by deleting some of the edges of G but keeping all vertices. If H is a spanning subgraph of G and H is a tree then we say that H is a spanning tree of G.

Chapter 4

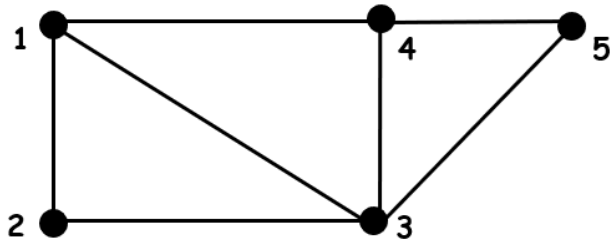
LAPLACIAN MATRIX OF A GRAPH

4.1 DEFINITIONS

The Laplacian is an alternative to the adjacency matrix for describing the adjacent vertices of a graph. Let G be a graph with V(G) = $\{v_1, v_2, \ldots.., v_n\}$ and E(G) = $\{ e_1, e_2, \ldots\ldots., e_m\}$.The Laplacian matrix L(G) of a graph G is the square matrix (n x n) that corresponds to the vertices of a graph. The main diagonal of the matrix represents the degree of the vertex while the other entries are as follows:

$$Aij = \begin{cases} -1\ , & \text{if } v_i \text{ and } v_j \text{ are adjacent} \\ 0\ , & \text{otherwise} \end{cases}$$

Example : Consider a graph G



$$L(G) = \begin{pmatrix} 3 & -1 & -1 & -1 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 4 & -1 & -1 \\ -1 & 0 & -1 & 3 & -1 \\ 0 & 0 & -1 & -1 & 2 \end{pmatrix}$$

The Laplacian matrix can also be derived from L(G) = D(G) – A(G), where D(G) is the diagonal matrix whose entries represent the degrees of the vertices, and A(G) is the adjacency matrix.

DEFINITION : *ALGEBRAIC CONNECTIVITY*

The Laplacian of a connected graph has eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots\ldots \leq \lambda_n$. Then the algebraic connectivity of a graph is defined to be 2, the second smallest eigenvalue. The name is a result of its connection to the vertex connectivity and the edge connectivity of a graph. It is the most important information contained within the spectrum of a graph.

REMARK 1: The oldest result about the Laplacian matrix concerns about the number of spanning trees of a graph. The Matrix Tree Theorem is one of the most significant applications of the Laplacian and is usually contributed to Kirchhoff.

DEFINITION : *POSITIVE SEMI-DEFINITE MATRIX*

A positive semidefinite matrix is one that is Hermitian, and whose eigenvalues are all non-negative. A Hermitian matrix is one which equals its conjugate transpose. This is usually written:

$\overline{A^H} = A^T = \mathrm{A}$

DEFINITION : *CHARACTERISTIC FUNCTION*

The characteristic function is the function for which every subset N of X, has a value of 1 at points of N, and 0 at points of X – N. In other words, it takes the value of 1 for numbers in the set, and 0 for numbers not in the set.

PROPERTY 1:

The smallest eigenvalue of L is 0.

PROPERTY 2:

The multiplicity of 0 as an eigenvalue of L is the number of connected components in the graph.

PROPERTY 3:

The algebraic connectivity is positive if and only if the graph is connected.

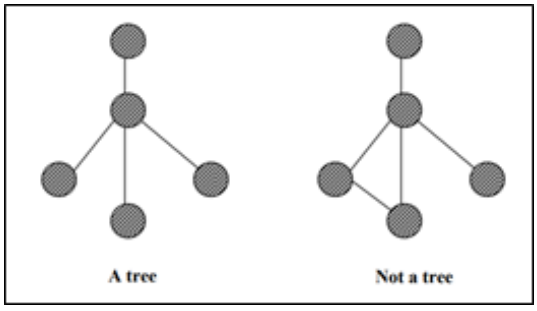
NOTE:

Eigenvalues of Laplace matrices of graphs are related to various combinatorial properties of graphs. They can be used to extract information about some parameters that are hard to compute or estimate, most notably those related to expansion and vertex partitions. The smallest and largest eigenvalue can be expressed as solutions to a quadratic optimization problem. It turns out that the right generalized setting for this is semi-definite programming, where duality theory leads to powerful applications. A more general setting of weighted graphs is presented, which brings us also to the study of simple random walks whose transition matrix can be expressed via the related Laplacian

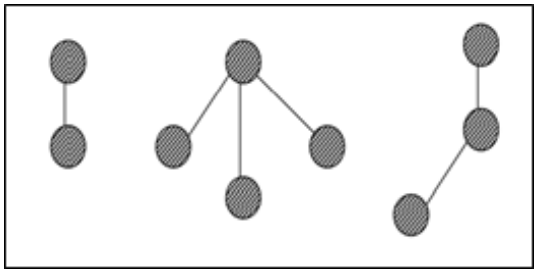
4.2 SPANNING TREES

In order to discuss spanning trees, we must first cover a few definitions. A tree is a connected graph that has no cycles.

Example:

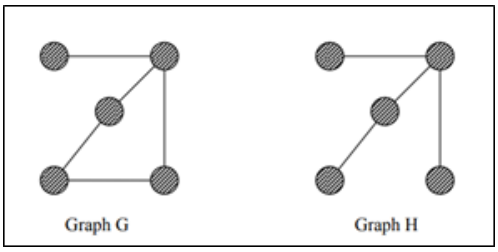


A subgraph H of a graph G is a graph whose vertex and edge sets are subset of V(G) and E(G) in that order. A few subgraphs of the tree above are:



A subgraph H is said to span a graph G if V(H) = V(G). A spanning tree of a

graph is a spanning subgraph that is a tree. Given graph G below, graph H is a spanning tree of G.



Before we go into the next few properties, we need to understand a co-factor of a matrix, which begins with the minor of a matrix. A minor M_{ij} of a matrix B is determined by removing row i and column j from B, and then calculating the determinant of the resulting matrix. The co-factor of a matrix is $(-1)^{i+j}M_{ij}$.

LEMMA:

Let G be a graph with $V(G) = \{v_1, v_2, \ldots.., v_n\}$ and $E(G) = \{ e_1, e_2, \ldots\ldots., e_m\}$. Then the following assertions holds:

(i) L(G) is a symmetric, positive semi definite matrix.

(ii) The row and the column sums of L(G) are zero.

(iii) The smallest non-zero eigenvalue of L is called the spectral gap.

(iv) The second smallest eigenvalue of L is the algebraic connectivity of G

PROPERTY 4:

The Matrix Tree Theorem

Given a graph G, its adjacency matrix A, and its degree matrix C, the number of non identical spanning trees of G is equal to the value of any co-factor of the matrix C – A.

COROLLARY - Cayley's Tree Formula:

The number of different trees on n labeled vertices is n^{n-2}.

4.3 AN ALTERNATIVE APPROACH

There is an alternative way of defining eigenvalues. We can define them in their "normalized" form. One advantage to this definition is that it is consistent with eigenvalues in spectral geometry and in stochastic processes.It also allows results which were only known for regular graphs to be generalized to all graphs. We will use NL to represent the Laplacian calculated using this definition.

In a graph where d_v represents the degree of vertex v, the Laplacian would be defined to be the matrix,

$$\mathbf{NL}(u, v) \begin{cases} 1 & \text{if } u = v \text{ and } d_v \neq 0 \\ \dfrac{-1}{\sqrt{d_u d_v}} & \text{if } u \text{ and } v \text{ and adjacent} \\ 0 & \text{otherwise} \end{cases}$$

Chapter 5

APPLICATIONS OF ALGEBRAIC GRAPH THEORY

Algebraic graph theory is the branch of mathematics in which algebraic methods are used to solve problems about graphs. It is an interesting subject concerned with the interplay between algebra and graph theory. In the last several decades, in investigating complex graph structures algebraic techniques are being increasingly used. Algebraic tools helps to give surprising and elegant proofs of graph theoretic facts, and there are many interesting algebraic objects associated with graphs. Graph theory is rapidly moving into the mainstream of mathematics mainly because of its application in diverse fields which include:

(i) Biochemistry (genomics)

(ii) Electrical engineering (Communications, network and Coding theory) -Identifying clusters is an important aspect in the field of electrical network connections.

(iii) Computer Science (Algorithms and computations) -In computer science graphs are used to represent networks of communication, data organization, the flow of computation etc. Also Computer scientists use it in load balancing algorithms.

(iv) Operation Research (Scheduling).

There are numerous applications of algebraic graph theory, specifically spectral graph theory, within the sciences and many other fields. It is an important sub branch of algebraic graph theory. Spectral graph theory is used in the study of chemical compounds, DNA, Brain network analysis etc. Laplacian eigenvalues determine the kinematic behavior of a liquid flowing through a system of communicating pipes. The graphs are useful in geometry and certain parts of topology

such as knot theory. Graph spectral method is extremely helpful in finding the needed results with minimal computations. Algebraic graph theory has close relation with group theory. Algebraic graph theory has many application and potential applications to problems in quantum computing, although the connection has become apparent very recently. Work in quantum information theory is leading to a wide range of questions which can be successfully studied using ideas and tools from algebraic graph theory. Contemporary scientific literature offers ample evidence that the algebraic methods which have revolutionized pure mathematics are now in process of having a similar impact in the physical sciences. The algebraic approach to statistical mechanics and quantum theory is an example of this new orientation. Another significant application of algebraic graph theory is the design and analysis of topologies of interconnection networks. The topologies that are used to connect processors in a supercomputer have a high degree of symmetry and are usually Cayley graphs. Some of the applications of algebraic graph theory are briefly described as viz:

5.1 GOOGLE PAGE RANK

Page Rank is a way of measuring the importance of a particular web pages. It was invented by Larry Page and Sergey Brin while they were graduate students at Stanford, and it became a Google trademark in 1998. This was named after Larry Page, one of the founders of Google. Page Rank (PR) is an algorithm used by Google Search to rank web pages in their search engine results. A Page Rank results from a mathematical algorithm based on the web graph, created by all World Wide Web pages as nodes and hyperlinks as edges.

According to Google, Page Rank works by counting the number and quality of links to a page to determine a rough estimate of how important the website is. The usefulness of a search engine depends on the relevance of the result set it gives back. There may of course be millions of web pages that include a particular word or phrase. However some of them will be more relevant, popular, or authoritative than others. A user may not have the ability or patience to scan through all pages that contain the given query words. One expects the relevant pages to be displayed within the top 10-20 pages returned by the search engine.

Page Rank algorithm is one of the most known and influential algorithms for

computing the relevance of web pages that is used by the Google search engine. The idea behind the Page Rank is that, the importance of any web page can be judged by looking at the pages that link to it. If we create a web page a and include a hyperlink to the web page b, this means that we consider b important and relevant for our topic. If there are a lot of pages that link to b, this means that the common belief is that page b is important. If on the other hand, b has only one backlink, but that comes from an authoritative site c, (like www.google.com, www.cnn.com, www.cornell.edu) we say that c transfers its authority to b, in other words, c asserts that b is important. Whether we talk about popularity or authority, we can iteratively assign a rank to each web page, based on the ranks of the pages that point to it.

Page Rank is a link analysis algorithm and it assigns a numerical weighting to each element of a hyperlinked set of documents, with the purpose of "measuring" its relative importance within the set. The algorithm may be applied to any collection of entities with reciprocal quotations and references. The numerical weight that it assigns to any given element E is referred to as the PageRank of E and denoted by PR(E).

The rank value indicates an importance of a particular page. A hyperlink to a page counts as a vote of support. The Page Rank of a page is defined recursively and depends on the number and Page Rank metric of all pages that link to it ("incoming links"). A page that is linked to by many pages with high Page Rank receives a high rank itself.

Page Rank works by counting the number and quality of links to a page and determine a rough evaluation of how important the website is. The underlying assumption is that more important websites are likely to receive more links from other websites.Numerous academic papers concerning Page Rank have been published. In practice, the Page Rank concept may be vulnerable to manipulation. But knowing about the advantages and disadvantages of Google page rank, one can try to avoid the possible errors that may reduce the rank value of an existing website or while creating new website.

5.2 SPECTRAL CLUSTREING

Spectral clustering is a technique with roots in graph theory.Here the approach is used to identify communities of nodes in a graph based on the edges connecting them. The method is flexible and allows us to cluster non graph data as well.Clustering algorithms have their application in many areas including detection, image segmentation, search result grouping, market segmentation and social network analysis. Clustering is one of the initial steps done in exploratory data analysis to visualize the similarity and to identify the pattern lying hidden in data points. The aim of clustering is to find the similarity within a cluster and the difference between two clusters. There are two major approaches in clustering. They are Compactness and Connectivity.

Spectral clustering helps us overcome two major problems in clustering. First being the shape of the cluster and the other is determining the cluster centroid. K-means algorithm generally assumes that the clusters are spherical or round i.e. within k-radius from the cluster centroid. In K means, many iterations are required to determine the cluster centroid. In spectral, the clusters do not follow a fixed shape or pattern. Points that are far away but connected belong to the same cluster and the points which are less distant from each other could belong to different clusters if they are not connected. This implies that the algorithm could be effective for data of different shapes and sizes.When compared with other algorithms, it is computationally fast for sparse data sets of several thousand data points. You don't need the actual data set to work with. Distance or though it might be costly to compute for large data sets as eigen values and eigen vectors need to be computed and then clustering is to be done. But the algorithms try to cut the cost.

In spectral clustering there are three major steps involved: Constructing a similarity graph, projecting data onto a lower-dimensional space, and clustering the data. Given a set of points S in a higher-dimensional space, it can be elaborated as follows:

1. Form a distance matrix
2. Transform the distance matrix into an affinity matrix A
3. Compute the degree matrix D and the Laplacian matrix L = D – A.

4. Find the eigenvalues and eigenvectors of L.

5. With the eigenvectors of k largest eigenvalues computed from the previous step form a matrix.

6. Normalize the vectors.

7. Cluster the data points in k-dimensional space.

Spectral clustering has its application in many areas which includes: image segmentation, educational data mining, entity resolution, speech separation, spectral clustering of protein sequences, text image segmentation. Though spectral clustering is a technique based on graph theory, the approach is used to identify communities of vertices in a graph based on the edges connecting them. This method is flexible and allows us to cluster non-graph data as well either with or without the original data.

5.3 PROTEIN STRUCTURE ANALYSIS

There are an enormous number of proteins present in our body cells. The three dimensional structure of proteins is the key to understanding their function and evolution. A protein is formed by the binding of amino acids by peptide bond to form a long chain-like molecules called poly-peptides. There are four major protein classes, shown below in Figure 5.1. The cylinders represent helix and the arrows represent the strands. A helix is a spiral molecule formed from the benzene rings.

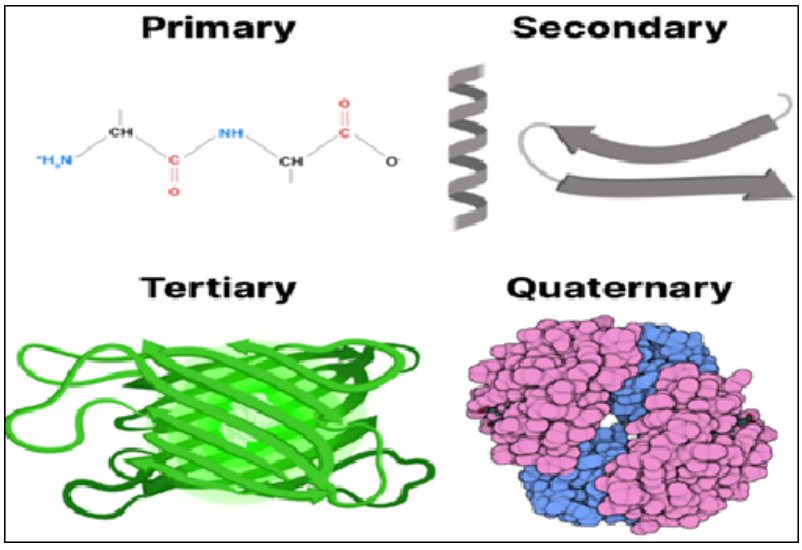


Figure 5.1: DIFFERENT STRUCTURES OF PROTEINS

The analysis of stable folded three-dimensional structures provide an insight into the protein structures for amino acid sequences and drug design studies. The composition of the protein backbone and side-chains makes the geometry of a protein structure. The Protein structure can have the same gross shape but have different geometric structure. Graphs help us to represent the topology of protein structures, regardless of their complexity. The main problem is to identify and define the vertices and edges. Fig 5.2 are a few simple examples of proteins with their graphs below them.

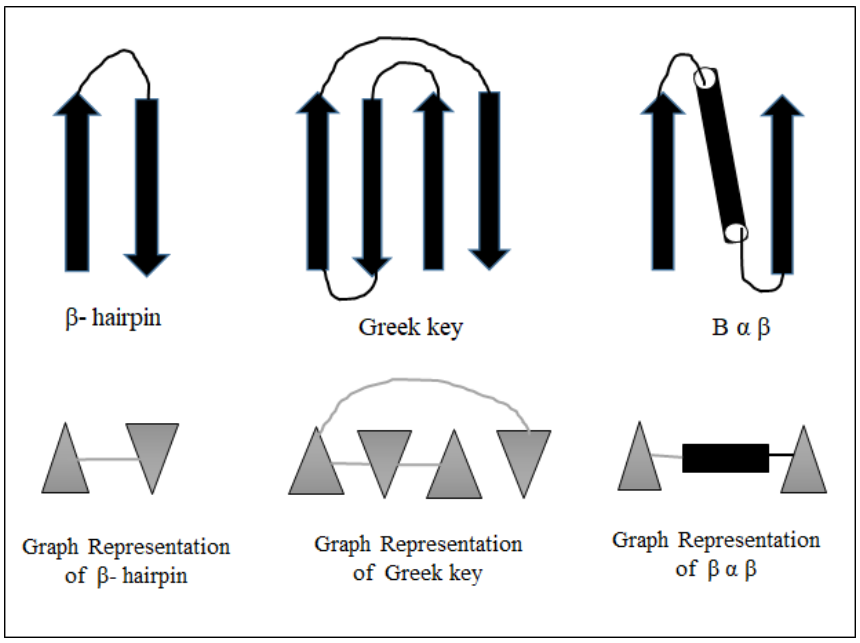


Figure 5.2: CONVERSION OF PROTEIN STRUCTURES TO GRAPH

Properties of graphs and their spectral graph give information about protein structure, depending on how the vertices and edges are defined. The basic unit of a protein is its amino acid sub units. To study cluster identification, the amino acids represent the vertices and the three-dimensional connectivity between them is represented by the edges. To study fold and pattern identification and the folding rules of proteins, α-helixes and β-strands are used for vertices and spatially closed structures are used for edges. To identify proteins with similar folds, the poly-peptide chain are the vertices and the side chains within a certain radius are the edges.

Mathematical graphs are used to represent β structures, which is much more advantageous than drawing three-dimensional figures. The vertices represent the

single β-strands and the two edge sets represent the sequential and hydrogen bond between them. Connected graphs are used to represent α-helical structures. The vertices represent secondary structures and the edges represent contacts between helices. The main reason for setting the structures up this way is to gain information about the folding process of protein structures and understand better organization and patterns within these structures. Since the comparison of protein structures is important for revealing the evolutionary relationship among proteins, predicting protein structures and function.

The protein connectivity is determined by identifying the main chain atoms within a prescribed distance. This comes from identifying clusters. Two protein graphs can be compared to check whether they have some common features, and thus provide insight into structural overlaps of proteins. One way of finding this is by tree searching algorithm, which is a series of matrix permutations. Through it, sub-graph isomorphism are detected to determine the largest sub-graph that is common in a pair of graphs. This can highlight the areas of structural overlap and therefore show structural and functional similarities not found through other methods. But this method requires a very high number of computations. Now a few heuristic methods have been discovered to reduce the time and cost of the computations. Structural biologists are finding promising applications of graph theory with optimism that this field of mathematics will continually contribute to the understanding of protein structure, folding stability, function and dynamics.

Many methods have been developed in the past to support two or multiple protein structures. In spite of the importance of this problem, many mathematical or statistical frameworks have rarely been pursued for general protein structure comparison. One of the major issue in this field is that with many different distances used to measure the similarity between protein structures, none of them are in proper distances when protein structures of different sequences are compared.

Using an elastic Riemannian metric on spaces of curves, geodesic distance, a proper distance on spaces of curves, can be calculated for any two protein structures. In this protein structures can be treated as random variables on the shape manifold, and means and covariance can be computed for populations of protein structures. It can be used to build Gaussian-type probability distributions of protein structures for use in hypothesis testing. The covariance of a population of

protein structures can be used to reveal the population-specific variations in it and it can be helpful in improving structure classification. With curves demonstrating protein structures, the matching is performed using elastic shape analysis of curves, which can effectively model the conformational variations and insertions/deletions.

The protein structure arrangement problem addresses the problem of measuring the degree of resemblance in three–dimensional structure of two proteins. The representation of each protein using a simple contact map permits the correspondence graph for the protein pair to be produced and the maximum clique within this graph provides a measure of the structural similarity among the two proteins. This study uses a recently developed method called maximum clique algorithm.

Having a protein structure helps us to provide a greater level of understanding of how a protein works, which can allow us to create hypotheses about how we can control it, or modify it. For example, knowing a protein's structure could allow you to design site-directed mutations with the intent of changing function that causes abnormality. And this abnormality may result in certain diseases like Thalassemia, Sickle cell anemia, Cystic Fibrosis etc.

For Example - *In the case of Thalassemia*

To diagnose Thalassemia, the usual procedure is likely taking a blood sample. They'll send this sample to a lab to be tested for anaemia and abnormal haemoglobin. A lab technician will also look at the blood under a microscope to see if the red blood cells are oddly shaped.

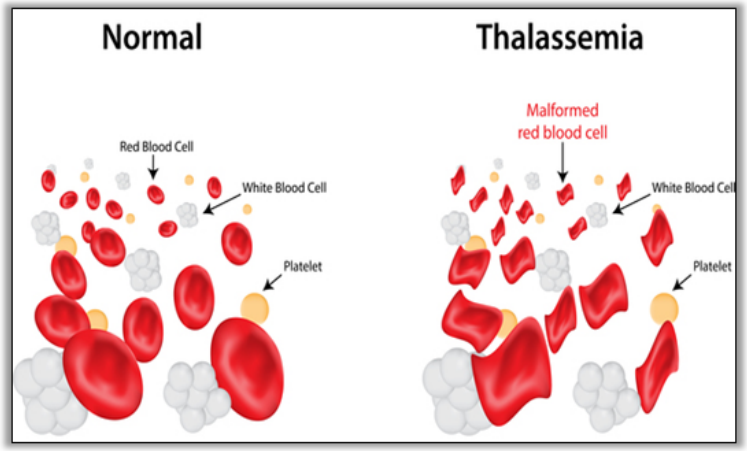


Figure 5.3: NORMAL BLOOD CELL AND THALSSEMIC BLOOD CELL

Abnormally shaped red blood cells are a sign of Thalassemia as shown in the Figure 5.3. The lab technician may also perform a test known as haemoglobin electrophoresis. This test separates out the different molecules in the red blood cells, allowing them to identify the abnormal type.

If we are observing the Figure 5.3, we can just identify that one is normal and other is abnormal. And we conclude that the person is Thalassemic. We won't be able to explain the defect in detail, since the structure is complex and vague.

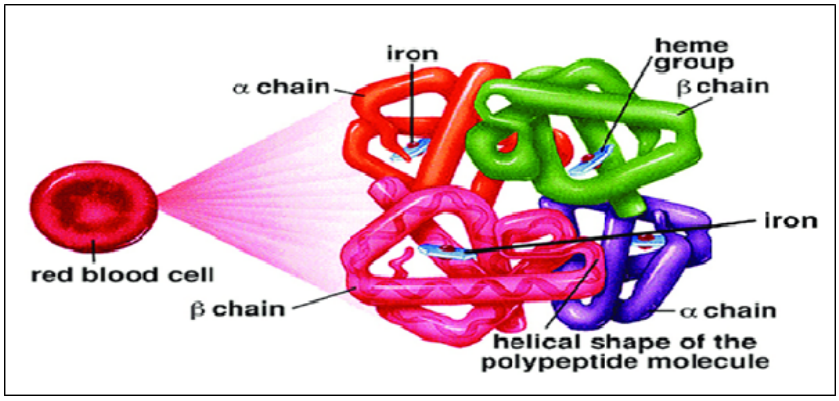


Figure 5.4: PROTEIN STRUCTURE OF HAEMOGLOBIN

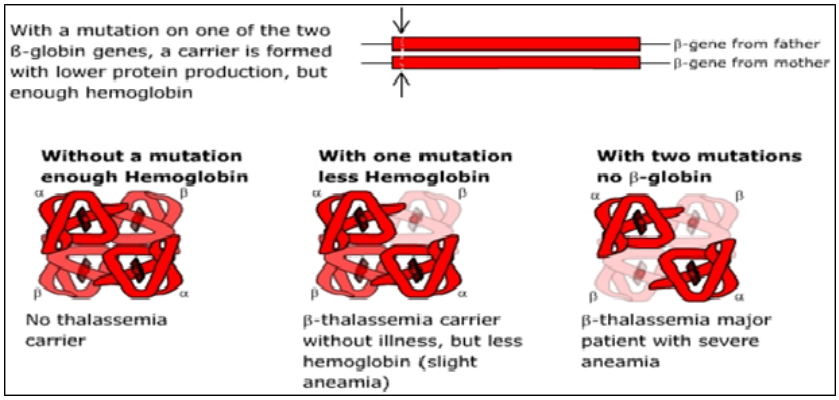


Figure 5.5: PROTEIN STRUCTURE OF ABNORMAL HAEMOGLOBIN

The Figure 5.4 is the protein structure of normal haemoglobin and Figure 5.5 is the protein structure of abnormal haemoglobin of Thalassemic patient. From the Figure 5.4, we can construct the graph of protein structure of normal haemoglobin as we discussed above in the protein structure (refer Figure 5.6).

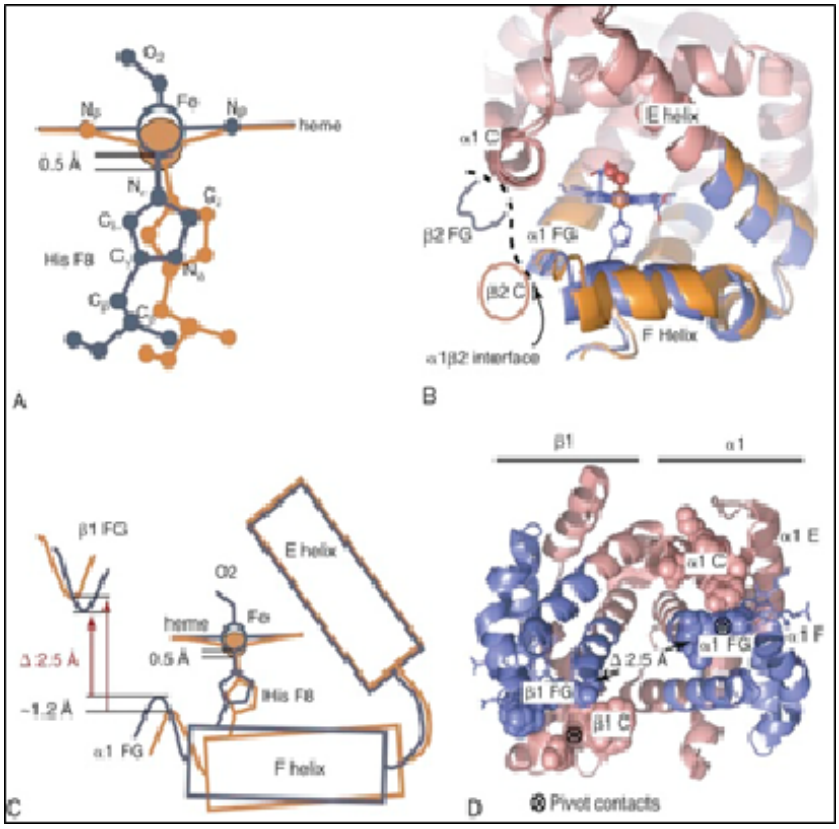


Figure 5.6: GRAPH OF NORMAL HAEMOGLOBIN

Similarly we can construct the graph of protein structure of abnormal haemoglobin of a Thalassemic patient. Then by comparing the graph structure of both haemoglobin we can identify and analyze the defectiveness in the structure. And we will be able to give a detailed explanation of the defective structure of the haemoglobin. The successful implementation of this approach could allow more sophisticated and novel clinical applications that may lead to a permanent cure for Thalassemia.

Chapter 6

CONCLUSION

Algebraic graph theory is concerned with the use of algebraic techniques in the study of graphs. Algebraic graph theory is a fascinating subject concerned with the interplay between algebra and graph theory. Algebraic tools can be used to give surprising and elegant proofs of graph theoretic facts and there are many interesting algebraic objects associated with graphs. Algebraic graph theory can be viewed as an extension to graph theory. The first part of algebraic graph theory involves the applications of linear algebra and matrix theory to the study of graphs. The adjacency matrix completely determines the graphs, and its spectral properties are shown to be related to the properties of the graph. For example, if a graph is regular then the eigenvalues of its adjacency matrix are bounded in absolute value by the valency of the graph.

The applications of linear algebra, graph theory, and the spectral of a graph remains forever through the various sciences and other fields. We know, mathematics is found everywhere, and is required in many situations in order to evaluate, process, and for the better understanding of the world around us.

Algebraic graph theory, particularly spectral graph theory, studies the algebraic connectivity via characteristic polynomial, eigenvalues, and eigenvectors of matrices associated with graphs, such as adjacency matrix or Laplacian matrix. Topological graph theory concerns the embedding and immersions of graphs, and the association of graphs with topological spaces, such as abstract simplicial complexes. Mathematically, graphs are useful tools in geometry and certain parts of

topology such as knot theory and algebraic topology.

Several discovery systems in graph theory have been very successful in helping mathematicians to formulate and explore speculations, or to suggest interesting speculations in an entirely automated way. Moreover, new systems sometimes based on new principles are being developed. The underlying paradigms, i.e., enumeration, interactive computing, formula manipulation, generation and selection, heuristic optimization are varied. They appear to be largely complementary. So one may expect much activity and the arrival of more comprehensive systems in the near future.

Finally, knowing about different application of algebraic graph theory one will be able to understand how simpler it is to use algebraic method in illustrating complicated problems. Application of algebraic method will help us to minimize the time consumption when compared with other methods. Also we will be able to give a clear cut result with concrete explanation for a particular problem.

REFERENCES

[1] Wikipedia

[2] Norman Biggs, *Algebraic graph theory,* 2^{nd} *Edition.* Cambridge University Press, Great Britain, 1993.

[3] Bojan Mohar, *The Laplacian Spectrum of Graphs*, Graph Theory, Combinatorics,and Applications, Vol 2, 1991.

[4] Beineke L.W and Wilson R. J (Ed.)(2005), *Topics in Algebraic Graph Theory.*

[5] Bapat R.B. , *Graphs and Matrices*, Springer International Edition, 2010.

[6] Michelle L. Rittenhouse, *Properties and Recent Applications in Spectral Graph Theory*, Virginia Commonwealth University, 2008.

[7] Cesar O. Aguilar, *An Introduction to Algebraic Graph Theory.* State University of New York at Geneseo, 2021.

[8] V. Saraswathi , K.V. Brinda, N. Kannan *Protein Structures: Insights from Graph Theory*, Journal of Theoretical and Computational Chemistry, Vol. 1, No.1(2002), World Scientific Publishing Company, 2002.